Understanding the regulation of acid resistance in *E. coli* using whole genome techniques

by

Matthew David Johnson

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
The University of Birmingham
September 2011

ABSTRACT

The ability of bacteria to thrive in a variety of host environments depends on their capacity to sense and respond to a wide array of stressors. *E. coli* encounters many stresses during transit through the gastro-intestinal tract, including acid stress. Acid stress response in *E. coli* is regulated by a complex network called AR2. The AR2 network comprises several local regulators that collate signals from multiple two-component systems (TCS) including RcsBD, EvgAS and PhoPQ.

We combined lab-based evolution and whole genome re-sequencing to generate and identify mutations that confer increased acid resistance in *E. coli* K-12. All of these mutations map in the gene encoding EvgS, the sensor kinase of the EvgAS TCS. Using a luciferase reporter system and phenotypic assays we characterised the nature of these *evgS* mutations and their contribution to acid resistance. We also used high-temporal resolution luciferase reporter assays to uncover novel aspects of this network and implicate PhoP in the repression of acid resistance. Finally, we used our *evgS* mutants to characterise novel interactions within the AR2 network between the two component systems RcsBD and EvgAS. These results are discussed in relation to the role of regulatory networks in bacteria.

Dedicated to the loving memory of my nephew
Harry J. Tree
2004 - 2010

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor Pete Lund for his excellent supervision during my PhD. Pete's support and advice for talks and the preparation of this thesis have been invaluable. I would also like to extend my gratitude to Chris Thomas, Steve Busby and Ian Henderson, for their advice and constructive criticism during this project. I would like to thank all those past and present who worked in T101 and T102 for making my experience in the lab thoroughly enjoyable, including, Neil and Olga, Tara, Tabish, Elsa, Andy, Martin, Rhiddi, Faye, Karina, Cathy, Tim and Doug. I acknowledge the Francesco Falciani and Robin May labs for their collaboration; in particular, I would like to thank Anna and Kerstin for all their help and patience.

I would like to make a special thank you to Rebecca, Denisse, Sean, James, Chris, Amanda H and Dan for their friendship, which has been so supportive during my studies. Of course, I would also like to thank my parents, Dave and Barbara, for their support both financially and emotionally, and for keeping my feet on the ground; and my sister, Charmaine, whose courage and perseverance during times of extreme adversity has been inspirational. Finally, I would like thank Amanda, to whom I owe so much for her support as a scientist, a friend, and as a companion.

TABLE OF CONTENTS

List of figures

List of figures (continued)

List of tables

List of Abbreviations

Abbreviation	Augmentation
ADI	Argenine dependent acid resistance mechanism
AFI	Acid fitness island
Ap	Ampicillin
AR	Acid resistance
bp	Base pair
CAD	Lysine dependent acid resistance mechanism
cAMP	Cyclic adenosine monophosphate
Cb	Carbenicillin
CRP	cAMP repressor protein
EAEC	Enteroaggregative *E. coli*
FRT	FLP recombination site
GABA	Gamma-amino butyric acid
GAD	Glutamate dependent acid resistance mechanism
HK	Histidine kinase
Km	Kanamycin
LB	Luria Bertani
Lux	Light emitting reaction requiring components from the *Photohabdus luminescens luxCDABE* operon
OD	Optical density
PAS	Per-Arnt-Sim
RBS	Ribosome binding site
RNAP	RNA polymerase
RR	Response regulator
SDW	Sterile distilled water
TCS	Two-component system
TF	Transcription factor

CHAPTER 1: Introduction

1.1 Introduction

Escherichia coli proliferates optimally in the mammalian gut. However, in order to reach its preferred niche *E. coli* must deal with a wide array of stresses. To cope with these stresses *E. coli* has many sophisticated stress response mechanisms. The diversion of resources to stress response mechanisms must be accurate in order for *E. coli* to remain competitive. Consequently, complex regulatory networks control these responses, which function to sense a particular stress and orchestrate an appropriate and measured response. We investigated the response of *E. coli* to acid stress. Our understanding of the regulatory networks that control acid stress response genes is discussed below.

1.2. *E. coli* acid resistance

E. coli is a member of the *Enterobacteriaceae* family that can colonise the gut of warm blooded mammals. Many *E. coli* strains are commensal and cause no harm to the host, in fact, *E. coli* supplies its host with vitamin K2 (Bentley *et al.*, 1982). However, *E. coli* strains that contain virulence factors can become dangerous pathogens capable of causing severe illness. Micro-organisms that inhabit the pH neutral environment of the mammalian gut must first pass through the acidic environment of the stomach. The stomach can have a pH as low as 1.5 during starvation (Smith, 2003). Despite the lack of any clear correlation between acid resistance and pathogenicity (Foster, 2004), acid resistance is a clinically important phenotype. This is because the greater an organism's acid resistance, the lower the infective dose of that organism. For example, *Shigella* species can cause dysentery in healthy adults with a dose of 10 to 500 cells (Dupont *et al.*, 1971), and between 25 and 100% of *Shigella* cells survive after 2 hours at pH 2.5 (Gorden and Small, 1993). In contrast, *Salmonella* species have an infective dose between 10^5 and 10^8 cells (Blaser and Newman, 1982), and less

than 0.001% of *Salmonella* cells survive after 2 hours at pH 2.5 (Gorden and Small, 1993). *E. coli* is also remarkably resistant to strong acid challenges and can survive pH 2.5 conditions for more than 2 hours (Foster, 2004). This level of acid resistance confers an infective dose ranging between 10 and 500 cells depending on the strain (Benjamin and Datta, 1995; Gorden and Small, 1993; Lin *et al.*, 1995). Once in the gut, *E. coli* must also cope with acid shock from volatile fatty acids produced by fermentation. *E. coli* is considered to be a neutrophile and yet can withstand acid shock levels equivalent to acidophiles. These attributes are unusual, therefore *E. coli* resistance mechanisms are an interesting subject to research.

1.2.1. Acid resistance mechanisms of *E. coli*

In 1995, Lin *et al.* were the first to show evidence that *E. coli* could utilise amino acids to survive extreme acid stress (< pH 2.5). Bacteria grown to stationary phase, in pH 5 medium containing glucose, were very sensitive to acid shock in pH 2.5 minimal medium (Lin *et al.*, 1995). However, Lin *et al.* also observed that *E. coli* could survive significantly better when the acid shock medium was supplemented with the amino acids glutamate or arginine. Additionally, work by Lin *et al.* also showed that *E. coli* survived better when challenged in minimal medium without amino acids. This resistance was observed when cultures were grown in rich medium that was mildly acidic and lacking glucose. From these observations three acid resistance (AR) systems were defined: AR1 or oxidative acid resistance, the glutamate dependent (GAD) and arginine dependent (ADI) respectively (Lin *et al.*, 1995). A fourth amino acid dependent system has also been characterised, which is dependent on lysine (CAD). However, CAD was found to be quite ineffective in enabling *E. coli* to survive extended periods of acid stress (Iyer *et al.*, 2003). More recent work on CAD has indicated

that it has a more essential role helping the cell survive endogenous stress caused by phosphate starvation (Moreau, 2007).

In extremely acidic conditions, the protection conferred by GAD was significantly higher than the protection conferred by other AR mechanisms (Lin *et al.*, 1996). As a consequence of this superior protective capability, GAD is the most studied of the acid resistance systems (Lin *et al.*, 1996). In addition to the amino acid dependent systems, which have been characterised in detail, genes located on an Acid fitness island (AFI) have also been implicated in acid resistance (Masuda and Church, 2003; Mates *et al.*, 2007). Transcriptomic analysis of cells over-expressing local acid resistance regulators GadE and GadX implicated 13 genes coding 12 proteins in an acid fitness island (Hommais *et al.*, 2004). Included in this island were the GAD genes *gadE/A/X/W/Y* and *hdeA/B/D* known to be involved in acid resistance (Gajiwala and Burley, 2000; Ma *et al.*, 2003a; Masuda and Church, 2003). Later, mutagenesis of other genes found in the AFI would confirm their role in acid resistance (Mates *et al.*, 2007). As some of the core GAD genes are located on the AFI, the terminology for these systems is at times confusing. For the purpose of this review, GAD will refer to the core GAD genes only (*gadA*, *gadB* and *gadC*). All other genes that contribute to acid resistance and are located on the AFI will be termed AFI genes.

All of the aforementioned mechanisms and genes are important for acid resistance in different conditions. Understanding how these mechanisms work, and in what conditions they are effective, is important for the understanding of the acid resistance phenotype as a whole.

1.2.2. Components of amino acid dependent acid resistance mechanisms

The most robust and best characterised mechanisms that *E. coli* uses for acid resistance are those dependent on amino acids. In 1942 Epps and Gale were the first to suggest the relationship between biodegradative amino acid decarboxylase activity and survival at low pH. As the reaction resulted in the net consumption of a proton, and had an acidic optimum pH, they hypothesised that these enzymes were involved in pH homeostasis. However, it would be nearly half a century before a definitive mechanism for this process was outlined. The discussion below summarises the characterisation of the components of amino acid dependent acid resistance mechanisms.

1.2.2.1. Characterisation of the acid dependent amino acid decarboxylases

The decarboxylase components of amino acid resistance mechanisms have been extensively characterised. The main role of these decarboxylase enzymes is to catalyse the reaction of their cognate amino acid into a reduced molecule. Without decarboxylases, the amino acid dependent systems are ineffective (Castanie-Cornet *et al.*, 1999). The GAD system utilises two decarboxylases, GadA and GadB, ADI has one decarboxylase, AdiA and CAD has two decarboxylases, CadA and Ldc. This section will discuss the characterisation of these proteins.

Following its discovery by Epps and Gale in 1942, the biochemical characterisation of the Adi decarboxylase (later to be known as AdiA) was done by Blethen *et al.* in 1968. The optimum working pH, characterised using purified, acid induced, arginine decarboxylase, was 5.2. The main substrate of AdiA was determined to be L-arginine (Blethen *et al.*, 1968). Meanwhile, the lysine decaboxylase gene, named *cadA*, was characterised by Tabor *et al.* in

1980. The *cadA* locus was mapped and mutated to show that a *cadA* strain was deficient in the production of cadaverine (Tabor *et al.*, 1980).

The arginine and lysine decarboxylases are both regulated by acid at a transcriptional level. When the promoters of *cadA* and *adiA* were fused to a *lac* reporter system, β-galactosidase activity was increased in acidic conditions (Auger *et al.*, 1989). Detailed analysis of the *cadA* locus revealed that *cadA* was transcribed in a dicistronic operon with an upstream gene *cadB*. The *cadBA* promoter was found to be controlled by a third component, *cadC*. In addition, the C-terminus of *cadC* was found to have a domain similar to other characterised environmental sensing domains (Watson *et al.*, 1992).

Despite extensive functional characterisation of the glutamate decarbxylase from *E. coli* by Gale *et al.* in 1946, the genes that code for these proteins were not characterised until 1992. Two glutamate decarboxylase genes, named *gadA* and *gadB*, were characterised at two different loci in the *E. coli* genome. The sequences of these genes were found to be 98% identical. Later, functional studies revealed that both decarboxylases had identical reaction rates and optimum pH (De Biase *et al.*, 1996; Small and Waterman, 1998).

1.2.2.2. Activation of decarboxylase enzymes by low pH

The structures of both arginine and glutamate decarboxylase enzymes have been solved by X-ray crystallography. Both enzymes were found to be inactive in neutral pH conditions and activated by acidic pH conditions (Andrell *et al.*, 2009; Capitani *et al.*, 2003). However, mechanisms of activation for both proteins are quite different.

Glutamate decarboxylase forms a hexamer made from three dimers that join to form a hexamer (Capitani *et al.*, 2003). Comparison of the pH neutral structure to the structure found in acidic conditions revealed that although a hexamer was formed, the protein was in an inactive conformation. Transformation to an active state involves minor changes in the protein conformation (Capitani *et al.*, 2003).

The arginine decarboxylase forms a similar structure to that of the glutamate decarboxylase, except with five dimers constituting an active decamer (Andrell *et al.*, 2009). However, unlike the glutamate decarboxylase, it's the assembly of the decamer that is subject to pH. During neutral pH conditions, only the dimer form of AdiA is found. Upon activation by low pH a decamer/active form is more abundant (Andrell *et al.*, 2009).

The involvement of these decarboxylases in acid resistance is supported at a structural level. Both enzymes show altered conformations from inactive to active forms in response to low pH and they are both involved in a reducing reaction. How the action of these enzymes actually aids de-acidification of the cytoplasm is not fully understood. However, it is understood that their role in acid resistance is completely dependent on the membrane bound antiporters. These antiporters, and the proposed amino acid dependant decarboxylase/antiporter mechanism, are described below.

1.2.2.3. Characterisation of the membrane bound antiporters and the amino acid dependant acid resistance mechanism

In 1992 Meng and Bennet characterised *cadB* as a gene that coded for a putative membrane bound transport protein. This was based on the homology of *cadB* with ArcD, which is an

arginine/ornithine antiporter found in *P. aeruginosa*. The expression of *cadBA* had already been linked to low pH (Meng and Bennett, 1992b; Watson *et al.*, 1992). Thus, the coupling of *cadAB* transcription and response to pH, suggested a joint role of CadA and CadB in pH homeostasis. Based on this logic, Meng and Bennett were the first to describe a mechanism for neutralising low extracellular pH using the combined functions of a lysine decarboxylase (CadA) and lysine/cadaverine antiporter (CadB) (Meng and Bennett, 1992b; Meng and Bennett, 1992a). In 1995, research by Lin *et al.*, which linked survival in extreme acid conditions with the presence of amino acids, suggested that *E. coli* had acid resistance mechanisms that relied on amino acids. This concept was similar to that of Meng and Bennet's CadA/CadB mechanism. However, it appears that the findings and suggested mechanism by Meng and Bennet 3 years earlier went unnoticed; perhaps because lysine dependent acid protection was low compared to the protection provided by both arginine and glutamate dependent systems (Iyer *et al.*, 2003; Lin *et al.*, 1995).

The first mechanism where a link was demonstrated between an amino acid dependent acid resistance phenotype and an acid resistance system was the GAD mechanism. Characterised by Castanie-Cornet in 1999, the role of two glutamate decarboxylases (GadA and GadB) in glutamate dependent acid resistance was elucidated using mutagenesis. Strains containing *gadA* and *gadB* deletions were extremely sensitive to acid shock at pH 2.5 (Castanie-Cornet *et al.*, 1999). The work by Castanie-Cornet *et al.*, on *gadA, gadB* and the GAD mechanism, built on work by Hersh *et al.* in 1996 that characterised the glutamate/GABA anti-porter. The membrane bound antiporter was discovered by transposon mutagenesis. Clones that exhibited acid sensitivity had a transposon insertion into a region between *ompC* and *gadB* (Hersh *et al.*, 1996). Further analysis of this locus revealed an open reading frame, which was named

xasA (extreme acid sensitive A). Comparative genomics indicated that the predicted transcript encoded a homologue of *gadC*. The *gadC* gene was previously characterised as an acid inducible gene in *Shigella*, which functioned as a glutamate and GABA anti-porter (Waterman and Small, 1996). Thus, the *xasA* gene was renamed *gadC*. A role in glutamate mediumted acid resistance was suggested due to the decreased resistance of a *gadC* mutant to acid shock in minimal medium specifically containing glutamate (Hersh *et al.*, 1996). The *gadA*, *gadB* and *gadC* genes (each encoding a protein of the same name), form the core set of genes required for the function of GAD.

Despite the early hypothesis by Meng and Bennett in 1992, regarding the action of the *cadBA* genes and numerous studies linking *cadBA* to acid resistance (Neely *et al.*, 1994; Park *et al.*, 1996; Watson *et al.*, 1992), it would be 11 years before phenotypic evidence for the role of the lysine dependent system in *E. coli* was obtained. The role of the CAD system in acid resistance was finally confirmed by Iyer *et al.* in 2003. Acid resistance assays isolated a lysine dependent AR phenotype. The level of protection conferred by the lysine dependent system was significantly lower than that of the arginine and glutamate systems (Iyer *et al.*, 2003). Mutagenesis of the *cadBA* operon proved beyond doubt that the protection observed was due to the action of these genes (Iyer *et al.*, 2003).

The ADI mechanism was the last to be fully characterised. The *adiA* gene was characterised in 1993 by Stim and Bennett, who noted that the organisation at the *adiA* locus was different from that of the *cad* locus as there was no antiporter located with *adiA*. This meant that despite *adiA* being induced by low pH its transcriptional unit was not coupled with a *cadB* homologue (antiporter component) (Stim and Bennett, 1993). The antiporter component of

the arginine dependent acid resistance mechanism, suggested by Lin *et al.* in 1995, was not characterised until 2003 by Gong *et al*. The gene *adiC* was characterised by its identity with *gadC*. Previously known as *yjdE*, *adiC* had 22% identity with *gadC.* Additionally, AdiC mutants were found to be deficient in arginine dependant acid resistance (Gong *et al.*, 2003).

The components of three amino acid dependent mechanisms, AR2, AR3 and AR4 were all characterised by 2003. The AR4 mechanism consisted of two genes, *cadB and cadA*, transcribed from a single promoter. The AR2 mechanism comprises three genes *gadA*, *gadB* and *gadC* coded in two separate operons, and the AR3 mechanism comprises two genes, *adiA* and *adiC*, found at two separate loci, controlled by different promoters. How these mechanisms actually operate to reduce the effect of acid shock is not well understood. Details of the mechanisms of these systems are discussed below.

1.2.3. Mechanism of amino acid dependent acid resistance systems

The actions of components of the amino acid dependent AR mechanisms are seemingly quite simple. Figure 1.1 shows a diagrammatic representation of these mechanisms and their predicted function. Briefly, extreme acidic conditions (< pH 2.5) cause the acidification of the cytoplasm, stimulating the expression of lysine, arginine and glutamate decarboxylases.

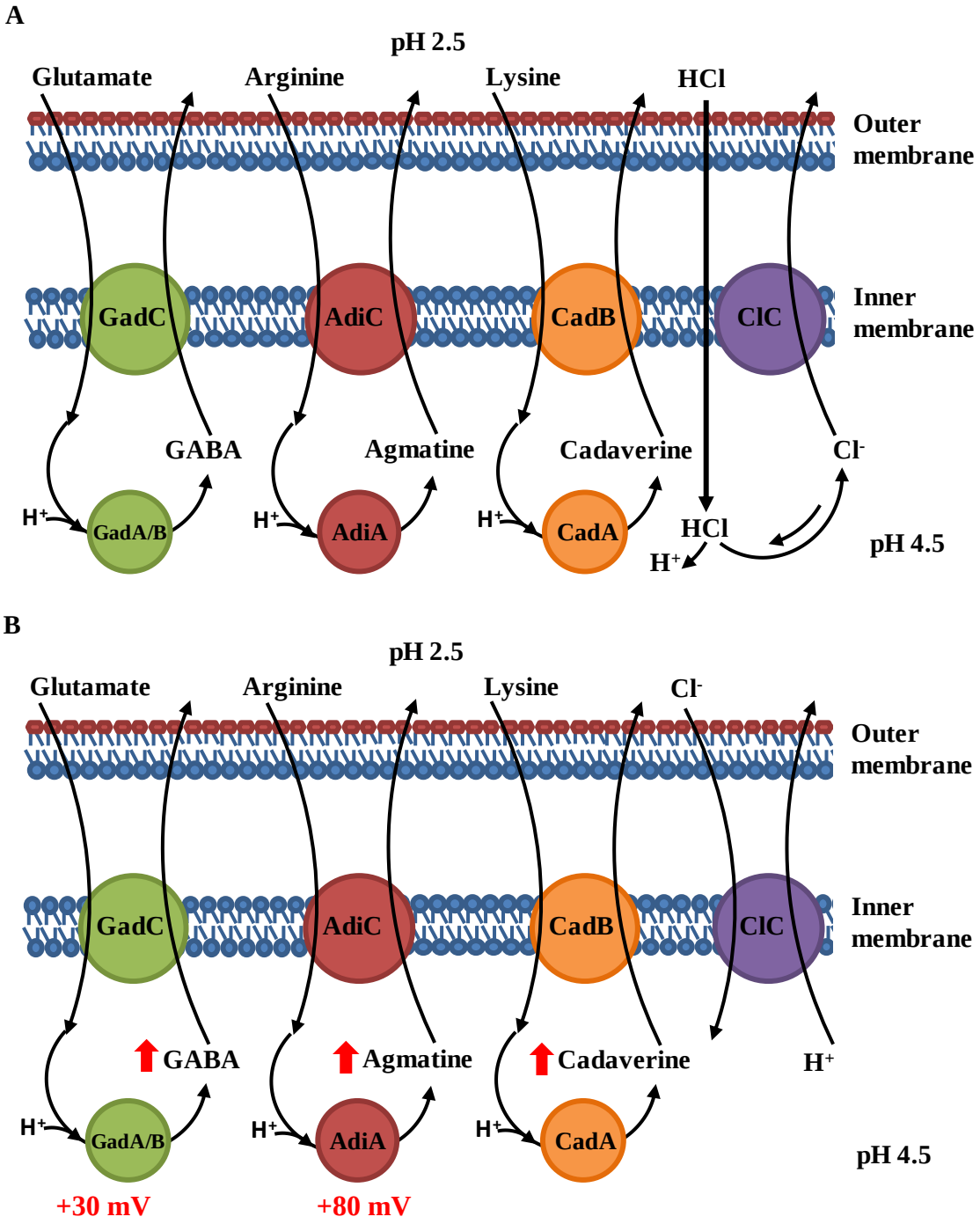


Figure 1.1. The amino acid-dependent acid resistance mechanisms. (A) The deacidification mechanism by the combined action of decarboxylase and antiporter enzymes pumping reduced products out to the cytoplasm. The antiporters, shown in the inner membrane, import their respective amino acids. Once in the cytoplasm the cognate decarboxylase catalyses the reaction to produce a reduced product, which is then exported from the cytoplasm.(B) Mechanism of protection by a membrane potential flip. The reduced products of the decarboxylase reaction accumulate causing positive charge to build in the cytoplasm. As a result the membrane potential would repel the extermal protons (adapted from Foster 2004).

The respective antiporters import glutamate, lysine and arginine from the extracellular environment. Cytoplasmic arginine, glutamate and lysine are converted into agmatine, GABA (gamma-amino butyric acid) and cadaverine by their respective decarboxylases. Protons are consumed as part of this process. The products of the decarboxylation reactions are then exported across the inner membrane by the antiporters in exchange for more amino acids (see figure 1.1 A) (Richard and Foster, 2004). This mechanism described poses two main quandaries. Firstly, it is not clear how effectively the AR systems can function in this way. In the case of GAD, glutamate has a functional side chain with a pKa of 4.1. Thus, during acid challenge at pH 2.5, this side chain will be greater than 50% protonated. When it enters the less acidic environment of the cytoplasm, through the GadC antiporter, a substantial portion of these protons will dissociate causing a reduction in cytoplasmic pH. Thus, before GadA/B activity incorporates protons into the GABA product, the glutamate has already acidified the cytoplasm. Secondly, during extreme acid challenge at pH 2.5 the pH in the cytoplasm drops to 3.6 (measured without any induced acid resistance mechanisms) (Richard and Foster, 2004), the acidification of the cytoplasm lowers the pH within the optimum range of the glutamate decarboxylase (3.6 to 4.6) (Capitani *et al.*, 2003). However, the optimum pH for the arginine and lysine decarboxylases is 5.25 and 5.5 respectively (Andrell *et al.*, 2009; Blethen *et al.*, 1968). Therefore, this environment could render the arginine and lysine decarboxylases ineffective. Additionally, the effectiveness of the glutamate and arginine dependent systems, when exposed to pH 2.5, is unexpectedly low. When the arginine and glutamate systems are active, they raise the cytoplasmic pH from pH 3.6 to only 4.7 and 4.6 respectively (Richard and Foster, 2004). As a potential resolution of this paradox, it has been suggested that *E. coli* is attempting to repel extracellular protons by reversing the membrane potential from a negative inside, relative to outside, to a positive inside relative to outside (see figure 1.1 B)

(Richard and Foster, 2004). The positive polarity could repel or slow the flow of protons into the cell. The positive charge inside the cell could be created by an increase in the pool of protonated products of decarboxylation reactions. This is a tactic is employed by acidophiles and could explain why the pH of the cytoplasm remains below 5 even when glutamate and/or arginine are present in the medium (Richard and Foster, 2004).

1.2.4. Amino acid independent acid resistance mechanisms – AR1

Three acid resistance mechanisms were originally characterised, two amino acid dependent systems (GAD and ADI) and an amino acid independent system (AR1) (Lin *et al.*, 1995). AR1, or the oxidative acid resistance system, is induced by growth in minimal medium without glucose at pH 5.5. After growth in these conditions *E. coli* is capable of surviving extreme acid shock (< 2.5) without exogenous amino acids (Lin *et al.*, 1995). The mechanism of AR1 is poorly understood. Only a Fo/F1 proton-translocating ATPase has been shown to be essential to its function (Richard and Foster, 2003). However, it has not been confirmed whether the aforementioned ATPase is in fact part of the mechanism or simply supplying the components of AR1 with ATP (Foster, 2004). The components of the AR1 system are regulated by CRP and RpoS. However, details of the relationship between CRP, RpoS and AR1 are currently unknown (Lin *et al.*, 1995; Richard and Foster, 2004).

1.2.5. Amino acid independent acid resistance mechanisms – The AFI

Transcriptomic analysis of strains over-expressing the regulator protein GadE (YhiE) implicated an additional locus of genes in acid resistance (Hommais *et al.*, 2004). Functional analysis of the genes found at this locus revealed that they were involved in acid resistance. The locus became known as the acid fitness island or AFI. This section discusses the

characterisation of the AFI, and the proposed roles of genes found at this locus in acid resistance.

The AFI is found in all sequenced strains of *E. coli* and *Shigella*, but is not found in *Salmonella* spp. shown in figure 1.2. The locus contains 13 ORFs coding for 12 proteins and 1 sRNA. Genes located in the AFI include: *slp*, *dctE*, *mtdE*, *mtdF*, *hdeA*, *hdeB*, *hdeD* and *yhiD*. These genes are organised into the operons shown in figure 1.2. Operon 1 contains *slp* and *dctR* (Tucker *et al.,* 2003). Slp, or starvation lipoprotein, is located in the outer membrane. Slp was first reported to be up-regulated in stationary phase, independently from *rpoS,* when cells were deprived of carbon (Alexander and St John, 1994). Additionally, analysis of the *slp* promoter suggests that it is bound by sigma-70 (Alexander and St John, 1994). Subsequent work showed that the *slp* promoter could be bound by RpoS, as expression of the *slp* gene in different conditions was reduced in an *rpoS* knockout (Shimada *et al.*, 2004). Currently, *slp* has a proposed role in acid resistance as its expression is induced by growth at pH 5.5. Mutagenesis of *slp* reduced the survival of cells shocked in spent medium at pH 2.5 compared to the parental strain. After 2 hours at pH 2.5 in spent medium *slp* mutants survived 10,000 fold less than the parental strain. However, the role of *slp* in acid resistance is not fully understood (Mates *et al.*, 2007).

AFI genes, *hdeA/B* and *yhiD*, are coded on operon 2, shown in figure 1.2. The product of *yhiD* is not characterised, but based on sequence homology it is classed as an ATP dependent transporter. Knock-out mutagenesis of *yhiD* resulted in reduced resistance to extreme acid stress at pH 2.1 (Mates *et al.*, 2007).

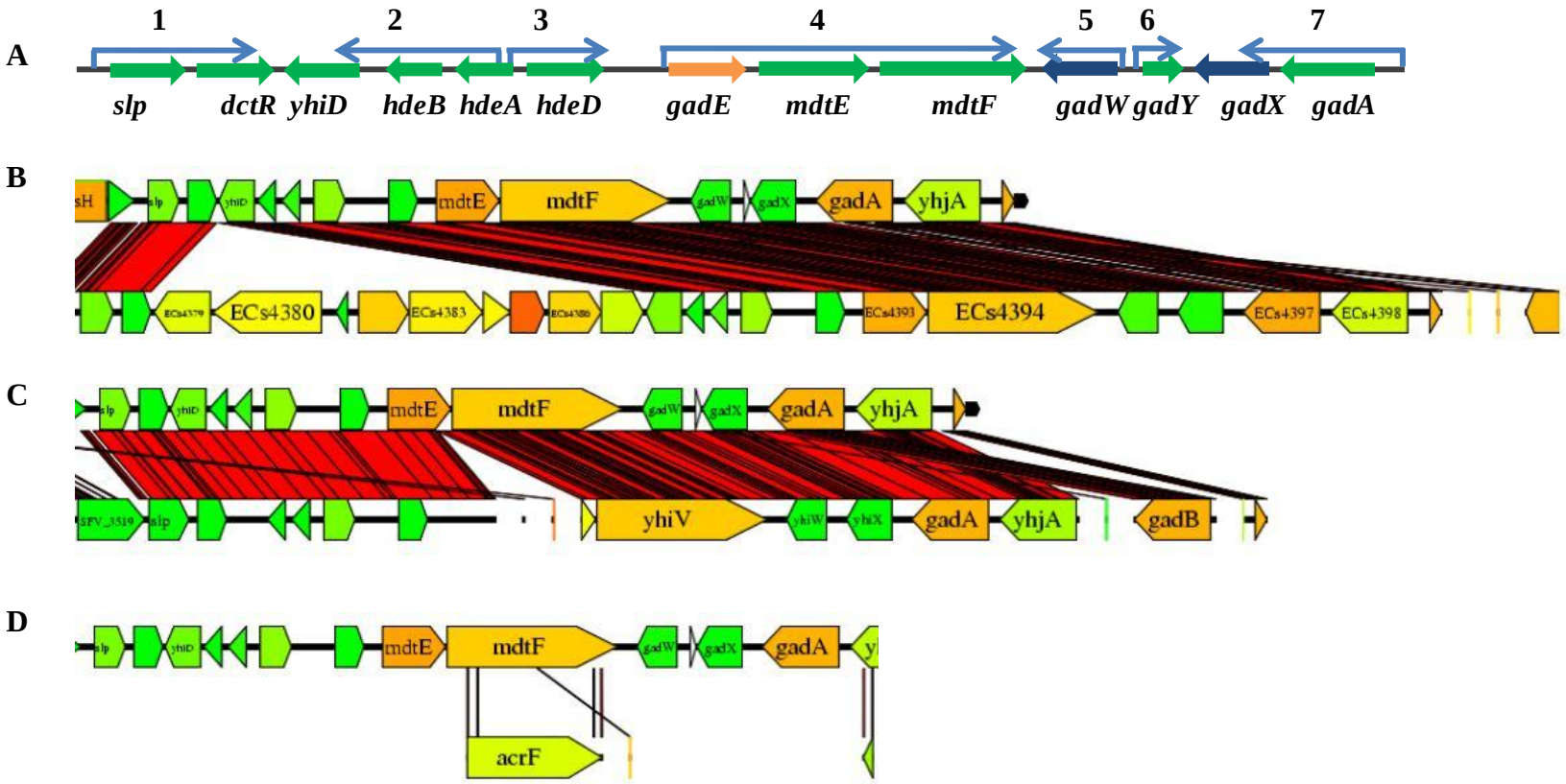


Figure 1.2. Comparison of the AFI locus from *E. coli* K-12 with Sakai, *Shigella* and *Salmonella*. Characterised by Hommais *et al* in 2004, the AFI contains many genes responsible for acid resistance. Figure A represents the AFI and the 7 main promoters/operons found in the AFI (not to scale). Alignments, based on nucleotide homology, of the AFI from K-12 MG1655 (Top) with O157:H7 Sakai (Bottom) (B), *Shigella flexinera* (Bottom) (C) and *Salmonella* (Bottom) (D), show that the AFI is present in all sequenced strains of *E. coli* and *Shigella* but not in any strains of *Salmonella*. xbase.ac.uk.

HdeA and HdeB were predicted to have a similar structure to each other and function as periplasmic chaperones, possibly acting as heterodimers (Gajiwala and Burley, 2000). HdeA monomers were found to dissociate *in vitro* in response to acid. HdeA monomers have been shown to protect periplasmic proteins from acid stress by binding proteins and stopping denaturation and aggregation at low pH. As the pH rises back to neutral, HdeA dissociates from its substrates in a gradual manner ensuring that the proteins do not aggregate (Gajiwala and Burley, 2000; Tapley *et al.*, 2010; Tapley *et al.*, 2009). The role of HdeB is very similar to that of HdeA, except HdeB has a pH optimum of 3 and is therefore ineffective in more acidic conditions (Kern *et al.*, 2007).

The *hdeD* gene product is an inner membrane protein, which is important for acid resistance in cells over-expressing *ydeO* (Daley *et al.*, 2005; Masuda and Church, 2003). The *hdeD* gene is indicated in figure 1.2 in operon 3, no other ORFs are controlled by this promoter.

Operon 4 contains genes *gadE*, *mdtE* and *mdtF*. The role of GadE as a central regulator of the GAD system will be discussed in more detail in section 1.5. The gene products of *mdtE/F* form trimeric inner membrane transport proteins, which form a complex with TolC to create a multidrug efflux pump (Nishino and Yamaguchi, 2002). So far, the *mdtEF-tolC* pump has not been shown to be required for acid resistance. In fact, the action of such a pump during acid shock would only acidify the cytoplasm, as antimicrobials are pumped out in exchange for protons.

Transcriptional units 5, 6, and 7 make up a complex regulatory sub-network that includes *gadW*, *gadY* and *gadX*. The roles of each of these regulators in the control of the GAD system

will be discussed in more detail later in the review. Briefly, GadX and GadW are AraC/XylS-like transcriptional regulators that regulate the AFI promoters including their own (Gallegos *et al.*, 1993; Ma *et al.*, 2002). Regulation of the AFI promoters by *gadWXY* occurs by relieving repression of this region by H-NS (Tucker *et al.*, 2002). The product of *gadY* is a small regulatory RNA which interacts with the *gadXW* transcripts (Opdyke *et al.*, 2004; Sayed *et al.*, 2007; Tramonti *et al.*, 2006).

The genes of the AFI have been shown to be important for acid resistance. However, the exact mechanisms by which these genes contribute to acid resistance are not fully understood. As these genes have been shown to partake in a similar process (i.e. regulated in kind by *gadE*), have a similar G+C content, and are repressed by global regulator H-NS, they were called an acid fitness island.

1.2.6. Amino acid independent acid resistance mechanisms – YdeP

EvgAS is a TCS that has been shown to activate acid resistance in response to a mild acid shock. Details of EvgAS and its role in acid resistance will be explained in section 1.4.3. The regulon of EvgAS includes two operons that are important for acid resistance. These are the *safA-ydeO*, which will be described in detail in section 1.4.3, and *ydeP*. The role of YdeP in acid resistance was first demonstrated when resistance conferred by over-expression of EvgA was shown to be YdeP-dependent (Masuda and Church, 2002). In fact, survival of the EvgA overexpressing strain was more dependent on YdeP than GadE (known as YhiE in this study) (Masuda and Church, 2002). The ability of YdeP to confer acid resistance has also been shown by over-expression (Masuda and Church, 2003). In summary, these results show that YdeP has the ability to confer high levels of acid resistance. However, the function of YdeP is

currently unknown. Based on sequence similarity the putative function is as an oxidoreductase. YdeP is the only known structural acid resistance gene to be directly activated by EvgA. The role of YdeP in acid resistance is not understood.

1.3. Control and induction of acid resistance mechanisms

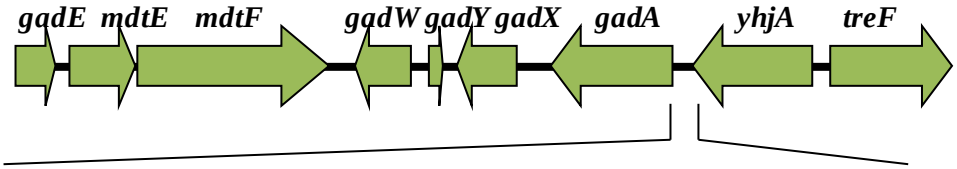
1.3.1. Induction of acid resistance mechanisms

Acid resistance in *E. coli* is induced by two main conditions, entry into stationary phase and mild acid shock (pH 4.8 – 5.8) (Castanie-Cornet *et al.*, 1999; Lin *et al.*, 1995). Regulating the acid resistance genes in accordance to these conditions is a complex regulatory network that will be discussed below. The main acid resistance network, which for the purpose of this review will be called the AR2 network, was initially characterised as a network that controlled GAD (Ma *et al.*, 2004). However, it is now understood that the AR2 network controls other acid resistance mechanisms, such as AR4 and the AFI. In addition to the AR2 network, acid resistance is also regulated by global transcription factors such as H-NS and CRP. Due to the complexity of acid resistance regulation, this review will focus on the local regulation of individual AR mechanisms before relating these systems with the AR2 network components, the alternative sigma factor RpoS, and the global regulators.

1.3.2. Promoter organisation of the GAD genes

The promoters of GAD structural genes are well characterised, figure 1.3 shows the organisation of the GAD genes and their transcriptional units. The AR2 genes are split into two operons, one coding for *gadB* and *gadC*, the other coding *gadA* and *gadX*.

A



yhjA Stop

taaacaataattaatttgatcgcccgaacagcaatgtttgggcgatttttattacgataa

-159 *GadXW binding* *-116*

taaagtctgtttttaatattatcatgttaaatgtttatattataaaaagtcgtttttctg

-99 GAD box *-22 RcsB box*

cttaggattttgttatttaaattaagcctgtaatgccttgcttccattgcggataaatcc

ttgaca t

-35

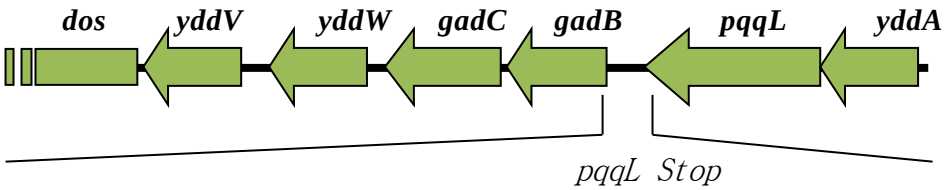
+1(-27) *gadA start*

actttttattgccttcaaataaatttaaggagttcgaaatggaccagaagctgttaac

ataat

-10

B



pqqL Stop

ccggtaaatacttataccggagtattattgccaaaataataacagccccgtcaacacatc

gttggcggggattttagcaatattcgctatttttatgtaataatttttataaatgcgttca

-269 *gadXW binding* *-226*

aaataataatcaagtactaatagtgatattttaaggtctgatttttacgtgataattcag

gagacacagaatgcgcataaaaataacagcataaaacaccttaccaccacccaagaattt

catattgtattgttttcaatgaaaaaatattattcgcgtaatatctcacgataaataac

-99 *GAD box* *-22 RcsB box*

attaggattttgttatttaaacacgagtcctttgcacttgcttactttatcgataaatcc

ttgaca

-35

+1 *gadB Start*

tactttttaatgcgatccaatcattttaaggagtttaaaatggataagaagcaagtaac

tataat

-10

Figure 1.3. Promoter regions of the *gadA* and *gadB* genes. The locus surrounding the *gadA* (A) and *gadB* (B) genes is shown above each promoter sequence, fat arrows, which are not to scale, indicate coding regions. In the sequence part of the figures, the intergenic region upstream of the *gadA* and *gadB* genes is shown in black font. Numbers above sequences refer to distances to translational start site. The mark, +1, indicates the site of the transcript start and is located in the same relative position for each promoter. The core promoter motifs are underlined, consensus sequences are written below. The conserved GAD box site is marked as well as the different positions of the 42bp GadXW binsing region and the RcsB box. The data presented in this figure is based on work by: Giangrossi *et al.*, 2005, Typas *et al.*, 2007, De Biase *et al.*, 1999, Castanie-Cornet *et al* 2010.

Studies monitoring the expression of the GAD promoters indicated that regulation occurred at the level of transcription initiation in the following conditions: un-buffered pH 5.5 minimal medium, buffered pH 5.5 complex medium, and during stationary phase growth. (Castanie-Cornet and Foster, 2001; Giangrossi *et al.*, 2005; Tramonti *et al.*, 2006). Analysis of the *gadAX* and *gadBC* promoters by primer extension detected one transcription start site for each gene (De Biase *et al.*, 1999). The *gadAX* and *gadBC* -35 and -10 hexamers (TTGCTT- 17n -TACTTT) are show only 50% identity to the vegetative sigma factor RpoD binding consensus sequence. Additionally, the core promoter has the structure of one which could be bound by the alternative sigma factor RpoS (Castanie-Cornet and Foster, 2001; De Biase *et al.*, 1999). The promoters of *gadAX* and *gadBC* both have a conserved 20bp AT rich sequence known as the GAD box. The location of the GAD box, shown in figure 1.3, is between -52 and -72 base pairs upstream of the transcriptional start site in both promoters (Castanie-Cornet and Foster, 2001). The GAD box is a binding site for GadE, the local regulator of the GAD genes (Ma *et al.*, 2003a). Recently the binding of GadE to the GAD box has been shown to be dependent on RcsB (for a detailed analysis of RcsB in the AR2 network see 1.4.3.4) (Castanie-Cornet *et al.*, 2010; Krin *et al.*, 2010a). RcsB also makes an additional interaction with the *gadA* promoter. When activated, RcsB binds an RcsB box centred at -18.5, and inhibits *gadA* promoter activity (Castanie-Cornet *et al.*, 2010). In addition to the degenerate core promoter, RcsB box and GAD box, each promoter has a -42bp consensus site that is bound by the GadX GadW activator/repressor proteins (Ma *et al.*, 2002; Tramonti *et al.*, 2008; Tramonti *et al.*, 2006).

The GAD genes are under the control of many local regulators that bind well-characterised operator sequences. Due to extensive analysis of these operons, a complex network of

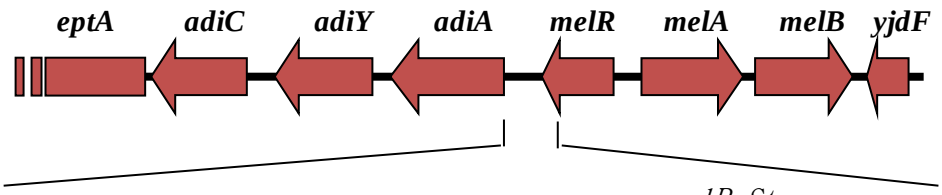
interactions has been uncovered. The regulation of other AR mechanisms is not so well understood.

1.3.3. Promoter organisation and local regulation of ADI

The ADI mechanism does not share the same genomic organisation as GAD described above. Figure 1.4 (A) shows the *adi* locus which consists of three genes *adiC*, *adiA* and *adiY*. Unlike the GAD system, where the *gadB* and *gadC* genes are organised into one operon, the ADI genes are organised into three separate transcriptional units. The regulation at these promoters is poorly understood. However, it is known that the whole locus is repressed by H-NS and activated by IHF but no binding sites have been mapped to the promoters (Hommais *et al.*, 2004; Shi *et al.*, 1993).

The central regulator of the ADI system is AdiY, which is an AraC-like regulator coded by *adiY*. AdiY is thought to regulate the expression of both the *adiA* and *adiC* genes. This is based on mutational studies, which show that an AdiY knock out has similar levels of survival as an *adiA* or *adiC* knock out (Kieboom and Abee, 2006). The actual binding sites for the AdiY regulator have not been mapped to these promoters. Expression of *adiY* has been shown to be increased in buffered LB medium shifted from pH 7.6 to pH 5.5 by the addition of 1 M HCl (Kannan *et al.*, 2008). The regulation of the Adi genes is poorly understood; however, it is likely that the *adiY* gene is activated by low pH (either directly or indirectly) and codes for a regulator which controls the decarboxylase and antiporter genes.

A



melR Stop

taccgcaaacttagccaacagcgccgccagacgtttcccggctaaagacgatatcagtat

-127 H-NS *-117* *-91*

cagccaaaaaaatagtttagccgtcgataataagcagtgaaaataatttaaaataatcac

IHF *-75* *-59* IHF *-43*

ataacgtattgtttataaaactaaatatttcaaaaatgtttgttttcacgcgctttaca

ttgaca

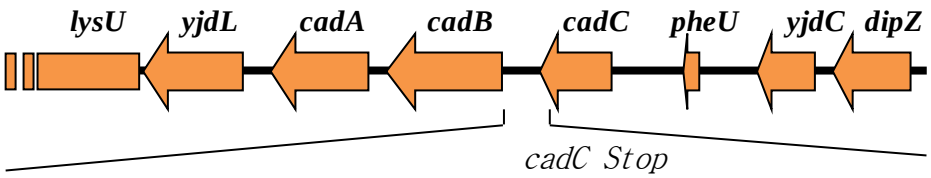
gcccgaaaaggccggaagatacttgcccgcaacgaagattccttcataaccgggtaagca

tataat

Start adiA

atgatgaaagtattaattgttgaaagcgagtttctccatcaagacacctgggtcggtaac

B



cadC Stop

gtaccttatctcgacaaatttcttgcttcagaataagtaactccgggttgatttatgctc

ggaaatatttgttgttgagttttttgtatgttcctgttggtataatatgttgcggcaattt

atttgccgcataatttttattacataaatttaaccagagaatgtcacgcaatccattgta

-144 CadC1 *-112* *-89*

aacattaaatgtttatcttttcatgatatcaacttgcgatcctgatgtgttaataaaaaa

CadC2 *- 59*

actcaagttctcacttacagaaacttttgtgttatttcacctaatcttaggattaatcc

ttgaca

-35

ttttttcgtgagtaatcttatcgccagtttggtctggtcaggaaatagttatacatcatg

tataat *+1*

-10 *cadB Start*

acccggactccaaattcaaaaatgaaattaggagaagagcatgagttctgccaagaagat

Figure 1.4. The promoter regions of the *adiA* and *cadBA* operons. The diagrams above each sequence represent the locus of genes which contain either *adiA* (A) or *cadBA* (B), arrows represent coding regions. The annotated sequences consist of the upstream intergenic regions. The translation starts are indicated in bold font. On each sequence the -10 and -35 elements are underlined and labelled. The consensus sequence is displayed below each sequence for comparison. Predicted H-NS and IHF binding sites are indicated by underlined bases (A) as are confirmed CadC binding sites (B). Numbers represent the amount of residues from the respective translation start site. These diagrams are based on work by: Kruper *et al*., 2005; Hommais *et al*., 2004; Stim-Herndo *et al*., 1996 and Shi *et al*., 1993.

1.3.4. Promoter organisation and local regulation of CAD

The main genes of CAD are located at a single locus. Figure 1.4 (B) shows *cadA*, *cadB* and *cadC* in close proximity to each other (Watson *et al.*, 1992). The CAD genes are transcribed from three promoters, one for each gene. The lack of a transcriptional terminator between the *cadA* and *cadB* genes results in a *cadBA* transcription unit (Watson *et al.*, 1992). The local regulator CadC binds to the *cadB* promoter *in vitro* and has been shown to regulate the expression of the *cadB* promoter *in vivo*. Direct binding of CadC at the *cadA* promoter, using *in vitro* assays, has not been shown (Kuper and Jung, 2005). CadC binds to two sites at the *cadBA* promoter Cad1 [-144 to -112, see figure 1.4 (B)] and Cad2 [-89 to -59, see figure 1.4 (B)] and it thought to de-repress the promoter by binding to Cad1 and displacing H-NS. RNAP is recruited to the *cadBA* promoter when CadC has bound both sites (Kuper and Jung, 2005). The *cadBA* operon is induced by low pH only in the presence of exogenous lysine. Repression of *cadBA* expression is due to a second regulator LysP (CadR). Strains that contained mutations in LysP are inducible by low pH without the requirement of exogenous lysine (Neely *et al.*, 1994). The affect of LysP on *cadBA* expression was found to be indirect, as strains over-expressing *lysP* did not affect *cadBA* expression in a *cadC* KO background. This result indicated that LysP represses *cadC* (Neely *et al*., 1994).

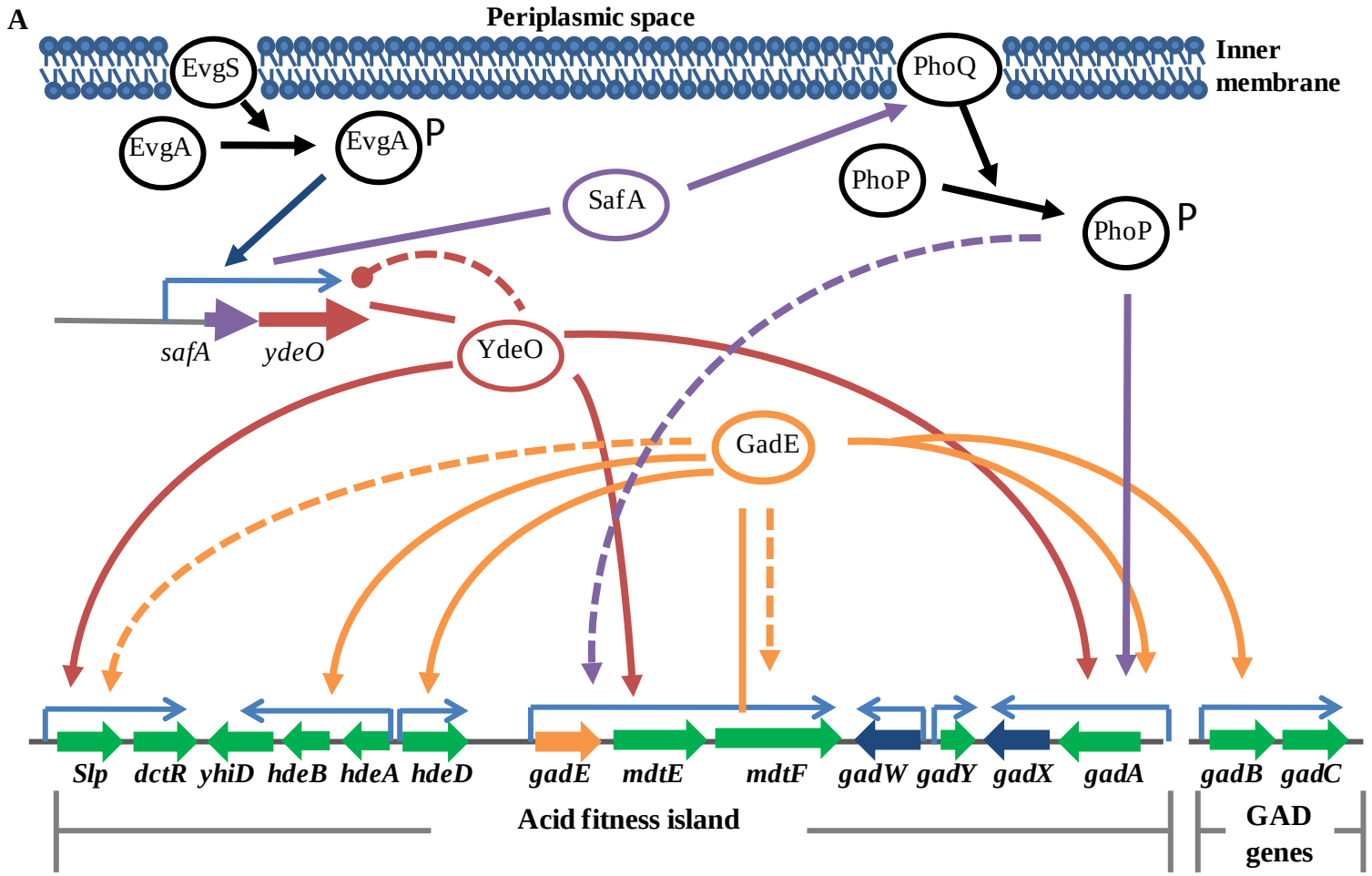
In summary, the CAD genes are regulated at a local level by CadC and LysP, which ensure that the genes are only expressed during growth at low pH in medium supplemented lysine. The CAD genes are also regulated, either directly or indirectly, by the AR2 network, which will be discussed in section 1.6.3.

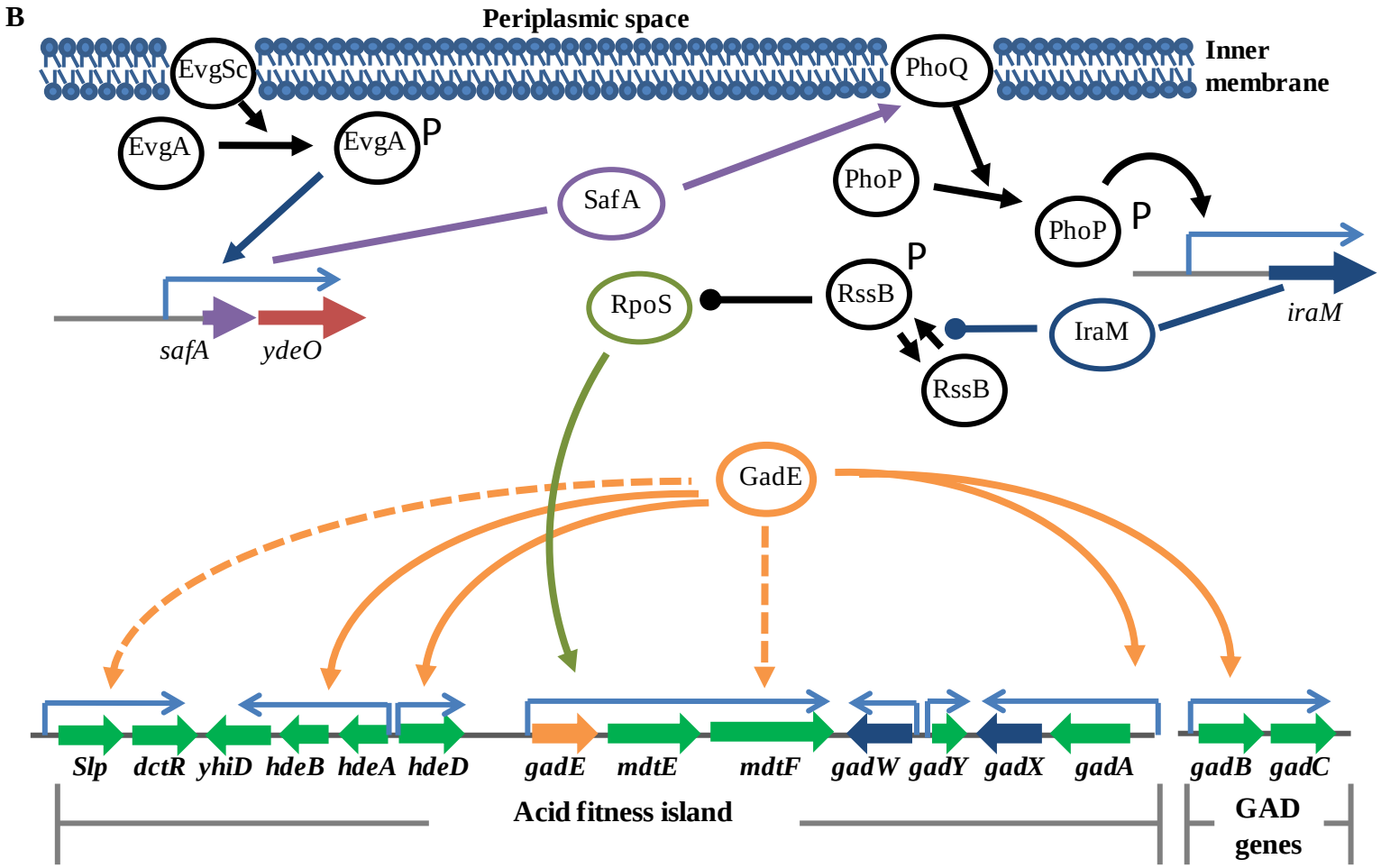
1.4. Control of Acid Resistance mechanisms by the AR2 network

The AR2 network is made up of many components including; two-component systems, global and local regulators, proteases and sRNAs. The key players in the AR2 network are the two-component system EvgAS, the local regulators GadE and YdeO, alternative sigma factor RpoS, and the regulatory sub-circuit of GadX-GadW-GadY. These regulators increase the reach of the AR2 network beyond the regulation of the GAD genes and the AFI to other acid resistance systems and other cellular processes. The AR2 network is activated either by growth in mildly acidic medium pH (4.8 – 5.9, maximally at pH 5.7) during exponential phase, or by entry into stationary phase. Induction by mild acid shock and entry into stationary phase is activated by two separate induction circuits, the EvgAS-YdeO-PhoPQ (figure 1.5 A) and RpoS-GadX-GadY-GadW (figure 1.5 C) circuits respectively. In response to low pH, or stationary phase growth, a cascade of activation occurs converging on the central regulators GadE and YdeO. This results in the activation of acid resistance genes at many different loci. This section will discuss the contributions of these AR2 network components in acid resistance regulation.

1.4.1. Repression of the AR2 network by CRP and H-NS

During exponential phase of cells grown at neutral pH the expression levels of genes involved in acid resistance is low. This is due to, among other things, repression by H-NS and CRP. H-NS (histone-like nucleoid structuring) is a nucleoid associated protein which has the propensity to bind and bridge DNA to cause compaction of the bacterial chromosome. H-NS primarily binds AT rich sequences such as those found at *E. coli* promoters.


A
Periplasmic space
Inner membrane
EvgS
EvgA
EvgA P
PhoQ
PhoP
PhoP P
SafA
safA
ydeO
YdeO
GadE
Slp
dctR
yhiD
hdeB
hdeA
hdeD
gadE
mdtE
mdtF
gadW
gadY
gadX
gadA
gadB
gadC
Acid fitness island
GAD genes


B
Periplasmic space
Inner membrane
EvgSc
EvgA
EvgA P
PhoQ
PhoP
PhoP P
SafA
RpoS
RssB P
RssB
IraM
iraM
safA
ydeO
GadE
Slp
dctR
yhiD
hdeB
hdeA
hdeD
gadE
mdtE
mdtF
gadW
gadY
gadX
gadA
gadB
gadC
Acid fitness island
GAD genes

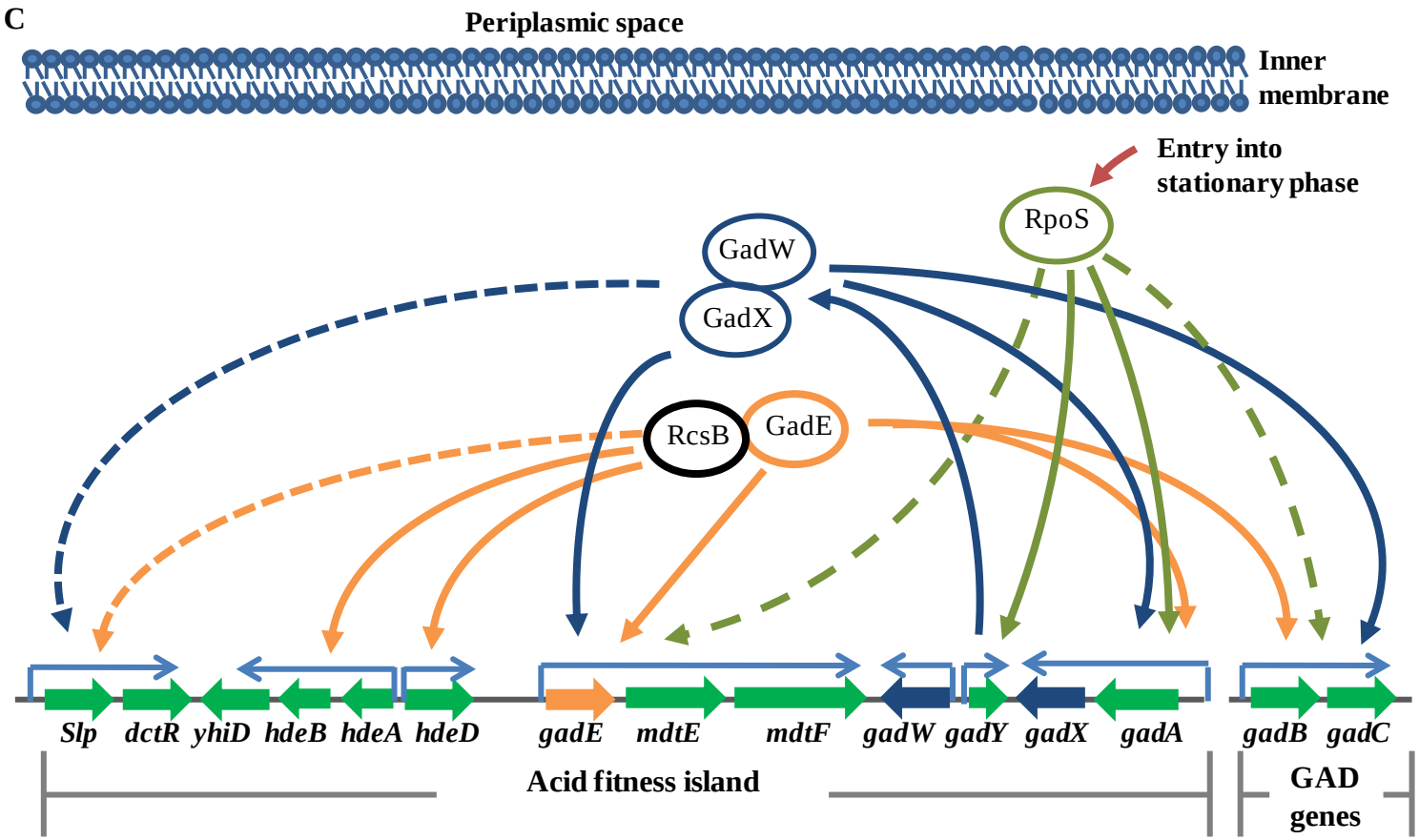


Figure 1.5. Induction of the AFI and GAD genes by the AR2 network . The induction of the AFI and GAD by mild acid shock during exponential phase **A**, by a constitutive on EvgSc mutation **B** (modified from Eguchi *et al*., 2010) and by growth in to stationary phase **C**. Solid arrows indicate confirmed direct interactions between the regulator proteins (coloured circles) and the promoters of a cid resistance genes (large green arrows). Dotted arrows represent interactions that are not confirmed as direct and lines with circled ends indicate repression.

When H-NS binds DNA, it induces bending that causes reduced expression of proximal genes (Dame *et al.*, 2000; Grainger *et al.*, 2006). CRP, or cyclic adenosine mono-phosphate receptor protein, is a global TF. When allosterically activated by cAMP, CRP may bind DNA to repress transcription initiation (Busby and Ebright, 1999). Both of these regulators have been implicated in the control of acid resistance. The contributions of each regulator will be discussed below.

1.4.1.1. Control of the AR2 network by cAMP receptor protein

The levels of AR2 network activators, RpoS and GadE, are increased during growth in minimal medium supplemented with glucose compared to growth in rich medium. Under these same conditions, levels of cAMP-CRP complex are also lower than in rich medium (Heuveling *et al.*, 2008). These results suggested that the AR2 network is under the control of CRP. Knock-out mutagenesis of the *crp* gene in *E. coli* resulted in an increase in expression of AR2 (Castanie-Cornet and Foster, 2001). The interactions that cause AR2 genes to become repressed are not clear. However, it is known that the increase in expression is dependent on RpoS. One possible explanation is that CRP mutants grow much slower than wild type cells, which would also increase RpoS levels (Hengge-Aronis, 2002). Although the effect of CRP seems to be dependent on RpoS, recent work has characterised a CRP binding site at the promoter of central regulator *gadE*. This finding suggests that at least on some level the repression of AR2 network by CRP is direct (Hirakawa *et al.*, 2006). There is no doubt that CRP is involved in the regulation of acid resistance, or that it represses the AR2 network. However, the physiological importance of CRP's role in acid resistance is yet to be discovered.

1.4.1.2. Control of the AR2 network by H-NS

H-NS is a global regulator that is normally associated with promoter repression. The level of activity at AR2 promoters in pH neutral medium increased in an *hns* knockout background (De Biase *et al.*, 1996; Hommais *et al.*, 2001; Waterman and Small, 2003b). Hommais *et al* compared the *hns* knock out transcriptome to the wild type and found that many genes in the AR2 network were up-regulated in this background. These genes, which will be discussed below include: *gadX*, *gadW*, *gadE*, *evgA*, and *ydeO*.

Recent *in vitro* work by Giangrossi *et al.* in 2005 has characterised H-NS binding to two regions of the AFI. The binding of H-NS to these elements resulted in the repression of *gadA* and *gadX* (GadX is an activator of AFI transcription, see section 1.4.4. below) (Giangrossi *et al.*, 2005). The AFI is an AT-rich region which is predicted to be transcribed when not actively repressed. A plausible model for repression, of the AFI by the H-NS, would be the binding of H-NS to the two sites characterised by Giangrossi *et al*, which would cause a hairpin loop in the DNA and a repression effect on the whole AFI (Giangrossi *et al.*, 2005).

In summary, under conditions where the growth rate is high, and pH is neutral, H-NS represses the AFI, the GAD genes, and the AR2 network. When the growth rate of the cells slow, or when the pH drops, these genes become de-repressed by the dissociation of H-NS from the locus. This allows other factors to bind the locus and activate transcription. The regulatory circuits that cause this de-repression and activate the acid resistance genes are discussed below.

1.4.2 Two-component systems involved in the AR2 network

The ability of bacteria to sense the environment and respond accordingly requires sensors and response regulators. The gap between environmental stimulus and gene regulation is often bridged by two-component systems (TCSs). There are at least three TCSs involved in the AR2 network, which are EvgAS, PhoPQ and RcsB. Details of the regulation by these TCSs are described below. This section will briefly describe the general characteristics of TCSs and their roles in bacteria.

1.4.2.1. The mechanism of TCSs

TCSs are composed of a histidine kinase (HK) and a response regulator (RR). A diagrammatic representation of a TCS is shown in figure 1.6. The role of the HK is to sense a signal or bind a ligand. This causes phosphorylation of the C-terminal transmitter domain. Either the phosphorylation event occurs via additional kinases or the HK can autophosphorylate (Mascher *et al.*, 2006). Next, the RR is activated by phosphorylation of the aspartate residue in the receiver domain. The RR is often phosphorylated by the HK, alternatively the RR can autophosphorylate or the RR is phosphorylated by a cytoplasmic kinase (Mascher *et al.*, 2006). The C-terminal domain of the RR then either interacts with DNA to regulate genes, or with other proteins. Finally, the system is reset by the dephosphorylation of the RR (Stock and Mowbray, 1995). In some cases the RR is capable of phosphorylating the HK. However, this process is often poorly understood. Recently, additional auxiliary proteins have been characterised in many TCSs, which interact with TCSs and increase the complexity of their regulation (Buelow and Raivio, 2010).

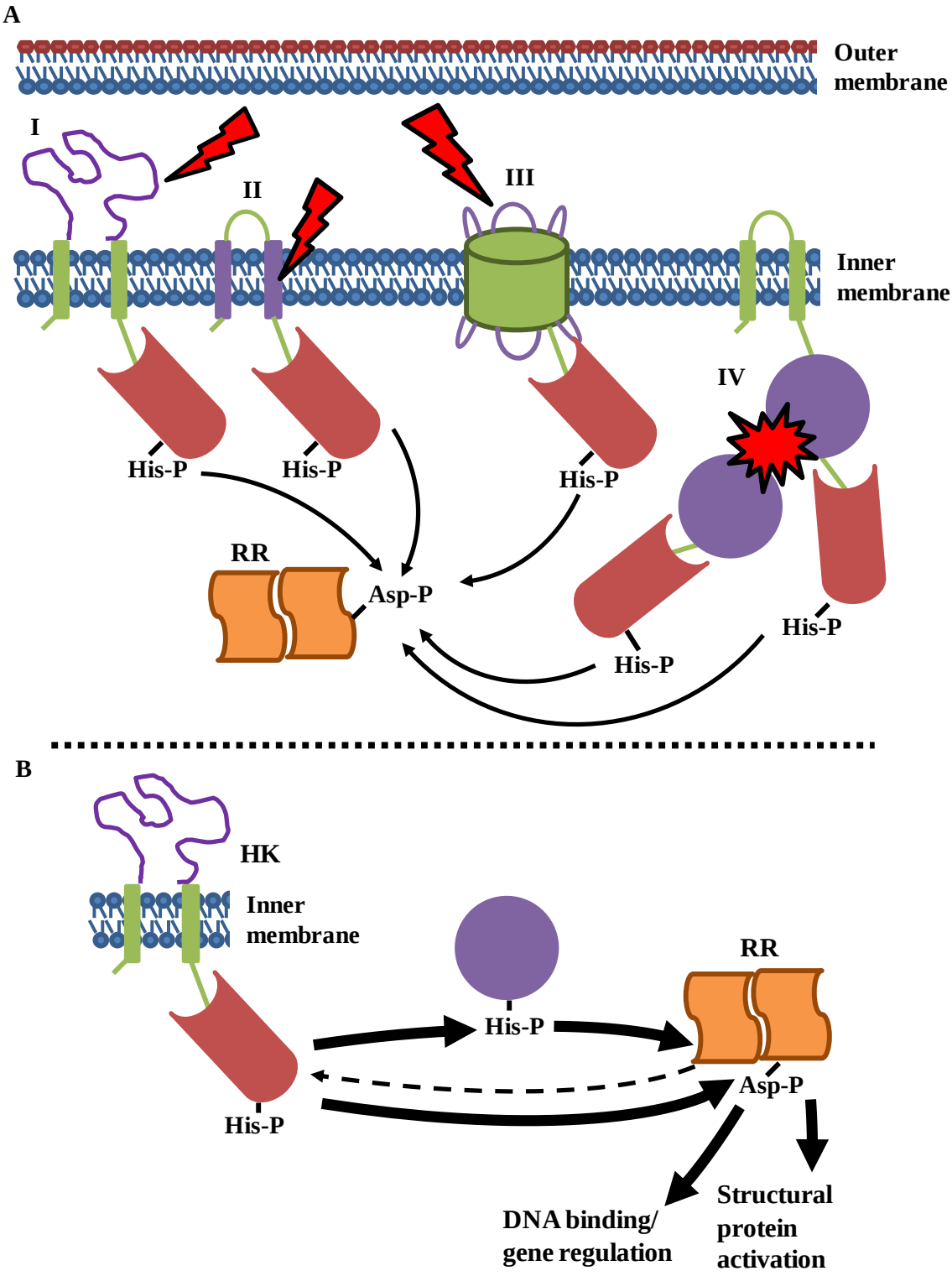


Figure 1.6. Activation of two-component systems. A The four main types of HK sensor. HKs with periplasmic (I), transmembrane (II), extracellular loop/transmembrane (III) and cytoplasmic (IV) sensing domains. **B** The activation of the response regulator by phosphorylation. The figure shows possible inclusion of an accessory factor (purple circle) and the role of some HKs in the dephosphorylation of the RR (dotted thin arrow). The activated response regulator activates structural proteins or regulates promoters. Adapted from Mascher *et al.*, 2006.

1.4.2.2. Activation of TCSs

There are four main methods of activation of HKs. Figure 1.6 A shows a diagrammatic representation of these methods. A HK can sense periplasmic signals using a periplasmic domain, which is normally flanked by two transmembrane domains. This architecture is found in the HK BvgS from *B. pertussis* (Figure 1.6 A I) (Bock and Gross, 2002). HK can also have sensing domains within the transmembrane regions of the protein. Sensing of a signal can occur within the transmembrane components themselves (figure 1.6 A II) or within the extracellular loops, as in the LuxN system (figure 1.6 A III) (Freeman *et al.*, 2000). Alternatively, the HK can receive signals from the cytoplasm, as found in the ArcB and FixL systems (Figure 1.6 A IV) (Malpica *et al.*, 2004; Miyatake *et al.*, 2000). The methods of activation described above are often found in combination. As a consequence, the ability of TCSs to integrate multiple signals is increased, which in turn increases the complexity of the response.

1.4.3. The EvgAS-YdeO-PhoPQ-GadE control network

1.4.3.1. Characterisation of the EvgAS two-component system as an AR regulator

The two-component system, EvgAS, was originally characterised as a homologue of the *B. pertussis* virulence regulator BvgAS. Extensive studies implicated EvgAS as a regulator of multidrug efflux pump EmrKY. EvgAS was implicated in acid resistance in 2002 by Masuda and Church. Increased acid resistance was shown by strains over-expressing the response regulator EvgA compared to the parental strain (Masuda and Church, 2002). EvgA over-expression was found to up-regulate many genes known to be involved in acid resistance. This included the AR2 and AFI genes (Masuda and Church, 2002; Nishino and Yamaguchi, 2001). The up-regulated genes were similar to those found when the transcriptome of a

constitutively active EvgS mutant strain was compared to that of a parental strain (Eguchi and Utsumi, 2005; Kato *et al.*, 2000). Subsequent mutagenesis of up-regulated genes revealed that only mutations in *ydeO*, *ydeP* and *yhiE* (now *gadE*) reduced the survival of the artificially induced acid resistance phenotype. These results implicated *ydeO, ydeP and gadE* in acid resistance (Masuda and Church, 2002). The regulators GadE and YdeO will be discussed in more detail later in this review. Briefly, YdeO is a member of the AraC/XylS super-family of transcriptional regulators, GadE is a member of the LysR family of transcriptional regulator and YdeP is a putative oxidoreductase that has homology to the alpha subunit of *E. coli* formate dehydrogenase H (Masuda and Church, 2003; Schell, 1993; Senda and Ogawa, 2005). When over-expressed, these genes can independently induce an acid resistant phenotype. YdeP and its role in acid resistance has been discussed in 1.2.6. Over-expression of *gadE* had the lowest impact. The role of YdeP in acid resistance has not been researched any further (Masuda and Church, 2003). In contrast, the role of transcriptional activators, EvgA, GadE and YdeO, has been extensively characterised.

1.4.3.2. Characterisation of the EvgA-YdeO-GadE activation cascade

The regulons of EvgA and YdeO were analysed by transcriptomic studies in which each of the regulators were over-expressed (Masuda and Church, 2002; Nishino and Yamaguchi, 2001). Among the genes found to be controlled by EvgA and EvgS were the GAD genes. As previously mentioned, YdeO and GadE were up-regulated by over-expression of EvgA (Masuda and Church, 2002). However, it was not clear from this data whether the activation of YdeO or GadE was directly by EvgA. In addition, over-expression of YdeO also increased GadE expression (Masuda and Church, 2003). It was suggested that these three genes were part of an induction cascade with EvgA activating YdeO and YdeO in turn activating GadE

and all of the AR2 genes and the AFI. Further study of the expression of *gadE* using a *gadE-lacZ* operon fusion in *ydeO*, *evgA* and *gadE* backgrounds revealed the nature of this circuit (Ma *et al.*, 2004). The expression of *gadE* in an *evgA* KO background was more than 3-fold less than the parent strain. This result was complementary to the transcriptomic studies (Ma *et al.*, 2004). A 2-fold reduction of *gadE* expression was observed in a *ydeO* KO background. The remaining expression at the *gadE* promoter was hypothesised to be due to direct EvgA activation. The expression of *gadE* was reduced over 3-fold in a *gadE* KO background, which suggested that GadE autoregulated its own expression (Ma *et al.*, 2004). Finally, the expression of *gadE* was measured in a *ydeO/evgA* double knockout. The activity of the *gadE* promoter was reduced nearly 6-fold, which was the largest reduction. This result indicated that both EvgA and YdeO activate *gadE* in an additive manner (Ma *et al.*, 2004). All three regulators were found to bind the *gadE* promoter, GadE itself bound specifically to three putative GAD box sequences.

Ma *et al.*, in 2004 provided conclusive evidence that *ydeO*, under transcriptional control of EvgA, is subject to acid induction. By contrast, transcription of the *evgAS* operon was shown not to be acid induced despite previous reports that it was subject to auto-regulation (Masuda and Church, 2003). In addition, a constitutively active *evgS* mutant was used to show that EvgA binds strongly to specific sites on the following promoters: *ydeP, safA, yfdX, frc*, *yegR* and *gadE* (Itou *et al.*, 2009). The role of *safA* in this pathway will be discussed later in this review; the functions of *ydeP*, *yfdX*, and *yegR* are not known. Phenotypic analysis of the acid resistance of the constitutive-on mutant confirmed that activation of the EvgAS pathway, induced acid resistance independently from acid or stationary phase growth (Itou *et al.*, 2009).

Work by Burton *et al.*, using a *lux* reporter system to measure promoter activity at high temporal resolution, confirmed the induction cascade from EvgAS to the AR2 genes. Additionally, the high temporal resolution data produced by Burton *et al.* captured the relative timings of induction and proved that the activity of the *ydeO*, *gadE* and *gadB* occurred in succession (Burton *et al.*, 2010).

1.4.3.3. Role of PhoPQ in the GAD Network

The signal transduction cascade of the EvgAS TCS has been shown to cross-talk to the PhoPQ TCS (Eguchi *et al.*, 2007; Eguchi *et al.*, 2004; Eguchi and Utsumi, 2005). This interaction was found to be dependent on a small protein SafA (sensor-associating factor A). Previously known as SMP (small membrane protein), SafA has been shown to interact with the PhoQ histidine kinase. As a result of this interaction the response regulator PhoP is activated. This activation has been observed by monitoring the expression of PhoP dependant genes such as *mgtA* (Eguchi *et al.*, 2007). SafA is encoded by the *b1500* gene (now referred to as *safA*) which forms an operon with *ydeO*. The *safA-ydeO* promoter is under the direct control of EvgA (Masuda and Church, 2003). Initially the activation of the PhoPQ TCS by EvgAS and SafA had only been shown by artificially over-expressing the aforementioned genes (Eguchi *et al.*, 2004; Kato *et al.*, 2000). More recently the activation of PhoP regulated genes has been shown by acid induction and without any artificial over-expression (Burton *et al.*, 2010). In conditions of low Mg^{2+}, which activates the PhoPQ TCS, *gadE* and *gadW* were found to be up-regulated. This suggests that PhoP also regulates AR2 genes in addition to its originally proposed role as a Mg^{2+} starvation regulator (Groisman, 2001; Zwir *et al.*, 2005).

PhoPQ has recently been implicated in a completely new activation arm of the AR2 network. Survival of a EvgSc (constitutively-on EvgS) strain was shown to be completely dependent on *phoP* (see figure 1.5 B) (Eguchi *et al*., 2011). The *iraM* gene, which is activated by PhoP, was required for EvgSc conferred acid resistance (Eguchi *et al.*, 2011). In addition, the activation of the *gadE* promoter by an EvgSc mutant was shown to be completely dependent on RpoS. IraM (inhibitor of RssB activity) controls the levels of RpoS in the cells by interacting with RssB. When activated RssB binds to RpoS, near amino acid L173, and the amino-terminus of ClpX (Studemann *et al.*, 2003). This interaction connects substrate (RpoS) and protease (ClpX) and allows degradation (Eguchi *et al.*, 2011). The model of activation by an EvgSc mutant, depicted in figure 1.5 B, is as follows. EvgA activates the *safA-ydeO* promoter, SafA expression is increased leading to the activation of PhoPQ. PhoP activates *iraM* transcription and IraM interacts with RssB. As a result RssB cannot interact with RpoS and RpoS levels rise. The increase in RpoS levels causes RNAP to be recruited to RpoS regulated promoters, including the *gadE* promoter (Eguchi *et al.*, 2011).

The resulting network involves EvgAS TCS cross-talking to PhoPQ via SafA. The PhoP activation results in PhoP dependent activation of *gadE* via IraM and RpoS (See Figure1.5 B). The cross talk from EvgAS to PhoPQ therefore also connects EvgAS to RpoS and the RpoS-GadW-GadX-GadY circuit of activation. It should be noted that the investigation of the PhoP-IraM-RpoS cascade was done during late exponential phase (OD_{600} 0.6) (Eguchi *et al.*, 2011). During this phase of growth, the levels of RpoS are increasing, as growth slows into stationary phase. The importance of RpoS in the activation of GadE is therefore increased in these conditions compared to exponential phase conditions. This variation makes comparison of this circuitry with the exponential phase AR2 network problematic.

1.4.3.4. Role of RcsB in the GAD Network: An essential inactive activator

RcsB is a response regulator that is activated by phosphorylation by the RcsCDF histidine-aspartate phosphorelay. Recently, RcsB has been reported to be essential for *gadB* and *gadA* promoter activity during stationary phase (Castanie-Cornet *et al.*, 2007). Interestingly, an *rcsB* mutation had no affect on *gadE* transcription and over-expression of GadE could not compensate for loss of RcsB (Castanie-Cornet *et al.*, 2007). Furthermore, activation of the Rcs phosphorelay by over-expression of either RcsB, RcsA or DjlA (a factor which activates RcsC kinase activity), resulted in a reduction of survival in extreme acid conditions. Therefore, RcsB was hypothesised to co-operate with GadE in the activation of AR2 (Castanie-Cornet *et al.*, 2010). It was shown that basal levels of un-phosphorylated RcsB are an essential requirement for *gadA/B* promoter activity. Recently, the role of RcsB as a regulator of AR2 structural genes was confirmed. A GadE-RcsB heterodimer was purified and used to show that activation of the *gadA* promoter requires GadE and RcsB binding (Castanie-Cornet *et al.*, 2010; Krin *et al.*, 2010b). Additionally, an RcsB box was identified just upstream of the -10 element of the *gadA* promoter (Castanie-Cornet *et al.*, 2010). Mutational analysis of the *gadA* promoter revealed two binding sites for RcsB. It was shown that the binding of RcsB to the second low affinity site caused repression of *gadA* promoter activity. However, results reported by another group indicated that the heterodimer formation required phosphorylated RcsB (Krin *et al.*, 2010b). An RcsB mutant that contained a histidine to glutamate mutation that mimicked phosphorylated RcsB was required for heterodimer formation. These results showed that RcsB was required to be active. In support of this model, RcsB can be phosphorylated by acetyl-P, and the Rcs histidine kinases form a dynamic equilibrium, which under inactive conditions actually function to reduce the levels of phosphorylated RcsB (Clarke, 2010). It is therefore possible that the deletion of RcsC/D/F

could increase the levels of phosphorylated RcsB. It should be noted that the group did not publish a control to confirm the mimic, and the phosphorylation state of RcsB was not measured. The effect of RcsB on other GadE regulated promoters has also been analysed (Krin *et al.*, 2010a). Currently all acid resistance promoters that are activated by GadE are also dependent on RcsB.

To summarise, the response regulator RcsB has been confirmed to play a key role in the AR2 network, both as an essential activator and a repressor. However, there is still some confusion as to whether RcsB is required in an active form. In addition all of the investigations into the role of RcsB have only been done during stationary phase. The role of RcsB in exponential phase and acid induction of resistance has not been investigated. Overall, the full extent to which RcsB interacts with the AR2 network is still unclear.

1.4.4. The RpoS-GadX-GadW-GadY regulatory circuit

1.4.4.1. Characterisation of the components of the RpoS-GadY-GadX-GadW regulatory circuit

The second major circuit controlling the expression of AR is the RpoS-GadX-GadW-GadY regulatory circuit. This circuit is activated during stationary phase by the alternative sigma factor RpoS. The main constituents of this pathway are the GadX and GadW regulators. The roles of GadX and GadW are hard to pin down due to the variation of function observed between different strains of *E. coli* and during different growth conditions (See figure 1.5 C).

RpoS is an alternative sigma factor, which becomes more abundant in stationary phase cells. The regulation of RpoS is very complex, and beyond the scope of this review. Many

comprehensive reviews have been written on this subject (Hengge-Aronis, 2002; Typas *et al.*, 2007). There is no question that RpoS is involved in acid resistance. Analysis of *rpoS* KO strains showed that survival was reduced during stationary phase (Castanie-Cornet *et al.*, 1999; De Biase *et al.*, 1999). During exponential phase growth the most prominent sigma factor, RpoD, is competent to recruit RNA polymerase to the promoters of acid resistance genes (Giangrossi *et al.*, 2005; Waterman and Small, 2003a). However, during exponential phase growth, or in an *rpoS* KO background, no transcript from the GadX promoter can be detected. This is due to the affect of RpoS levels on the expression of the *gadY*. The small regulatory RNA, *gadY*, up-regulates the amount of GadX in two ways. Firstly, *gadY* base pairs to the 3' end of the *gadXW* transcript, stabilising the RNA. Secondly, *gadY* aids the processing of the dicistronic *gadXW* transcript (Opdyke *et al.*, 2004; Tramonti *et al.*, 2008; Tramonti *et al.*, 2002). GadX is an AraC-like regulator that has been shown to function as an important activator of acid resistance genes (Ma *et al.*, 2002; Tramonti *et al.*, 2002). GadW is also an AraC-like regulator. The role of GadW is more diverse and the precise conditions and strain seem to dictate its role (Ma *et al.*, 2002; Tucker *et al.*, 2003). Recent evidence, from work on *E. coli* K-12 MC4100 and MG1655, has suggested that GadW acts as an activator of the AR2 genes (Sayed *et al.*, 2007; Tramonti *et al.*, 2006). The GadX and GadW proteins then regulate the transcription of many genes involved in acid resistance. Recently, the role of GadX and GadW in exponential phase induction of acid resistance was investigated. The induction of acid resistance by mild acid shock was shown to be completely independent from GadX and GadW (Burton *et al.*, 2010).

1.4.4.2. Characterisation of GadX and GadW regulation

The mechanisms by which GadX and GadW are thought to activate the expression of genes are difficult to interpret because their interactions differ between promoters. Much of what is known about GadX/W regulation was characterised while investigating the *gadBC* and *gadAX* promoters. At the *gadBC* and *gadAX* promoters, GadX and GadW can bind to a 42-bp GadX/W consensus site found in both promoter regions (Richard and Foster, 2007; Tramonti *et al.*, 2008; Tramonti *et al.*, 2006). However, the proposed mechanism of activation at the *gadA* promoter is different from that of the *gadBC* promoter. When GadX binds the *gadAX* promoter, *in vitro*, GadX antagonises H-NS mediumted repression (Giangrossi *et al.*, 2005). In contrast, at the *gadBC* promoter, where the GadX/W binding site is in the distal upstream promoter region, the proposed mechanism of activation by these factors invokes a DNA looping mechanism (Tramonti *et al.*, 2006). In addition to the *gadAX* and *gadBC* promoters, it is now apparent that each of these factors can also bind to, and activate, the *gadE* promoter region (Sayed and Foster, 2009; Sayed *et al.*, 2007; Tramonti *et al.*, 2008). Although there is no doubt that GadX and GadW activate the AR2 genes, the fact that GadX binds the *gadE* promoter and activates expression of GadE casts doubt as to which interactions take precedence over the others.

1.4.4.3. Characterisation of the RpoS-GadY-GadX-GadW regulatory circuit

The components of the RpoS-GadY-GadX-GadW circuit combine to regulate acid resistance in response to stationary phase growth. Regulation by this network is as follows. During exponential growth the levels of local acid resistance regulators, GadX and GadW, are very low in the cell. This is due to inhibition by global regulator H-NS (see section 1.4.1.2.) (Tramonti *et al.*, 2008; Tramonti *et al.*, 2002). Entry into stationary phase causes the levels of

RpoS to increase in the cell. RpoS recruits RNA polymerase to a subset of promoters, one of which is the *gadY* promoter. Up-regulation of *gadY* causes the *gadX* and *gadXW* transcripts to be stabilised and more effectively processed, thus increasing the abundance of activators GadX and GadW in the cell. A 42 bp GadX/GadW binding sequence has been characterised at five of the eight main promoters of the AFI that GadX and GadW bind and remove H-NS. Thus, the main action of the circuit, when induced, is to relieve the repression of AFI and GAD promoters from H-NS.

1.4.5. Role of TrmE in the AR2 network

TrmE, or MnmE, is a molecular switch that undergoes a conformational change when bound by GTP (Cabedo *et al.*, 1999). TrmE was found to be involved in GAD regulation by Gong *et al.*, in 2004. However, how TrmE interacts with the GAD promoters or the AR2 network is currently unknown. The bulk of evidence that supports the role of TrmE in the AR2 network was revealed by a *trmE* mutation in the EK227 K-12 strain of *E. coli.* This strain had dramatically reduced survival in rich medium supplemented with glucose and glutamate (Gong *et al.*, 2004). The mutation had no effect on survival in un-buffered LB grown cultures. The *trmE* deletion strain was acid sensitive compared to the parent strain when glucose was added to the medium (Gong *et al.*, 2004).

Transcription of key acid resistance genes was examined in the *trmE* mutant background. Mutants lacking *trmE* had reduced *gadE* transcription, during stationary phase growth, in complex medium containing glucose. However, there was no difference in expression between a *trmE* deletion strain and the parent in medium without glucose.

1.5. Role of GadE in the AR2 network

1.5.1. Characterisation of GadE

GadE is a member of the LysR family of transcriptional regulators and is the central regulator of the AR2 network. GadE was first implicated in acid resistance when it was shown to be essential to acid resistance conferred by EvgA over-expression (Masuda and Church, 2002). Later, it was found that acid resistance levels, similar to that shown by *evgA* over-expression, could be achieved by over-expression of *gadE* (Masuda and Church, 2003).

GadE was originally characterised as a regulator of the GAD genes because acid resistance conferred by GadE was observed in conditions favourable to GAD (Tucker *et al.*, 2002). However, a *gadE* deletion mutant also had significant effects on survival conferred by the AR1 and AR3 systems (Ma *et al.*, 2003a). The regulation of GAD by GadE has been extensively characterised. EMSAs of purified GadE with DNA from the promoters of *gadBC* and *gadA* revealed a 20 bp GadE binding site now known as the GAD box (Ma *et al.*, 2003a). The GAD box is a conserved sequence found at all promoters that GadE binds. GadE also has a GAD box at its own promoter, which GadE binds to activate its own expression.

1.5.2. Post-transcriptional control of GadE

The sections above have described the transcriptional control of GadE. In addition to transcriptional regulation, GadE is also regulated post-transcription. There are two main factors that regulate GadE post-transcription. RNA stability is regulated by a small RNA, and protein stability is regulated by Lon protease. This section will discuss the interaction of each of these post-transcriptional regulators with *gadE* RNA and protein.

The transcript of *gadE* is regulated by a small RNA named acid resistance-related sRNA. This sRNA is transcribed in the opposite direction from *gadE* (Aiso *et al.*, 2011). Acid resistance-related sRNA is an antisense RNA to the *gadE* transcript. This sRNA was shown to bind and stabilise the GadE transcript (Aiso *et al.*, 2011). The acid resistance-related sRNA is transcribed from the 6H57 gene during conditions of mild acid shock (pH 5.5) and growth into stationary phase (Aiso *et al.*, 2011).

GadE protein is also actively degraded by Lon protease. The degradation rate of GadE is 3 minutes in a K-12 MG1655 strain. However, in a *lon* KO mutant background, GadE half life is increased to 50 minutes (Heuveling *et al.*, 2008). The biological significance of this is not fully understood. However, it has been suggested that Lon protease is important for resetting the GAD system once a cell moves from acidic to pH neutral conditions (Heuveling *et al.*, 2008).

1.5.3. GadE is a central regulator of acid resistance in *E. coli*

The role of GadE in acid resistance is important for the induction of the AR2 network. As previously explained, two main stimuli cause the up-regulation of acid resistance genes and therefore greater survival at extreme pH. The first is entry into stationary phase, which activates the RpoS-GadY-GadX-GadW circuit described in 1.4.4. Briefly, RpoS levels increase in the cell which activates the GadE promoter. Additionally, RpoS activates the GadX-GadY-GadW mini circuit which results in high levels of GadX. GadX can activate genes involved in acid resistance and GadE. This activation is independent from the EvgAS-YdeO regulatory circuit described in 1.4.3.1. The second stimulus which activates the AR2 network is a mild acid shock (pH 4.8 to pH 5.8). This causes response regulator EvgA to

become active and bind the *safA-ydeO* and *gadE* promoters, resulting in the activation of these promoters. YdeO activates GadE, further increasing GadE levels and therefore activation of acid resistance genes. Regardless of the stimuli each circuit results in the activation of GadE and the AR2 genes.

1.6. Regulatory reach of the AR2 network

The current range of regulators implicated in the AR2 network is extensive; as a result, AR2 forms a wide array of interactions with systems that are not all involved in acid resistance. The AR2 network comprises of the activators EvgAS, RcsB, YdeO, GadE, PhoPQ, GadW and GadX. Although many of these regulators were characterised by their regulation of the AR2 mechanism their reach also affects other acid resistance mechanisms. The action of these regulators increases acid resistance in *E. coli*. This section will discuss the interactions of these regulators with acid resistance genes.

1.6.1. AR2 network control of AR2

The AR2 genes, GadA, GadB and GadC, are regulated directly by the central AR2 network regulator GadE. During conditions of acid shock or during stationary phase growth GadE is up regulated. GadE in turn binds to GAD boxes located in the promoters of GadBC and GadA and activates their transcription. Currently it is understood that GadE binds to the GadA promoter only in a heterodimer with RcsB. In addition to up regulation by GadE, the *gadAX* promoter is also activated by GadX in stationary phase conditions (see figure 1.5 B). The activation of the *gadA* promoter occurs by removal of H-NS. The activation of *gadAX* promoter by GadX could be direct or via GadE.

1.6.2. AR2 network control of AFI

The AFI includes the promoters of central regulators GadE and GadX. During exponential phase growth induction occurs via mild pH shock (4.8 to 5.8). In response to low pH the EvgA-RcsB-YdeO circuit is activated. Both EvgA and YdeO activate the *gadE* promoter additively. GadE then up regulates the expression of *hdeAB* and *hdeD* promoters the products of which reduce protein aggregation in the periplasm. In addition to activating the *gadE* promoter, YdeO also activates *slp*, found to be important in acid resistance (see figure 1.5 A)

During stationary phase growth the RpoS-GadX-GadW-GadY circuit is induced. As a result the local regulators GadX and GadW activate the AFI. The GadX and GadW regulators directly control components of the AFI including *hdeA-hdeB-yhiD, hdeD,* and *slp-dctR*. However, the aforementioned promoters contain binding sites for both GadW/X and GadE (Tramonti *et al.*, 2008). The fact that binding sites for both activators are present suggests that both can probably influence promoter activity directly at these sites. One explanation for this regulatory redundancy is that as the AFI is predicted to be a horizontally acquired region of DNA, it has been proposed that GadX and GadW evolved to function as an H-NS anti-repression element in the AFI. The role of GadX and GadW is to prevent the silencing of this region by H-NS as it is acquired by an organism (Dorman, 2007; Tramonti *et al.*, 2008). Once GadE is freed from H-NS repression it can assist in generating a higher level of transcriptional activation across the AFI (Tramonti *et al.*, 2008).

1.6.3. Regulation of AR4 by the AR2 network

The AR4 genes, *cadA* and *cadB*, are controlled by local regulators CadC and LysP. LysP regulates AR4 in response to the abundance of exogenous lysine; *cadC* activates AR4 genes

in response to low pH stress. The mechanism by which CadC senses low pH is unknown; it is known however, that *cadC* and *cadB* are regulated by both GadE and GadX (Hommais *et al.*, 2004). The regulation of the *cadC* and *cadB* promoters by GadE or GadX is not well understood, no binding sites have been characterised at these promoters nor have the GadE or GadX proteins been shown to activate *cadC* and *cadB* directly.

1.7. Experimental evolution experiments

Bacteria are ideal for studying evolution, due to their short generation times and size. In addition, studying evolution using bacteria in a laboratory environment allows the control of many variables during an evolution experiment (Elena and Lenski, 2003). This section will outline the current methods for lab-based evolution experiments using microbes, and explain the rationale for the evolution experiment used in this study.

1.7.1 Evolving microbes

The basic form of an evolution experiment is to propagate a culture from a single colony over a series of generations while applying some selective pressure. The culture is often sampled at regular intervals and stored at -80°C. After a period of time the fitness of the "evolved" strains can be compared to their "lineal ancestor". The ancestor strain is normally compared to the evolved strains using a competition assay. Competitions enable the direct comparison of the ancestor strain with the evolved strains, as opposed to measuring growth rate, where neither strain competes for resources. In many cases, the conditions of the competition assay will emulate the conditions in which the evolution experiment was done. In some cases the competition assay is done in other conditions to test for trade-offs. In order to do a competition assay, the ancestor and evolved strains must be marked to distinguish between

them. Markers can take many forms including; differing PCR products using universal primers, the ability to bind a dye, fluorescent markers or antibiotic resistance (Elena and Lenski, 2003).

1.7.2 Applications of lab based evolution

Evolution experiments have been done using microbes over long periods of time, in some cases over 40,000 generations (Barrick *et al.*, 2009; Pelosi *et al.*, 2006; Woods *et al.*, 2006), These experiments can be used to test fundamental concepts in evolution. For example, a study by Barrick *et al.* in 2009 monitored the mutation rate of cultures growing in a constant laboratory environment. The study looked at cultures that had been growing for 40,000 generations. The genomes of cultures from throughout the 40,000 generations were sequenced. This study showed that genomic evolution was almost constant, but the proportion of adaptive mutations was high (Barrick *et al.*, 2009). The work by Barrick *et al.* showed that there was a clear difference in genomic and adaptive evolution of these strains during the course of the experiment. The Barrick *et al.* study is one of many that use lab-based evolution to understand evolution and not the organism. For a comprehensive review on such studies see Elena and Lenski 2003.

Research using short-term evolution experiments is also important, especially when trying to understand the relatively fast emergence and transfer of antibiotic resistance and pathogenic phenotypes. Short-term evolution experiments can also be used to select mutations that are important for a particular phenotype. This differs from other evolution experiments as the emphasis is on understanding the mutants or the phenotype and not the process. Using evolution to select for mutations has many advantages over traditional mutagenesis strategies.

Firstly, the evolution experiment is discovery-based and not biased by any particular hypothesis. Consequently, there is a possibility of serendipitous results. The evolution experiment will also select mutations that are important to the phenotype. The main caveat of lab-based evolution is that the mutations can be hard to identify. Due to these difficulties, lab-based evolution experiments traditionally focused on acquired or lost phenotypes (Pelosi *et al.*, 2006; Woods *et al.*, 2006). In most cases very little was understood about the underlying genetic changes that were responsible for the evolved phenotype. However, the application of whole genome techniques such as re-sequencing has largely solved this issue. To date several studies have used experimental evolution and high throughput techniques to understand both evolution in general, and specific pathways and phenotypes (Barrick *et al.*, 2010; Crozat *et al.*, 2011; Stanek *et al.*, 2009). However, to date, no studies have used lab-based evolution and high-throughput sequencing, to select for mutant strains under extreme stress conditions and to identify all mutations acquired during the process.

We have utilised lab-based evolution to select for mutations that are advantageous to survival in extreme acid conditions. This method has been combined with whole genome re-sequencing to investigate the mutations. Finally the results from these experiments have been used to understand the regulation of acid resistance.

1.7.2 Evolution of acid resistant *E. coli* K-12

Prior to the start of this project, used lab based evolution to evolve 8 separate *E. coli* K-12 cultures by selection with extreme acid shock (Russell and Lund, unpublished data). The cell lines were labelled A-H and a single acid resistant clone was isolated from each culture. These clonal isolates were named Aa-Ha. The evolution experiment is described in detail in

Chapter 3. The acid resistant phenotype of these strains and the mutations that have been selected for during the evolution experiment were the subject of this study (Russell and Lund, unpublished data).

1.8. Aims

This review has described the various mechanisms that *E. coli* uses to survive acid stress. There are many regulatory factors that control these genes and as a consequence the number of possible mutations that could cause an acid resistance phenotype is large. In this study we have investigated the regulation of acid resistance by analysing strains of *E. coli* that have evolved high levels of acid resistance in the lab. The primary goal of this work was to understand the genetic changes that caused previously acid sensitive strains to become acid resistant. In order to achieve this, the following sub objectives were completed.

- Full characterisation of the acid resistance phenotype of the evolved strains using standardised assays and molecular techniques.
- Whole genome analysis of the evolved strains to find the mutation(s) responsible for the evolved resistance.
- Characterisation of mutations found in evolved strains and their role in acid resistance.

Characterisation of the evolved strains could increase our understanding of acid resistance in *E. coli*. Acid resistance is regulated by a complex network; mutations in this network could give rise to an acid resistant phenotype. Studying the accumulation and tolerance of such mutations in regulatory networks could provide insights into the robustness and evolvability of bacterial networks.

CHAPTER 2: Materials and methods

2.1. Bacterial strains and plasmids

The bacterial strains used in this study are listed in Table 2.1. Gene deletion mutant strains were constructed as described in 2.6.2 and single base chromosomal mutants were made as described in 2.6.3. The plasmids used in this study for mutagenesis, transcriptional analysis and complementation, are listed in Tables 2.2, 2.3 and 2.4 respectively. Plasmid maps of the reporter vector, pLUX; gene doctoring vector, pDOC-C; and the low copy vector, pZC320, are indicated in figure 2.1. A, B and C.

Table 2.1. Bacterial strains used in this study.

Strain	Relevant genotype	Source / reference
DH5α	*E. coli* DH5α (F-, φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk-, mk+), *phoA*, *supE*44, λ-, *thi*-1, *gyrA*96, *relA*1)	Invitrogen, Paisley, UK
MG1655	*E. coli* K-12 MG1655 (F-, lambda-, *ilvG-*, *rfb*-50, *rph*-1)	Blattner *et al.*, 1997
EvoAa	Evolved MG1655	P. B. Russell
EvoBa	Evolved MG1655	P. B. Russell
EvoEa	Evolved MG1655	P. B. Russell
EvoGa	Evolved MG1655	P. B. Russell
EvoHa	Evolved MG1655	P. B. Russell
EvoAa Δ*rpoS*	Evolved MG1655 *rpoS::*Km^R	This study
EvoBa Δ*rpoS*	Evolved MG1655 *rpoS::*Km^R	This study
EvoEa Δ*rpoS*	Evolved MG1655 *rpoS::*Km^R	This study
EvoGa Δ*rpoS*	Evolved MG1655 *rpoS::*Km^R	This study
EvoHa Δ*rpoS*	Evolved MG1655 *rpoS::*Km^R	This study
EvoAa Δ*gadC*	Evolved MG1655 *gadC::*Cm^R	This study
EvoBa Δ*gadC*	Evolved MG1655 *gadC::*Cm^R	This study
EvoEa Δ*gadC*	Evolved MG1655 *gadC::*Cm^R	This study
EvoGa Δ*gadC*	Evolved MG1655 *gadC::*Cm^R	This study
EvoHa Δ*gadC*	Evolved MG1655 *gadC::*Cm^R	This study
*evgSc*G658A	MG1655 *evgS::*G658A[1]	This study
*evgSc*G658A Δ*evgA*	MG1655 *evgS::*G658A[1] *evgA*::Km^R	This study
*evgSc*S600I	MG1655 *evgS::*S600I[1]	This study
*evgSc*S600I Δ*evgA*	MG1655 *evgS::*S600I[1]	This study
*evgSc*S600I Δ*rcsB*	MG1655 *evgS::*S600I[1]	This study
Δ*rpoS*	MG1655 *rpoS::*FRT	This study
Δ*gadE*	MG1655 *gadE::*FRT	P1 x MC4100 *gadE*::Km^R (Heuveling *et al.*, 2008) x MG1655 (done by Lesley Griffiths); x pCP20
Δ*B1500*	MG1655 *b1500::*Cm^R	This study
Δ*ydeO*	MG1655 *ydeO::*FRT	P1 BW25113*evgA*::Km^R *(Baba et al., 2006)*x*MG1655* x*pCP20*
Δ*ydeO rpoS*	MG1655 *ydeO::*FRT *rpoS*::Cm^R	N. Burton (Burton *et al.*, 2010)

Table 2.1. Bacterial strains used in this study (continued.)

Strain	Relevant genotype	Source/ reference
Δ*phoP*	MG1655 *phoP*::Cm^R	N. Burton (Burton *et al*., 2010)
Δ*evgS*snv	MG1655 *evgS* (SNV region)*::*Km^R	This study
Δ*ydeO* Δ*phoP*	MG1655 *ydeO*::FRT *phoP*::Cm^R	N. Burton (Burton *et al*., 2010)
Δ*rpoS* Δ*phoP*	MG1655 *rpoS*::FRT *phoP*::Cm^R	N. Burton (Burton *et al*., 2010)
Δ*rcsB*	MG1655 *rcsB*::Cm^R	This study
Δ*gadC*	MG1655 *gadC*::Cm^R	This study
Δ*gadE* Δ*ydeO*	MG1655 *gadE::*Km^R *ydeO::*FRT	This study
Δ*gadE* Δ*ydeO* Δ*phoP*	MG1655 *gadE::*Km^R *ydeO::*FRT *phoP::*Cm^R	This study
Δ*ydeP*	MG1655 *ydeP*::Cm^R	This study

[1] Resulting amino acid substitution
Km^R – kanamycin resistant
Cm^R – chloramphenicol resistant
FRT – scar sequence left after cassette removal

Table 2.2. Plasmids used in this study for mutagenesis.

Plasmid	Description	Source / reference
pKD3	Gene mutagenesis plasmid; Cm^R; carries *cat* gene between two flanking FLP recognition target sites (FRT)	Datsenko and Wanner, 2000
pKD46	Gene mutagenesis plasmid; Ap**6**R; repA101ts operon curable at 37°C; carries Red recombinase genes under the control of P*araB*	Datsenko and Wanner, 2000
pCP20	Gene mutagenesis plasmid; Ap^R and Cm^R; ts replicon curable at 43°C; encodes FLP recombinase for FLP mediumted excision of DNA between FRT sites	Cherepanov and Wackernagel, 1995
pDOC-C	Gene mutagenesis plasmid; Ap^R; carries multi-cloning site in between I-*Sce*I digest sites, parent vector to gorging plasmids[1]	Dave Lee (Lee *et al.*, 2009)
pDOCevgASG658A	pDEX; EvgAS operon -607 to 3600 containing EvgS mutation G658A	This study
pDOCevgASS600I	pDEX; EvgAS operon -607 to 3600 containing EvgS mutation S600I	This study
pABCSR	Gene mutagenesis plasmid; Cm^R; carries I-*sce*I gene and Red recombinase genes under the control of P*araB*	Herring *et al*., 2003

Table 2.3 Promoter probe plasmids used in this study.

Plasmid	Description[1]	Source / reference
pLUX	pCS26 derivative; STOP codons, ribosome binding site, and *Nco*I restriction site switch; Parent plasmid for pLUX series2	N. Burton (Burton *et al* 2010)
pLUX*acp*p	pLUX; *acp* -340 to 108	N. Burton (Burton *et al* 2010)
pLUX*b1500p*	pLUX; *b1500* -335 to 119	N. Burton (Burton *et al* 2010)
pLUX*csrAp*	pLUX; *csrA* -390 to 136	N. Burton (Burton *et al* 2010)
pLUX*evgAp*	pLUX *evgA* -600 to 125	N. Burton (Burton *et al* 2010)
pLUX*gadAp*	pLUX; *gadA* -288 to 273	N. Burton (Burton *et al* 2010)
pLUX*gadBp*	pLUX; *gadB* -553 to 273	N. Burton (Burton *et al* 2010)
pLUX*gadEp*	pLUX; *gadE* -868 to 94	N. Burton (Burton *et al* 2010)
pLUX*gadWp*	pLUX; *gadW* -446 to 141	N. Burton (Burton *et al* 2010)
pLUX*gadXp*	pLUX; *gadX* -514 to 114	N. Burton (Burton *et al* 2010)
pLUX*gadYp*	pLUX; *gadY* -326 to 69	N. Burton (Burton *et al* 2010)
pLUX*hdeAp*	pLUX; *hdeA* -323 to 69	N. Burton (Burton *et al* 2010)
pLUX*hdeDp*	pLUX; *hdeD* -363 to 237	N. Burton (Burton *et al* 2010)
pLUX*mgtAp*	pLUX; *mgtA* -446 to 98	N. Burton (Burton *et al* 2010)
pLUX*slpp*	pLUX; *slp* -314 to 70	N. Burton (Burton *et al* 2010)
pLUX*ydePp*	pLUX; *ydeP*-405 to 63	N. Burton (Burton *et al* 2010)

[1]bp from the translation start site of the respective gene

Table 2.4 Vectors used for complementation in this study

Plasmid	Description	Source reference
pZC320[1]	Single copy F replicon cloning vector; Ap^R, parent vector to all complementation plasmid[1]	Shi and Biek, 1995
p*evgAS*cS584F	pZC320 complementation plasmid containing the *evgAS* operon and promoter -607 to 3679[1] from evolved strain Aa. Contains S584F EvgS mutation	This study
p*evgAS*cN573L	pZC320 complementation plasmid containing the *evgAS* operon and promoter -607 to 3679[1] from evolved strain Ba. Contains N573L EvgS mutation	This study
p*evgAS*	pZC320 complementation plasmid containing the *evgAS* operon and promoter -607 to 3679[1] from MGA.	This study
p*evgA*	pZC320 complementation plasmid containing the *evgA* gene and promoter region from -607 to 615[1]	This study
p*rcsB*	pZC320 complementation plasmid containing the *rcsB* gene and promoter region from -1144[2] to 3350[3]	This study

[1] bp relative to the translation start of *evgA*
[2] bp relative to the translation start of *rcsB*
[3] bp relative to the translation strat of *rcsD*

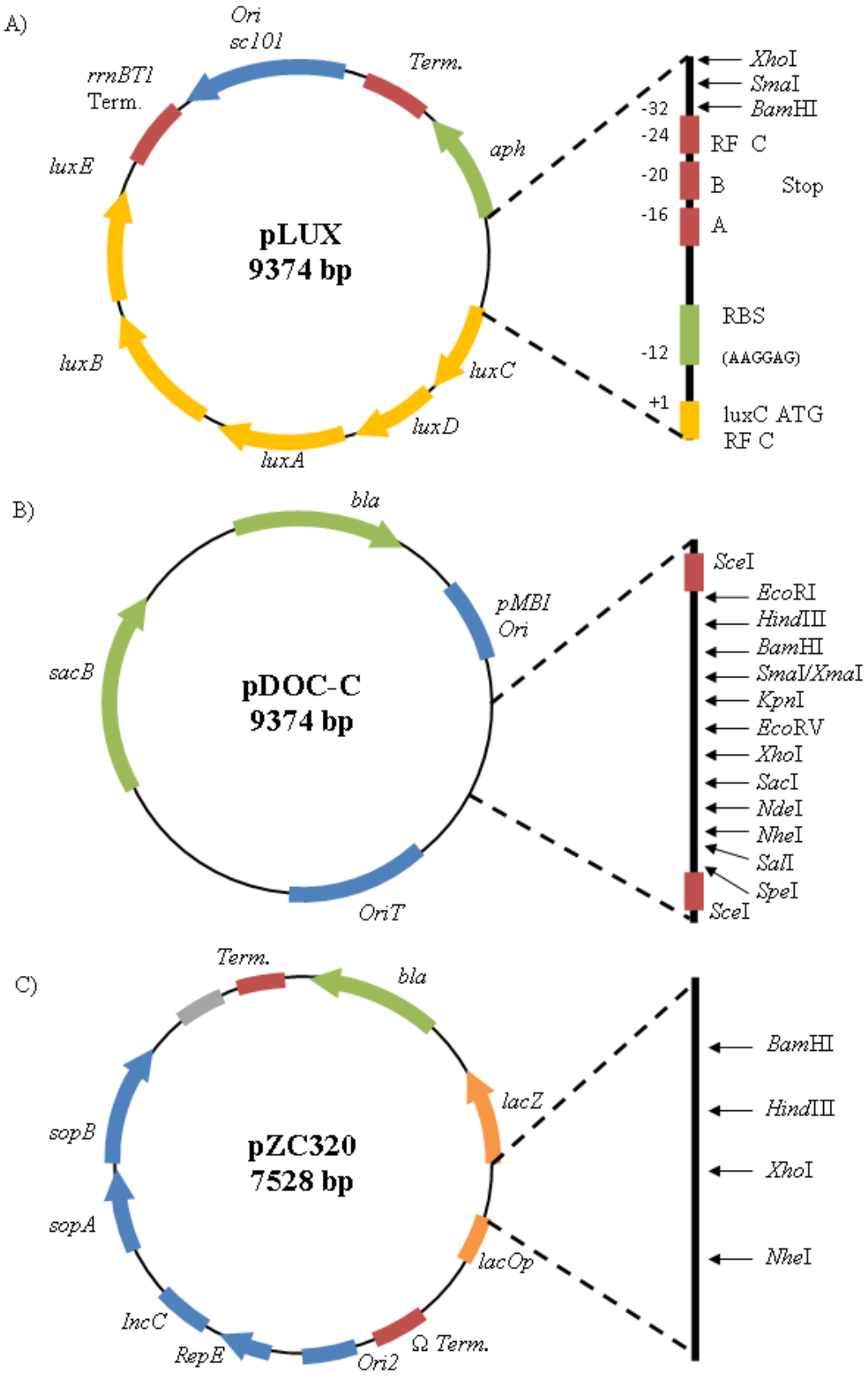

A)
Ori
sc101
Term.
rrnBT1
Term.
aph
luxE
pLUX
9374 bp
luxB
luxC
luxD
luxA
XhoI
SmaI
BamHI
-32
-24
RF C
-20
B
Stop
-16
A
RBS
-12
(AAGGAG)
+1
luxC ATG
RF C
B)
bla
pMBI
Ori
sacB
pDOC-C
9374 bp
OriT
SceI
EcoRI
HindIII
BamHI
SmaI/XmaI
KpnI
EcoRV
XhoI
SacI
NdeI
NheI
SalI
SpeI
SceI
C)
Term.
bla
lacZ
sopB
pZC320
7528 bp
sopA
lacOp
IncC
RepE
Ori2
Ω Term.
BamHI
HindIII
XhoI
NheI

Figure 2.1 Plasmid maps for pLUX, pDOC-C and pZC320. The name and size of each plasmid is shown in the centre of each map. A) Plasmid pLUX, numbers within the multiple cloning site indicate the distance in bp to from the 3' end of each sequence element relative to the start of the *luxC* translation (labelled +1). B) the pDOC-C plasmid, red blocks shown in the multiple cloning site indicate the presence of unique *Sce*I restriction sites. C) Plasmid pZC320.Open reading frames are shown by large coloured arrows, which are labelled with their respective gene names. Restriction sites used for cloning are enlarged and marked with a small black arrow.

2.2 Growth conditions

Unless otherwise stated, *E. coli* was grown in Luria-Bertani broth (LB; 1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v salt; pH 7), at 37°C and with aeration using a shaking incubator at 180 rpm. Luria-Bertani agar (LBA; 1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v salt, 1.5% w/v agar; pH 7) was used as solid medium and where plating of bacteria was required. Both LB and LB agar were made up in distilled water and sterilised by autoclaving by an in house medium and glassware service. M9supp (a rich semi–defined medium based on M9 salts; Table 2.5; pH 7 with KOH; or pH < 5.8 with HCl; cold filtered, Table 2.3) was used for acid resistance assays, high temporal resolution promoter probe assays and single time point promoter probe assays described in sections 2.7, 2.8 and 2.9 respectively. Where required, the amino acid complement of M9supp was altered for assays in the presence of only glutamate (M9E; M9supp containing 0.5 mM glutamic acid in place of cas-amino acids), only lysine (M9K; M9supp containing 0.6 mM lysine in place of cas-amino acids) and only arginine (M9R; M9supp containing 0.6 mM lysine in place of cas-amino acids). For the purpose of selection of strains by resistance to antibiotics, antibiotics were added to growth medium at the following concentrations: ampicillin (Ap), 100 μg/ml; carbenicillin (Cb), 100 μg/ml; kanamycin (Km), 50 μg/ml; chloramphenicol (Cm), 25 μg/ml. Other agents that were added to medium for more specific applications are indicted in the appropriate methods section.

2.3. Custom oligonucleotides

All oligonucleotides used in this study are listed in Tables 2.6, 2.7 and 2.8 and were synthesised by Alta Biosciences (Birmingham, UK) or Eurogentec (Southampton, UK).

Table 2.5. Constituents of M9-supp growth medium.

Compound	Chemical formula	Final working concentration
di-Sodium hydrogen orthophosphate (anhydrous)	Na_2HPO_4	42.3 mM
Potassium dihydrogen orthophosphate	KH_2PO_4	22.1 mM
Sodium chloride	NaCl	8.56 mM
Ammonium chloride	NH_4Cl	18.7 mM
D-Glucose (anhydrous)	$C_6H_{12}O_6$	0.4% w/v; 22.2 mM
BactoTM cas-amino acids (Acid-Hydrolysed casein, low sodium chloride and iron concentrations)	Undefined	0.2% w/v
MOPS	$C_7H_{15}NO_4S$	100 mM
MES Hydrate	$C_6H_{13}NO_4S\ xH_2O$	100 mM (anhydrous basis)
Magnesium sulphate 7-Hydrate[1]	$MgSO_4\text{-}7H_2O$	2 mM
Calcium chloride[1]	$CaCl_2$	0.1 mM
Thiamine (Vitamin B1) Hydrochloride[1]	$C_{12}H_{17}CiH_4O_5HCl$	0.001% w/v (0.03 mM)

[1] Made up as solutions prior to addition: $MgSO_4\text{-}7H_2O$ 1M; $CaCl_2$ 1 M; thiamine 1% w/v

Table 2.6. Primers used in the construction of plasmids made in this study.

Primer name[1]	**Sequence (5' to 3')**	**Restriction site**
evgAS-607	CCGAT**GGATCC**ACTACCTGAGACTTGTGCAG	*Bam*HI
evgAS_R+3679+RS3	CCGAT**GCTAGC**CCACATTTGAACATTGTGGG	*Sal*I
rcsBF-1144	GGTCC**GGATCC**GCGGATGATAGCTGGAAAAGT	*Xho*I
rcsBD+3350R	GGTCC**CTCGAG**TGCGTCTTATCTGGCCTAC	*Bham*HI
evgA+615R	GGTTA**GCTAGC**TTAGCCGATTTTGTTACG	*Nhe*I

[1] Number in the primer name corresponds to the location of the 5'end of the primer relative to the translation start site of that gene.
Restriction sites are indicated in **bold**.

Table 2.7. PCR screening and/or sequencing primers.

Name	Sequence	pos[1].
pCS26_F	GCAATTCCGACGTCTAAG	-118
pCS26_R	GGAAAGATTTCAACCTGG	+59
pKD3_674_R (mut_644_R)	TAACCAGACCGTTCAGCTG	+674
Km_R_+535	CGGAGAACCTGCGTGCAAT	+535
MJrpoS1F	GATTACCTGAGTGCCTAC	-243
MJrpoS1R	ACGACCTACATCTTCCAG	+879
MJrpoS2F	GAGGAACTGTTATCGCAG	-127
MJrpoS2R	TCAGGGTTCTGGATTGTG	+1218
*gadC*_+1616_R	CGTTAAGTCAATACGACAGC	+1616
*evgAp*_-607_F	ACTACTCGAGACTTGTGCAG	-607
*evgAp*_+140_R	TCAGGCTTAAGTGGATCCAC	+140
*evgA*_+697_R	CTACTGATGCCACGATATTC	+697
*b1500p*_-347_F	TCCAGACTCGAGTCACATGC	-347
*ydeO*_+106_R	CGTTATCTACCATCAGGATC	+106
ydeO_+813_R	CCCATTTAATTCTTACGCAGC	+813
evgS_F-44	GGATCTTTACACATTCGC	-44
evgS_F+1081	GAGGATGGGATATAATAC	+1080
evgS_F+2278	CACTCCTCGGCTTAATTG	+2278
evgS_R+1330	CCTTGTGAAATGCAGCGC	+1330
evgS_R+2461	GGTAATGTTCAGGAAACG	+2461
evgS_R+3679	GAACAAATTCGCCAGGAG	+855
rcsB-122_F	CAAGCAGTTATGTGAAACGC	+122
rcsB+84_R	CCACTCAATTTGCTCAAGTG	+84
rcsB+699_R	GCAAATGCCAGATAAGACAC	+699
evgS-110_F	CAGCAACAAAACTGTCAGC	+110
evgS+79_R	CGATGTAATCTTCGTCTGC	+79
evgS_R+2821	GCTGGCTAATTTCTACCG	+2821
F_A0500-113	GGTGGCCAGCGAGAAAGC	O42[2]
R_A0500-944	CAACTCCACATCCTTGCC	O42[2]
EvgS.R.Ha.Ha	GTTCAAAAGCACTATTATGAA	+1139
Ea.evgS.Ea.R	GCATCTTTTCTCGATGG	+1991

[1] 5' primer annealing relative to the start of the gene within the primer name

[2] Used in O42 screen, no homology to MG1655

Table 2.8. Chromosomal gene deletion mutagenesis primers.

Gene	Primer Name	Sequence (5' - 3')[1]	Position[3]
rpoS	rpoSKOF	TGAATGTTCCGTCAATTTATCACGGGTAGG AGCCACCTTGTGTAGGCTGGAGCTGCTTC	-39
	rpoSKOR	CAGCCTCGCTTGAGACTGGCCTTTCTGACA GATGCTTACCATATGAATATCCTCCTTA	+789
gadC	gadCKOF	GAACAAAACAGGTGCGGTTCCGACAGGAAT ACCTGTGTAGGCTGGAGCTGCTTC	-49
	gadCKOR	ATCGCTCCCTTGTCTTATAACCATTCAGACA TGCATATGAATATCCTCCTTA	+1570
b1500	b1500_KO_F-53	TTTTTAACGTTATCCGCTAAATAAACATATT TGAATAGGCTGGAGCTGCTTCG	-35
	b1500_KO_R+237	CTTTTTTAACATTTCATATTTATAATTTGCTG TTTGATGAATATCTCTCTTAG	+544
evgS[2]	evgSsnpKOF	CAGTTCGTCGTCGTAAAGTCATTCAGGGTG ATTTAGGTGTAGGCTGGAGCTGCTTC	+1678
	evgSsnpKOR	CTACCTCGAGTGCATTAATTAGATCACGCG TTTCAGATGAATATCCTCCTTAGTTC	+2100
evgA[4]	evgAMUTF	GATATCGTCATCATTGAGGTCGATATCCCC GGAGT	+138
	evgAMUTR	ACTCCGGGGATATCGACCTCAATGATGACG ATATC	+172
ydeP	ydeP-0KOF	CGCTATTACAAATCCTAATAATTCATTTCCA CACAGGTGTAGGCTGGAGCTGCTTC	-45
	ydeP+2280.KOR	CTGTGCGGATGACAGCAGAAGAAATGAGA AGAGGCAATATGAATATCCTCCTTA	+2315

[1] Underlined bases anneal to pKD3 or pKD4, non-underlined bases anneal to *E. coli* K-12 MG1655 DNA

[2]Knockout region is a region of DNA central to the *evgS* gene containing *evgS* mutations, flanking *evgS* regions remain

[3] bp from the 5' end of the primer to the translational start site of stated gene.

[4] Mutation was made by moving a mutation created by Baba *et al.*, 2006.

2.4. Molecular biology techniques

2.4.1. Preparation of genomic DNA

To prepare genomic DNA from bacterial cells, an overnight culture was grown from a single well isolated colony; 1.5 ml of overnight culture was harvested for lysis and purification using invitrogen PureLinkTM Genomic DNA Mini Kit (Carlsbad, CA, USA) according to the manufacturer's instructions. Briefly, this kit used a proteinase K induced lysis followed by treatment by RNase to digest RNA then purification by selective binding to a silica based membrane.

2.4.2. Preparation of plasmid DNA

In all circumstances a mini preparation was used to isolate plasmid DNA. To prepare plasmid DNA from bacterial cells an overnight culture was grown from a single well-isolated colony and 1.5 – 10 ml of overnight culture (depending on the copy number of the plasmid) was harvested for lysis and purification using QIAprep Spin Miniprep Kit (Qiagen, Crawley, UK) was used to prepare plasmid DNA according to the manufacturer's recommendations. Briefly, this kit works by permeabilising the cells by SDS induced lysis followed by digestion of RNA using RNase and treatment with a chaotropic agent (guanidine thiocyanate) and then purification of plasmid DNA by selective binding to silica gel membranes. DNA was typically eluted into a 35 µl volume for maximal concentration of plasmid product.

2.4.3. Amplification of DNA by Polymerase Chain Reaction (PCR)

PCR used for all purposes other than sequencing and cloning was done using the following protocol. An over-night culture was grown up from a single colony, 200 µl of overnight

culture was harvested by centrifugation and resuspended in 50 µl sterile H_2O. The re-suspension was boiled at 100°C for 10 minutes to lyse the cells. Large cell debris was removed by centrifugation in a micro-centrifuge at 13500 rpm (15000 x g) for 5 minutes and 5 µl of the supernatant was added to a PCR master mix (1x ReddyMixTM PCR Master Mix (Abgene, Epsom, UK; 0.5 units Thermoprime Plus DNA Polymerase; 75 mM Tris-HCl (pH 8 at 25°C); 20 mM $(NH_4)_2SO_4$; 1.5 mM $MgCl_2$; 0.01% (v/v) Tween® 20; 0.2 mM each of dNTPs; precipitant and red dye); 0.25 µM Forward primer; 0. 25 µM reverse primer; SDW) to a final volume of 50 µl. Reactions were loaded into a thermal cycler set to the following programme:

Step 1	Initial denaturation	98°C	5 min	1 cycle
Step 2	Denaturation	98°C	20 sec	18 cycle
Step 3	Annealing	45-60°C[1]	30 sec	
Step 4	Extension	72°C	1 min/ kb	
Step 5	Final extension	72°C	10 min	1 cycle

[1]annealing temperature 5°C less than the lower of the two primer melting temperatures, as recommended by the manufacturer.

2.4.4. PCR for cloning and sequencing procedures

Template DNA was prepared using one of the DNA isolation methods outlined above. To prepare the PCR reaction, 5 µl of template DNA was used per 50 µl PCR reaction, primers were added to reactions at a final concentration of 0.25 µM. PhusionTM High-Fidelity DNA polymerase master mix (Finnzymes, Espoo, Finland) was used to amplify DNA as it has exonuclease activity (proof reading), due to this activity Phusion polymerase was added to the reaction immediumtely before the start of the thermal cycler program. Reaction tubes were then loaded into a thermal cycler set to the following program:

Step 1	Initial denaturation	98°C	1 min	1 cycle
Step 2	Denaturation	98°C	20 sec	18 cycle
Step 3	Annealing	45-62°C[1]	30 sec	
Step 4	Extension	72°C	20 sec/ kb	
Step 5	Final extension	72°C	10 min	1 cycle

[1] annealing temperature 3°C greater than highest of the two primer melting temperatures, as recommended by manufacturer

2.4.5. PCR purification

To remove unused PCR reagents, small olionucleotides (< 30 bp) and protein, a QIAquick PCR Purification Kit (Qiagen, Crawley, UK) was used according to the manufacturer's recommendations. This kit purifies DNA by adhesion to a silica membrane, similar in principle to the QIAprep Spin Miniprep Kit (section 2.4.2 above), the main exception, apart from the absence of a lysis step, is that the silica gel membranes have been optimised for the binding of small DNA fragments. DNA was eluted in 35 µl elution buffer for maximal concentration of PCR product. Where more concentrated samples of DNA were required a Zymo Research DNA Clean & ConcentratorTM kit was used. This kit utilises the same methods as the QIAquick kit and was used per the manufacturer's instructions. This kit differs from the QIAquick PCR Purification Kit as it has a smaller column and elution volume of 6 µl.

2.4.6. DNA analysis by agarose gel electrophoresis

Different percentage agarose gels were used depending on the size of the product to be visualised (typically 1.5% w/v for fragments <1 kb, 1% w/v for fragments >1 kb). Gels were made up in and electrophoresed in TAE buffer for between 25 minutes and 1 hour 20 minutes

at a constant voltage of between 90 and 120 V depending on the size of the gel. A standard marker (Bioline, HyperladderTM I, UK) containing DNA of known sizes and concentrations, was loaded into the gel to allow size measurement and quantification.

2.4.7. Quantifying DNA

The NanoDropTM ND-1000 Spectrophotometer (Labtech Int., Ringmer, UK) was used according to the manufacturer's recommendations to quantify DNA.

2.4.8. Digestion of DNA by restriction enzymes

All enzymes were purchased from NEB (Hitchin, UK) or Fermentas (York, UK). Digests of plasmid DNA were carried out at a final volume of between 30 µl and 50 µl (typically 50 ng/µl – 200 ng/µl, depending on plasmid prep) using manufacturer's recommended buffers at the recommended temperature for typically 3 hours. For fast digest enzymes (Fermentas, York, UK) reactions were incubated as described by the manufacturer for 5 minutes. Where necessary, enzymes were heat inactivated by incubation at 65°C for 20 minutes after digestion. DNA was analysed by gel electrophoresis to verify the correct digest pattern and to check for star activity or degradation.

2.4.9. Ligation of DNA

All ligations including blunt ended and sticky ended DNA fragments were done using a Quick-Stick (QS) Ligase kit (Bioline, London, UK). Ligations were done in 20 µl volumes containing 1x quick stick ligase buffer (stored at 4x concentration in 20 µl aliquots), 1 µl quick stick ligase and 14 µl of vector and insert DNA mixture. The molarity of vector DNA used in ligation varied depending on the size of the vector and the efficiency of transformation

of uncut vector in a control reaction. Insert DNA was added to the range of 3x to 10x molar excess over vector when sub-cloned from another vector, for inserts synthesised by PCR and digested as linear fragments the molarity was in excess of 100x that of the vector. Ligations were left for 15 minutes (cohesive end ligation) or 30 minutes (blunt end ligation) at room temperature. Ligations were checked by agarose gel electrophoresis after heat inactivation at 65°C for 20 minutes and 5 µl was transformed into chemically competent DH5α cells.

2.4.10. Preparation of chemically competent cells

A 5 ml overnight culture was set up by inoculating 5 ml of LB broth containing any appropriate antibiotic selection with a single well isolated colony. A 100 ml day culture was started by inoculating 100 ml of antibiotic free LB broth, pre-warmed to 37°C in a 500 ml flask, with 1 ml of the overnight culture. The day culture was incubated at 37°C, 200 rpm until the culture reached an optical density (absorbance at λ 600 nm) of 0.45. The culture was incubated on ice for 5 minutes, before being aliquoted into four 50 ml falcon tubes (50 ml per tube). The cells were harvested by centrifugation (5,000 x g, 4°C, 10 minutes, swing out rotor, eppendorf centrifuge) and the supernatant was removed. Each cell precipitate was then re-suspended in 40 ml, ice cold 0.1 M $CaCl_2$ before repeating the precipitation by centrifugation. After removing the supernatant each cell precipitate was then re-suspended in 25 ml ice cold 0.1 M $CaCl_2$ and incubated on ice for 30 minutes. After this period the cells were re-precipitated by centrifugation (5,000 x g, 4°C, 10 minutes, swing out rotor, eppendorf centrifuge), the supernatant removed and then the precipitated was re-suspended in 0.5 ml of ice cold 0.1 M $CaCl_2$ 15% v/v glycerol and aliquoted into sterile eppendorf tubes (0.5 ml per tube). Aliquots were and either used directly for transformation (see below), or transferred to the -80°C for long term storage.

2.4.11. Transformation of chemically competent cells by heat shock

To transform chemically competent cells with plasmid DNA, DNA was mixed with competent cells, with a maximum ratio of 1:10 (DNA:competent cells), 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 120 seconds. The DNA-cell mixture was then immediumtely returned to room temperature and 1 ml of LB broth was added. The cells were mixed well by inverting the polypropylelene tube 4-5 times and then incubated for 1-2 hours at 37°C (unless otherwise stated). This gave the cells time to express any relevant antibiotic resistance genes. Depending on the expected frequency of transformation different quantities of cells were then plated out onto selective solid medium.

2.4.12. Preparation of electrocompetent cells

An overnight culture containing 5 ml LB broth in a 20 ml universal tube, inoculated with a single well isolated colony of choice was incubated overnight at 37°C (unless otherwise specified for temperature sensitive plasmids), 180 rpm. A day culture was prepared using a 500 ml conical flask containing 100 ml of LB both, pre-warmed to 37°C, inoculated with 1 ml of the overnight culture. This was grown at 37°C (unless otherwise specified for temperature sensitive plasmids), 180 rpm until the culture reached an OD_{600} of 0.6. At this point the flask was incubated on ice for 5 minutes, then the culture was aliquoted into two pre-cooled 50 ml polypropylene tubes. The cells were precipitated by centrifugation (6,000 x g, 4°C, for 10 minutes), then the clear supernatant was removed and the cells were re-suspended in 40 ml of ice cold, sterile, 10% glycerol. The cells were immediumtely re-precipitated by centrifugation (6,000 x g, 4°C, for 10 minutes) and the supernatant was removed. The cells were then re-suspended in 20 ml of ice cold, sterile, 10% glycerol and immediumtely re-precipitated by

centrifugation (6000 x g, 4°C, for 10 minutes). The previous re-suspension and precipitation steps were repeated with a re-suspension volume of 2 ml and then 0.2 ml. Finally, when the cells were in a final concentration of 2 x 10^{10} cells/ml, they were transferred into 4 sterile 1.5 ml polypropylene tubes (50 µl per tube) for storage at -80°C or immediumte use.

2.4.13. Electroporation

To transform cells by electroporation with plasmid or linear DNA, the DNA was mixed with competent cells, with a maximum ratio of 1:10 µl (DNA:competent cells), on ice by light pipetting. Electroporation cuvettes (Geneflow, Fradley, UK; 1mm gap) were pre-chilled by incubation at -20°C for at least 20 minutes prior to use. The cell/DNA mix was pipetted into the electroporation cuvette and the samples were then electroporated at 1800V (Geneflow electroporator). To allow for adequate expression of antibiotic resistance genes 1 ml of LB broth was added to the cells immediumtely after electroporation, the cells were then incubated at 37°C for 1-2 hours. Depending on the expected frequency of transformation different quantities of cells were then plated out onto selective solid medium.

2.4.14. Dideoxy terminator based DNA sequencing

Plasmid DNA was sequenced in the University of Birmingham Functional Genomics Laboratory, Plasmid to Profile sequencing (Birmingham, UK). In most cases the DNA to be sequenced was amplified using PCR with a proof reading DNA polymerase at 50°C (in line with the conditions used by the sequencing facility). This provided cleaner templates especially for low copy number plasmids.

2.5. High-throughput sequencing

2.5.1 Sequencing

High-throughput sequencing was done by The Genome Institute, Washington University, School of Medicine (St. Louis, USA). Genomic DNA was isolated as previously described and analysed using IlluminaTM high throughput sequencing methods as per the manufacture's guidelines. Genomic DNA isolated from the *E. coli* K12 MG1655 ancestor strain was pooled in equal amounts with each Evolved strain (Ea, Aa and Ba) before sequencing analysis. Sequencing of evolved stain Ga was done separately by the University of Nottingham. The read length was set at 35 bp and paired end runs were used.

2.5.2 Data analysis

Raw sequence reads were analysed by The Genome Institute, Washington University, School of Medicine (St. Louis, USA). FASTA files containing each 35 bp read were then analysed using the my.xbase web font end facility, each read was automatically aligned to the *E. coli* K12 MG1655 reference genome (www.my.xbase.ac.uk). When analysing strains that had been pooled (MG1655, Aa, Ba and Ea) the aligned sequences were used to provide a table of single nucleotide variants that were either present in both the MG1655 and evolved strains (all reads at that position had the same base change) or present in only one strain (half the reads at that position had the same base change). Analysis used the mapping tools incorporated in to the my.xbase.ac.uk facility. When analysing non-pooled samples (Ga) the sequence reads were aligned to the same reference genome as described above, the presence of a single nucleotide variant was present in all reads. The presence of mutations in particular strains was later confirmed by dideoxy terminator sequencing methods.

2.6. Mutagenesis

2.6.1. Site directed mutagenesis

Site directed mutagenesis on plasmid DNA template was done using the QuikChangeTM Site-directed Mutagenesis Kit (Stratagene, Cambridge, UK) according to the manufacturer's instructions. Mutagenic primers were designed according to Stratagene's guidelines. Reactions were set up as follows: 5 µl of 10 x reaction buffer, 60 ng of dsDNA template, 125 ng of each oligonucleotide primer, 1 µl of dNTP mix, 1 µl of *PfuTurbo* DNA polymerase (2.5 U/µl) and SDW to a final volume of 50 µl. The reactions were then run in a thermal cycler set to the following programme:

Step 1	Initial denaturation	95°C	30 sec	1 cycle
Step 2	Denaturation	95°C	30 sec	18 cycle
Step 3	Annealing	55°C	1 min	
Step 4	Extension	68°C	x min*	

* One minute per 1kb of plasmid length.

The PCR product was treated with a *Dpn*I restriction enzyme for 1 hour at 37°C, and heat inactivated for 20 minutes at 65°C, which removed any methylated DNA (DNA not produced by the PCR reaction). The *Dpn*I treated PCR product was then cleaned using a Qiagen QiaquickTM clean-up kit (see above) and 5 µl of PCR product was then transformed into chemically competent DH5α cells by heat shock (see above). Plasmids were isolated from single colony transformants and the presence of site directed mutations confirmed by sequencing.

2.6.2. Gene replacement mutagenesis

Gene replacement mutagenesis was done as described by Datsenko and Wanner (Datsenko and Wanner, 2000). The full process is described below and any modifications from the original protocol are included.

2.6.2.1. Preparation of a mutagenic PCR product

A mutagenic linear fragment was amplified by PCR from one of three templates, either the plasmid pKD3 (containing a chloramphenicol resistance cassette) or the plasmid pKD4 (containing a Kanamycin resistance cassette) or a chromosomal template of a representative mutant from the KEIO library. The primers used are included in Table 2.5. The resulting fragment contained either a Cm cassette (pKD3 template) or a Km cassette (KEIO library template or pKD4 template) flanked by 36 bp homology regions (up to 200 bp when the KEIO library strain was used as a template), which are homologous to regions either side of the gene to be replaced. To obtain a final concentration of mutagenic PCR product, four 50 µl PCR reactions were typically purified using the PCR protocol outlined above and eluted in a 35 µl volume. To remove template pKD3 plasmid the reactions were digested with 1 µl *Dpn*I (20 units) and incubated for 3-4 hours at 37°C. Following this digestion fragments were re-purified.

2.6.2.2. Single gene knock-out mutagenesis protocol

The strain to be mutated was transformed with pKD46 by heat shock (see above), transformants were selected for at 30°C on ampicillin plates. An over-night culture was made with a 20 ml universal containing 5 ml LB + ampicillin was inoculated with a single transformant colony and incubated overnight at 30°C, 180 rpm. A day culture, made with a 100 ml conical flask containing 20 ml of LB, was inoculated with 1 ml of overnight culture and incubated at 30°C, 180 rpm until the optical density of the culture reached 0.3 (A600 nm). The culture was then split into four 1.4 ml aliquots in 1.5 ml polypropylene tubes. Tubes 1 to 3 contained 25 µl of 20% w/v L-arabinose, an equivalent volume of SDW added to the fourth aliquot. These cultures were then incubated at 37°C, 180 rpm, for 80 minutes. This allowed

the expression of the λ red genes from pKD46 in aliquots 1, 2 and 3 the cells for recombination of an incoming linear DNA fragment, whereas aliquot 4, a negative control, does not express λ red genes. Cells were prepared for electroporation as described above. The electrocompetent cells in tubes 1, 2 and 4 were electroporated with 200 ng of purified PCR product (prepared as described above). Tube 3 served as a negative control, containing no DNA. Transformants were selected over-night by growth at 37°C on LB plates with the relative antibiotic to select for the resistance marker. Colonies from plates 1 and 2 were screened for the correct replacement of the gene by PCR using the relevant primer set in table 2.6. An example of the PCR screening procedure is given in Fig. 2.2.

2.6.2.3 Removing antibiotic resistance cassettes

Antibiotic resistance cassettes flanked by FRT sites were removed as described (Datsenko and Wanner, 2000). The plasmid pCP20 was transformed by heat shock into the strain and transformants selected at 37°C on LB + ampicillin agar plates. Ten single colony transformants were grown over-night at 43°C, 180 rpm to induce FLP recombinase synthesis and the curing of pCP20. Loss of the resistance cassette and pCP20 was confirmed by replica plating; loss of resistance to ampicillin confirmed the curing of pCP20 and sensitivity to the antibiotic that was contained on the excised cassette indicated successful removal. Finally, candidates were screened and confirmed by colony PCR using the procedure outlined in Fig. 2.2, using the relevant primer set in Table 2.6.

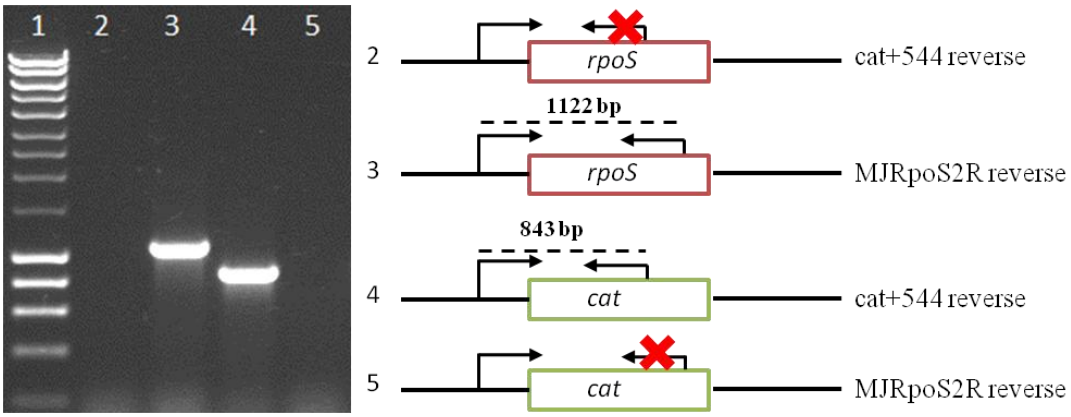


Fig. 2.2. PCR screening for mutations in the MG1655 *rpoS* gene. A chloramphenicol (Cm) cassette was interted into the *rpoS* gene of MG1655. A common forward primer, MJrpoS1F, which annealed -243 bp from the *rpoS* translation start site, was used with two different reverse primers, the first, MJRpoS2R, annealed +879 bp from the *rpoS* translational start site and the second, cat+544 annealed +544 bp from the *cat* translational start site (600 bp from the start of the mutation cassette). The presence of a 1122 bp fragment, when the wild type MG1655 strain was used as the template, and when using the MJRpoS2R reverse primer, confirmed a wild type *rpoS* gene (lane 3). The absence of this fragment, when the mutant candidate was used as the template, indicated the loss of the *rpoS* gene (lane 5). The presence of a 843 bp fragment, when the candidate was used as a template and when the cat+544 reverse primer was used, confirmed the presence and location of the *cat* cassette (lane 4).

2.6.3 Single nucleotide chromosomal mutations

2.6.3.1 Strategy

To make single nucleotide a mutation on the chromosome of strains, without the use of markers, a strategy was devised that incorporated methods of mutagenesis from Datsenko and Wanner 2000, and Lee *et al.,* 2009. Each stage of this process is described in detail below. Briefly, the method involved the generation of a gene knock out, at the locus where the SNV was to be inserted, using Datsenko and Wanner mutagenesis (Datsenko and Wanner, 2000). The cassette introduced was then replaced with DNA homologous to that locus, which contained the single nucleotide mutation. Successful mutations were screened by the loss of antibiotic resistance and sequencing.

2.6.3.2 Gorging plasmid contruction

Template plasmid pDOC-C was used to introduce a linear fragment into the chromosome of the recipient strain. The linear fragments, which contained the *evgAS* locus from either evolved stain Ea or evolved strain Ha, were generated using PCR with the Phusion polymerase and primers evgAS-607 and evgAS_R+3679+RS3 (2.4.4; Table 2.4). These fragments were cloned into the multiple cloning site on the pDOC-C plasmid using the *Bam*HI and *Sal*I restriction sites. This created plasmids pDOC*evgAS*G658A and pDOC*evgAS*S600I, which contain copies of the Ea and Ha *evgAS* loci respectively. Cloning methods are described in section 2.4.9. Selection was by growth on LB +ampicillin.

2.6.3.3 Mutagenesis of the SNV region

The parental strain, *E. coli* K-12 MG1655 was mutated using Datsenko and Wanner mutagenesis as described in 2.6.2. The primers evgSsnpKOF and evgSsnpKOR were used to

make the mutagenic fragment. The region of *evgS* from +1099 to +2821 (bp relative to the *evgS* translation start site) was replaced with a kanamycin resistance cassette. Colonies were identified by the methods described in 2.6.2.2. The strain generated by this mutagenesis contained a Km cassette in place of the SNV region and was named *evgS*snv-.

2.6.3.4 Gene Doctoring protocol

Prior to gene doctoring the *evgS*snv- strain was transformed with either the pDOC*evgAS*cG658A or the pDOC*evgAScS*600I gorging plasmids and with pABCSR. An overnight culture was made by inoculating 5 ml of LB +Km, Cm and Cb with a single well isolated colony of an *evgS*snv- strain containing a gorging plasmid and the pACBSR helper plasmid. A 20 ml LB day culture was inoculated with 1 ml of overnight culture and grown to an OD_{600} of 0.3 with shaking at 180 rpm and 37°C. At this point 200 µl of 20% (w/v) L-arabinose was added to the culture. This induced the expression of the λ-RED genes and the *Sce*I from the pABCSR plasmid. The gorging plasmids were cut by *Sce*I either side of the multiple cloning site creating a linear fragment. The λ-RED enzymes increased the homologous recombination events in the cell. As a result the Km cassette was replaced by a copy of *evgS* containing the single nucleotide mutation. The culture was grown overnight at 37°C, 180 rpm and plated onto LB+Sucrose (0.01 % w/v) agar plates. Any cells containing the pDOC plasmid will also contain the *Sac*B gene rendering the cells Suc^S. Colonies found to be Suc^R were then screened for Km^S. Colonies found to be Suc^R Km^S had successfully replaced the Km cassette with the *evgS* gene. The presence of the *evgS* gene and the single nucleotide mutation were confirmed by PCR and standard Sanger sequencing.

2.7. Acid resistance assay

2.7.1 Plate setup

A clear polypropylene 96-well microtitre plate was used to dilute cultures into pH 7 and pH 2.5 medium. The plate was set up by adding 180 µl of M9supp pH 7 into 6 wells in a row. This was repeated for the M9supp pH 2.5 medium and each strain tested required 1 pH 7 row and 1 pH 2.5 row. The plate and its contents were warmed to 37°C for 30 minutes prior to the start of the assay.

2.7.2 Assay protocol

Stains to be assayed were used to inoculate 5 ml M9supp (with appropriate antibiotic). Cultures were incubated overnight at 37°C, 180 rpm. The overnight cultures were diluted into 5 ml M9supp (with appropriate antibiotics) to a starting OD_{600} of 0.005. Cultures were incubated at 37°C, 180 rpm to an OD_{600} of 0.2. At this point the cultures were diluted 10-fold 7 times in pH 7 M9supp and in pH 2.5 M9supp (no antibiotics) using the 96-well plate set up in section 2.6.1. Each dilution required to transfer of 20 µl of culture into 180 µl of M9supp medium. The 10^{-2} to 10^{-6} pH 7 dilutions were immediumtely spotted onto square LB plates. Once spotted, the plates were tilted to spread the spot into a line down the plate (Figure 2.3). This represented the time zero pH 7 reading. The 96-well plate was incubated at 37°C for 2 hours. After incubation the spotting procedure was repeated with the pH 2.5 wells. The square LB plates were incubated overnight at 30°C (to prevent colonies from growing too large to count). The survival of each strain was expressed as a percentage, which was calculated by dividing the number of colonies scored from the pH 2.5 dilutions by the number of colonies scored from the pH 7 dilutions and multiplying by 100. The most concentrated dilution that could be counted was used in each case.

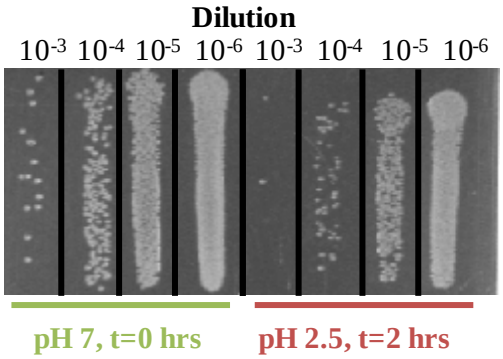


Figure 2.3 Example of an acid resistance assay. The figure shows the colonies from an acid resistance assay. The dilution factors are shown in to top of the figure and the conditions at the bottom. The lines are created by tilting the plate once the cultures have been spotted. Colonies are counted by eye.

2.7.3. Acid-induction acid resistance assay

The protocol described in section 2.7.2 was modified to assay the survival of strain that had been induced by a mild acid shock. A second 96-well microtitre plate was used to acidify the cultures before acid shock. An acidification well and a pH 7 well was set up for each culture to be assayed. The acidification well contained 15 µl of 2 M HCl and 25 µl of sterilised H_2O. The pH 7 well contained 40 µl of H_2O. When the cultures to be tested reached an OD_{600} of 0.2, 240 µl of each culture was transferred to both the acidification well and the pH 7 well. The plate was incubated at 37°C for 70 minutes. After incubation both the pH 7 and the acidified cultures were assayed for survival by dilution into pH 7 and pH 2.5 M9supp as described in 2.7.2.

2.7.4. Alterations made to assay cultures in stationary phase

To measure the acid resistance of strains during stationary phase, the following alterations were made. Overnight cultures were not sub-cultured into day culture, instead the overnight cultures were immediumtely diluted into M9supp at pH 2.5 and pH 7. The cultures were diluted to 10^{-8} and the dilutions 10^{-4} to 10^{-8} were plated as previously described.

2.8. High temporal resolution promoter probe assays

2.8.1. Overview

In preparation for this assay the strains to be tested were transformed with the reporter pLUX plasmids as required. The assay to measure the promoter activity of using the pLUX plasmids can be simplified in to the following steps, outgrowth into exponential phase, measurement of exponential phase promoter activity, acidification by dilution into HCl, measurement of promoter activity during induction by acidification and finally data analysis. The protocol

requires two 96-well polypropylene plates with white walls and clear bottoms, Fisher Scientific, Loughborough, UK. The sections below will cover the preparation of these plates, the assay itself and the data analysis.

2.8.2. Outgrowth

A single colony of a strain that has been transformed with a reporter plasmid, which was less than one week old, was used to inoculate 5 ml M9supp medium in a 20 ml Sterilin tube. This culture was incubated at 37°C, 180 rpm overnight. The overnight culture was used to inoculate a 25 ml conical flask containing 5 ml M9cas, which had been pre-warmed to 37°C. The overnight was diluted in this flask to a starting OD_{600} of 0.005. The culture was incubated in a shaking water bath at 37°C, 180 rpm. The incubation was stopped when the cultures reached an OD_{600} nm of 0.18.

2.8.3. 96-well polypropylene plate set up

During the outgrowth of cultures to be tested, two 96-well plates were set up. Plate 1 was prepared by adding 280 µl of H_2O to all wells not used by the assay. Each culture to be tested required 6 wells in a column, other samples could not be measured in adjacent wells and the outside wells were not used to avoid edge effects. As a result only 6 individual cultures could be assayed at one time (5 strains and 1 control). This plate was pre-warmed at 37°C, 30 minutes prior to use. Plate 2 was prepared the same as plate 1 except for the following additions. The top three wells (B-D) of each column are to contain acidification solution, which consists of 15 µl of HCl and 25 µl of H_2O that will shift the pH of the culture to 5.7. The bottom three wells (E-G) of each column were prepared by adding 40 µl of H_2O. Cultures added to these wells were diluted, as in the acidification wells, but not acidified. The bottom

three wells therefore act as a non-acidified comparison. Plate 2 was incubated at 37°C, 30 minutes prior to use.

2.8.4. Measuring promoter activity

Once the cultures had grown to OD_{600} 0.18 the cultures were transferred to plate 1. A multichannel pipette was used to transfer 240 µl of culture into wells B-G of each empty column. This was repeated for each culture to be assayed. Plate 1 was put into the Fluoroskan Ascent, Thermo Scientific, Basingstoke. The plate reader incubated the cultures at 37°C with shaking and measured the lux values from each well every 40 seconds. After every 12 readings the plate was removed from the Fluoroskan and placed in the Multiscan Plate reader to measure the OD_{600}. After 24 readings in the Fluoroskan, the cultures were acidified to pH 5.7 by transferring 280 µl from each well of plate 1 into plate 2. Each strain was transferred at a set time point in order to account for the delay in acidification between strains. Due to the preparation of plate 2, the wells B to D of each column contained acidified cultures and wells E to G of each column are non-acidified. The plate 2 was then returned to the Fluoroskan for continued measurements. A further 9 cycles of 12 lux measurements and 1 OD_{600} measurement were completed before the assay was stopped. The assay measures the lux of each well 108 times and the OD_{600} of each well 9 times over the course of 2.75 hours. Details of how the lux values are corrected for OD_{600} and time discrepancies will be explained below.

2.8.5. Data analysis

Luciferase values from the at each time point pre- and post-induction were exported from the Fluoroskan Ascent to Microsoft Excel for analysis. The Fluoroskan recorded both the lux measurement and the timings, which aids in the analysis. The 9 OD_{600} measurements were

exported from the Multiskan MS. The values were corrected for the blank, which was determined as the average OD_{600} measurement from wells containing fresh M9supp. The 9 corrected OD_{600} readings were used as a standard by the R statistical software to generate predicted OD_{600} readings for all time points based on best fit parameters of a logistical growth curve equation. The lux reads were divided by the respective OD_{600} reads to give lux per cell, per well. Each strain is measured in triplicate so the mean and standard deviation of the 3 technical replicates of each culture in both acidified and non-acidified conditions were calculated. The technical repeats were used as an indication of the variance between wells on the plate and technical reproducibility. All assays were repeated 3 times and the values and error bars in this thesis represent the mean and standard deviation of 3 biological repeats.

2.9. Single time point promoter probe assays

The preparation and outgrowth of these assays was done as described in section 2.8.2 except for the following. The outgrowth was done in a clear 96-well polypropylene plate which was prepared by adding 280 µl of H_2O to all outside wells and 270 µl of M9supp to all other wells. The plate was pre-warmed at 37°C for 30 minutes prior to use. The overnight cultures were used to inoculate the each well of the plate. This allowed up to 60 strains to be compared on one plate. The cultures were diluted to a starting OD_{600} of 0.005 and covered with a gas-permeable adhesive seal. The cultures were incubated at 200 rpm, 37 °C until the OD_{600} of all cultures was between 0.15 and 0.20.

The OD_{600} reading was measured in the clear plate before 200 µl of each well was transferred into a second plate for the lux reading. Plate 2 was prepared in advance by adding sterile H_2O to all outside wells and pre-warmed at 37°C 30 minutes prior to use. Plate 2 was a white

walled, white bottomed 96-well polypropylene plate that was used to reduce carry-over of signal from adjacent wells. The use of these plates also amplified the signal, resulting in much higher values compared to the high temporal resolution promoter probe assays, which use a clear bottomed white walled plate. The lux values were corrected for both OD_{600} and volume by dividing the lux value first by the OD_{600} and then each value by 0.2 ml. Each assay was repeated a minimum of 3 times and the values and error bars presented in this thesis represent the mean and standard deviation of these repeats respectively.

2.9.1 Alterations of the single time point promoter probe assay to measure stationary phase activity

To measure the promoter activity during stationary phase, the following alterations were made. Overnight cultures were not diluted into day cultures, instead the overnight cultures were diluted 10 fold into pre-warmed medium in a 96-well polypropylene plate. The cultures were aerated by incubation with shaking for 30 seconds. This process ensured that the cultures were aerated enough to allow the luciferase reaction to occur. The cultures were then read using the Fluoroskan Ascent. Data analysis was as described in 2.9.

CHAPTER 3: RESULTS (1): Characterisation of evolved acid resistance in *Escherichia coli*

3.1 *E. coli* acid resistance and the evolution of an acid resistant phenotype

E. coli K-12 MG1655 is resistant to extreme acid challenge (pH 2.5) during stationary phase, in rich and in minimal medium (De Biase *et al.*, 1999; Lin *et al.*, 1995). However, during exponential phase *E. coli* K-12 MG1655 is sensitive to extreme acid shock. This sensitivity can be overcome with pre-induction with a mild acid shock (pH 5.7) (Burton *et al.*, 2010; Tucker *et al.*, 2002). Underpinning these phenotypic phenomena is a complex regulatory network that controls many of the genes responsible for acid resistance. Acid resistance is regulated by local regulators GadE, YdeO, GadX, AdiY, CadC and GadW (De Biase *et al.*, 1999; Tramonti *et al.*, 2006), global regulators H-NS and CRP (Giangrossi *et al.*, 2005; Ma *et al.*, 2003b), two-component systems such as EvgAS, PhoPQ and RcsBD (Burton *et al.*, 2010; Castanie-Cornet *et al.*, 2010; Eguchi *et al.*, 2011; Eguchi *et al.*, 2004; Johnson *et al.*, 2011; Krin *et al.*, 2010b; Masuda and Church, 2003), alternative sigma factor RpoS, and Lon protease (Heuveling *et al.*, 2008; Ling *et al.*, 2008). The acid resistance (AR) phenotype of *E. coli* varies greatly between strains (Gorden and Small, 1993). Structural genes implicated in acid resistance are conserved amongst all *E. coli*, which suggests that the variation between the levels of *E. coli* acid resistance is due to the regulatory factors controlling the acid resistance genes. The regulation of acid resistance genes by this network is not fully understood.

Recently, several studies have used short term lab-based evolution experiments to evolve new phenotypes using micro-organisms, for a recent review see Conrad *et al.*, 2011. The mutations generated by evolution experiments can reveal the function of novel genes, implicate genes of known function in new processes and perturb regulatory networks. Such approaches require high throughput technologies such as whole genome re-sequencing, transcriptomics and

phenomics to find and understand the relevance of the acquired mutations. In this study we aimed to understand the regulation of *E. coli* acid resistance using lab-based evolution and whole genome techniques. In this first chapter, data is presented that confirms the acid resistant phenotype of five cell lines that were generated by repeated selection by extreme acid shock. The evolved phenotype of all five cells lines was dissected using phenotypic assays. The general aim was to define the limits and conditions of the evolved acid resistant phenotype(s), and to understand what causative genetic changes have occurred. Particular attention was paid to the differences, if any, between the evolved cell lines. In addition, while characterising the evolved strains, the resistance of the ancestor strain will also be analysed.

3.1.1. Lab-based evolution experiment

E. coli K-12 is sensitive to acid shock during exponential phase growth without acid induction. A lab-based evolution experiment was used to evolve acid resistant strains in these conditions. Eight cell lines were evolved by iterative exposure to extreme acid challenge. This resulted in more survival after extreme acid challenge (unpublished work by P.B Russell). This section will discuss the method used to generate the evolved strains. All results described in 3.1.1 were done by P.B. Russell (unpublished work) prior to the start of this study. To evolve strains in the lab, an artificial selection was applied, the method for selecting acid resistance was as follows. The medium used in the evolution experiment was rich LB. Eight different cell lines were grown in shake flasks at 180 rpm, 37°C until they reached an OD_{600} of 0.3 (predetermined to be exponential phase in LB). Once at OD_{600} 0.3, the cells were serially diluted 10-fold, seven times, into pH 2.5 LB (called shock cultures). After the cultures were shocked at pH 2.5 for 20 minutes, they were recovered overnight in pH 7 medium. The next day, each dilution was analysed for growth. The most dilute overnight was used to seed a

day culture. This process was then repeated until a significant increase in acid resistance was observed. To track the acid resistance during the evolution experiment, the cells were tested at set intervals. The strains were tested for acid resistance by the standard acid resistance assay described in Chapter 2. The evolution experiment was stopped once a significant (greater than 10-fold increase in acid resistance) acid resistant phenotype was observed. A single clonies were isolated from each cell line and tested for acid resistance. Finally, an acid resistant clone from each cell line was stored at -80°C.

Figure 3.1 shows a graph of the progress of acid resistance during the evolution experiment. The resistance of the eight cell lines is higher compared to the ancestor (un-evolved *E. coli* K-12 MG1655). It was also noted that the acid resistant phenotype increased rapidly. As a result, the evolution experiment was halted after 17 days. Finding and understanding the mutations that confer this increased resistance could enhance our understanding of acid resistance regulation in *E. coli*. This chapter will continue by characterising the acid resistance phenotype of the ancestor and evolved strains.

3.2. Comparison of ancestor and evolved strains by growth

The acid resistance phenotype depends on the phase of growth that the cells are in. In *E. coli* K-12 MG1655, cells are more sensitive to acid shock at extreme pH during exponential phase. This sensitivity is not found in stationary phase cells (De Biase *et al.*, 1999). As the evolved strains may have mutations that affect growth, it is important to assay the growth characteristics of each strain for two reasons. Firstly, to gain insights into the mutations which may be causing the acid resistant phenotype; do the evolved strains grow slower during exponential phase?

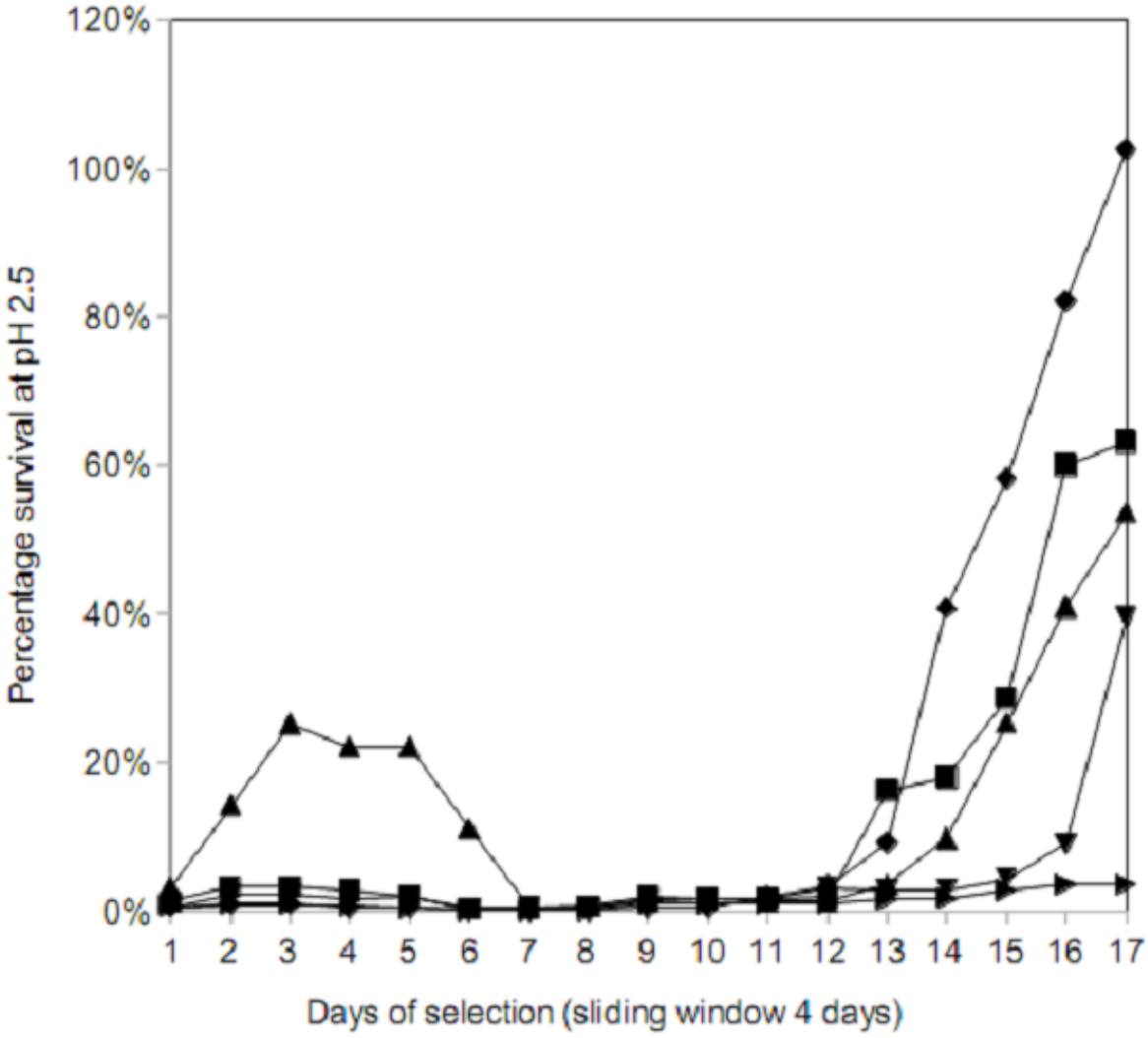
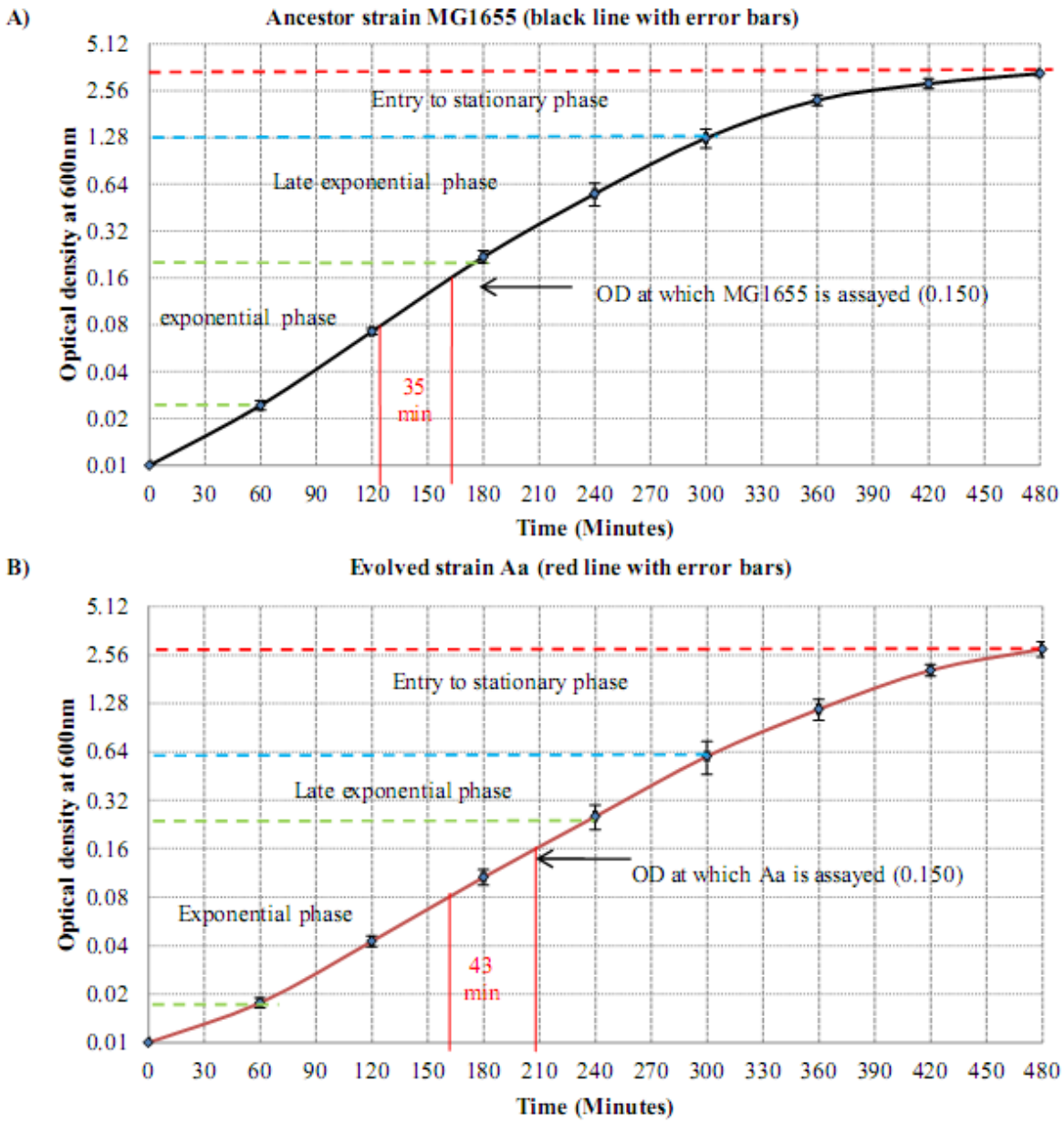


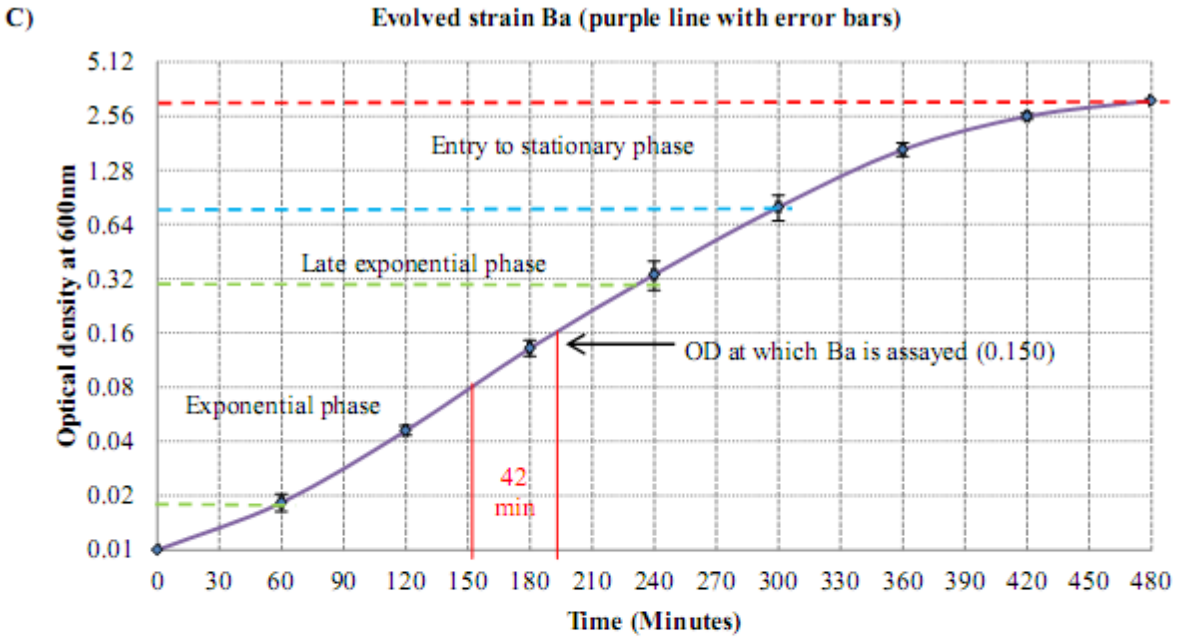
Figure 3.1. The evolution experiment. Survival of five individual cell lines evolving over 20 days of iterative exposure to extreme acid (pH 2.5). Each cell line was exposed to pH 2.5 LB for 20 minutes during exponential growth in LB media; survivors were rescued in pH 7 LB medium and grown overnight, this process was repeated using the overnight cultures to seed new day cultures. The five individually evolving cultures are shown as follows: A (squares), B (diamonds), E (inverted triangles), G (forward triangles), H (upright triangles). Resistance levels for each culture were measured every day for twenty days, and averaged over a four day sliding window to remove small fluctuations. (Thus the level shown for day 1 is the average for days 1 to 4). Survival was calculated by dividing the colony count of cultures subjected to 2 hours of pH 2.5 acid shock by the colony count of the same culture after 0 hours at pH 7 (Russell, unpublished data)

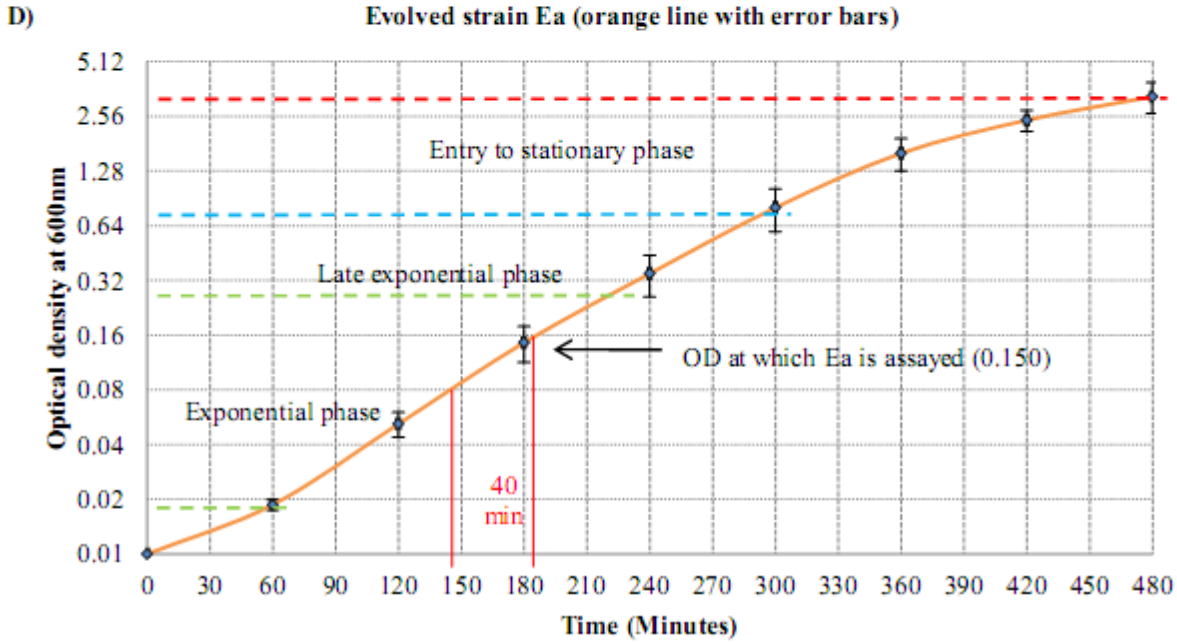
Do the evolved strains enter stationary phase earlier or later than the ancestor? Secondly, as different strains maybe at exponential phase at different optical densities, the OD_{600} that acid resistance is measured must be standardised so that all strains are measured during exponential phase.

Figure 3.2 shows the growth curve analysis of the ancestor (A) and the evolved strains Aa to Ha (B-F) in M9supp (minimal medium supplemented with casamino acids, MOPS and MES; see Chapter 2 for details). The exponential phase doubling time varies between the ancestor, which was 35 minutes, and evolved strains, between 40 and 43 minutes. The reduced doubling time of the ancestor strain in M9supp medium, compared to doubling rates in rich LB medium, suggests that despite containing glucose and casamino acids, the M9supp medium is sub-optimal for *E. coli* growth. The difference in optimum growth rate between ancestor and evolved strains is suggestive of a reduction of fitness at pH 7. No significant difference is seen in final absorbance after 8 hours of growth, although entry into stationary phase does vary between strains. The ancestor enters late exponential phase at 0.2 OD_{600} and stationary phase at 1.28 OD_{600}. The evolved strains are all in exponential phase growth between OD_{600} 0.15 and 0.2.

In summary, all the evolved strains grow slower than the ancestor during exponential phase. In addition, they appear to enter stationary phase growth earlier. The mutations have affected growth in all cases.


A)
Ancestor strain MG1655 (black line with error bars)
Optical density at 600nm
5.12
2.56
1.28
0.64
0.32
0.16
0.08
0.04
0.02
0.01
Entry to stationary phase
Late exponential phase
exponential phase
OD at which MG1655 is assayed (0.150)
35 min
0 30 60 90 120 150 180 210 240 270 300 330 360 390 420 450 480
Time (Minutes)
B)
Evolved strain Aa (red line with error bars)
Optical density at 600nm
5.12
2.56
1.28
0.64
0.32
0.16
0.08
0.04
0.02
0.01
Entry to stationary phase
Late exponential phase
Exponential phase
OD at which Aa is assayed (0.150)
43 min
0 30 60 90 120 150 180 210 240 270 300 330 360 390 420 450 480
Time (Minutes)


C)
Evolved strain Ba (purple line with error bars)
Optical density at 600nm
5.12
2.56
1.28
0.64
0.32
0.16
0.08
0.04
0.02
0.01
Entry to stationary phase
Late exponential phase
OD at which Ba is assayed (0.150)
Exponential phase
42 min
0 30 60 90 120 150 180 210 240 270 300 330 360 390 420 450 480
Time (Minutes)


D)
Evolved strain Ea (orange line with error bars)
Optical density at 600nm
5.12
2.56
1.28
0.64
0.32
0.16
0.08
0.04
0.02
0.01
Entry to stationary phase
Late exponential phase
OD at which Ea is assayed (0.150)
Exponential phase
40 min
0 30 60 90 120 150 180 210 240 270 300 330 360 390 420 450 480
Time (Minutes)

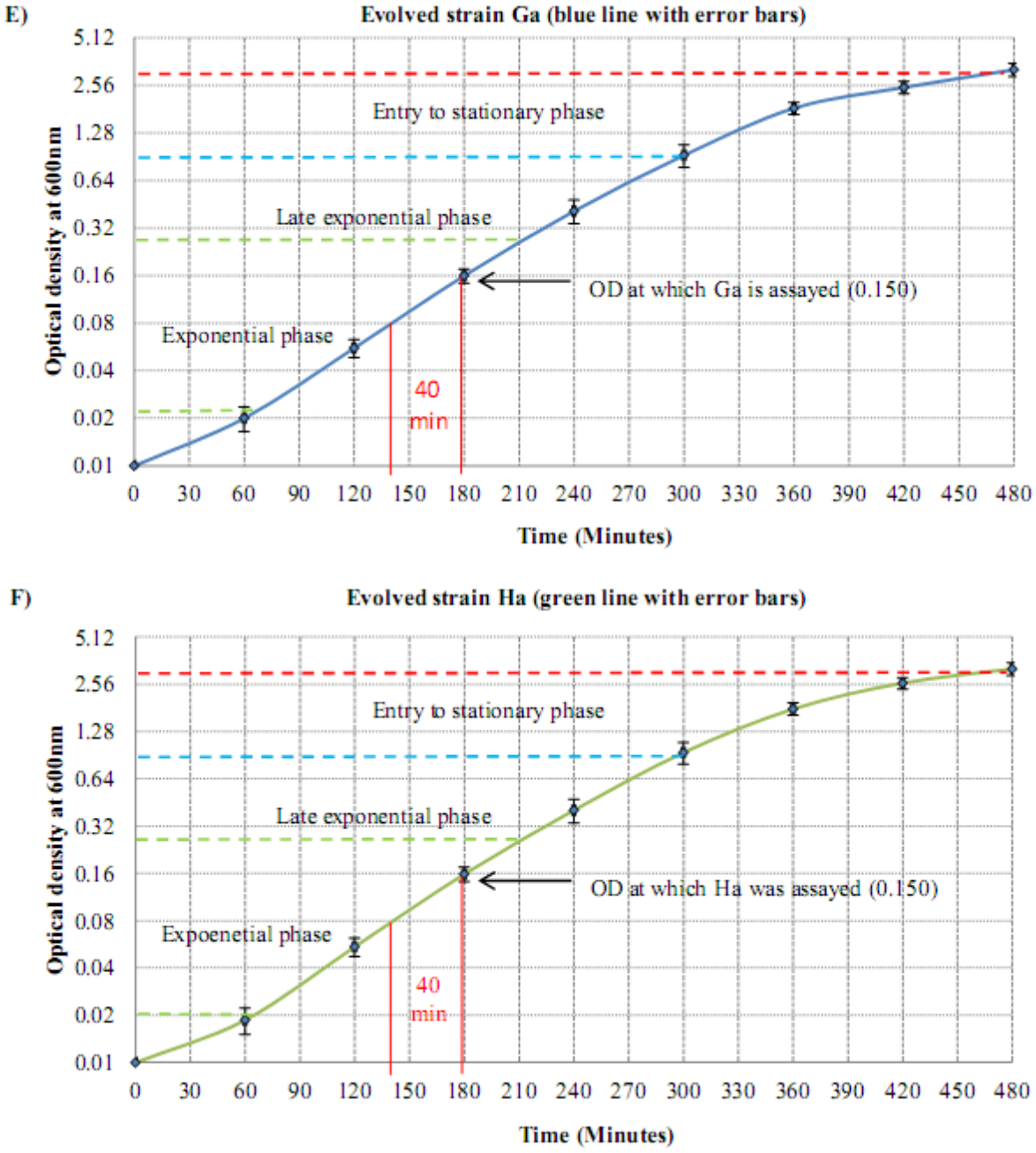


Figure 3.1. Growth curves of Ancestor and evolved strains. Growth of the ancestor strain (A) and five evolved strains (B-F) in M9supp media. Each strain was grown overnight in M9supp and then diluted to a starting OD_{600} of 0.005 in 50ml M9supp in a 250ml glass conical flask and grown with shaking at 37°C for 8 hours. Values represent the average of three independent biological repeats and error bars represent the standard deviation of these repeats plotted on a log base 2 x-axis. Coloured dashed lines indicate different growth phases as marked on the figure. The doubling times and the optimal exponential phase OD_{600} are indicated on the figures for each strain.

As previously stated, the growth phase of *E. coli* is intertwined with the acid resistance phenotype. Therefore, it is important to ensure that all of the acid resistance assays are done during exponential phase and with enough out-growth from overnight culture. This is needed to ensure that no stationary phase factors affect acid resistance. At least 4 generations are required to ensure that the cells have enough out-growth from a stationary phase overnight culture (Burton *et al.*, 2010). The standard acid resistance assay starts at 0.005 OD_{600} (see Acid resistance assay Chapter 2) and the ancestor and evolved strains are all in exponential phase of growth between 0.150 and 0.2 OD_{600}, indicated in figure 3.2. This is 4.5 generations of out-growth, thus the optimum absorbance for all assays will be 0.15 OD_{600}.

3.3. Resistance of the ancestor and evolved strains to extreme acid challenge in LB and M9supp medium

In order to assay the acid resistance of the evolved strains, and compare their phenotype with the ancestor, the % survival of each strain was measured using an acid resistance assay. As the lab based evolution experiment was done using rich, undefined, LB medium the strains were originally compared in this medium. However, to reduce technical variation, acid resistance assays were done using M9supp, which is a defined medium. The full constitution of M9supp is shown in Chapter 2. This section will compare the acid resistance of the evolved strains and the ancestor in both medium types.

Acid resistance was tested using a standard acid resistance assay, as described in Chapter 2. In order to assess the resistance of the evolved strains, cells were initially tested in pH conditions that emulated the selection pressure used in the evolution experiment (pH 2.5). The acid shock time period was chosen as 2 hours as this is comparable with published acid resistance

assays (Castanie-Cornet *et al.*, 1999). Briefly, strains to be assayed were grown from OD_{600} 0.005 to OD_{600} 0.15 (OD_{600} 0.15 was predetermined to be exponential phase for all of the strains, see section 3.2). At this point each culture was serially diluted 10-fold 7 times into pH 7 and pH 2.5 M9supp. A time zero sample was taken immediumtely after dilution by spotting each dilution onto LB agar. After 2 hours the cells incubated in pH 2.5 medium were also plated on LB agar. The plates were incubated at 30°C overnight after which % survival was calculated by dividing the scored colonies after two hours at pH 2.5 by the scored colonies at pH 7, time zero.

Figure 3.3 shows the survival of the ancestor and evolved strains after 2 hours in M9supp at pH 2.5. The evolved strains Aa, Ba, Ea, Ga and Ha show 63, 53, 21, 42 and 44% survival respectively compared to the ancestor that has 1% survival in the same conditions. Values represent the average of three biological repeats. The survival of the evolved strains was compared to the ancestor strain by t-test, which indicated that the difference was significant. When the evolved strains are compared to each other no differences in survival levels are observed in these conditions. These values are averages of three biological repeats and are expressed on a graph with a log base 10 scale on the vertical axis to enable comparison on the same graph. The results show that survival levels of the ancestor strain are elevated by growth in M9supp compared to published levels of acid resistance of *E. coli* K-12 MG1655 (Castanie-Cornet *et al.*, 1999; Lin *et al.*, 1995; Lin *et al.*, 1996). Increased survival in nutrient poor medium has been shown previously (Lin *et al.*, 1995).

The results in figure 3.3 also show the survival of the Ancestor and evolved strains in LB. The survival level of the Ancestor in LB was below the limit of detection for this assay.

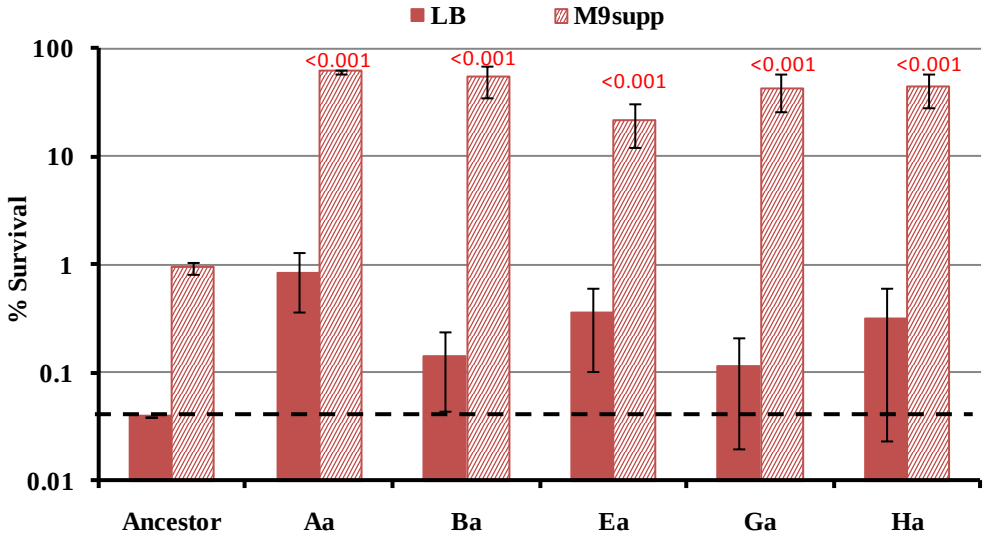


Figure 3.3. Survival of the ancestor and evolved strains in LB and M9supp media after 2 hours at pH 2.5. **S**urvival represents the scored cells after 2 hours at pH 2.5 in LB (solid red bars) and M9supp (stripped red bars) expressed as a % of total cells at time zero. Dotted line represents the limit of detection predetermined as 0.04%. Values and error bars represent the average and standard deviation respectively. Calculated from three independent biological repeats. Comparison of the survival of the evolved strains to the ancestor in M9supp media was made by T-test, numbers in red represent the p values from these comparisons.

This level of survival is comparable to published levels for *E. coli* K-12 MG1655 (Castanie-Cornet *et al.*, 1999). The survival of the evolved strains is also affected by the growth medium and is significantly lower in rich medium (Figure 3.3).

Analysis of the acid resistance of five clonal isolates from the lab based evolution experiment, in conditions similar to the selection pressure applied during the evolution of acid resistance, shows that the five isolates, Aa-Ha, are significantly more resistant to extreme acid stress compared to the ancestor. The resistance of the ancestor and the evolved strains is elevated by growth in M9supp medium. It should be noted that the use of a defined medium increases the reproducibility of the acid resistance assay and allows more quantifiable comparisons between strains (based on the comparison of standard deviations). As the ancestor and evolved acid resistance have been quantified, a more comprehensive analysis of acid resistance was possible. Specific details about the range of acidities that the evolved strains can resist, and dependence on particular components of the medium or specific acid resistance mechanism, could reveal differences between evolved resistance phenotypes and give insights as to which mechanisms underlie the evolved phenotype(s).

3.4. Analysis of evolved resistance

3.4.1. Resistance to other levels of acidity

The lab-based evolution experiment exposed the cells to medium at pH 2.5 as a selective pressure. This selective pressure increased the prevalence of mutations that conferred resistance to these conditions. To investigate the specificity of the evolved resistance to a particular pH range, the survival of the five evolved strains and the ancestor were compared over a range of acidities. The mechanisms involved in acid resistance provide protection at

different acidities (Castanie-Cornet *et al.*, 1999; Foster, 2004; Lin *et al.*, 1996). Thus, comparing the strains over a range of acidities may reveal which acid resistance mechanisms have been affected by the evolution process. Considering, that the evolution experiment selected for mutations that aid survival at pH 2.5 we hypothesise that the evolved resistance is specific to this pH. The cells were assayed using the acid resistance assay as described above.

Figure 3.4 shows the survival of the evolved and ancestor strains after two hours of exposure to medium ranging in pH from 5 to 1. The ancestor strain showed a reduction in survival as the pH of the challenge medium is lowered from 5 to 2.5 (present as a dotted line on each graph A to E). In addition, the ancestor showed no detectable survival at pH 2 or below. In contrast, the evolved strains showed no significant reduction in survival between pH levels 5 to 2. At pH 1.5 the survival of the evolved strains decreased rapidly to levels more than 20-fold less than at pH 2 and above. There was no detectable survival of strains at pH 1. These results have indicated that the evolved resistance is not limited to pH 2.5 and that the resistance mechanism can protect the cells at higher pH levels. However, the survival levels are significantly reduced at acidities lower than pH 1.5 indicating that the evolved resistance is limited to pH 2.0 and above. It is also important to note that there is no difference between the evolved strains when comparing their resistance phenotypes over a range of pH levels. Previous work on the resistance mechanisms of *E. coli* has shown that the GAD system is the most effective at low pH (Lin *et al.*, 1995). However, the level of survival is also high during less acidic conditions, which could indicate a role for other AR systems.

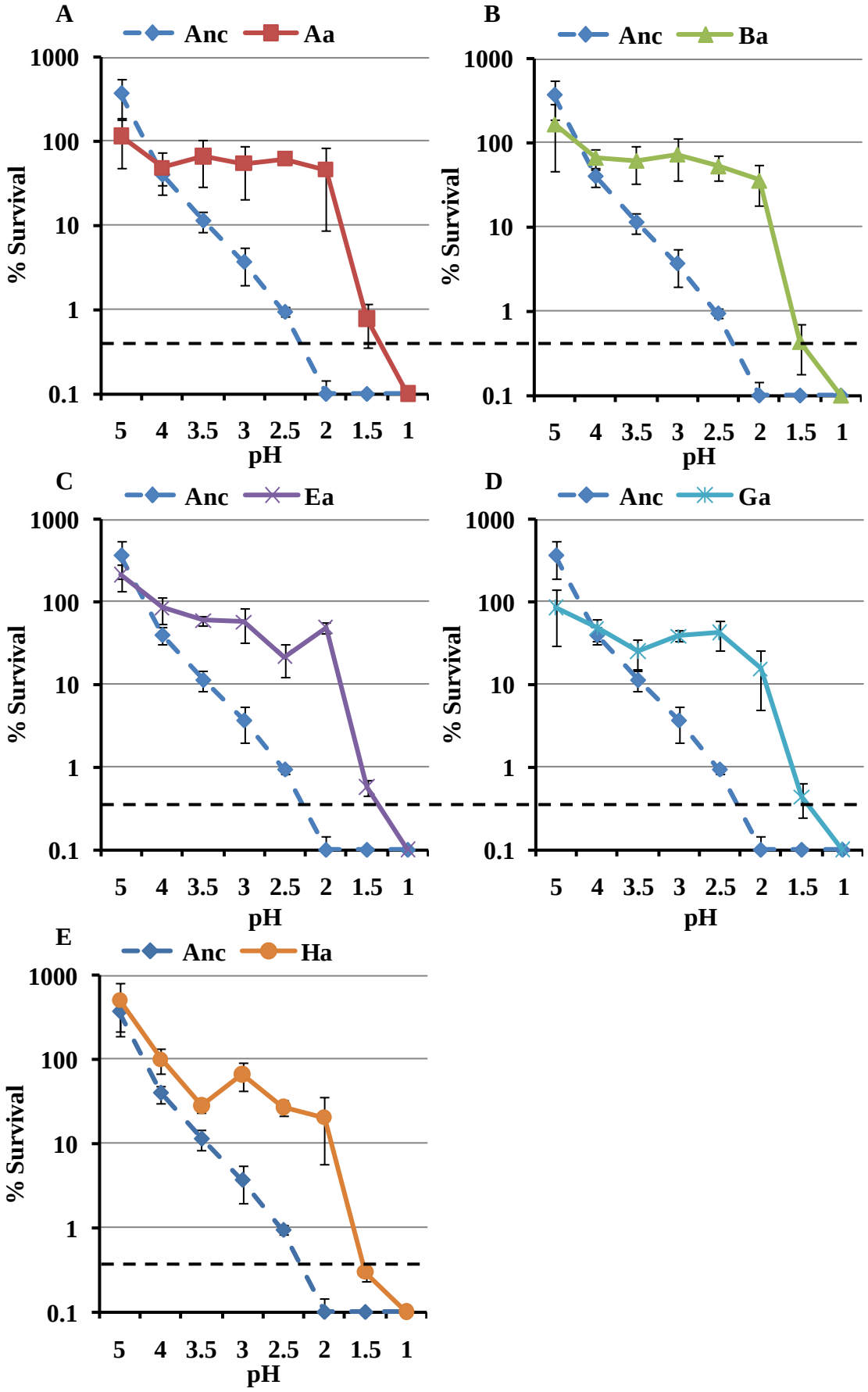

A
Anc
Aa
B
Anc
Ba
C
Anc
Ea
D
Anc
Ga
E
Anc
Ha
% Survival
1000
100
10
1
0.1
5 4 3.5 3 2.5 2 1.5 1
pH

Figure 3.4. Evolved resistance over a range of acidities during exponential phase growth. Graphs A to E show the survival of evolved strains Aa to Ha respectively after 2 hours at a range of acidities in M9supp media during exponential phase growth. The dotted line on each graph indicates the % survival of the ancestor strain (MG1655). Survival was determined by scoring the cells after 2 hours at the indicated pH expressed as a % of total cells scored at time zero. Each point represents the average of three independent biological repeats and error bars represent the standard deviation between repeats. The dotted line represents the limit of detection of this assay predetermined to be 0.04%.

3.4.2 Resistance to extreme acid challenge during different phases of growth.

The ability of *E. coli* to survive extreme acid stress is completely dependent on its phase of growth (De Biase *et al.*, 1999). During stationary phase acid resistance is greatly increased in *E. coli* K-12. The increase of stationary phase factors in exponential phase cells could raise the level of survival. One hypothesis to explain the evolved phenotype, is that the evolved strains could have high levels of one or more stationary phase factors present during exponential phase growth. RpoS, the stationary phase sigma factor, is up regulated during slow growth and is also responsible for the activation of stationary phase acid resistance (De Biase *et al.*, 1999; Hengge-Aronis, 2002; Small *et al.*, 1994). Mutations that slow growth or increase the concentration of stationary phase proteins such as RpoS could be contributing to the exponential phase resistance. To test the impact of different growth phases in the survival of the ancestor and evolved strains, the acid resistance assay was adjusted to measure survival during lag phase, exponential phase, extended growth at exponential phase and stationary phase. This was done by testing extreme acid resistance at an OD_{600} of 0.01 (shortly after sub-culturing the cells into a day culture), 0.15 (4.5 generations of outgrowth), 0.15 (9 generations of outgrowth), early stationary phase (0.6) and after overnight growth (stationary phase) for two hours at pH 2.5 (Figure 3.5 B).

The results shown in figure 3.5A indicate that growth phase does not significantly affect the resistance of the evolved strains to extreme acid shock. In addition, the repeated outgrowth of the evolved strains into exponential phase growth (Exponential 2) did not affect survival. These results demonstrate that the evolved resistance is not due to a carryover of RpoS, or other stationary phase factors, into exponential phase cultures.

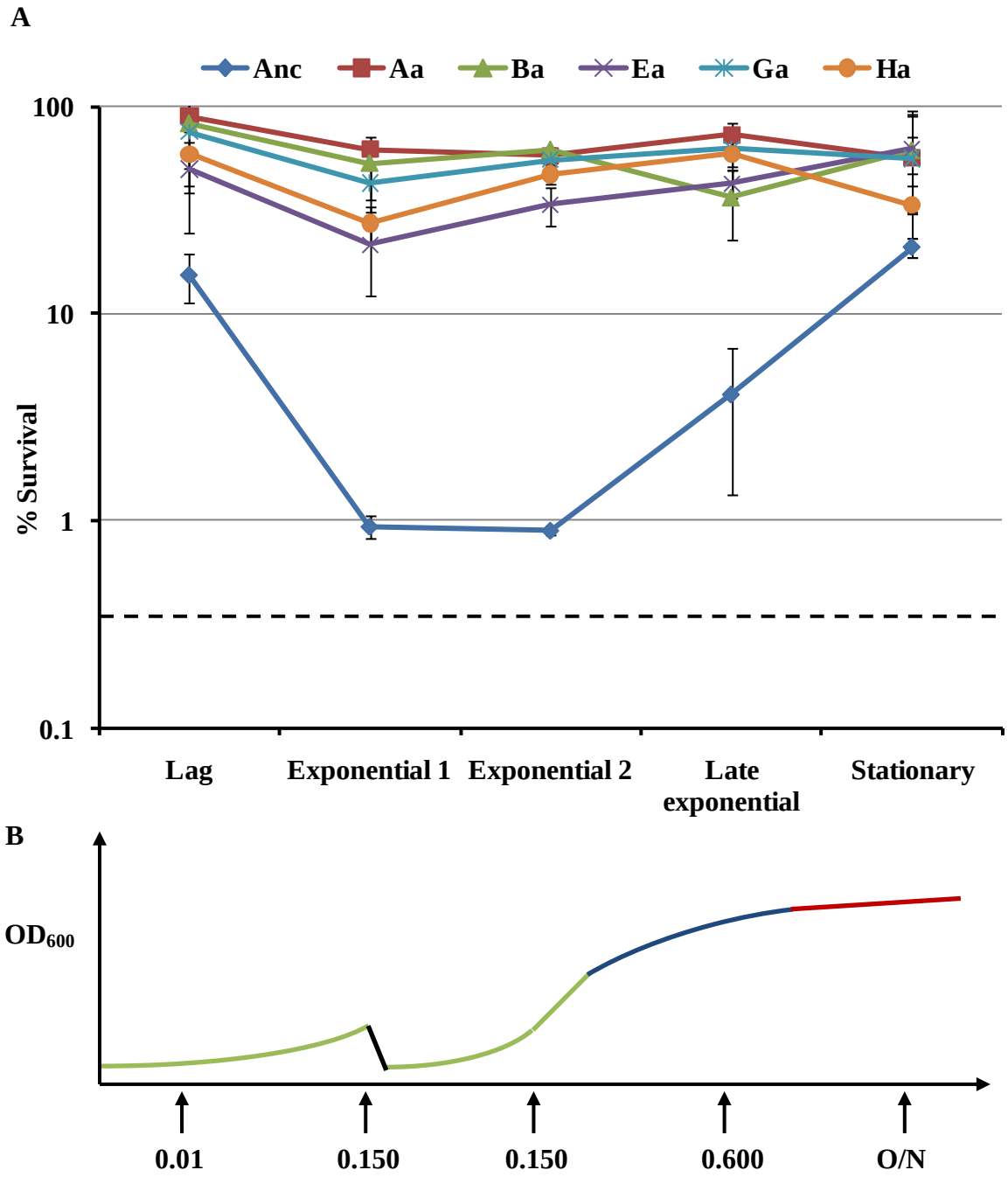


Figure 3.5. Survival of the ancestor and evolved strains during different phases of growth at pH 2.5. (A) survival of the ancestor strain (blue) and evolved strains Aa, Ba, Ea, Ga and Ha (red, green, purple, light blue and orange respectively). Survival is determined as the scored cells after 2 hours at pH 2.5 in M9supp expressed as a percentage of cells at time zero. Values and error bars represent the average and standard deviation of three biological repeats respectively. The dotted line represents the limit of detection predetermined to 0.04% for this assay. **(B)** A diagrammatic representation of the growth of the cultures during the experiment. Optical densities are indicated on the y axis.

In contrast, the ancestor is significantly more resistant to extreme acid challenge during lag and stationary phase growth, compared to survival during log phase. The survival observed during stationary phase is due to the induction of the acid resistance genes by increased RpoS levels. Presumably, lag phase resistance is due to factors remaining in the cells from stationary phase. The survival of the evolved strains was higher than the ancestor strain in all phases of growth. This suggests that the induction of acid resistance isn't simply an early switch to stationary phase growth. Instead, it is clear that the evolved resistance is independent from growth phase, and that the all of the evolved strains showed similar levels of resistance over all growth phases. However, these results do not conclusively show that the evolved resistance is not due to an increased level of RpoS.

3.4.3. Dependence of evolved acid resistance on RpoS

Promoters of genes found in the AFI and the GAD system are under the control of alternative sigma factor RpoS, which recruits RNA polymerase to a subset of promoters during entry into stationary phase. RpoS levels are regulated in the cell by many factors acting at different levels of regulation that act in response to a variety of signals (Hengge-Aronis, 2002). The affect of increased RpoS levels on the GAD and AFI systems is dependent on the regulators GadE, GadW and GadX (De Biase *et al.*, 1999; Giangrossi *et al.*, 2005; Tramonti *et al.*, 2006). As a result the acid resistance phenotype would be affected by RpoS levels. Stationary phase cells are more resistant to extreme acid challenge compared to exponential phase cells. This is due to the induction of the AFI and GAD systems in an RpoS dependant manner (Figure 3.5A) (Coldewey *et al.*, 2007; De Biase *et al.*, 1999). One hypothesis, which would explain the increased resistance in the evolved strains, is that mutations in RpoS, or indeed factors which regulate RpoS levels, could increase RpoS during exponential phase, and cause

acid resistance. This hypothesis is consistent with the fact that all of the evolved strains have a longer doubling time during optimal growth in exponential phase.

To see if any mutations could be found in the *rpoS* gene and upstream promoter region, the *rpoS* genes of each of the evolved strains were sequenced using low-throughput Sanger based sequencing methods. No mutations were found in the *rpoS* gene or intergenic regions located upstream and downstream of the coding sequence. The lack of mutations in the *rpoS* locus only shows that there is no promoter based alterations to transcription. The absence of mutations in the *rpoS* gene and downstream region suggest that RNA stability is not affected.

As previously stated, RpoS protein levels are regulated at many stages and by many factors. To test the dependence of the evolved acid resistance on RpoS, the *rpoS* gene was knocked out using standard Datsenko and Wanner mutagenesis (see Chapter 2). If RpoS is required, either in part or in full, the survival of the evolved strains will be affected. Δ*rpoS* KO strains were assayed using the standard acid resistance assay (see chapter 2, during exponential phase growth, after 2 hours at pH 2.5). Figure 3.6 shows the survival of the ancestor and evolved strains during exponential phase growth, at pH 2.5 for 2 hours in *rpoS*$^{+}$ and Δ*rpoS* backgrounds. The survival of an Δ*rpoS* mutant of the ancestor strain was below the level of detection for this assay (predetermined to 0.04%, Burton *et al.*, 2010). The reduction in survival was consistent with previously published levels and indicates that RpoS does provide some protection during exponential phase (Castanie-Cornet *et al.*, 1999). It is possible that RpoS is activating the AFI and GAD systems in response to stress, or basal levels of RpoS are sufficient to activate acid resistance which provides a low level of protection.

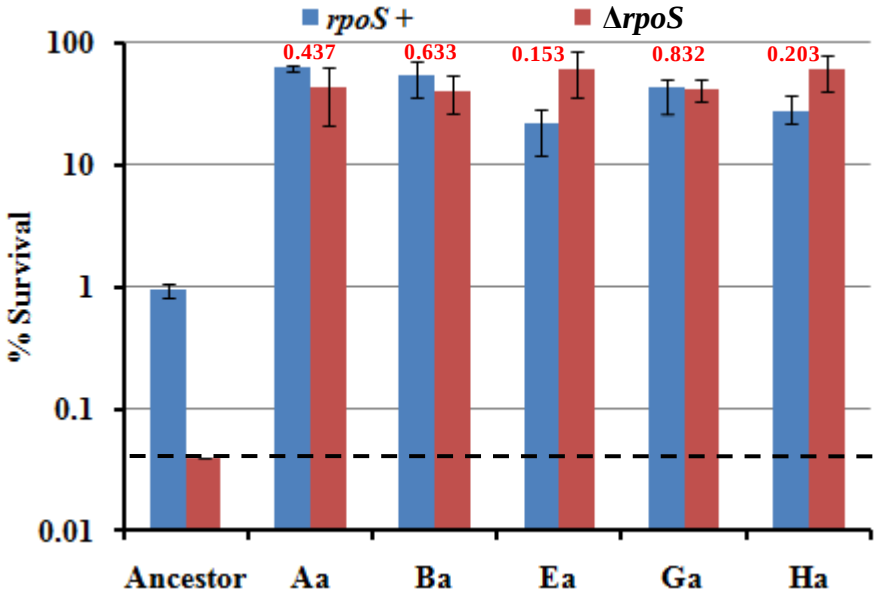


Figure 3.6. Survival of the Ancestor and evolved strains after 2 hours at pH 2.5 during exponential phase growth in *rpoS*+ (blue) and Δ*rpoS* (red) backgrounds. Survival is determined as the scored cells after 2 hours at pH 2.5 in M9supp expressed as a percentage of cells at time zero. Values and error bars represent the average and standard deviation of three biological repeats respectively. The dotted line represents the limit of detection predetermined to 0.04% for this assay. P-values from comparisons of the *rpoS*+ and Δ*rpoS* strains by T-test are shown by red numbers.

The evolved strains showed no significant difference in survival between *rpoS*$^{+}$ and Δ*rpoS* backgrounds. Any affect that *rpoS* may have had on survival is clearly being masked by the evolved resistance. It is important to note that all of the strains showed the same independence from RpoS.

3.4.4. Dependence of evolved resistance on specific amino acids

E. coli has 3 main acid resistance mechanisms and each mechanism requires a different amino acid to function. GAD, AR3 and AR4 utilise glutamate, arginine and lysine respectively (Castanie-Cornet *et al.*, 1999; Lin *et al.*, 1995; Lin *et al.*, 1996). Previous studies have shown that it is possible to isolate the contribution of each of these systems to acid resistance by providing each amino acid in the medium individually (De Biase *et al.*, 1999).

The resistance of the evolved strains was tested in M9 medium, M9 supplemented with arginine, M9 supplemented with glutamate and M9 supplemented with lysine. The acid resistance assay was otherwise unaltered and survival after 2 hours at pH 2.5 was calculated. The ancestor strain showed no significant difference in survival in M9 supplemented with any of the amino acids compared to M9 medium alone (Figure 3.7A). The survival of the evolved strains was significantly lower in all types of medium compared to the levels recorded in M9supp. However, no significant difference in survival was seen between the different types of medium. It is important to note that all of the evolved strains are showing the same level of survival in all of the medium. Once again, this is suggestive of a similar mechanism which is conferring the evolved acid resistance. However, it was not possible to distinguish the acid resistance phenotype into distinct amino acid dependent systems.

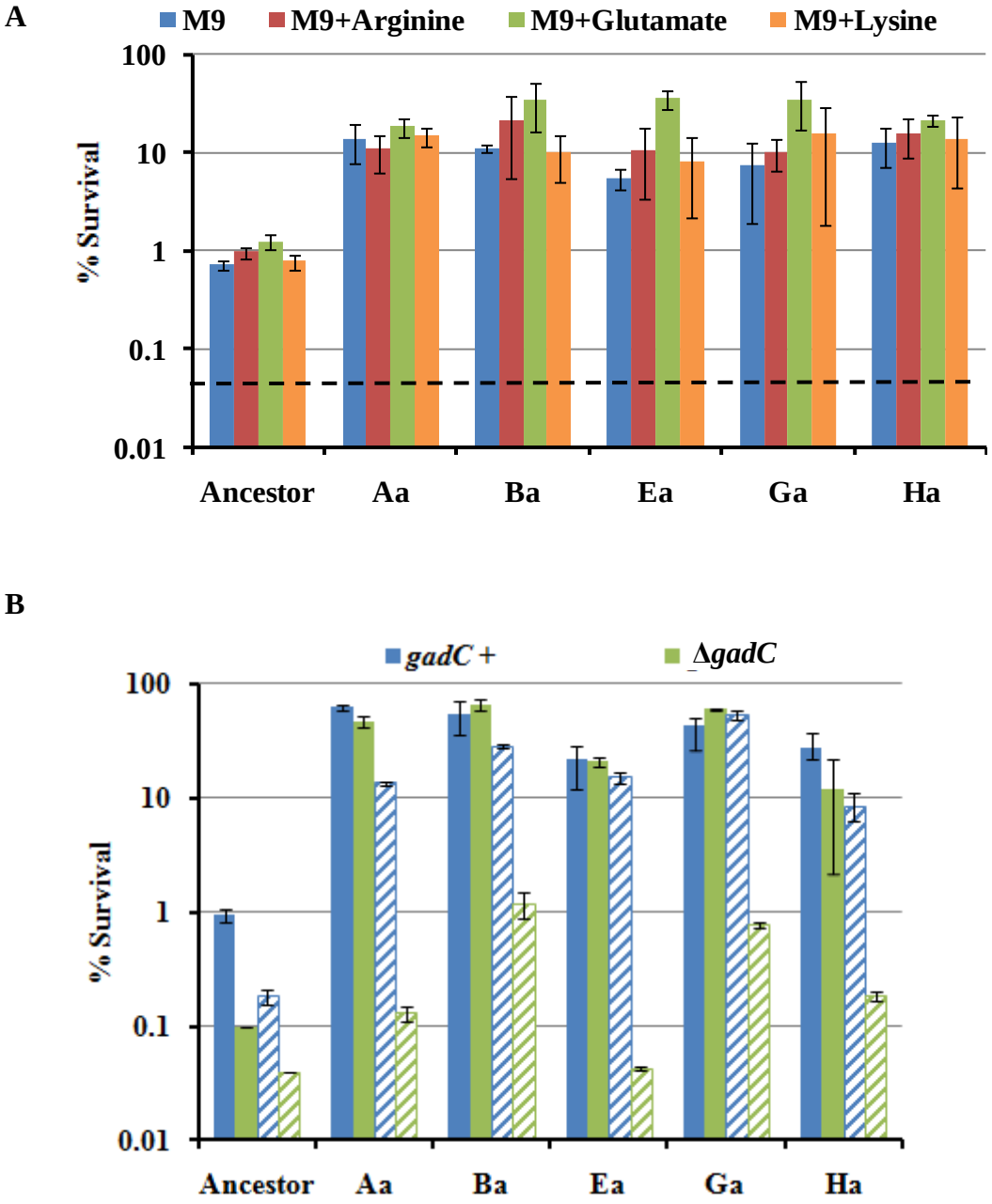


Figure 3.7. Dependence of evolved acid resistance on specific amino acid dependent systems. **(A)** Survival of the Ancestor and evolved strains after 2 hours at pH 2.5 during exponential phase in M9 (blue), M9+Arginine (red), M9+Glutamate (green) and M9+Lysine (orange). Concentrations of amino acids are described in chapter 2. **(B)** Survival of the *gadC*+ (blue bars) and Δ*gadC* (green bars) Ancestor and evolved strains after 2 hours in M9supp. Survival is determined as the scored cells after 2 hours at pH 2.5 (solid bars) or 2 hours at pH 2 (hatched bars) expressed as a percentage of cells at time zero. Values and error bars represent the average and standard deviation of three biological repeats respectively. The dotted line represents the limit of detection predetermined to 0.04% for this assay.

3.4.5. Dependence of evolved resistance on the GAD system.

The best characterised and most robust acid resistance mechanism found in *E. coli* is the GAD system. For a full description of the action of GAD see Chapter 1. Briefly, the GAD system comprises three components; the decarboxylases GadB and GadA, which converts cytoplasmic glutamate into GABA, and the GABA/glutamate antiporter GadC which exports GABA from the cytoplasm to the periplasm in exchange for glutamate (Bearson *et al.*, 2009; Diez-Gonzalez and Karaibrahimoglu, 2004; Hersh *et al.*, 1996). The pH at which the GAD system is most effective is between 2 and 2.5 (Capitani *et al.*, 2003; De Biase *et al.*, 1996). The minimum requirement for a functioning GAD system is a single decarboxylase (GadA or GadB) and the GABA/glutamate antiporter (GadC). Removing *gadC* makes *E. coli* K-12 more sensitive to pH 2.5 acid shock during stationary phase (Hersh *et al.*, 1996).

The evolution experiment selected cells that had increased survival at pH 2.5, and the evolved resistance is capable of extending to pH 2. This is consistent with the hypothesis that mutations that increased the expression of the GAD genes could have been selected for during the evolution experiment. To test this hypothesis we analysed the acid resistance of the evolved strains in a *gadC*- background.

To assay the dependency of the evolved strains on GAD, ancestor and evolved strains with the *gadC* gene knocked out were constructed using standard Datsenko and Wanner mutagenesis (see Chapter 2). The survival of the ancestor and evolved strains after two hours at pH 2.5 was assayed. The results are shown in figure 3.7B (solid bars are pH 2.5, hatch bars will be discussed below). Residual survival of the ancestor is below the limit of detection of this assay in a Δ*gadC* background. This confirms the reliance of *E. coli* on the GAD system to

survive acid challenge at pH 2.5. None of the evolved strains showed a significant difference in survival at pH 2.5 in Δ*gadC* and *gadC*$^+$ backgrounds. This indicates that GAD independent acid resistance mechanisms are involved in the evolved acid resistance.

The apparent lack of the dependence of the evolved strains on the GAD system was surprising as this is the most robust acid resistance mechanism at pH 2.5-2.0. The evolved strains have been shown to survive at pH 2. Do the GAD independent mechanisms, which confer resistance at pH 2.5, confer resistance at pH 2? To investigate the dependence of the evolved acid resistance on the GAD system at pH 2 the evolved strains in Δ*gadC* and *gadC*$^+$ backgrounds were shocked at pH 2.0. The level of survival of the *gadC*$^+$ strains had already been assayed in the gradient assay shown in figure 3.4 above. There was no significant decrease in survival at pH 2 compared to pH 2.5 in a *gadC*$^+$ background. However, in a Δ*gadC* background (figure 3.7B) survival levels were significantly lower after two hours at pH 2.0. This indicates that the GAD system is required for acid resistance at pH 2 and that any other mechanisms affected by the mutation(s) accumulated in the evolution experiment are not sufficient to protect the cells to a pH 2 acid shock. These results show that the GAD system has been affected by the evolution experiment and is providing protection in conditions under which normally it is not active (in exponential phase without induction).

3.4.6 Summary of mutagenesis and low throughput sequencing approaches

Mutagenesis and phenotypic assays have been useful for understanding the limitations and characteristics of the acid resistance phenotypes. The results have shown that certain systems, like the GAD system, are not crucial for the evolved phenotype. However, these assays have given no real insight into the genetic components that confer the evolved resistance. Several

attempts were made to find mutations in the evolved strains using low-throughput sequencing and current knowledge of acid resistance regulation.

Many acid resistance genes of *E. coli* are activated by GadE (Hommais *et al.,* 2004). Mutations that increase the expression of *gadE* would give an acid resistance phenotype. Mutations in the promoter of the *gadE* gene could cause increased expression. Mutations found in the coding region could affect the function of the regulator (e.g. increase of binding affinity to promoters). Mutations in the coding region could also affect the RNA stability. The *gadE* gene and the intergenic regions upstream and downstream of the evolved strains were sequenced using low throughput methods. No mutations were found in any of the evolved strains. A similar approach was used on the *safA-ydeO* and *hns* genes. Products of the *safA-ydeO* genes are involved in activation of the AR2 network, which activates the acid GAD mechanism and the AFI (Eguchi *et al.*, 2011; Ma *et al.*, 2004; Masuda and Church, 2003). H-NS is a nucleoid associated protein that represses the acid fitness island (Giangrossi *et al.*, 2005). No mutations were found in all cases stated above. Low throughput sequencing guided by the current understanding of acid resistance has been unfruitful. To find the genetic elements responsible for the evolved acid resistance a genome-wide method is required.

3.5. Whole genome re-sequencing of the evolved strains

3.5.1. Rationale for sequencing methodology

There are many ways of identifying the genetic elements responsible for a phenotype. The sections above describe attempts to use phenotypic assays, current knowledge and the understanding of *E. coli* acid resistance to identify the mechanisms involved in the evolved acid resistance. More traditional molecular techniques for finding genetic components of

phenotypes include genomic libraries and genomic hybridisation arrays. However, both of these methods have their caveats.

Genomic libraries are a method of screening genetic elements for a particular property. Briefly, the method involves fragmenting the genome of the organism of interest and cloning these fragments into a vector by ligation. The ligations are transformed into recipient bacteria and colonies are screened for the particular phenotype. The main problems with this method are centred on the requirement for a cloning step. Cloning often results in an under representation of large genes. The method is also very labour intensive and may not actually yield a result, especially if a particular phenotype requires a mix of mutations that are at distant loci and are therefore will not be found in the same clone. Additionally, the information about other genetic changes, which may not be essential for the evolved phenotype but maybe interesting for other investigations, will not be obtained with this method.

Genomic hybridisation arrays are a method for mutation screening. The genome from a strain of particular interest is fragmented, labelled with a fluorescent dye, and then hybridised to an array. If large areas of chromosome are missing due to deletion mutations then there will be no presence of a signal on the array. Strong signals can be attributed to duplication events. This method does remove the requirement for cloning and therefore all of the inherent problems with this step but genomic hybridisation arrays have a very low resolution. This problem has been reduced with the latest arrays which are capable of identifying single base substitutions and deletions but this also this comes at a high price. Possibly the biggest

problem with this technique is that arrays, especially high resolution arrays, are very expensive to print and use.

In order to make the concept of doing discovery based evolution experiments a reality, a method for finding mutations which is not biased, time consuming or expensive needs to be used. Fortunately, cheap whole genome re-sequencing is now a reality. The development of non-Sanger based sequencing technologies has exploded in the last 5 to 10 years. Currently, the most widely available technologies are the Illumina sequence by synthesis system and the Roche 454 pyrosequencing system. These technologies are complicated and varied, for the purposes of these experiments this section will only be concerned with the general advantages and disadvantages of these methods.

Roche 454 pyrosequencing is a method that uses the DNA synthesis reaction to sequence DNA fragments of up to 1 kb (Hert *et al.*, 2008). The main advantage of this method is that the read lengths are between 400 and 1000 bps. The main disadvantage is the low level of genome coverage compared to the Illumina method. Illumina sequencing uses a molecular colony created by PCR which can complicate the process. However, it is capable of producing high levels of genome coverage but with very short read lengths (20 – 100 bps) (Hert *et al.*, 2008). As a result these technologies are well suited to two separate tasks. 454 sequencing, which is capable of long read lengths, is best suited to pioneer sequencing of previously un-sequenced organisms. 454 would also be preferred when sequencing strains with long repeat sequences. However, Illumina sequencing is well suited to the re-sequencing of genomes due to its high coverage and the lower cost per base read.

The ancestor strain for this study is *E. coli* K-12 MG1655, which has a published sequence. This makes *de novo* assembly and alignments straightforward with small read lengths. The main aim is to compare the genome sequences of the ancestor and evolved strains for mutations, which could be resulting in an acid resistant phenotype. This does not require complete *de novo* assembly but does require high levels of coverage. Considering all of these factors, the best technology to use for this study was Illumina high-thoughput sequencing. The results from re-sequencing the ancestor and evolved strains are discussed below.

3.5.2. General results of whole genome re-sequencing analysis of the ancestor and evolved strains Aa, Ba, Ea and Ga

The strains sequenced by whole genome re-sequencing were the ancestor strain and evolved strains Aa, Ba, Ea and Ga. The sequencing of Aa, Ba, Ea and the ancestor was done by George Wienstock's group at Washington university St Louis. The Ga strain was sequenced by Mark Pallen's group. All sequences were analysed by Nick Loman from the University of Birmingham. Evolved strains Aa, Ba and Ea were sequenced with the ancestor in the same sample. This essentially allowed two strains to be sequenced in one sample and effectively halved the cost of the sequencing. Evolved strain Ga was sequenced individually. The analysis of all strains was done using the online software supplied by xbase.ac.uk. The sequence reads were uploaded in fasta format and aligned to the reference genome *E. coli* K-12 MG1655 (Blattner *et al.*, 1997). A summary of the sequencing vital statistics is shown in table 3.1. The results shown in table 3.1 indicate that an average coverage for all strains was 17 fold or above and that the majority of reads were aligned to the reference sequence.

Table 3.1. General results of high-throughput sequencing analysis

Strain	Total Reads	Read Length	Total Mb	Ref Genome Size Mb	Mapped Reads	% Mapped	Depth of coverage[a]	Depth of coverage[b]
Aa	4332550	35	151	4.6	4021048	92.8	35.215	17.607
Ba	8636870	35	302	4.6	8378338	97.0	67.163	33.581
Ea	12668168	35	443	4.6	12296182	97.0	98.455	49.227
Ga	13872659	35	485	4.6	13627292	98.2	N/A	162.626

[a] Ancestor and evolved strain fold coverage
[b] Evolved strain fold coverage

In addition, the level of coverage for the evolved strain is near 50% in each case, indicating that the concentrations of the ancestor and evolved genomes were relatively equal when they were sequenced. The proportions of the sample are very important when analysing the results, if a particular genome is under-represented in the sample, the reliability of the mutation calls made by the sequence analysis software will be reduced. The output clearly shows that it is possible to include more than one strain in the sample and that coverage can be maintained. The result of the Ga alignment indicated the high level of coverage that can be achieved when only one strain is present in the sample.

3.5.3. Detection of single nucleotide variants

Once the sequence reads had been aligned to the reference genome, a list of single nucleotide variants was generated. As the evolved strains Aa, Ba and Ea were sequenced with the ancestor in each case, a mutation present in approximately 50% of the reads must be present in the evolved strain but not in the ancestor. Equally, a mutation present in 100% of the reads is present in both the ancestor and the evolved strains but is not in the reference strain. In the sequence analysis of Ga, which was sequenced individually and then aligned to the *E. coli* K-12 MG1655 reference strain, all mutations found would be present in all of the reads. The results of this analysis are shown in table 3.2. The locations of the mutations are shown along with any annotation and the nature of the mutation, which was manually assigned. The number of mutations varies between the strains but it is interesting to note that in each case the mutations are few. In the analysis of the Aa, Ba and Ea SNV tables, two mutations were shown in each strain that were present in all reads. This indicated that these mutations were in the ancestor strain. One mutation was in an intergenic region non-proximal to any gene. The second, was found in the *ylbE* operon.

Table 3.2. Single nucleotide variants found by high throughput sequencing

Strain	Gene	Annotation	Mutation	Base change; AA change	Position
Aa	*yeeV*	CP4-44 prophage; toxin of the YeeV-YeeU toxin-antitoxin system	missense	CCC > TCC; Pro > Ser	2075614
	evgS	hybrid sensory histidine kinase in two-component regulatory system with EvgA	missense	TCC > TTC; Ser > Phe	2484146
	arcB	hybrid sensory histidine kinase in two-component regulatory system with ArcA	missense	CGT > CTT; Arg > Leu	3348730
	yhcE	None	silent	AGT >AGC; Ser > Ser	3364777
	eptB	predicted metal dependent hydrolase	missense	TAC > TCC; Tyr > Ser	3708410
Ba	*hofB*	conserved protein with nucleoside triphosphate hydrolase domain	silent	GTG > GTA; Val > Val	116446
	crl	DNA-binding transcriptional regulator	missense	AAG > AGG; Lys > Arg	257908
	ybfQ	None	missense	GAA > TAA; Glu > stop	735860
	poxB	pyruvate dehydrogenase (pyruvate oxidase), thiamin-dependent, FAD-binding	missense	ACC > CCC; Thr > Pro	901034
	narZ	nitrate reductase 2 (NRZ), alpha subunit	missense	GTG > GGG; Val > Gly	1537952
	yeeV	CP4-44 prophage; toxin of the YeeV-YeeU toxin-antitoxin system	missense	CCC > TCC; Pro > Ser	2075614
	yfcU	None	missense	AGT > AGG; Arg > Ser	2451224
	evgS	hybrid sensory histidine kinase in two-component regulatory system with EvgA	missense	AAC > AAA; Asn > Lys	2484114
	gltX	glutamyl-tRNA synthetase	missense	CAC > CCC; His > Pro	2518315
	ypjC	None	missense	TGC > GGC; Cys > Gly	2781995
	yhcE	None	silent	AGT >AGC; Ser > Ser	3364777
	ggt	Gamma- glutamyltranspeptidase	missense	ACT > CCT; Thr > Pro	3584678
	priA	Primosome factor n' (replication factor Y)	missense	AAC > ACC; Asn > Thr	4123557
	rpoC	RNA polymerase, beta prime subunit	missense	CCG > TCG; Pro > Ser	4184630
Ea	*dcp*	dipeptidyl carboxypeptidase II	missense	GAG > GGG; Glu > Gly	1624365
	yeeV	CP4-44 prophage; toxin of the YeeV-YeeU toxin-antitoxin system	missense	CCC > TCC; Pro > Ser	2075614
	evgS	hybrid sensory histidine kinase in two-component regulatory system with EvgA	missense	GGC > GCC; Gly > Ala	2484368
	iscR	DNA-binding transcriptional repressor	missense	TAT > CAT; Tyr > His	2660033
	priA	Primosome factor n' (replication factor Y)	missense	AAC > ACC; Asn > Thr	4123557
	yjjA	Conserved protein	missense	GTA > ATA; Val > Ile	4597945
Ga	*crl*	DNA-binding transcriptional regulator	missense	AAG > AGG; Lys > Arg	257908
	ylbE	Predicted protein, C terminal fragment (Pseudogene)	silent	GAA > GAG; Glu > Glu	547694
	evgS	hybrid sensory histidine kinase in two-component regulatory system with EvgA	missense	AGT >ATT; Ser > Ilu	2484194
	fimD	outer membrane usher protein, type 1 fimbrial synthesis	silent	GCC >GCT; Ala > Ala	4543253

This is a predicted protein with an unknown function. Due to the subtle differences between the ancestor strain and the published reference strain, the ancestor was renamed MGA. As these SNVs are present in the ancestor strain they are not included in table 3.2. The evolved strains Aa, Ba, Ea and Ga have 5, 15, 6 and 4 SNVs respectively. It is also important to note that most SNVs are missense mutations. Unfortunately, it is impossible to calculate whether the mutation rate is similar to those that have been previously reported due to the nature of the evolution experiment, as this experiment didn't account for the amount of generations the cell lines have gone through.

3.5.4. Analysis of single nucleotide variants

There are many interesting mutations in each evolved strain including: *arcB*, in evolved strain Aa; *rpoC*, in evolved strain Ba; *iscR*, in evolved strain Ea and *fimD*, in evolved strain Ga. These mutations are discussed in detail below. However, the most interesting observation is that all of the evolved strains have a mutation in the same gene, EvgS, which is a histidine kinase involved in the induction of acid resistance (Masuda and Church, 2002). Further analysis confirmed that these mutations are all different and that they are missense mutations. In addition, the mutations in EvgS are located in the PAS domain of the cytoplasmic region of the protein. For detailed analysis of these mutations see Chapter 4.

The identification of mutations in *evgS* in four of the evolved strains by whole genome re-sequencing justified the Sanger method based sequencing (low throughput sequencing) of the EvgAS locus (upstream promoter region and the *evgA* and *evgS* coding regions) in the fifth evolved strain Ha. The evolved strain Ha also had a mutation in *evgS* gene. However, this mutation was the same as the *evgS* mutation found in evolved strain Ga. To confirm that, in

fact, two separate evolved cell lines had acquired the exact same mutation, and that they weren't the same strain, the *fimD* locus, which contains a mutation in the evolved strain Ga, was also sequenced in the Ha strain. No mutation was found in the *fimD* locus in evolved strain Ha. This confirmed that Ha is a separate strain and that the same mutation has been acquired in parallel.

3.5.5 Other mutations found by whole genome re-sequencing

In addition to the mutations found in *evgS,* the strains also have also acquired some other interesting mutations, which could be involved in the acid resistance phenotype, e.g. *arcB,* or they could be involved in suppressing the effects of other mutations e.g. *rpoC*. This section will briefly discuss these mutations and the roles of the genes which they are in.

3.5.5.1 *arcB*

ArcB is the histidine kinase of the ArcAB TCS, named after its role in the regulation of anoxic redox control (Gunsalus and Park, 1994; Iuchi and Lin, 1988; Iuchi and Lin, 1993). The response regulator of this system acts as a transcriptional repressor of many processes involved in metabolism, such as glyoxylate shunt and fatty acid degradation (Gunsalus and Park, 1994; Lin and Iuchi, 1991). Interestingly, ArcAB is also involved in the repression of RpoS transcription (Mika and Hengge, 2005).

The mutation is in the transmembrane domain of the ArcB protein. The amino acid change is from an arginine to a leucine. This would cause a long polar hydrocarbon side group to be changed with a non-polar hydrocarbon side chain. It is unclear what affect this alteration would have in the context of the inner membrane. If the function of ArcB was affected, this

mutation could be increasing the levels of RpoS in the cells and so conferring acid resistance. However, preliminary experiments on the evolved strains, which have been mentioned above, suggest that this is unlikely as RpoS is not required for the evolved acid resistance phenotype.

3.5.5.2. *rpoC*

The product of *rpoC* is the β' subunit of RNA polymerase, which is involved in the formation of the holoenzyme. β' acts by binding to the sigma factor (Chenchik *et al.*, 1982). The β' subunit binds DNA non-specifically and is not thought to interact with the +1 (Naryshkina *et al.*, 2001). However, it is required for the formation of the "jaw" with the sigma factor, which is required for open complex stabilisation (Young *et al.*, 2004).

The mutation, found in strain Ba, is not located in a domain known to be important for any interactions made by β'. The proline to serine mutation is at residue 420 in the protein. Proline is a cyclic amino acid with a large non polar R group. Serine has a polar hydroxyl R group that is small in comparison to a proline. Considering these factors, this amino acid change could confer a significant alteration in the protein structure.

Obviously, the presence of mutations in such an important component of the cell's machinery is interesting and the recent findings by Conrad *et al.* in 2010 demonstrate the impact of *rpoC* mutations on the global transcription of the cell. It has been shown that deletion mutations in *rpoC* dramatically reduced the open complex stability and the distribution of RNAP over the cell. Specifically, some of the mutations were found to increase the transcription of acid resistance genes (Conrad *et al.*, 2010). In addition, the mutations found in *rpoC* were shown

to increase growth in minimal medium (Conrad *et al.*, 2010); it is possible that the *rpoC* mutation in Ba is involved in the counteraction of the affects of other mutations in the strain.

3.5.5.3 *iscR*

The product of *iscR* is a transcriptional regulator IscR, which is a repressor of the iron sulphur cluster genes (Schwartz *et al.*, 2001). The IscR regulon includes genes that are involved in metabolism and biofilm formation (Giel *et al.*, 2006; Schwartz *et al.*, 2001; Wu and Outten, 2009).

The mutation found in *iscR* is located in the DNA binding domain, which is a Helix-Turn-Helix motif. The base change is a tyrosine to a histidine, which is an aromatic residue changed to a basic residue. The side chains of both these amino acids occupy similar space as they both contain ring structures. However, tyrosine has a hydroxyl group that could be making important secondary structure interactions. This could mean that the structure of the HTH motif and the ability of this regulator to repress transcription could be altered. However, there is no connection, based on bioinformatic analysis, to link this regulator with acid resistance genes. The meta analysis of 544 micro arrays does not suggest any interaction between IscR and any known acid resistance genes (http://genexpdb.ou.edu/main). It is possible that this mutation could be involved in aiding survival via other mechanisms.

3.6 Transcriptional analysis of the GAD network in the evolved strains

The presence of mutations in the *evgS* gene in each of the evolved strains suggests that the regulation of acid resistance genes by *evgAS* may be altered in the evolved strains compared to the MGA strain. This hypothesis is supported by studies that have confirmed a constitutive-

on mutation in the same domain of EvgS. This EvgS mutation was linked between this mutation and an acid resistant phenotype (Eguchi *et al.*, 2011; Itou *et al.*, 2009). As discussed in Chapter 1, the induction of the AR2 network only occurs under conditions of mild acid shock (pH 5.7) (Burton *et al.*, 2010; Itou *et al.*, 2009). The result of this activation is the production of proteins required for AR2 acid resistance. Specifically, promoters of the *ydeP*, *safA*, *ydeO* and *gadE* genes are activated directly by phosphorylated EvgA (Itou *et al.*, 2009). The product from *ydeP* is known to aid survival during extreme acid shock (Masuda and Church, 2003), while *safA*, *ydeO* and *gadE* products activate the *phoPQ* TCS, promoters of the acid fitness island (AFI) and the AR2 structural genes *gadA*, *gadB* and *gadC*, respectively (Eguchi *et al.*, 2007; Ma *et al.*, 2004; Masuda and Church, 2003). Considering that all of the evolved strains contain a mutation in *evgS*, that EvgAS activates acid resistance, and that a constitutive-on EvgS mutation (EvgSc) has been characterised, we hypothesise that the AR2 network will be activated in the evolved strains. To test this hypothesis, we measured the promoter activity in both evolved and MGA strains.

An *in vivo* luciferase based promoter probe approach was used to assay the promoter activity of the genes involved in the AR2 network. A detailed description of the luciferase based promoter constructs and assays can be found in Chapter 2. Briefly, the evolved strains were transformed with the lucifrease promoter probes and the promoter activity was measured in M9supp pH 7 during exponential phase growth. The results in figure 3.8.1 show the promoter activity at the *evgAS* promoter (A) in all of the evolved strains and MGA strain. There is no significant difference between the MGA strain and any of the evolved strains at the *evgAS* promoter. This result is important as it indicates that there is no affect on expression of the TCS.

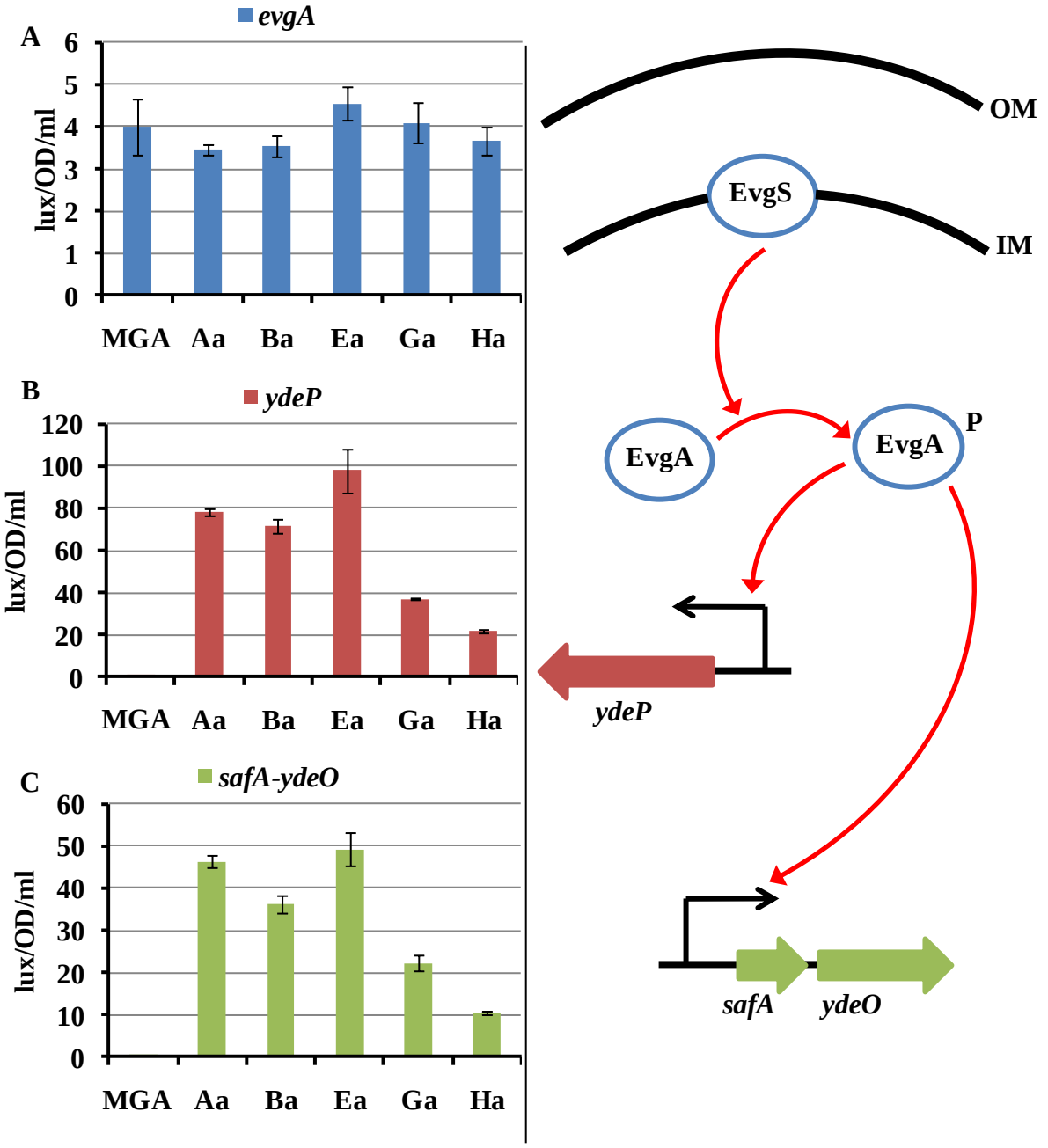


Figure 3.8.1. Promoter activity of the *evgA*, *ydeP* and *safA-ydeO* promoters. Promoter activity was determined at pH 7 in M9supp medium during exponential phase growth. Activity of the *evgA* promoter **(A)**, *ydeP* promoter **(B)**, and *safA-ydeO* promoter **(C)** is shown in the Ancestor (MGA) and evolved strains Aa, Ba, Ga, Ea and Ha. Values and error bars represent the average and the standard deviation of three biological repeats respectively. The activation cascade of the EvgAS TCS **(D)**. Coloured circles represent proteins, OM (outer membrane) IM (inner membrane), large coloured arrows represent coding regions, black arrows represent promoters and red arrows indicate activation.

Other studies have already shown that over-production of the response regulator EvgA can confer an acid resistant phenotype (Masuda and Church, 2002; Masuda and Church, 2003), but this explanation for the acid resistance of the evolved strains can be ruled out on the basis of these data.

The promoters of the *ydeP* and *safA-ydeO* genes, which are directly activated by EvgA (figure 3.8.1.D), shown in figure 3.8.1.B and C respectively are significantly active in the evolved strains compared to the MGA strain. This supports the hypothesis that an activated EvgAS could result in the activation of the rest of the pathway. The product of *ydeP* is itself a structural gene and can provide acid resistance at pH 2.5.

The activities of the promoters of the *slp* and *gadE* genes, which are regulated directly by *ydeO* (figure 3.8.2 D), shown in the figures A and B respectively, are also elevated in the evolved strain compared to the MGA strain. The product of *slp* has been shown to be important for acid resistance, and could be contributing to the acid resistant phenotype of the evolved strains.

The promoter activity of the *mgtA* gene is a direct indicator of the activity of the PhoPQ system (figure 3.8.2 D). PhoPQ, which is activated by low Mg^{2+} and low pH directly activates the *mgtA* promoter (Zwir *et al.*, 2005). If EvgS is constitutively-on then the cross-talk between the EvgAS and PhoPQ systems should cause the activation of the *mgtA* promoter. We tested this hypothesis by measuring the promoter activity of the *mgtA* promoter in the MGA and evolved strains without induction.

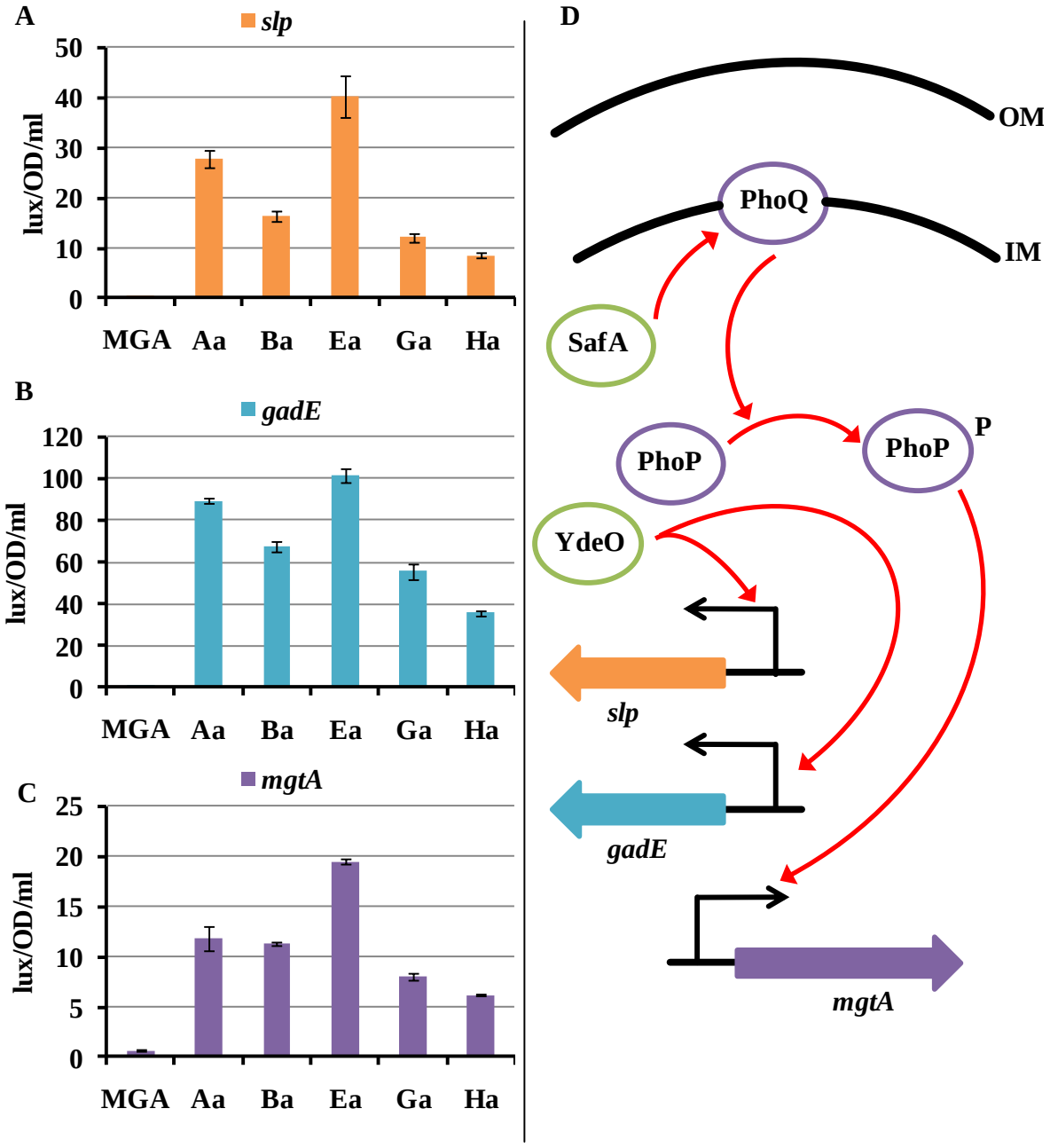


Figure 3.8.2. Promoter activity of the *slp*, *gadE* and *mgtA* promoters. Promoter activity was determined at pH 7 in M9supp medium during exponential phase growth. Activity of the *slp* promoter **(A)**, *gadE* promoter **(B)**, and *mgtA* promoter **(C)** is shown in the Ancestor (MGA) and evolved strains Aa, Ba, Ga, Ea and Ha. Values and error bars represent the average and the standard deviation of three biological repeats respectively. The activation cascade of SafA and YdeO via PhoPQ **(D)**. Coloured circles represent proteins, OM (outer membrane) IM (inner membrane), large coloured arrows represent coding regions, black arrows represent promoters and red arrows indicate activation.

The promoter activity of *mgtA* is presented in Figure 3.8.2 C, the *mgtA* promoter has significantly higher activity in the evolved strains compared to the MGA strain. This result suggests that the PhoPQ TCS and the *phoP* regulon are also active in the evolved strains. This result supports the hypothesis that the mutations in EvgS are causing the EvgS HK to be constitutively active. The promoter activities of the *gadA*, *gadB*, *hdeA* and *hdeD* genes, which are activated by the local regulator *gadE* (see figure 3.8.3 E), are shown in figures 3.8.3 A, B, C and D respectively. All of the genes, which code for products that are involved in acid resistance, are activated in the evolved strains compared to the MGA strain. These include *gadA*, *gadB* and *gadC* (*gadC* is transcribed from the *gadB* promoter), which code for the main constituents of the AR2 system for acid resistance.

The final component of the AR2 network is the GadXYW circuit. This component is a regulatory component, which is normally only activated by RpoS in stationary phase (figure 3.8.4. D). The role of this circuit in the activation of AR2 structural genes is not fully understood. The promoters of the *gadXYW* genes are activated by mild acid shock but this activation is completely independent of EvgA and completely dependent on RpoS (Burton *et al.*, 2010). However, more recent studies by Eguchi *et al.* in 2010 have shown that the level of RpoS is increased in the cell in a constitutively on EvgS mutant. If the mutations in EvgS are constitutive-on then the promoter activity would be increased. The promoter activity of the *gadX*, *gadY* and *gadW* promoters is shown in figure 3.8.4 A, B and C respectively. The promoters of these genes are also active in the evolved strain compared to the MGA strain. This supports the hypothesis of a constitutive-on EvgS. Additionaly, if the EvgS mutants are constitutively-on then this result also supports the model of RpoS regulation by EvgAS proposed by Eguchi *et al.* in 2010.

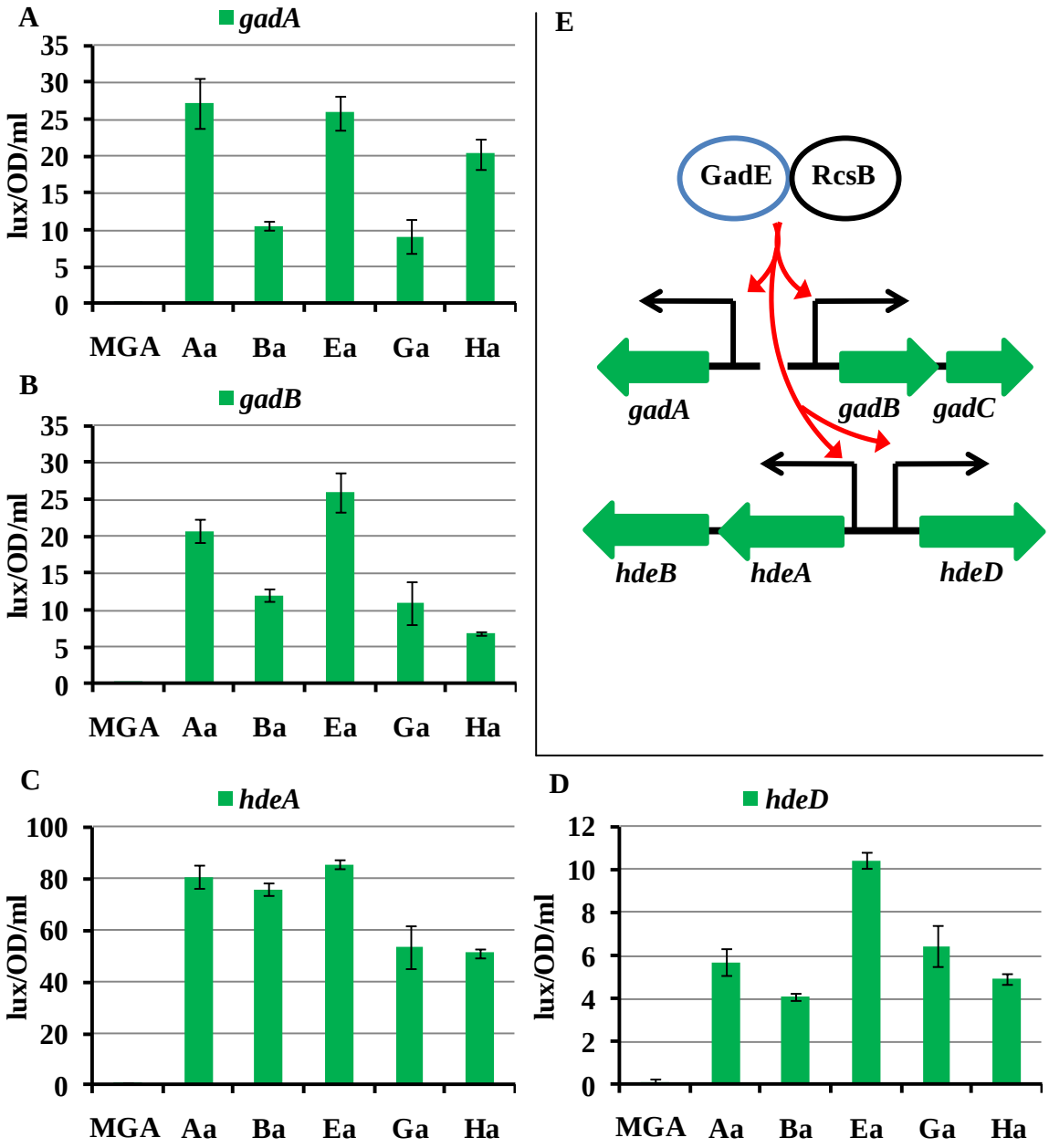


Figure 3.8.3. Promoter activity of the *gadA*, *gadB*, *hdeA* and *hdeD* promoters. Promoter activity was determined at pH 7 in M9supp media during exponential phase growth. Activity of the *gadA* promoter **(A)**, *gadB* promoter **(B)**, *hdeA* promoter **(C)** and *hdeD* promoter **(D)** is shown in the Ancestor (MGA) and evolved strains Aa, Ba, Ga, Ea and Ha. Values and error bars represent the average and the standard deviation of three biological repeats respectively. The GadE regulon **(E)**. Coloured circles represent proteins, large coloured arrows represent coding regions, black arrows represent promoters and red arrows indicate activation.

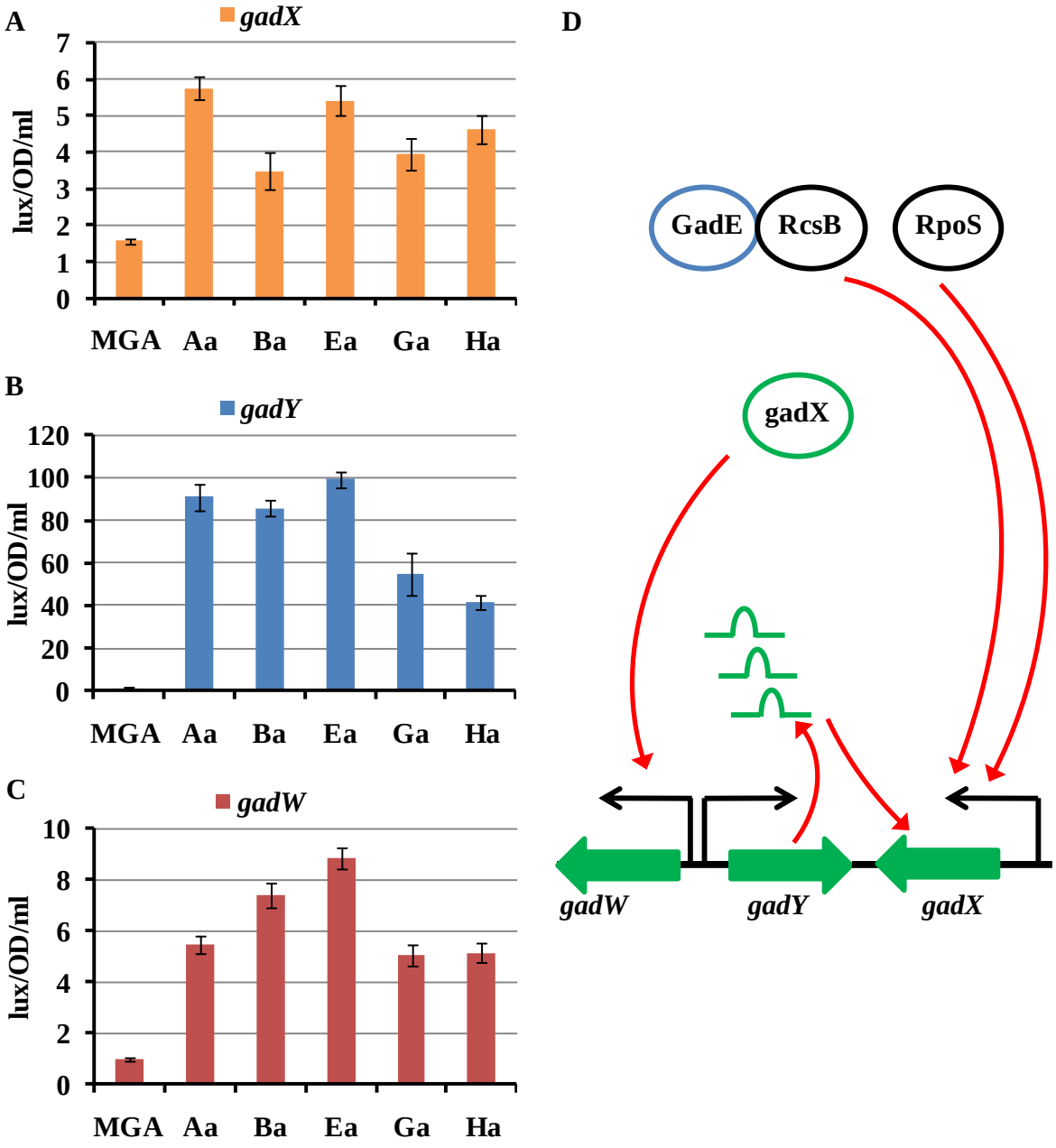


Figure 3.8.4. Promoter activity of the *gadX*, *gadY* and *gadW* promoters. Promoter activity was determined at pH 7 in M9supp medium during exponential phase growth. Activity of the *gadX* promoter **(A)**, *gadY* promoter **(B)** and *gadW* promoter **(C)** is shown in the Ancestor (Anc) and evolved strains Aa, Ba, Ga, Ea and Ha. Values and error bars represent the average and the standard deviation of three biological repeats respectively. The GadXYW mini-circiut **(D)**. Coloured circles represent proteins, large coloured arrows represent coding regions, black arrows represent promoters and red arrows indicate activation.

The general trend from the promoter probe assay is that all of the promoters that are involved in the activation of the AR2 network are activated under un-induced conditions. In all cases (except for *evgA*) the evolved strains show more activity but the amount of activity varies greatly between the strains. In particular, the evolved strains Ea and Aa and Ba show the most activity, with evolved strains Ga and Ha consistently showing less activity.

The results obtained using the promoter probes suggest that the strains have a constitutively activated AR2 network. Considering that EvgAS is the TCS that activates much of this network and that all of the evolved strains contain mutations in this gene it is reasonable to suggest that the *evgS* mutations are the likely contributories to the acid resistance phenotype. In addition, this reasoning is strongly supported by the fact that Ga and Ha have the same mutation in *evgS* and that they show similar levels of activation. However, in order to show that these mutations confer an acid resistant phenotype they will need to be assayed in isolation from the other mutations in the strains. Experiments doing this will be described in the next chapter

3.7. Discussion

3.7.1 Phenotype of the evolved strains

In this chapter the acid resistance phenotype of five evolved strains of *E. coli* K-12 MG1655 and their ancestor strain have been characterised. To summarise the above results, the evolved strains grow slower than the MGA strain, which suggests that their increased level of survival at low pH comes at a cost during optimal growth conditions. The evolved strains showed superior protective capabilities during extreme acid shock at acidities as low as pH 2.0 and a complete independence from RpoS and GadC. These results suggest that the acid resistance

phenotype is not an up-regulation of RpoS in exponential phase cells or due to the GAD system.

3.7.2 Coupling lab-based evolution with whole genome re-sequencing

Although these attempts to identify the systems or regulatory components required for the evolved phenotype have given a comprehensive insight into the acid resistance phenotype, they were unrealistic as a tool for finding the genetic mutations that confer the evolved resistance. This chapter has shown how whole genome re-sequencing can be used to capture all of the genetic information about a particular strain. The re-sequencing identified various mutations in each strain. The most interesting observation made using this technique was that all the evolved strains have a mutation in the *evgS* gene. In fact, one mutation was found in two separately evolved strains. Considering the various mutations that could result in acid resistance, this result was pivotal to this study. EvgS had already been shown to be involved in acid resistance regulation (Itou *et al.*, 2009; Ma *et al.*, 2004; Masuda and Church, 2002). In addition, an EvgS mutation, in the same domain as those found by this study, had been shown to confer acid resistance (Itou *et al.*, 2009; Kato *et al.*, 2000). Thus, the use of whole genome re-sequencing has revealed a candidate mutation in each strain that could be responsible for the acid resistance phenotype. Finally, this work can be looked upon as a successful example of how these techniques can be used together to produce interesting results for further investigation.

3.7.2 Transcriptional comparison of the evolved strains against the ancestor

The results described above showed that all the promoters of the AR2 network are activated in the evolved strain compared to the MGA strain. These results suggested that a mutation high

in the regulatory cascade is causing the activation. The *ydeP* and *safA-ydeO* promoters are among those that are activated. These promoters are directly activated by EvgA, which is the response regulator of the EvgAS TCS. When the whole genome re-sequencing data is taken into consideration, it can be hypothesised that the mutations found in EvgS are causing this TCS to be active in conditions where it is normally inactive. To test this hypothesis, the mutations will have to be assayed in isolation. These experiments will form the basis of the next chapter.

3.7.3 Summary

The important results of work presented in the above results chapter are summarised as follows:

1) The evolved strains have significantly decreased doubling times during conditions of optimal growth. However, at the OD_{600} of 0.150 all of the strains are in exponential phase growth. This is the OD_{600} which all strains were tested for all assays unless otherwise stated.
2) There is additional survival of all strains including MGA in M9supp medium compared to rich LB medium. However, the relative survival of the evolved strains compared to the MGA strain is consistent between conditions.
3) The evolved strains show between 20 and 50 fold more survival after 2 hours at pH 2.5 compared to the MGA strain. This shows that it is possible to evolve an acid resistant phenotype that has a significant effect on the survival of the strains to extreme acid challenge. The acid resistance assay outlined in Chapter 2 is accurate enough to resolve this phenotype.

4) The evolved strains show greater resistance compared to the MGA strain over a range of acidities. The MGA strain had no significant survival below pH 2.5, while the evolved strains had survival levels greater than 10% in medium with pH values as low as 2. The evolved strains had no protection against medium with a pH lower than 1.5.
5) The MGA strain is significantly more resistant to extreme acid challenge during exponential phase growth. Increased survival of the MGA strain was seen during lag phase, which is probably due to protection from stationary phase factors which have not been diluted out or degraded. There is no affect of growth phase on the evolved strains and the continual growth of cells in exponential phase didn't reduce survival revealing that survival is not due to the presence of stable stationary phase resistance proteins.
6) Despite the reduction of the MGA strain survival level below the level of detection in an Δ*rpoS* background, the evolved strains show no reduction in survival in Δ*rpoS* background. This indicates that the evolved acid resistant phenotype is independent from *rpoS*, and that any mutation(s) accumulated in the evolved strains are not affecting acid resistance via *rpoS*.
7) The attempt to isolate the contribution of specific amino acid dependent acid resistance systems was unsuccessful. Survival was notably greater in medium supplemented with glutamate but the difference was not significant.
8) When exposed to pH 2.5, which is the pH that the evolved cell lines were selected, all of the evolved strains in a showed *gadC* independent resistance.
9) At pH 2, where the GAD system is still effective, all evolved strains show a significant reduction in survival in a *gadC*$^{-}$ background. The evolution experiment has

affected the GAD system and other system that can protect the cells at pH 2.5 but not at pH 2.

10) Whole genome re-sequencing of four of the evolved strains Aa, Ba, Ea and Ga revealed a variety of mutations. However, all strains have a mutation in the histidine kinase EvgS. EvgS has already been implicated in acid resistance and is the inducer of exponential phase resistance during mild acid shock (pH 5.7).
11) Transcriptional analysis of the AR2 network, which is activated by EvgAS revealed that the pathway is activated in the evolved strains where in the MGA strain it is not. It appears that the evolved strains have an activated AR2 network during non-inducing conditions.

CHAPTER 4: RESULTS (2): Characterisation of *evgS* SNVs and their contribution to acid resistance in *Escherichia coli*

4.1 Characterisation of the *evgS* mutants: Location within EvgS

The previous chapter characterised the phenotype of five evolved strains, Aa-Ha, which are significantly more resistant to extreme acid stress compared to the MGA strain. Whole genome re-sequencing identified several single nucleotide mutations in each strain. Interestingly, all five evolved strains contained a mutation in the *evgS* gene that codes for the histidine kinase EvgS. Two strains, Ga and Ha, have the exact same mutation. This finding prompted investigation into the activity of the AR2 network, which is activated by EvgAS TCS. The promoter activity of genes in the AR2 network was measured by promoter probe assays, and the finding was that all of the promoters were active in the evolved strains. However, in a wild type strain no promoter activity was observed. This data was indicative of a constitutive-on EvgS, which is activating the AR2 network in uninducing (pH 7) conditions, and causing the acid resistant phenotype. To confirm the role of the *evgS* mutations in both the acid resistance phenotype and the activity of the AR2 network, the mutations must be assayed in isolation from the other mutations in the evolved background. This will rule out the roles of other mutations in the evolved phenotype.

The experiments discussed in this chapter characterise the *evgS* mutations. The location, nature, and biological significance of the mutations are analysed by bioinformatics methods. Finally, the *evgS* mutations are isolated and assayed for promoter activity and acid resistance.

All of the mutations found in *evgS* are located in the PAS (*Per Arnt Sim*) domain (Bock and Gross, 2002; Huang *et al.*, 1993). Figure 4.1 is a diagrammatic representation of the EvgS protein. Indicated on the diagram are the transmitter, receiver and Hpt (histidine

phosphotransfer) domains, which are required for activation by autophosphorylation and the phosphorylation of EvgA (Bock and Gross, 2002).

In figure 4.1 the PAS domain of EvgS is expanded and has been aligned to BvgS. BvgS is the homologue of EvgS found in *B. pertussis;* much of what is known about EvgS function is derived from BvgS (Bock and Gross, 2002). Figure 4.1 also shows the 2D structure of the PAS domain, which is broken down into 4 motifs including the N-terminal cap, PAS core and beta-scaffold (Taylor and Zhulin, 1999). The *evgS* mutations from strains Aa, Ba, Ea and Ga/Ha confer the following amino acid changes; serine at position 584 to phenylalanine (Aa), asparagine at position 573 to leucine (Ba), glycine at position 658 to alanine and serine at position 600 to isoleucine. The *evgS* mutations from evolved strains Aa and Ba are found in the N-terminal cap of the PAS domain, the S1 constitutive-on mutant characterised by Kato *et al* 2000 is also located in this same domain. EvgS mutations from strains Ha/Ga and Ea are found in the PAS core and beta-scaffold respectively.

The *evgS* mutation from evolved strain Aa, which has been named S584F, is a change of amino acid from a serine to a phenylalanine. The other mutation found in the N-terminal cap was found in the evolved strain Ba, which has been named N573L, is a change from an arsparagine to a leucine. Both of these mutations cause a polar hydrophilic residue to be replaced by a non-polar hydrophobic residue. Mutations Ea and Ha/Ga have been named G658A and S600I. Mutation S600I is located in the PAS core and results in a polar, hydrophilic mutation being replaced by a non-polar, hydrophobic residue, which is similar in nature to the S573F and N584L mutations.

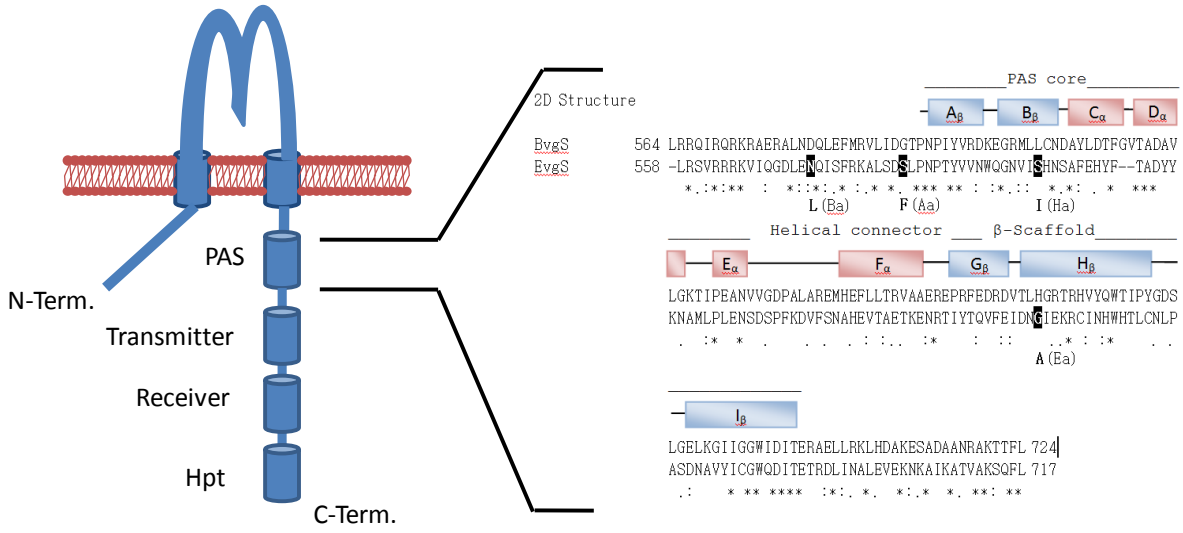


Figure 4.1. Locations of the *evgS* mutations and 2D structure of the PAS domain. A diagrammatic representation of the EvgS protein in the inner membrane of the cell. The required for the function of EvgS are indicated on the diagram and the PAS domain has been enlarged to show the locations of the mutations. A large venous flytrap-like domain is shown looping into the periplasm (adapted from *in silico* data by Dr Peter Winn). The enlarged PAS domain (right) shows the 2D structure of the protein (adapted from Taylor and Zhulin 1999). The blue and red boxes indicate beta strands and alpha helices respectively. The 2D structure is divided into N-terminal cap (located at the N-terminal end of the PAS domain) followed by the PAS core, helical connector and beta-scaffold. The PAS of EvgS is shown aligned to *B. purtussis* homologue BvgS, using ClustalW from EBI. Symbols . : and * indicate level of amino acid similarity. The location of each of the mutations found in the evolved strains is shown by a black box and the amino acid substitution is indicated along with the strain of origin.

The G658A mutation is not similar to the other mutations as the amino acid change replaces one aliphatic, non-polar residue with another. However, the impact of an alanine in a protein structure is dependent on its context. For example, the alanine side chain does have a methyl group, which in isolation is a soluble molecule. However, if the alanine is clustered in the structure with other methyl groups the alanine will contribute to the overall hydrophobicity.

In summary, the *evgS* mutations are distributed between three domains and three of the *evgS* mutations are substantial amino acid changes that could cause alterations in EvgS protein structure.

4.2. EvgA, EvgS and the EvgS mutations in Escherichia species

The level of acid resistance varies greatly between *Escherichia* species (Small *et al.*, 1994). A multiple alignment was done to see if the mutations isolated in the evolution experiment were present in any of the other sequenced strains of *E. coli*. The multiple alignment is also a powerful tool to assess how conserved the mutated residues are. We wanted to investigate the *evgS* genes from other *E. coli* strains to see if these mutations occur naturally. The occurrence of these mutations in other strains in *E. coli* could explain the variation in acid resistance.

The full *evgS* genes from 28 strains of *E. coli* and *Shigella* were aligned using ClustalW from EBI. As all of the mutations are located in the PAS domain this section of the alignment is shown in figure 4.2. The results of the alignment indicated that all of the mutations are in positions that are completely conserved (indicated by an asterisk). The location of the S1 mutation (characterised by Kato *et al* 2000) is also conserved. The PAS domain itself is also highly conserved with only 24 positions not completely conserved across all 28 strains.

We noted during the course of this analysis that *E. coli* strain 042 didn't have an *evgAS* locus. The region of the chromosome that normally contains the *evgAS* locus could provide insights as to how the locus was lost from *E. coli* 042. To investigate the lack of the *evgAS* locus, the *E. coli* K-12 *evgAS* locus was aligned to the 042 genome. The result of this alignment is shown in figure 4.3 A. (insert analysis of the loss of evgAS).The EvgAS system regulates the AR2 network, as *E. coli* 042 doesn't contain an *evgAS* locus we wanted to see if the other components were coded in this strain. Sequences of components of the AR2 network from *E. coli* K-12 were aligned to the *E. coli* 042 genome. Despite not containing an *evgAS* locus, *E. coli* 042 does have other components of the AR2 network. Figure 4.3 B shows the genes that were checked for their presence in *E. coli* 042.

In summary, the results of the alignment clearly show that there is high conservation of this domain between *E. coli* and *Shigella* species. None of the mutations are found naturally occurring in other strains of *E. coli*. Finally, the *E. coli* 042 strain has no EvgAS locus despite containing homologues for the other genes of the AR2 network.

4.3. Acid resistance conferred by *evgS* mutations in isolation

The evolved strains contain a variety of single nucleotide mutations that could be contributing to the acid resistant phenotype in full or in part. The data present in Chapter 3 suggests that the mutations found in EvgS are responsible for each strain's resistance. The EvgAS TCS has already been implicated in the control of the AR2 network and constitutive-on mutants of EvgS have been previously characterised (Kato *et al.*, 2000). In addition, all of the strains contain a mutation in this gene and two of the strains have the same mutation.

```
                                                  (Ba)#      (S1)~
UTI89           GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
S88             GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
APEC            GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
0127_H6         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
536             GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
CFT073          GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
SMS-3-5         GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1530
IAI39           GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
ED1a            GTTCGTCGTCGTAAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
Sd197           GTTCGTCGTCGTGAAGTTATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
UMN026          GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCA 1740
TW14359         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
EC4115          GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1530
Sakai           GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
EDL933          GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGGAAAGCG 1740
k12             GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
BL21_DE3_       GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
W3110           GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
HS              GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
Sf2a            GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
O26_H11         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
O111_H-         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
O103_H2         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
55989           GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
E24377A         GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
Sb227           GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
CDC             GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
SS046           GTTCGTCGTCGTAAAGTCATTCAGGGTGATTTAGAAAACCAAATATCATTCCGAAAAGCA 1740
                ************.**** **********************************.*****.

                    (Aa)  #                                          (Ha/Ga)  #
UTI89           CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
S88             CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
APEC            CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
0127_H6         CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
536             CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
CFT073          CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
SMS-3-5         CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1590
IAI39           CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
ED1a            CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
Sd197           CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
UMN026          CTCTCGGACTCCTTACCGAATCCAACTTATGTCGTAAACTGGCAAGGTAATGTCATTAGT 1800
TW14359         CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
EC4115          CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1590
Sakai           CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
EDL933          CTCTCGGACTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
k12             CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
BL21_DE3_       CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
W3110           CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
HS              CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
Sf2a            CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
O26_H11         CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
O111_H-         CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
O103_H2         CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
55989           CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
E24377A         CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
Sb227           CTCTCGGATTCCTTACCAAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
CDC             CTCTCGGATTCCTTACCAAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
SS046           CTCTCGGATTCCTTACCGAATCCAACTTATGTTGTAAACTGGCAAGGTAATGTCATTAGT 1800
                ******** ********.************** ****************************
```

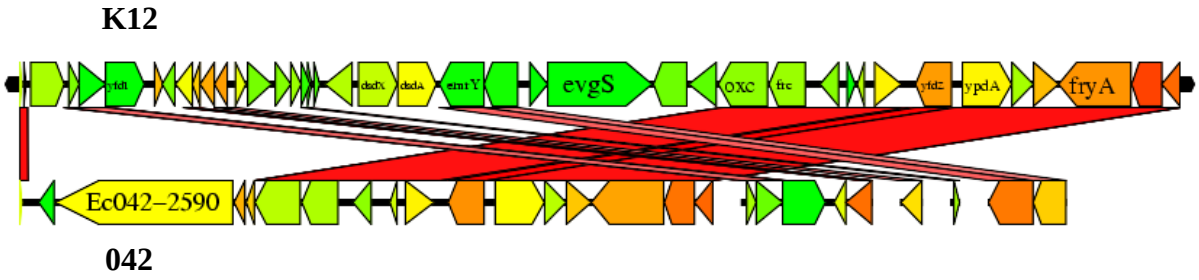
```
UTI89           CATAATAGTGCATTTGAACATTATTTCACTGATGATTACTACAAAAATGCAATGTTGCCA 1860
S88             CATAATAGTGCATTTGAACATTATTTCACTGATGATTACTACAAAAATGCAATGTTGCCA 1860
APEC            CATAATAGTGCATTTGAACATTATTTCACTGATGATTACTACAAAAATGCAATGTTGCCA 1860
0127_H6         CATAATAGTGCATTTGAACATTATTTCACTGATGATTACTACAAAAATGCAATGTTGCCA 1860
536             CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAATGCAATGTTACCA 1860
CFT073          CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAATGCAATGTTACCA 1860
SMS-3-5         CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAGTGCAATGTTACCA 1650
IAI39           CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAGTGCAATGTTACCA 1860
ED1a            CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAATGCAATGTTACCA 1860
Sd197           CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTATAAAAATGAAATGTTACCA 1860
UMN026          CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTACAAAAATGCAATGTTGCCA 1860
TW14359         CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTACAAAAATGCAATGTTGCCA 1860
EC4115          CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTACAAAAATGCAATGTTGCCA 1650
Sakai           CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTACAAAAATGCAATGTTGCCA 1860
EDL933          CACAATAGTGCATTTGAACATTATTTCACTGCTGATTACTACAAAAATGCAATGTTGCCA 1860
k12             CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
BL21_DE3_       CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
W3110           CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
HS              CATAATAGTGCTTTTGAACATTATTTCACTGTGGATTACTACAAAAATGCAATGTTACCA 1860
Sf2a            CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
O26_H11         CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
O111_H-         CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
O103_H2         CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
55989           CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
E24377A         CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
Sb227           CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
CDC             CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATATTACCA 1860
SS046           CATAATAGTGCTTTTGAACATTATTTCACTGCGGATTACTACAAAAATGCAATGTTACCA 1860
                ** ********:*******************  ******** ****.**.***.**.***

UTI89           TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
S88             TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
APEC            TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
0127_H6         TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
536             TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
CFT073          TTAGAAAATAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
SMS-3-5         TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1710
IAI39           TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
ED1a            TTAGAAAACAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
Sd197           TTAGAAAATAGTGAATCCCCCTTTAAAGATGTTTTTTCTAATACACATGAAGTCACAGCA 1920
UMN026          TTAGAAAACAGTGAATCACCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
TW14359         TTAGAAAACAGTGAATCACCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
EC4115          TTAGAAAACAGTGAATCACCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1710
Sakai           TTAGAAAACAGTGAATCACCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
EDL933          TTAGAAAACAGTGAATCACCCTTTAAAGATGTTTTTTCTAATACGCATGAAGTCACAGCA 1920
k12             TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
BL21_DE3_       TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
W3110           TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
HS              TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
Sf2a            TTAGAAAACAGTGATTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
O26_H11         TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
O111_H-         TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
O103_H2         TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
55989           TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
E24377A         TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
Sb227           TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
CDC             TTAGAAAACAGTGACTCACCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
SS046           TTAGAAAACAGTGACTCCCCCTTTAAAGATGTTTTTTCTAATGCGCATGAAGTCACAGCA 1920
                ******** ***** **.************************.*.***************
```

```
                                                                (Ea)#
UTI89           GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
S88             GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
APEC            GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
0127_H6         GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
536             GAAACGAAAGAAAACCGAACAATATACACACAGGTTTTTGAAATTGATAATGGCATCGAG 1980
CFT073          GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
SMS-3-5         GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1770
IAI39           GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
ED1a            GAAACGAAAGAAAACCGAACAATATACACACAGGTTTTTGAAATTGATAATGGCATCGAG 1980
Sd197           GAAACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
UMN026          GAGACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
TW14359         GAGACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
EC4115          GAGACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1770
Sakai           GAGACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
EDL933          GAGACGAAAGAAAACCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
k12             GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
BL21_DE3_       GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
W3110           GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
HS              GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
Sf2a            GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
O26_H11         GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
O111_H-         GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
O103_H2         GAGACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
55989           GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
E24377A         GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
Sb227           GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
CDC             GAAACGAAAGAAAATCGAACAATATACACACAGGTATTTGAAATTGATAATGGCATCGAG 1980
SS046           GAAACGAAAGAAAATCGAACAATATACACACAGGTTTTTGAAATTGATAATGGCATCGAG 1980
                **.*********** ********************:************************
```

Figure 4.2. Multiple alignment of the PAS domain of the *evgS* gene from sequenced *Escherichia* and *Shigella* species. The alignment was made using ClustalW web resource from EBI. Symbols . : and * indicate levels of conservation, an * marks complete conservation across all compared sequences. Symbols ~ and # indicate the S1 mutation characterised by Kato *et al* 2000 and the mutations characterised in this study respectively.

A



B

Gene	Presence (Y/N)
gadB	Y
gadE	Y
gadA	Y
ydeO	Y
ydeP	Y
safA	Y
slp	Y
hdeA	Y
hdeD	Y
gadX	Y

Figure 4.3. Presence of evgAS and the AR2 network in *E. coli* 042. The EvgAS locus of the *E. coli* K12 MG1655 (top) compared to the *E. coli* 042 genome (bottom) by alignment using the xbase.ac.uk web program **(A)**. Conserved regions are joined by red lines and G+C level is indicated by colour (Green = high G+C, Red = low) (xbase.ac.uk). **(B)** Table indicating the presence of components of the AR2 network that is controlled by EvgAS. Y indicates the presence of a homologue.

Finally, the activity of the AR2 network was shown to be elevated in the evolved strains in uninduced conditions. In order to assay the acid resistance conferred by each *evgS* mutation, each mutation was moved back into the MGA strain to be tested in isolation from the other mutations found in the evolved strains.

In order to move the *evgS* mutations into the MGA strain a strategy was devised that combined Gene Doctoring and Datsenko and Wanner mutagenesis techniques (Datsenko and Wanner, 2000; Lee *et al.*, 2009). This strategy introduced the G658A and S600I *evgS* mutations on to the MGA strain's chromosome with no selective markers. Details of the method for "gorging" the mutations onto the chromosome of the MGA strain are included in Chapter 2. The *evgS* mutations G658A and S600I were introduced into the MGA strain chromosome creating strains *evgSc*G658A and *evgSc*S600I. In order to move the *evgS* mutations between strains with ease, the *evgS* genes conntaining the S584F and N573L mutations were cloned onto an expression vector pZC320. Vector pZC320 is ideal for this purpose as its replication is controlled by an F plasmid replicon. This gives the plasmid a low copy number of 2-3 copies per cell. Details of the constructs and the plasmid pZC320 can be found in Chapter 2. The expression plasmids contained the entire *evgAS* locus and are under the control of the natural *evgAS* promoter. The MGA *evgAS* locus was also cloned into the same plasmid to aid as a control. The plasmids containing wild type *evgAS*, S584F and N573L were named p*evgAS*, p*evgAS*cS584F and p*evgAS*cN573L respectively.

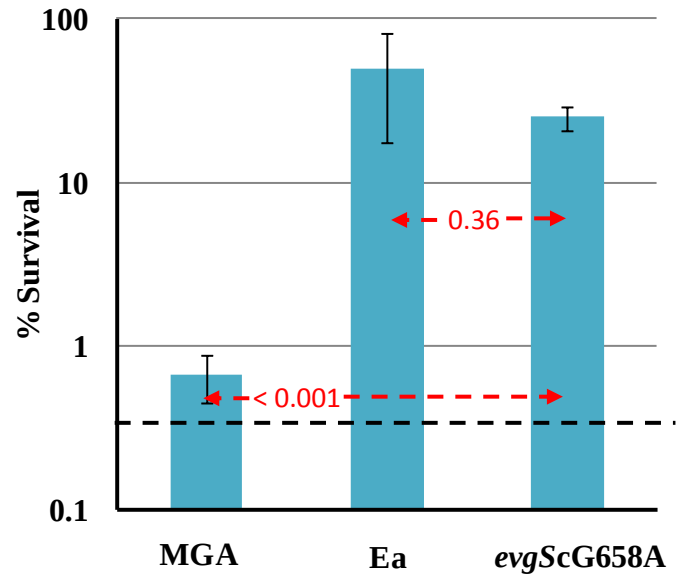
The EvgS mutations could be conferring the acid resistant phenotype. To test this hypothesis the survival of the gorged strains and strains carrying the wild type or mutated *evgAS* loci were assayed using the standard acid resistance assay during exponential phase growth.

Cultures were shocked for 2 hours at pH 2.5 and the colony counts from before and after the incubation were scored. Figure 4.4 A shows the acid resistance of the *evgS*cG658A mutant after 2 hours at pH 2.5. The resistance of evolved strain Ea and the MGA strain were also measured for comparison. The average survival of the *evgS*cG658A strain over 3 biological repeats was 25%, which is elevated compared to the MGA strain. The average survival of the original evolved Ea strain over 3 biological repeats was 49%. Statistical comparison of Ea survival with the *evgS*cG658A confirmed that the difference was insignificant (t-test, p value 0.36). Figure 4.4 B shows the survival of the *evgS*cS600I mutation after 2 hours at pH 2.5. The MGA and evolved strain Ga were also compared. The average survival of the *evgS*cS600I mutation over 3 biological repeats was 31.8%, which is much higher than the MGA strain. The average survival of evolved strain Ga over 3 biological repeats was 60%. The difference between the survival of Ga and *evgS*cS600I was determined to be insignificant (t-test, p value 0.53).

The presence of an additional plasmid encoded EvgAS may increase acid resistance as EvgA overexpression has previously been shown to confer resistance (Masuda and Church, 2002). To test the affect of an additional copy of the EvgAS, the survival of MGA containing p*evgAS* assayed for acid resistance.

The evolved strains, which contain a variety of mutations including EvgS mutations N573L and S584F, are acid resistant. One hypothesis that explains the acid resistance of the evolved strains is that the EvgS mutations are activating acid resistance genes. To test the affect of plasmid encoded EvgAS mutants N573L and S584F on acid resistance they were assayed using the standard acid resistnace assay.

A



B

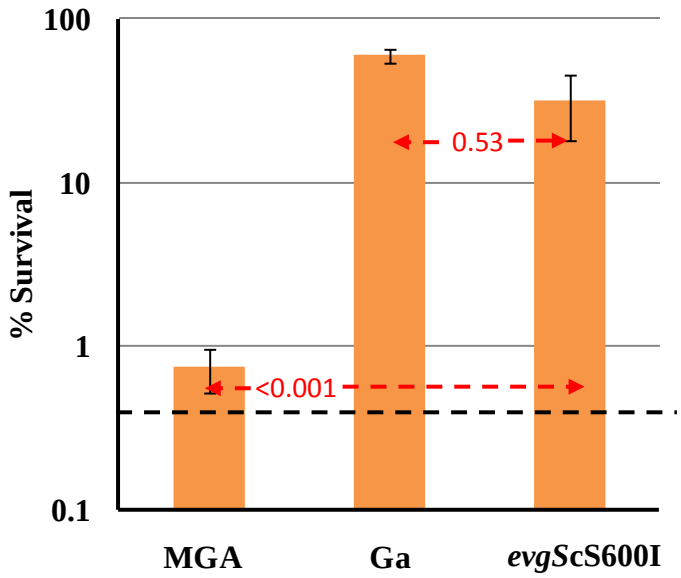


Figure 4.4. Survival of strains containing chromosomal *evgS* mutations. The acid resistance of the evgScG658A (A) and evgScs600I (B) strains was assayed by incubation in pH 2.5 M9supp medium for 2 hours. The parent evolved strain and MGA strain survival is provided for comparison. Survival was calculated as the colonies scored after 2 hours at pH 2.5 expressed as a percentage of the colonies scored in pH 7 after 0 hours. T-test p values are displayed in red, red dotted arrows indicate the values that have been compared. Each data point is an average of three independent biological replicates. Error bars represent the standard deviation of thee independent biological replicates. A dotted black line represents the limit of detection for the assay, predetermined to be 0.04%.

Survival was measured after 2 hours at pH 2.5 during exponential phase growth in M9supp medium. Figure 4.5 shows the levels of survival of the MGA +p*evgAS*cS584K and the MGA +p*evgAS*cN573L strains. The survival of MGA alone, MGA +pZC320 (empty vector) and MGA +p*evgAS* were also assayed as controls. The survival of the evolved strains Aa and Ba, which contained the *evgS* mutations N573L and S584F, are plotted for comparison. There is no difference in survival between the MGA and the MGA +pZC320 strains, and there is no difference in survival between MGA and the MGA strain +p*evgAS*. Thus, there is no effect on survival when the MGA strain contains the plasmid pZC320 or an extra copy of *evgAS*.

The average survival of MGA +p*evgAS*cS584F over 3 biological repeats was 93%. Statistical analysis of this strain versus MGA +p*evgAS* indicated that the difference in survival was significant (t test, p value <0.001). In addition, the introduction of p*evgAS*cS584F into MGA increases MGA survival to that of the evolved strain Aa. No significant difference in survival was observed between the evolved strain Aa and MGA +p*evgAS*cS584F (t test, p value 0.72). The average survival of MGA +p*evgAS*cN573L over 3 biological repeats was 89%. The difference in survival between this strains and MGA +p*evgAS* is statistically significant (t test, p value <0.001). The introduction of the plasmid p*evgAS*cN573L increased the survival of MGA to levels comparable to the evolved strain Ba, it was determined that there was no significant difference in survival (t test, p value 0.72).

In summary, the survival conferred by all of the *evgS* mutations, when introduced in isolation into MGA, were equivalent to levels of survival of the evolved strains. These results irrefutably show that the evolved acid resistance phenotype is completely due to the *evgS* mutations.

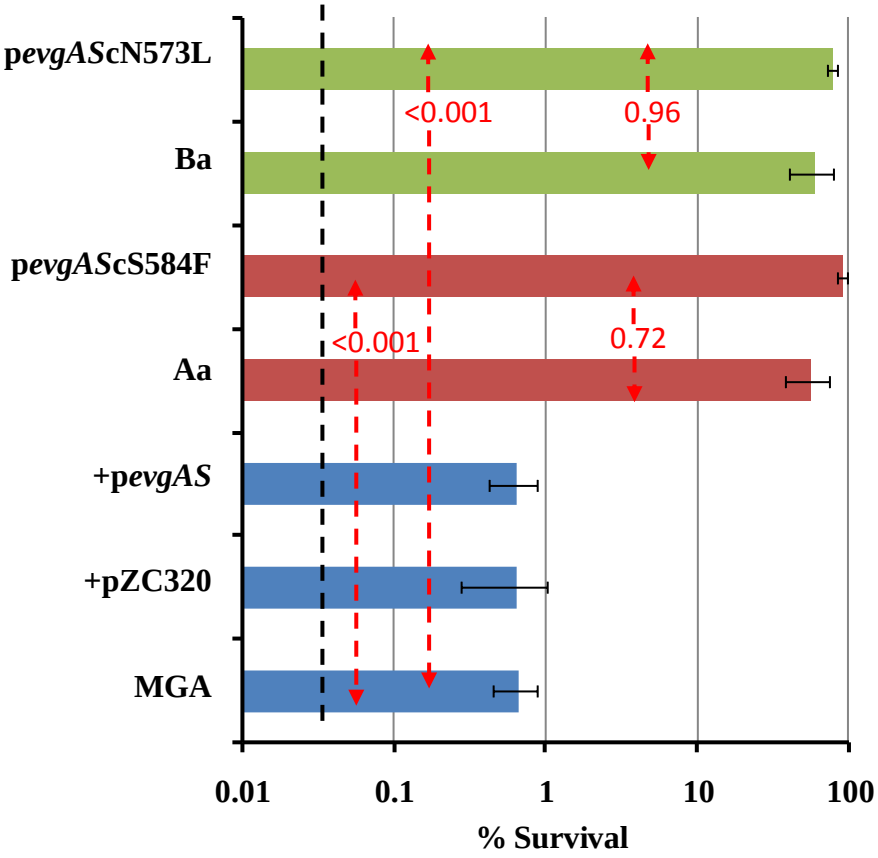


Figure 4.5. Survival of the MGA strain containing *evgASc* expression plasmids. Survival of the strains containing isolated *evgS* mutations were assayed by acid shock in pH 2.5 M9supp medium for 2 hours. MGA strains are shown in blue. The strains containing S584F and N573L are shown in red and green bars respectively. The plasmid p*evgAS*, which contains a wild type copy of *evgA* and *evgS* under the control of the native *evgA* promoter, was used as a control. Survival represents the scored colonies after 2 hours at pH 2.5 expressed as a percentage of the scored colonies after 0 hours at pH 7.t-test p-values are displayed in red, red dotted arrows indicate the values that have been compared. Data points and errors bars are an average or standard deviation of three biological repeats respectively. A dotted black line represents the limit of detection of the assay predetermined at 0.04%.

4.4 Effect of the *evgS* mutations on the AR2 network

The results described above indicate that the *evgS* mutations are conferring an acid resistant phenotype. Results from the Chapter 3 have shown that the AR2 network, which controls many acid resistance genes, is activated in the evolved strains in conditions where the MGA strain's AR2 network is inactive. Based on this information and the fact that in isolation the *evgS* mutations are conferring the acid resistance phenotype it is reasonable to hypothesise that the *evgS* mutations are also activating the AR2 network. Do the *evgS* mutations confer constitutive activation of this pathway when they are isolated from the other mutations found in the evolved strains?

In order to assay the affect of the *evgS* mutations on the activity of the AR2 network, the strains *evgSc*G658A, *evgSc*S600I and MGA containing either p*evgAS*, p*evgAS*cS584F or p*evgAS*cN573L were transformed with the promoter probe plasmids pLUX*evgA*p, pLUX*ydeP*p, pLUX*safA-ydeO*p, pLUX*slp*p, pLUX*gadE*p, pLUX*mgtA*p, pLUX*gadA*p, pLUX*gadB*p, pLUX*hdeA*p, pLUX*hdeD*p, pLUX*gadW*p, pLUX*gadX*p and pLUX*gadY*p. The strains were assayed for luciferase activity in M9supp during log phase and at pH 7, using the standard single time point promoter probe assay.

The results in figure 4.6.1 show the activity of the *evgA*, *safA-ydeO* and *ydeP* promoters in the *evgS* mutant strains. The activity of these promoters in the evolved and MGA strains is also shown for reference. As a control, the promoter activity in the MGA +p*evgAS* strain was also measured. The level of *evgA* promoter activity in the MGA +p*evgAS* strain confirmed that there is no activation of AR2 promoters as a consequence of including an additional plasmid encoded copy of the *evgAS* locus. As with the evolved strains, there was no affect of the *evgS*

mutations on the activity of the *evgA* promoter. The promoter activity of the *ydeP* and *safA-ydeO* promoters in MGA strains containing the *evgS* mutations S584F, S600I and N573L was comparable to the activity found in the evolved strains. However, the *ydeP* and *safA-ydeO* promoter activity in *evgS*cG658A was significantly lower than any of the evolved strain Ea (t test, p value <0.001). This cannot be an affect of the G658A mutation on the luciferase reporter system as the level of activity at the *evgA* promoter is unaffected. The *evgA* activated promoters are active in the *evgS* mutant strains, which is suggestive of a constitutive on EvgS.

If the promoters of the *ydeO-safA* genes are active, then it is reasonable to hypothesise that the promoters activated by YdeO (*slp* and *gadE*) will also be active. In addition, the PhoPQ TCS will also be activated by SafA. The *mgtA* promoter can be used as an indirect assay of PhoPQ activation. To test this hypothesis the activity of the YdeO activtated promoters and the *mgtA* promoter were assayed. The results are presented in figure 4.6.2. The MGA and evolved strains have been plotted for comparison. As was found for the *safA* promoter an additional plasmid encoded copy of wild type *evgAS* does not affect *slp*, *gadE* or *mgtA* promoter activity. The *slp*, *gadE* and *mgtA* promoter activities in the N573L, S600I and S584F *evgS* mutations were equivalent to the promoter activity observed in the evolved strains. As with the *evgA* regulated promoters, the level of activity of the *safA-ydeO* activated promoters conferred by the G658A *evgS* mutation was lower than the evolved strain Ea.

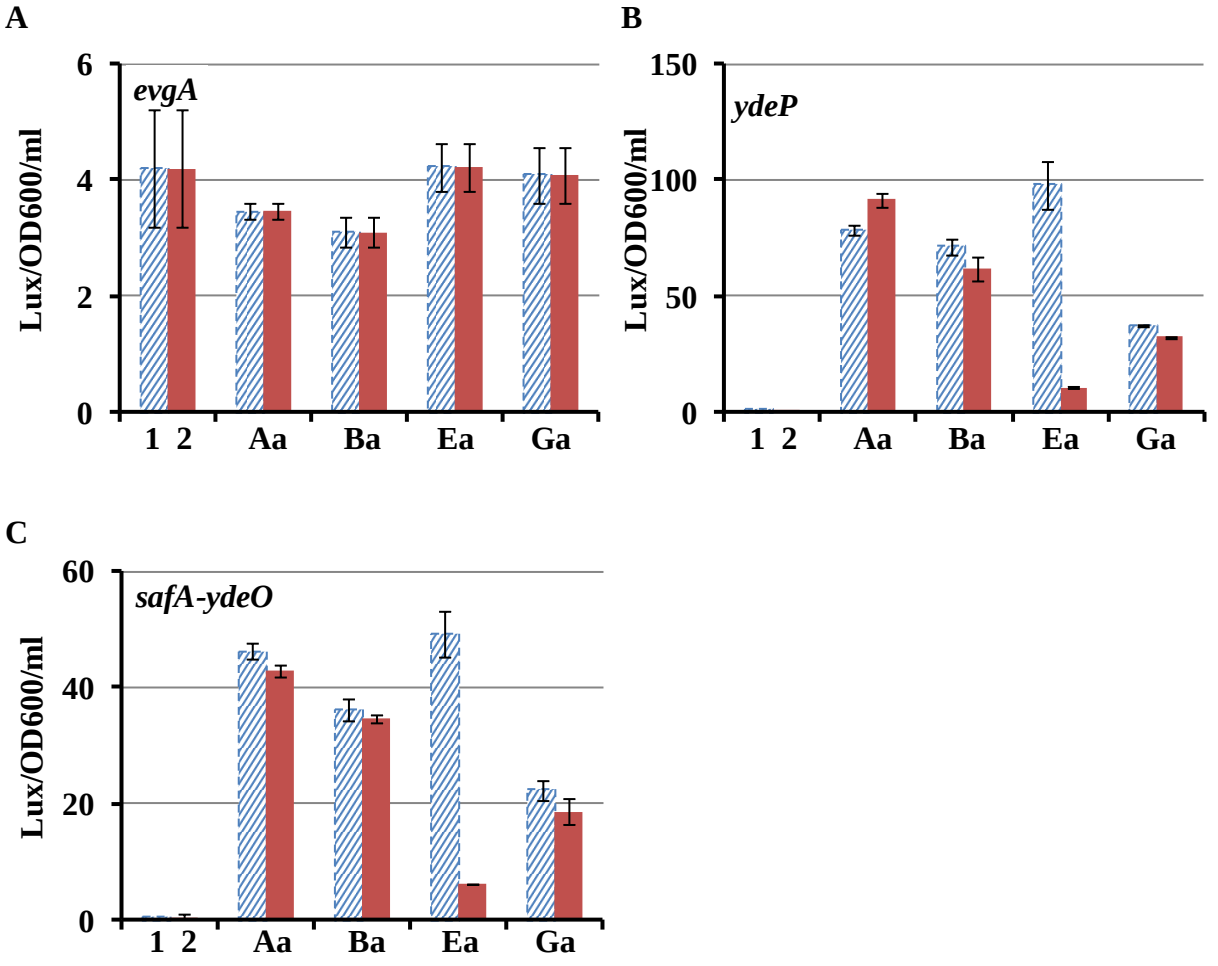


Figure 4.6.1. Activity of the *evgA* (A), *ydeP* (B) and *safA-ydeO* (C) promoters in the MGA, evolved and *evgS* mutant strains. The promoter activity was measured during exponential phase in M9supp medium at pH 7. Luciferase values were corrected for OD and culture volume. Lanes 1 and 2 represent the promoter activity in MGA and MGA containing the plasmid p*evgAS* (contains a wild type of *evgA* and *evgS* under the control of the *evgA* promoter) respectively. The promoter activity of the MGA strain containing the isolated *evgS* mutations are shown by red bars, the promoter activity of the evolved strains, which contained each *evgS* mutation are shown as blue striped bars. Each data point is the average of three independent biological repeats. The error bars represent the standard deviations between three biological repeats.

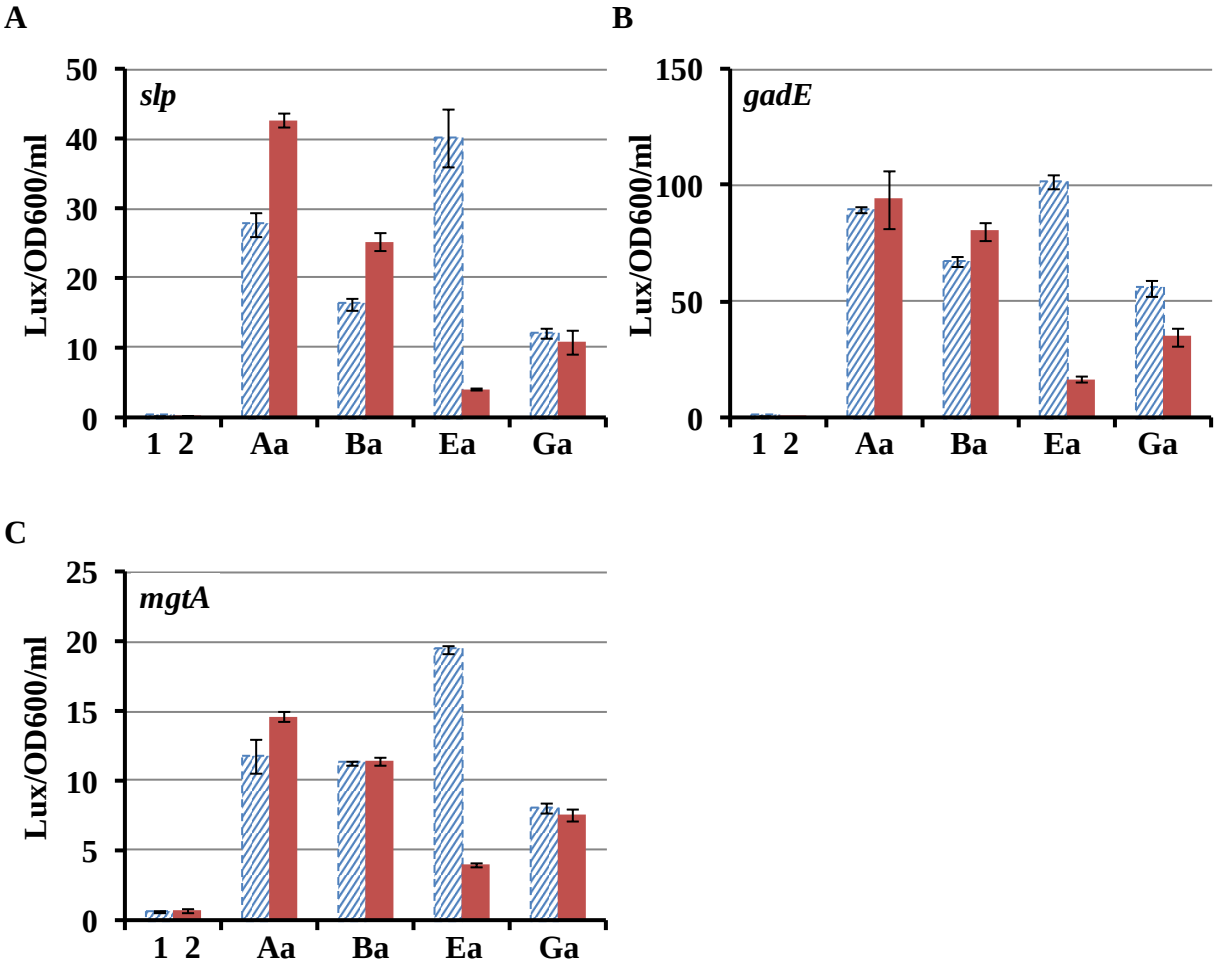


Figure 4.6.2. Activity of the *slp* (A), *gadE* (B) and *mgtA* (C) promoters in the MGA, evolved and *evgS* mutant strains. The promoter activity was measured during exponential phase in M9supp medium at pH 7. Luciferase values were corrected for OD and culture volume. Lanes 1 and 2 represent the promoter activity in MGA and in MGA containing the plasmid p*evgAS* (contains a wild type of *evgA* and *evgS* under the control of the *evgA* promoter) respectively. The promoter activity of the MGA strain containing the *evgS* mutations are shown by red bars, the promoter activity of the evolved strains. which contained each *evgS* mutation are shown as blue striped bars. Each data point is the average of three independent biological repeats. The error bars represent the standard deviations between three biological repeats.

The *evgS* mutations are all conferring activation for the *ydeO* regulated structural gene, *slp*, and the local AR2 regulator *gadE*. Finally, activity of the *mgtA* promoter suggests that the PhoPQ TCS is also activated by the *evgS* mutation. As shown by the activity of other promoters involved in the AR2 network, the *mgtA* promoter activity conferred by the G658A *evgS* mutation is much lower than the evolved strain Ea.

The consequence of activating the AR2 network fully is the activation of many structural genes involved in acid resistance. The level of resistance conferred by the *evgS* mutations is high and would suggest that all of the acid resistance genes regulated by GadE would also be active. To test this hypothesis, the promoter activity of the remaining structural genes was tested. The results are shown in figures 4.6.3. A similar pattern of activation was observed for the *evgS* mutations at the *gadA*, *gadB*, *hdeA* and *hdeD* promoters (figure 4.6.3). As shown for all aforementioned promoters, the addition of a plasmid encoded *evgAS* had no affect on promoter activity. All of the AR2 structural genes are activated in the *evgS* mutant strains. However, the G658A mutation does activate these promoters to a lesser extent.

Finally, work by Eguchi *et al*. in 2010 showed that their *evgS* constitutive on mutant conferred an increase in RpoS levels. Transcription of the GadXYW minicircuit has been shown to be RpoS dependent. If the *evgS* mutations are causing the EvgAS TCS to be constitutively on then these promoters should be active. To test this hypothesis, the promoter activities of the remaining regulatory components of the AR2 network (*gadX*, *gadY* and *gadW*) were assayed. The activity of the *gadX*, *gadY* and *gadW* promters was increased in all *evgS* mutant backgrounds relative to the MGA strain (figure 4.6.4).

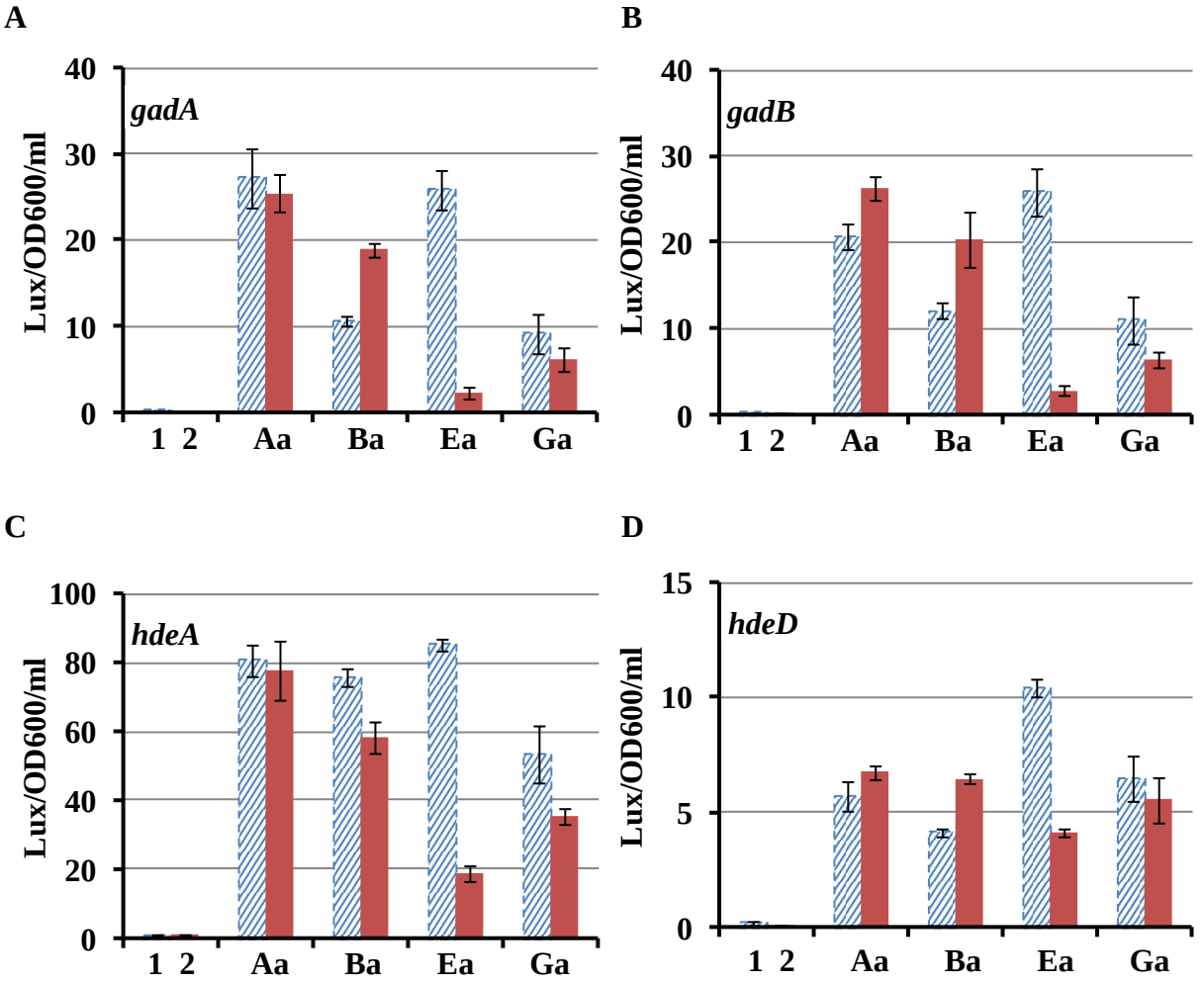


Figure 4.6.3. Activity of the *gadA* (A), *gadB* (B), *hdeA* (C) and *hdeD* (D) promoters in the MGA, evolved and *evgS* mutant strains. The promoter activity was measured during exponential phase in M9supp medium at pH 7. Luciferase values were corrected for OD and culture volume. Lanes 1 and 2 represent the promoter activity in MGA and in MGA containing the plasmid p*evgAS* (contains a wild type of *evgA* and *evgS* under the control of the *evgA* promoter) respectively. The promoter activity of the ancestor strain containing the *evgS* mutations are shown by red bars, the promoter activity of the evolved strains that contained that mutation are shown as blue striped bars. Each data point is the average of three independent biological repeats. The error bars represent the standard deviations between three biological repeats.

As with other promoters in the AR2 network, activation of each promoter was lower in the strain carrying the G658A *evgS* mutation compared to the evolved strain Ea

These results confirm that the promoter activity of the whole AR2 network is elevated in strains containing an *evgS* mutation. This observation supports the hypothesis that the *evgS* mutations are causing the protein to become constitutively-on. It is likely that the activation of this network and the structural acid resistance genes is conferring acid resistance in the *evgS* mutant strains. The promoter activity of the AR2 network in the *evgS* mutant strains is comparable to that of the evolved strains. The only exception being the G658A *evgS* mutation, which when present alone has significantly lower levels of AR2 promoter activity compared in the background of the evolved strain Ea. However, this activity is significantly higher than that of the MGA strain. Finally, the promoter activity in the MGA strain containing the control plasmid p*evgAS* was the same as that of MGA alone at all promoters, confirming that the presence of a plasmid encoded wild type *evgAS* locus had no affect on promoter activity. This latter result is particularly important as it confirms the validity of the results obtained using the p*evgAS*cN573L and p*evgAS*cS584F plasmids. The advantage of using a plasmid base system is that the S584F and N573L mutations can now be moved easily between different strains without the need for chromosomal mutagenesis.

4.5 Comparison of wild type and mutant *evgS* induction dynamics

The EvgS mutations have been shown to confer acid resistance and activation of the AR2 network, these results are indicative of a constitutive-on EvgS (Kato *et al.*, 2000). The wild type EvgAS two-component system can be activated by exposure to pH 5.7 (Burton *et al.*, 2010). In addition, the EvgS mutations are in the PAS domain.

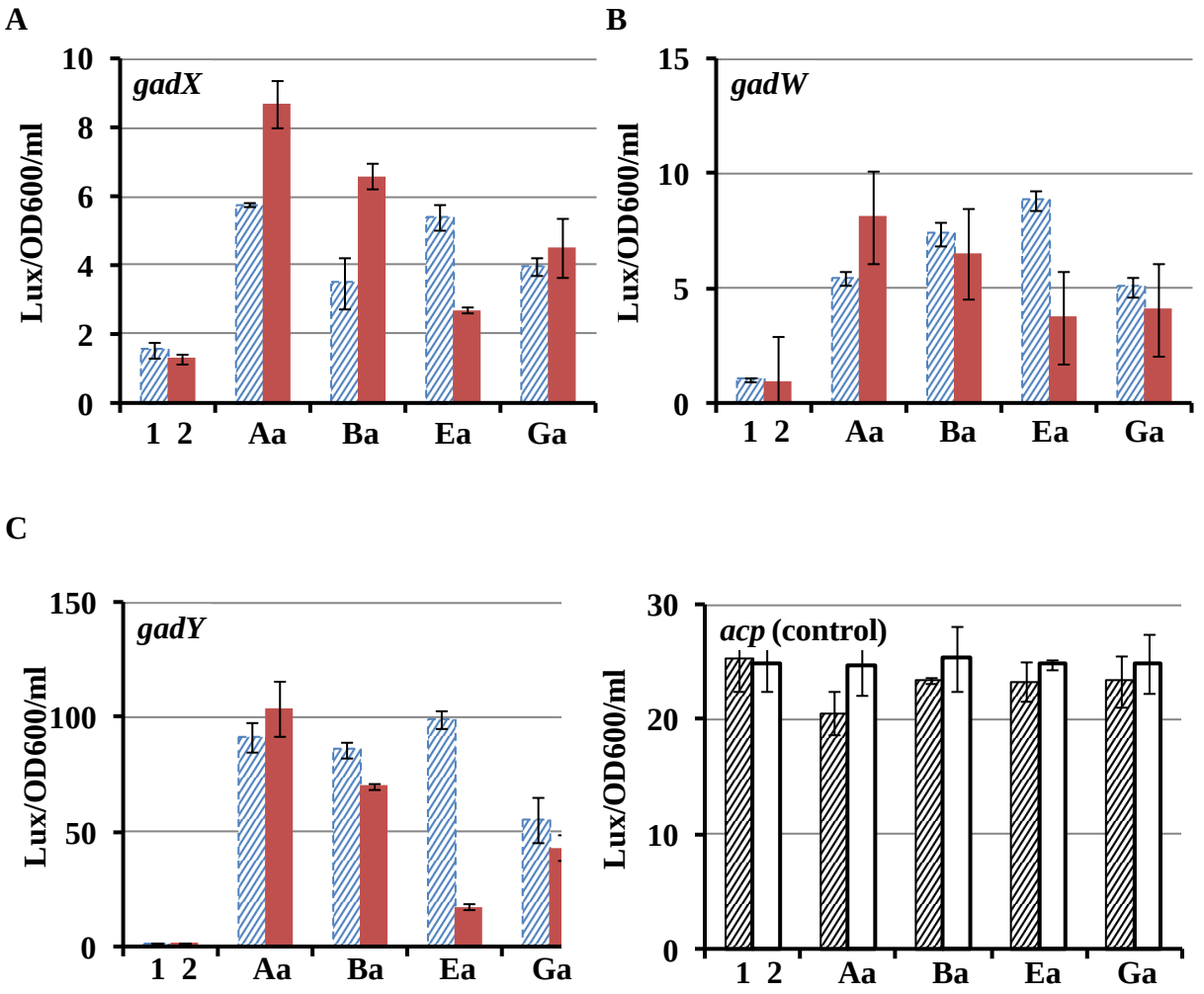


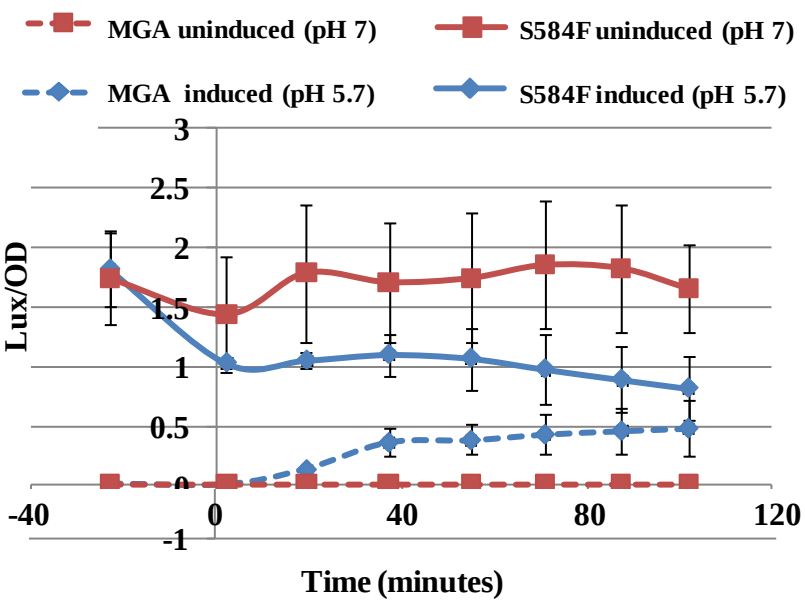
Figure 4.6.4. Activity of the *gadX* (A), *gadW* (B) and *gadY* (C) promoters in the ancestor, evolved and *evgS* mutant strains. The promoter activity was measured during exponential phase in M9supp medium at pH 7. Luciferase values were corrected for OD and culture volume. Lanes 1 and 2 represent the promoter activity in MGA and in MGA containing the plasmid p*evgAS* (contains a wild type of *evgA* and *evgS* under the control of the *evgA* promoter) respectively. The promoter activity of the MGA strain containing the *evgS* mutations are shown by red bars, the promoter activity of the evolved strains that contained that mutation are shown as blue striped bars. Each data point is the average of three independent biological repeats. The error bars represent the standard deviations between three biological repeats.

Mutations in the PAS domains of two-component systems can have an array of effects (Campbell *et al.*, 2010; Etzkorn *et al.*, 2008; Watts *et al.*, 2006). By looking at the induction of the *evgS* mutant strains and comparing their induction with the wild type strain we can answer two questions. Firstly, what level of activation do the EvgS mutations confer compared to acid induced wild type EvgS? Secondly, what is the affect of induction on the wild type and mutant EvgS?

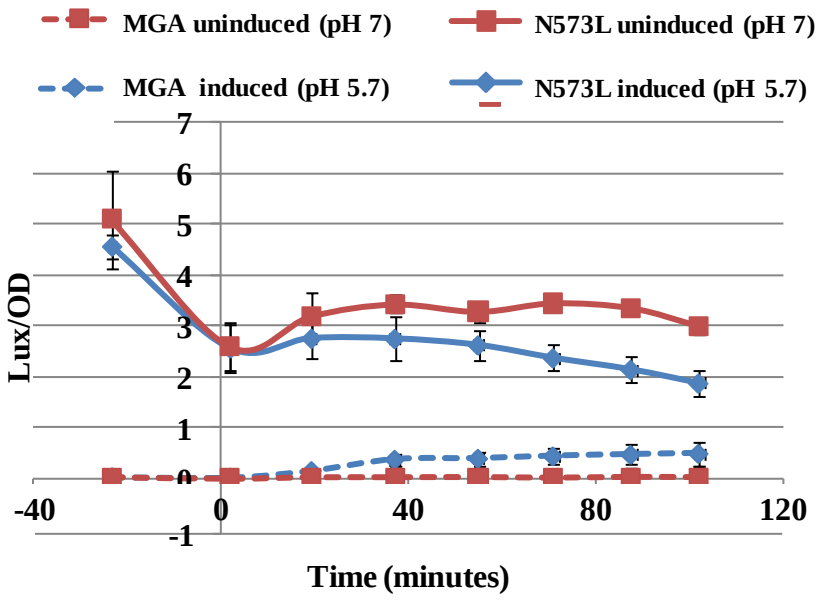
The *ydeP* promoter activity was used to assay the activity of the wild type *evgS* and *evgS* mutants with and without induction by mild acid. The *ydeP* promoter is directly bound by *evgA* and is the promoter that is most strongly activated by the EvgAS two-component system (Itou *et al.*, 2009). The promoter activity of *ydeP* can therefore be used as a representation of EvgS activity. The high-temporal resolution promoter probe assay was used to analyse the activation of the *ydeP* promoter by EvgAS with and without induction by acid (Burton *et al.*, 2010). This assay is described in full in Chapter 2. Briefly, cultures were grown to exponential phase and the promoter activity of the *ydeP* promoter was monitored before and after a pH shift from 7 to 5.7 in M9supp. The luciferase activity was corrected for the OD_{600} and plotted against the time relative to the point of induction.

The results are shown in figure 4.7, *evgS* mutants S584F, N573L, G658A and S600I are shown in figures A, B, C and D respectively. The *ydeP* promoter activity in the wild type EvgAS strain is shown in each figure for reference. The induction of the *ydeP* promoter by wild type EvgS is consistent with published results (Burton *et al.*, 2010). Wild type EvgS under uninduced conditions does not activate the *ydeP* promoter (figure 4.7 A-D). A shift to pH 5.7 causes the activation of the *ydeP* promoter by EvgAS.

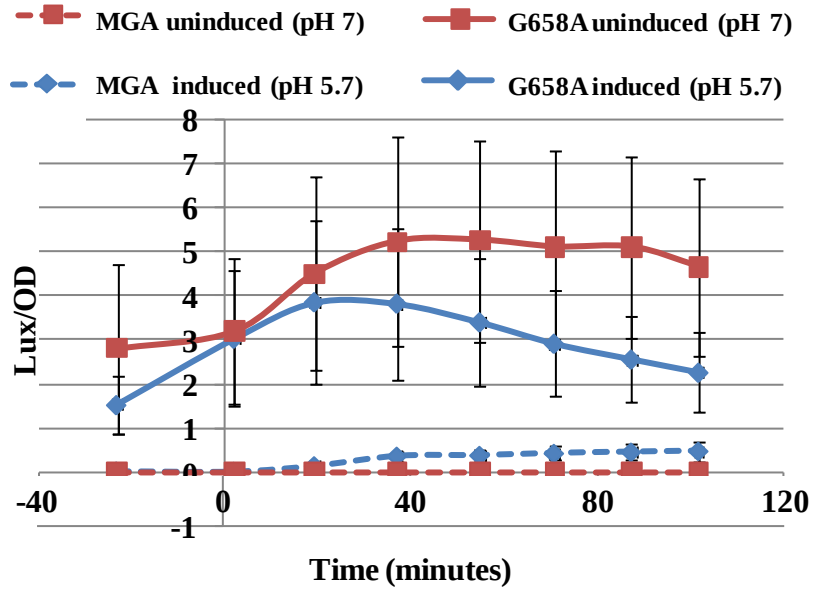
A


MGA uninduced (pH 7)
S584F uninduced (pH 7)
MGA induced (pH 5.7)
S584F induced (pH 5.7)
Lux/OD
Time (minutes)

B


MGA uninduced (pH 7)
N573L uninduced (pH 7)
MGA induced (pH 5.7)
N573L induced (pH 5.7)
Lux/OD
Time (minutes)

C


MGA uninduced (pH 7)
G658A uninduced (pH 7)
MGA induced (pH 5.7)
G658A induced (pH 5.7)
Lux/OD
Time (minutes)

D

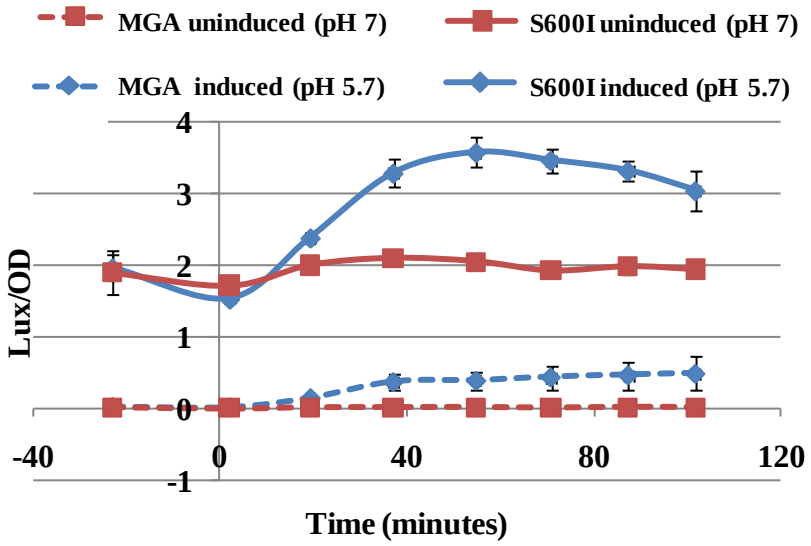


Figure 4.7. Dynamics of *ydeP* activation by EvgS mutants. Activity of the *ydeP* promoter in the EvgSc background with and without induction by mild acid shock (pH 5.7). The *ydeP* promoter activity in EvgSc mutants S573F, N584L, G658A and S600I is shown in figures A-D respectively. Induced cultures (Blue lines) were acidified to pH 5.7 at t = 0. Uninduced cultures (Red lines) were kept at pH 7 for the entire assay. The ancestor (wild type) *ydeP* promoter activity is shown in each figure for comparison (dotted lines). Each data point is the average of three biological repeats. The error bars represent the standard deviation of three biological repeats.

All of the EvgS mutants activate the *ydeP* promoter more than the wild type uninduced and induced. Interestingly for three of the EvgS mutants (S584F, N584L and G658A) the affect of acid induction actually reduces the promoter activity of the *ydeP* promoter. Conversely, mutant S600I is induced by exposure to mild acid, figure 4.7 D. The promoter activity of in the *evgS*cG658A strain is equivalent to the other *evgS* mutants; this result contradicts the previous observation regarding the low activity in this strain. However, the Ea strain was not assayed by this method for comparison. There are significant differences in the methods for single time point and high-temporal resolution promoter probe assays, which could explain this difference. The most significant difference is in the outgrowth of the cultures, in the single time point assays the cultures are grown in 96-well polypropylene plates, as opposed to the high-temporal resolution assays, which used conical flasks. Consequently the cultures are in a single time point assay are not aerated as well as those in a high-temporal resolution assay.

In summary, monitoring the induction of EvgS using the *ydeP* promoter probe revealed key differences between the EvgS mutations. The activation of *ydeP* by the EvgS mutants is much greater than that of the EvgS wild type even under inducing conditions. Induction by mild acid shock reduces the *ydeP* promoter activity in three of the EvgS mutant strains (S584F, N584L and G658A). However, even under these conditions the induction of the *ydeP* promoter is much higher in the EvgS mutant strains than in the induced wild type. Induction by mild acid shock increased the level of *ydeP* promoter activity in the S600I *evgS* mutant.

4.6. Acid resistance phenotype of *E. coli* 042

E. coli 042 is an EAEC (enteroaggregative *E. coli*) that causes dysentery (Chaudhuri *et al.*, 2010). Bioinformatic analysis of 28 strains of *E. coli* and *Shigella* revealed *E. coli* 042 was the only strain that did not contain an *evgAS* homologue. However, the other components of the AR2 network are present in this strain. As *E. coli* 042 has no *evgAS* homologue, and the EvgAS TCS is required for the induction of exponential phase acid resistance, it is reasonable to predict that this strain will have no inducible acid resistance. To test this hypothesis, we measured the acid resistance of the *E. coli* 042 over a time course at pH 2.5 during exponential phase and after 2 hours at pH 2.5 during stationary phase.

E. coli strain 042 was assayed for acid resistance using the standard acid resistance assay. The survival of *E. coli* K-12 MG1655 (MGA), *E. coli* 042 and *E. coli* Sakai were assayed during exponential phase growth for 2 hours at pH 2.5 in M9supp. Survival was measured at time zero, and every thirty minutes for 2 hours. The results are shown in figure 4.8 A. The survival of the K-12 strain was 0.8%. *E. coli* Sakai has 17% survival after 2 hours at pH 2.5, which is significantly higher than the *E. coli* K-12 strain. The survival of the *E. coli* 042 strain was below the level of detection of this assay, even after only 30 minutes at pH 2.5. *E. coli* 042 has significantly lower survival at extreme pH compared to the pathogenic strain *E. coli* sakai and the lab strain *E. coli* K-12 MG1655.

Acid resistance during stationary phase is independent of EvgAS. To see if the *E. coli* 042 strains are capable of surviving extreme acid stress the strains were tested during stationary phase. The results are shown in figure 4.8 B. The average survival of *E. coli* K-12 over 3 biological repeats was 57%. The average survival of *E. coli* Sakai over 3 biological repeats

was 42%. Finally, the average survival of *E. coli* 042 over 3 biological replicates was 11%. These results show that *E. coli* 042 can survive extreme acid shock in stationary phase. However, the level of survival is significantly lower than both *E. coli* K-12 and Sakai (t test, p values 0.0046 and 0.0025 respectively). In summary, *E. coli* 042 is a particularly acid sensitive strain of *E. coli*. During exponential phase growth no detectable survival was recorded. In fact, no recordable resistance could be observed even after only 30 minutes at pH 2.5. However, during stationary phase growth the 042 strain was capable of resisting extreme acid shock. It is therefore reasonable to predict that with a functional EvgAS the *E. coli* 042 strain will be able to survive extreme acid stress.

4.7. Acid resistance of *E. coli* 042 with *evgAS*

E. coli 042 has been shown to be sensitive to extreme acid shock and has no *evgAS* homologue. However, the mechanisms required for acid resistance are present in the strain. These mechanisms can also confer acid resistance during stationary phase. It is therefore reasonable to predict that, with a functional EvgAS to induce the AR2 system, *E. coli* 042 will be able to survive extreme acid shock in stationary phase.

To test this hypothesis, the acid resistance of *E. coli* 042 carrying the plasmid p*evgAS* was assayed using the standard acid resistance assay. The plasmid p*evgAS* expresses a copy of the *evgAS* locus from *E. coli* K-12 MG1655. The expression of *evgAS* is under the control of its native promoter, which makes the plasmid ideal for investigating the effect of *evgAS* in 042.

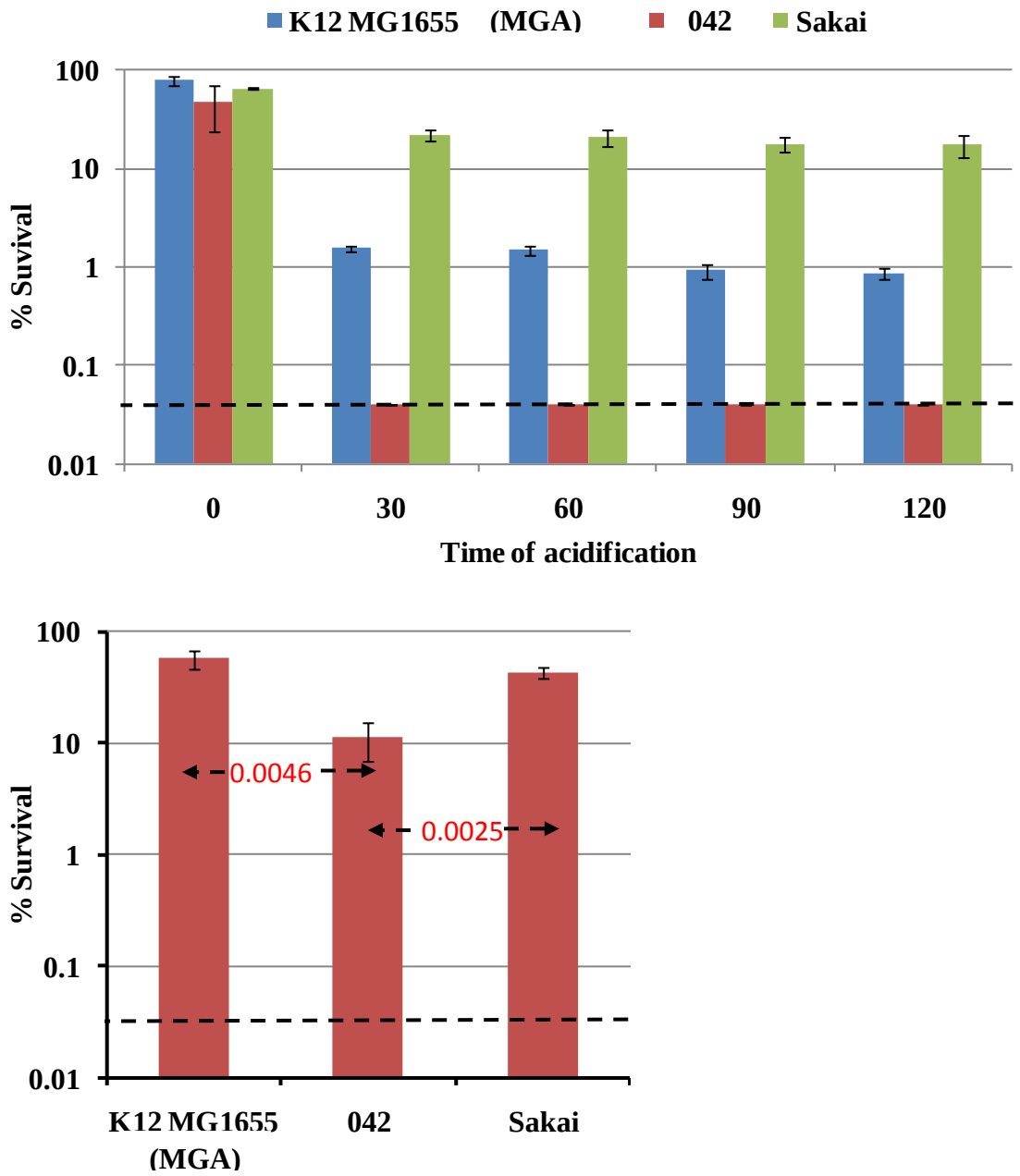


Figure 4.8. Acid resistance phenotype of *E. coli* K-12 MG1655, 042 and Sakai. Survival of *E. coli* K12 MG1655, 042 and Sakai after 0, 30, 60, 90 and 120 minutes in pH 2.5 M9supp medium (A). Survival of *E. coli* K12 MG1655 (MGA), 042 and Sakai after 2 hours in pH 2.5 M9supp medium during stationary phase (B). Survival represents the scored colonies after 2 hours at pH 2.5 expressed as a percentage of scored colonies at pH 7 after 0 hours. T-test p values are displayed in red, black dotted arrows indicate the values that have been compared. Each data point is an average of three independent biological replicates. Error bars represent the standard deviation of thee independent biological replicates. A dotted black line represents the limit of detection for the assay, predetermined to be 0.04%.

The survival of *E. coli* 042, *E. coli* 042 +pZC320, *E. coli* 042 + p*evgAS* are presented in figure 4.9. The survival of the *E. coli* 042 strain is below the limit of detection (predetermined to be 0.04%). The survival of this strain with either the pZC320 plasmid or the p*evgAS* plasmid was also below the limit of detection. These results indicate that the pZC320 plasmid does not affect the level of survival; any survival detected by the inclusion of the p*evgAS* plasmid would therefore be due to the expression of *evgAS*. However, the inclusion of p*evgAS* into *E. coli* 042 had no affect on survival. This result suggests that either basal levels of *evgAS* have no affect on survival or that *E. coli* K-12 *evgAS* is incapable of activating the AR2 network in *E. coli* 042. To test the latter hypothesis, the 042 +p*evgAS* strain was tested for acid resistance after 70 minutes of induction at pH 5.7. The survival of 042 +p*evgAS* after induction increased to 0.27%. This level of survival is not comparable to the acid resistance of *E. coli* K-12 after 70 minutes of induction, which was previously shown to be 15%.

Providing a plasmid encoded *evgAS* didn't confer any increase in acid resistance until the cells were induced at pH 5.7 for 70 minutes. This induction is known to activate the EvgAS two component system in *E. coli* K-12 and confers nearly 10-fold more resistance than *E. coli* 042. Acid induction of the pathway through EvgAS conferred a small amount of resistance. How much resistance could an *evgS* constitutive-on mutant confer in *E. coli* 042?

4.8. Acid resistance in *E. coli* 042 with *evgAS*c

The results shown above demonstrate that the addition of a plasmid encoded *evgAS* into *E. coli* 042 enabled acid resistance to be induced during exponential phase by mild acid shock. This suggests that the AR2 network can function in the same way in *E. coli* 042 as it does in *E. coli* K-12. Earlier in this chapter the plasmid p*evgAS*cS584F was characterised.

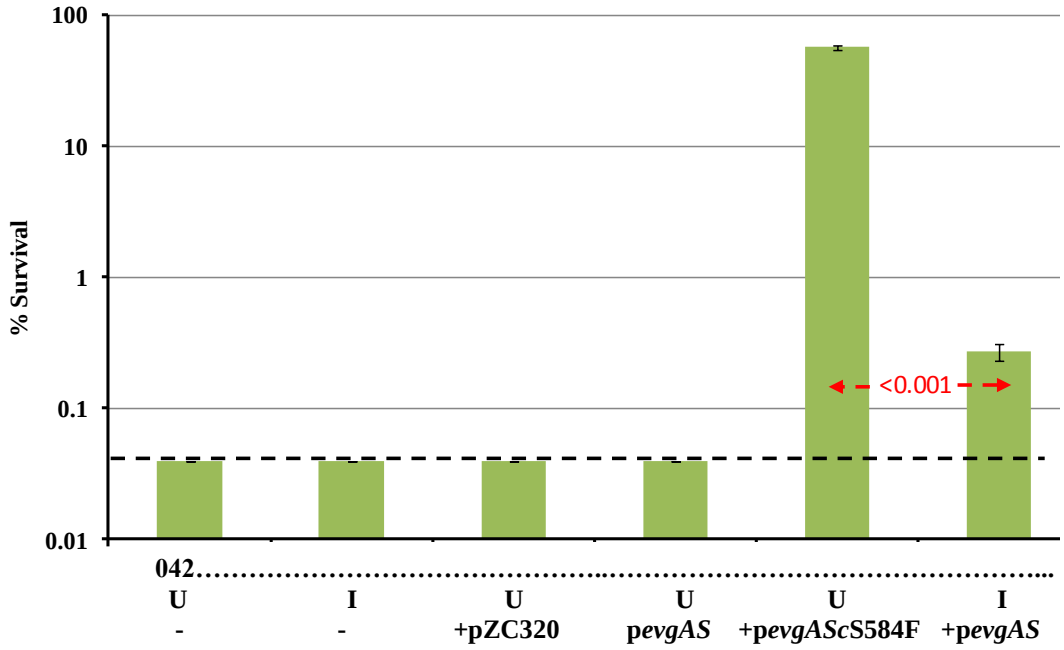


Figure 4.9. Acid resistance of *E. coli* 042. Acid resistance of 042 was measured with and without plasmids p*evgAS* or p*evgAS*cS584F in inducing and non-inducing conditions**.** Survival was assayed in M9supp medium, after 2 hours at pH 2.5. Survival represents the scored colonies after 2 hours at pH 2.5 represented as a percentage of scored colonies after 0 hours pH 7. U and I indicate uninduced and induced cultures respectively. t-test p-values are displayed in red, red dotted arrows indicate the values that have been compared. Each data point is an average of three independent biological replicates. Error bars represent the standard deviation of thee independent biological replicates. A dotted black line represents the limit of detection for the assay, predetermined to be 0.04%.

This plasmid contains a mutant copy of *evgS* that has been shown to constitutively activate the AR2 network and confer high levels of acid resistance. As the addition of a wild type *evgAS* in to *E. coli* 042 conferred inducible acid resistance, it is reasonable to predict that an EvgS constitutive-on mutant will confer an acid resistant phenotype. To test this hypothesis, we measured the affect of the p*evgS*cS584F plasmid on the acid resistance phenotype of *E. coli* 042. The plasmid was transformed into the strain and tested using the standard acid resistance assay. Cultures were shocked in pH 2.5 M9supp medium for 2 hours. The % survival is presented in figure 4.9. The survival of *E. coli* 042 +p*evgAS*cS584F was 56.9%, which is within the range of survival that was measured in *E. coli* K-12 +p*evgAS*cS584F. The survival conferred by the *evgS* constitutive mutant is also significantly higher than that of the *E. coli* 042 + p*evgAS* strain with induction. This experiment demonstrates the elevated level of induction conferred by the *evgS* constitutive mutants compared to the induced wild type *evgS*.

4.9. Discussion

4.9.1 The PAS domain of EvgS

Analysis of the EvgS mutations found in the evolved strains has shown that all of the mutated residues are conserved. Each EvgS mutation is located in the PAS domain. Isolation of each mutation indicates that the EvgS mutations are conferring activation of the AR2 network in uninducing conditions. The EvgS mutations can also confer acid resistance to the same level as the evolved strains.

PAS domains are ubiquitous across all kingdoms of life, in bacteria, they are often found in the histidine kinase (HK) of two-component systems. PAS domains have a range of functions

and the number of PAS domains varies between HKs. For a complete review of presence and function of PAS domains across an array of life see Taylor and Zhulin 1999. PAS domains have low sequence conservation despite their similar structure as a result many PAS sequences fall into the "twilight zone" (Finn *et al.*, 2010; Holm *et al.*, 1992; Vogt *et al.*, 1995). The sequence differences enable PAS domains that are structurally similar to interact with a variety of co-factors and sense a variety of signals. PAS domains have been shown to bind heme groups, FAD, divalent cations, 4-HCA, FMN and Di/tricarboxylic acids (Gilles-Gonzalez *et al.*, 1994; Qi *et al.*, 2009; Reinelt *et al.*, 2003; Vescovi *et al.*, 1997). PAS domains have also been shown to sense alterations in redox potential in the cell (Soderback *et al.*, 1998). The role of the PAS domain in EvgS is not understood. Structural modelling of the PAS domain of EvgS using other solved PAS domains as a template has revealed a possible pH sensing domain. Three histidine molecules, which have a favourable pKa for sensing pH changes in the cell (pKa of the imidazole side chain is 6.0 and 6.5 for the whole amino acid), are found buried in close proximity. In addition, the histidines are surrounded by acidic residues (Dr Peter Winn, unpublished communication). The PAS domain of EvgS is located to the cytoplasm very near to the inner membrane. PAS domains have been shown to interact with many molecules including lipids so it is possible that the EvgS PAS domain could be interacting with the inner membrane.

4.9.2 N-terminal cap mutations

Two of the four mutations are located in the N-terminal cap (N573L, S584F). This location has been shown to be poorly conserved among PAS domains. The N-terminal cap has been shown to be involved in dimerisation in DcuS using solid state NMR (Etzkorn *et al.*, 2008). The long alpha helix structure of the N-terminal cap forms a straight edge that facilitates

protein-protein interactions. The N-terminal cap is also the most highly disordered part of the PAS domain (Taylor and Zhulin, 1999; Watts *et al.*, 2006). Mutations that change the stability of this domain have already been shown to affect the function of the HK (Watts *et al.*, 2006). In addition, both N-terminal cap mutations are Polar to non polar residue exchanges; this could be a significant change in the structure or disorder of the protein. The N-terminal cap has also been found to be very important for signalling in Aer TCS (Watts *et al.*, 2006). A full mutagenic study of the N-terminal cap found that certain N-terminal cap mutants caused constitutive-on phenotypes. Other mutations conferred constitutive off and even reverse phenotypes that respond to the stimulus in the opposite way from the wild type protein (Watts *et al.*, 2006). Although the N573L and S584F confer constitutively on activity of the AR2 network, mild acid induction of these mutants conferred a reduction of *ydeP* promoter activity. An explanation of this phenomenon is that the mutations have shifted the pH at which the EvgS protein becomes activated. Investigation into the induction of the AR2 network by Burton *et al.*, 2010 revealed that the network was induced specifically at pH 5.7. Other levels of acidity actually reduced the induction of the network. These results suggest that EvgS responds to a specific pH. It is therefore conceivable that the *evgS* mutations have shifted the optimum pH towards pH 7, causing activity in normally uninducing conditions and the reduction in activity in more acidic conditions.

The EvgS mutation S1, characterised to be constitutively-on by Kato *et al.*, 2000, is also located in the N-terminal cap of the PAS domain. The location of the N573L and S584F mutations is clearly crucial for EvgS signalling it is possible that these mutations are effecting the dimerisation of the EvgS. In addition, the affect of induction on the activity of these EvgS mutations suggests that the optimum pH of EvgS has been altered. However, this effect was

indicated by the activity of the *ydeP* promoter. It is therefore possible that many other factors could be affecting the activity of this promoter. Conclusive characterisation of these mutations can only be obtained by thorough structural analysis, as a result, the role of the PAS domains and the N-terminal cap in EvgS and other proteins could be better understood.

4.9.3 PAS core mutation S600I

The S600I mutation is located in the PAS core of the protein. The PAS core is the active site of the PAS domain that is often responsible for binding cofactors (Reinelt *et al.*, 2003). The active site of the core is also mainly hydrophobic. The S600I mutation results in a hydrophilic to hydrophobic residue change. Such a mutation could alter the structure of the PAS core and effect cofactor binding. Interestingly, this *evgS* mutation is still inducible with a mild acid shock suggesting that the S600I mutation does not affect the optimum pH of the HK or the pH regulated component of the HK. Many two component systems respond to more than one stimulus, this could be the case for the EvgAS system. Comprehensive structural characterisation of the S600I mutation could uncover possible cofactors involved in EvgS sensing and other input signals.

4.9.4 Beta-scaffold mutation of EvgS

EvgS mutation G685A is located in the beta-scaffold of the PAS domain. The amino acid base change itself is conservative. However, the mutation is close enough to the putative pH sensing domain to significantly affect its function. The promoter activity of the *ydeP* promoter is reduced by mild acid induction in this mutant background, which is a similar affect of induction on promoter activity to the N-terminal cap mutations. Suggesting that the mutation could have altered the optimum pH at that EvgS activates.

4.9.5. Activation of the acid resistance genes in evolved strain Ea

The level of activation conferred by EvgS mutation G658A in strain *evgS*cG658A is much lower than that of the evolved strain Ea. This suggests that another mutation found in evolved strain Ea is affecting the activation of the AR2 network. However, despite the reduced activation of the network the level of survival of the *evgS*cG658A strain was not significantly different from the evolved strain Ea. The activity of the all of the promoters was significantly lower in the *evgS*cG658A strain compared to the Ea evolved except for the EvgA promoter. This suggested that another mutation or mutations, present in the evolved strain Ea, was affecting the activation of the network. The other mutations found in evolved strain Ea include; *dcp*, *iscR*, *priA* and *yijA*. None of these genes have been previously characterised to be involved in acid resistance. The mutation that is most likely to be involved in the regulation of AR2 would be *iscR*, as this is the only mutation found in a known regulator (Wu and Outten, 2009). In addition, the mutation in *iscR* is located in the DNA binding domain. However, results from the high-temporal resolution assays, shown in figure 4.7 D contradicted this finding. The conditions of the high-temporal resolution assays are different from that of the single time point assays. Specifically, the assays differ in the aeration of the cultures as the high-temporal resolution assays use flask culture, which provided better aeration. It is possible that the other mutations that are required for activation of the network in the Ea strain are required in micro-aerobic conditions.

4.9.6. EvgS in *E. coli* 042

A multiple alignment of the *evgS* genes from 28 sequenced strains of *E. coli* and *Shigella* species revealed that *E. coli* 042 did not have an *evgAS* locus. This result prompted

investigation of the acid resistance phenotype of *E. coli* 042 and the role of *evgAS* in 042 acid resistance.

Acid resistance assays on wild type *E. coli* 042 revealed that not only did this strain of *E. coli* have very low acid resistance but also the acid resistance could not be induced with a mild acid shock. Introduction of wild type *evgAS* into the *E. coli* 042 strain resulted in no additional resistance in non-inducing conditions. However, a small increase in resistance was observed when the *E. coli* 042 +p*evgAS* strain was induced with a mild acid shock. This indicated that when EvgAS are present, the rest of the pathway was capable of conferring acid resistance. Finally, when the constitutive-on *evgS* was introduced into the strain the survival increased significantly.

The ability of a constitutive on EvgS to confer such high levels of resistance is interesting. However, considering the level of activity of the *ydeP* promoter conferred by the constitutive on mutation this difference is understandable. In addition, the lack of survival of the p*evgAS* strain with induction could be due to differences in the EvgS environment. *E. coli* 042 expresses an LPS layer that could buffer pH changes and alter the internal pH of the cell. It is possible that the pH in the cytoplasm and periplasm does not drop as much in *E. coli* 042 as it does in *E. coli* K-12.

Overall, this chapter has demonstrated that the EvgS mutations are causing the acid resistant phenotype displayed by the evolved strains. These mutations are activating the AR2 network and causing the over-expression of the regulators and structural genes involved in acid resistance. This activation is powerful enough to confer acid resistance in *E. coli* 042, which

has been characterised as a very acid sensitive strain. The actual acid resistance genes that are conferring acid resistance to the *evgS* constitutive-on strains are still unknown. From this point EvgSc will refer to any strain contains an *evgS* constitutive-on mutation excluding the evolved strains. What components of AR2 acid resistance are required for the EvgSc based acid resistance phenotype?

4.9.7. Summary

1) All of the EvgSc mutations are located in the PAS domain of EvgS. The residues which have been mutated are conserved amongst all *E. coli evgS* genes. However, *E. coli* 042 does not have an *evgAS* locus.
2) When isolated the EvgSc mutations were all capable of conferring acid resistance to levels observed by the evolved strains. This indicates that the EvgS mutations alone can confer an acid resistant phenotype.
3) Each EvgSc mutation was capable of activating the AR2 network. Revealing that the EvgSc mutations were in fact causing a constitutive-on EvgS. However, EvgSc mutation G658A was not capable of conferring equivalent levels of survival as the evolved strain Ea suggesting that other mutations in Ea are important for induction of the AR2 network.
4) Comparison of the EvgSc mutants and the activity of the induced wild type EvgS revealed that the EvgSc mutations could all induce the *ydeP* promoter significantly more than the induced wild type.
5) The affect of induction by mild acid shock on the activity of *ydeP* in the EvgSc mutants varied. Mutations S573F, N584L and G658A all showed reduced activity after induction; whereas S600I further activated the *ydeP* promoter when induced.

6) *E. coli* 042 is very sensitive to extreme acid challenge. Survival was low even during stationary phase growth. The addition of a plasmid encoded *evgAS* locus did not increase acid resistance. However, when the strain was induced the *evgAS* locus did confer low levels of acid resistance.
7) Introduction of a plasmid encoded constitutive on *evgAS* locus into *E. coli* 042 increased acid resistance dramatically.

CHAPTER 5: RESULTS (3): Contributions of AR2 components to EvgSc acid resistance

5.1. Role of EvgA in EvgSc acid resistance

The analysis of EvgSc mutations, which were selected by an evolution experiment, has shown that these mutations can confer increased acid resistance of *E. coli* K-12 during exponential phase. These mutations also cause the activation of the AR2 network, which regulates the AFI (acid fitness island) and GAD (glutamate dependent) acid resistance systems. This chapter describes the contributions that each AR2 regulated mechanism made to acid resistance in EvgSc (EvgS constitutively active) strain. Throughout this chapter the plasmid encoded EvgSc mutations were used. For details of the construction of these plasmids see Chapter 2. The affect of the plasmid encoded EvgSc mutations on acid resistance and AR2 activity was covered in Chapter 4. These plasmids allowed fast and straightforward transfer of the EvgSc mutations into an array of mutant backgrounds.

5.1.1. Dependence of EvgSc phenotype on the response regulator EvgA

EvgS is a histidine kinase that forms a TCS with the response regulator EvgA (Bantscheff *et al.*, 2000). Much of what is known about the activation of the EvgAS TCS has been transferred from BvgAS. BvgS is an EvgS homologue from *B. purtussis* that regulates virulence genes (Bock and Gross, 2002). The current mechanism of activation of EvgAS TCS is as follows: the EvgS transmitter domain autophosphorylates upon receiving a signal (Bock and Gross, 2002). The transmitter domain then phosphorylates the receiver domain. Then either the receiver domain phosphorylates the Htp domain or after a second autophosphorylation, the transmitter domain phosphorylates the Htp domain (Bock and Gross, 2002). Finally, the phosphorylated Htp domain phosphorylates EvgA and EvgA binds its consensus DNA binding sites (Bock and Gross, 2002; Masuda and Church, 2002). Currently there is no account of EvgS activating any other response regulator other than EvgA. The

activation of AR2 by EvgA has been well characterised by *in vitro* and *in vivo* studies (Burton *et al.*, 2010; Itou *et al.*, 2009). Direct binding of EvgA to the *safA-ydeO* and *ydeP* promoters has been demonstrated (Itou *et al.*, 2009). In addition, the activation of the AR2 network by acid induction in exponential phase has been shown to be completely EvgA-dependent (Burton *et al.*, 2010). However, the activation of AR2 promoters by EvgSc could be independent of EvgA. For example, the activating mutations in EvgS could have altered the specificity of EvgS causing it to activate proteins other than EvgA. Based on the current understanding of EvgAS activation of the AR2 network, we hypothesise that the acid resistance conferred by the EvgSc mutations is dependent on EvgA. To test this hypothesis, the acid resistance phenotypes of EvgSc strains containing an Δ*evgA* deletion were assayed.

The *evgA* locus was mutated by Datsenko and Wanner mutagenesis in the MGA and *evgS*cG658A strains. The standard acid resistance assay was used to assay survival. For details of the mutagenesis and the acid resistance assay see Chapter 2. Briefly, cultures were exposed to pH 2.5 in M9supp medium for 2 hours. Induced cultures were exposed to pH 5.7 for 70 minutes prior to exposure to pH 2.5. The results are shown in figure 5.1. For comparison, the MGA and MGA Δ*evgA* strains were assayed in both uninduced (orange bars) and induced (blue bars) conditions. The average survival of MGA uninduced and MGA induced cultures over 3 biological repeats was 0.66% and 15.61% respectively. This is similar to previously reported survival levels for this strain. The average survival of the MGA Δ*evgA* strain in uninduced and induced conditions over 3 biological repeats was 0.11% and 0.69% respectively. The survival of the uninduced MGA Δ*evgA* strain shows that EvgA is required for survival even without induction. This could be due to the survival conferred by basal activation of EvgA regulated genes.

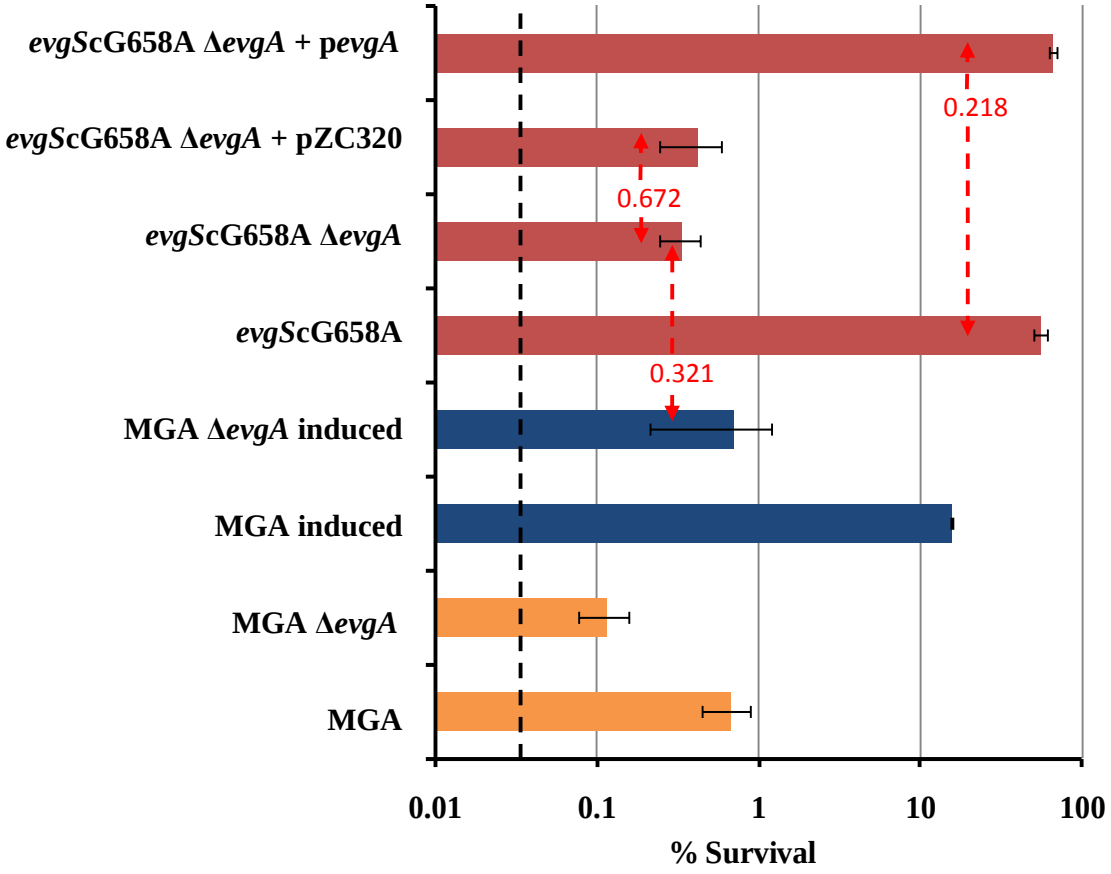


Figure 5.1. Survival of *E. coli* containing an EvgSc mutations in an *evgA* knockout background after 2 hours at pH 2.5. The survival of an *evgSc*G658A Δ*evgA* strain was compared to the survival of MGA and MGA Δ*evgA*, with and without induction, and *evgSc*G658A. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The Δ*evgA* mutation was complemented by supplying wild type *evgA* on plasmid p*evgA* .The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively. Red arrows and numbers represents comparisons of data by T-test and T-test p values resperctively.

The survival of the MGA Δ*evgA* strain after induction is equivalent to the survival of the uninduced MGA strain. This is consistent with previous findings that all induced resistance is dependent on EvgA (Burton *et al.*, 2010). The average survival of the *evgS*cG658A strain over 3 biological repeats was 55%. This result is consistent with previous survival levels for this strain described in Chapter 4. The average survival of *evgS*cG658A Δ*evgA* over 3 biological repeats was 0.48%. This level of survival is equivalent to that of the MGA Δ*evgA* strain when induced (0.69%, shown above). Statistical comparison of the *evgS*cG658A Δ*evgA* and MGA Δ*evgA* strains showed that they are not significantly different (t-test, p-value 0.321). The contribution of the EvgSc mutation is therefore completely dependent on the response regulator EvgA. Furthermore, the structural acid resistance genes that are conferring the evolved resistance are regulated either directly or indirectly by EvgA.

5.1.2 Complementation of *ΔevgA* phenotype

The results described in 5.1.1 show that the introduction of an *evgA* knockout mutation reduces both acid induced survival and EvgSc survival. In both cases, survival was lowered to levels equivalent to the uninduced wild type. This indicates that all of the EvgSc induced survival is dependent on EvgA. However, it is possible that other genes may have been affected when *evgA* was mutated. This is because the introduction of a resistance cassette into the chromosome requires a homologous recombination event. To increase the frequency of a homologous recombination lambda red genes are expressed in the strain. The frequency of other recombination events is also increased and this can lead to secondary mutations (Hobman *et al.*, 2007). To ensure that the phenotypes described in 5.1.1 are due to the removal of *evgA*, the *evgA* gene was reintroduced into the strain on a plasmid. If the

*evgS*cG658A Δ*evgA* phenotype is due to the deletion of *evgA* then the addition of a plasmid encoded *evgA* will complement the phenotype.

The plasmid p*evgA* was used to complement the *evgA* deletion. Details of the construction of p*evgA* can be found in Chapter 2. Briefly, the *evgA* operon and upstream intergenic region was cloned into pZC320. This plamsid was used as it has an F' plasmid replicon that confers a low copy number (Shi and Biek, 1995). It is essential to control the levels of EvgA as over-expression can confer acid resistance (Masuda and Church, 2002). The average survival of the *evgS*cG658A Δ*evgA* +p*evgA* and *evgS*cG658A Δ*evgA* +pZC320 strains is shown in figure 5.1. The average survival of *evgS*cG658A Δ*evgA* +pZC320 over 3 biological repeats was 0.56%. The survival of the *evgS*cG685A Δ*evgA* was 0.48% and the difference between them was not significant (t test, p value 0.672). This result shows that the survival of the *evgS*cG658A Δ*evgA* is not affected by the presence of the pZC320 plasmid. The average survival of the *evgS*cG658A Δ*evgA* +p*evgA* over 3 biological repeats was 61%. This was significantly higher survival than the *evgS*cG658A Δ*evgA* +pZC320 strain (t test, p value < 0.001). There was no significant difference between the survival of *evgS*cG658A Δ*evgA* +p*evgA* compared to the *evgS*cG658A strain (t test, p value 0.218). Therefore, addition of a plasmid encoded EvgA conferred comparable levels of survival to the *evgS*cG658A strain. This shows conclusively that the mutation in *evgA* is responsible for the reduced survival shown by the *evgS*cG658A Δ*evgA* strain.

To summarise, all of the EvgSc conferred acid resistance is dependent on the response regulator EvgA. The introduction of a plasmid encoded *evgA* restored survival to levels comparable to an EvgSc strain. The EvgAS TCS activates many acid resistance genes, some

are activated directly such as *ydeP,* but most are activated indirectly by other regulators. The contributions of all EvgAS regulated components of the AR2 network to EvgSc acid resistance are described below.

5.2. Role of RpoS-PhoPQ in the activation of EvgSc acid resistance

The *safA* promoter is activated by EvgA (Masuda and Church, 2003). The product of *safA* is a small membrane protein, also called SafA, which activates PhoQ. PhoQ is the histidine kinase of the PhoQP TCS that responds to low Mg^{2+} and low pH (Choi *et al.*, 2009; Eguchi *et al.*, 2007). The regulon of PhoP includes *iraM* that codes for IraM. IraM inactivates RssB, which is a protein that targets RpoS for degradation. As a result, the activation of PhoPQ by SafA causes an increase in RpoS levels (Eguchi *et al.*, 2011). RpoS is a sigma factor that recruits RNAP to a subset of promoters. Included in this subset are the *gadE, gadY* and *gadX* promoters (Ling *et al.*, 2008). Activation of these promoters causes the activation of the GAD system and parts of the AFI (*hdeA*, *hdeD* and h*deB*). Research by Eguchi *et al.* in 2011 reported that survival of a constitutive on EvgS mutant (EvgS1) was reduced from 2.65% to 0.022% by a *phoP* mutation (Eguchi *et al.*, 2011). This effect of the *phoP* mutation was explained by induction cascade described above.

This study has already addressed the role of RpoS in the survival of the evolved strains. The survival of the evolved strains, which was shown to be due to a constitutive-on EvgS, is RpoS independent (Chapter 3). However, the published data indicates that all survival conferred by a constitutive-on EvgS mutant is RpoS and PhoP dependent. Based on this data, we hypothesised that an Δ*rpoS* Δ*phoP* double mutant will eliminate all EvgSc resistance. To test this hypothesis an Δ*rpoS* Δ*phoP* double deletion strain was constructed. This strain was

transformed with plasmid p*evgAS*cN573L, which encodes a wild type EvgA and a constitutive-on EvgS. The acid resistance of this strain was assayed and compared to the MGA strain with the same double deletion in both induced and uninduced conditions.

The survival of the Δ*rpoS* Δ*phoP* double deletion strain is shown in figure 5.2. All strains were tested using the standard acid resistance assay described in Chapter 2. After shocking the cultures for 2 hours at pH 2.5 in M9supp survival was measured. The survival of the MGA, MGA induced and the MGA +p*evgAS*cN573L strains are shown in the figure for comparison.

The average survival of uninduced Δ*rpoS* Δ*phoP* over 3 biological repeats was below the limit of detection (predetermined as 0.04%, (Burton *et al.*, 2010)). This was a significant reduction in survival that is consistent with Δ*rpoS* mutants as shown in Chapter 3. The average survival of induced Δ*rpoS* Δ*phoP* over 3 biological repeats was 5.34%. The difference in survival between this double mutant strain and the induced MGA strain was significant (t test, p-value <0.005). Induction by mild acid shock is mediumted though EvgAS. This result shows that at least some activation of acid resistance requires *rpoS* and *phoP*. The average survival of MGA Δ*rpoS* Δ*phoP* +p*evgAS* over 3 biological repeats was below the limit of detection. This result shows that there is no protection conferred by the addition of a plasmid encoded copy of *evgAS* in this background. Finally, the average survival of the Δ*rpoS* Δ*phoP* +p*evgAS*cN573L strain over 3 biological repeats was 28.76%. This is a drop in survival of nearly 50% compared to MGA +p*evgAS*cN573L, which had an average survival of 55.39% over 3 biological repeats. However, statistical comparison of these values indicated that the difference was not significant (t test, p-value 0.171).

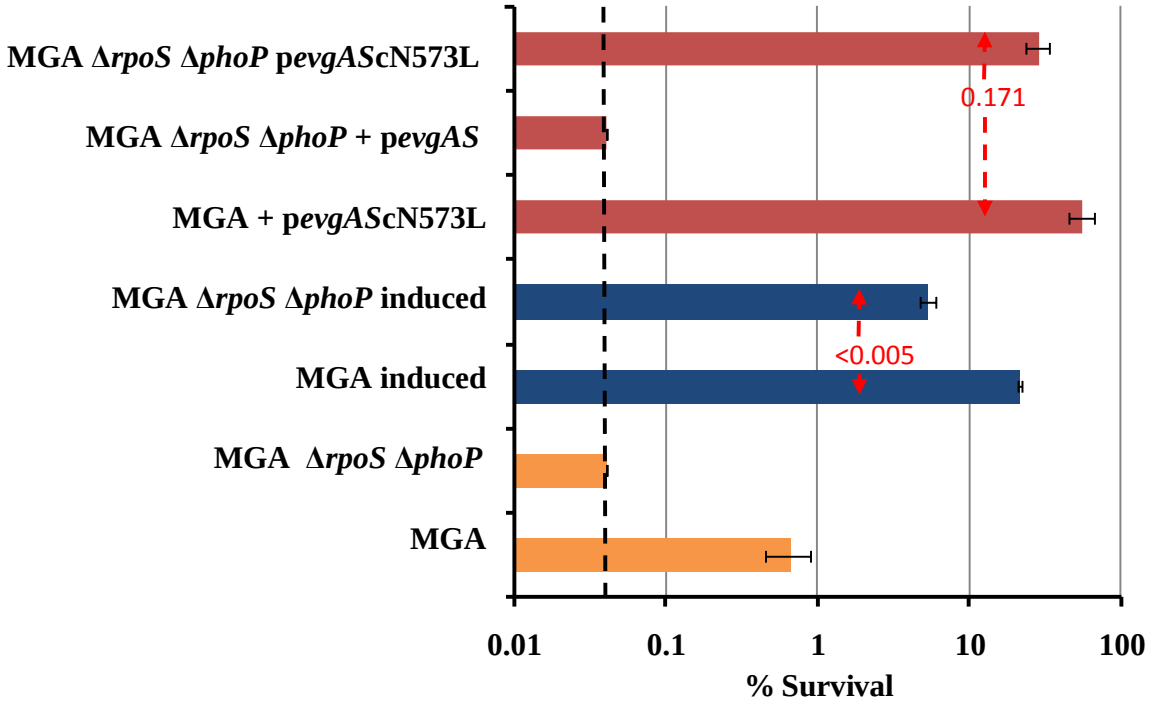


Figure 5.2. Survival of *E. coli* containing an EvgSc mutation in an *rpoS phoP* double knockout background after 2 hours at pH 2.5. The survival of MGA Δ*rpoS* Δ*phoP* containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA Δ*rpoS* Δ*phoP*, with and without induction, and MGA +p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively. Comparisons made by T-test are indicated by red arrows, p value from the test is indicated in red.

To summarise, the survival conferred by an EvgSc mutant is independent of PhoP and RpoS. This result is in contrast to published results that show that nearly all acid resistance of a different EvgSc mutant (EvgS1, Eguchi *et al.*, 2010) was dependent on a PhoP-RpoS activation cascade. A similar reduction in survival was measured when the Δ*rpoS* Δ*phoP* double deletion strain was tested in inducing conditions (survival of the Δ*rpoS* Δ*phoP* dropped to 33% of the wild type). However, this difference in survival, caused by the deletion of *rpoS* and *phoP*, was not statistically significant. In addition, even with both of these genes deleted the EvgSc mutation was able to confer 28.76% survival. This level of survival is still higher than the survival of MGA when induced.

5.3. Role of the GadE and YdeO in EvgSc acid resistance

5.3.1. The role of GadE in EvgSc acid resistance

GadE activates the acid resistance genes of the GAD system (*gadA/B* and *gadC*) and genes found in the AFI (Hommais *et al.*, 2004; Ma *et al.*, 2003a). The GadE promoter has been shown to be activated both directly and indirectly by EvgA (indirect via YdeO) (Ma *et al.*, 2004). The deletion of *gadC*, which codes the antiporter that is essential for GAD function, had little effect on survival at pH 2.5 (section 3.4.5.). This was a surprising result as the GAD system is the most robust acid resistance system at pH 2.5 (Lin *et al.*, 1996). However, genes found in the AFI can also confer high levels of acid resistance, in particular, the *hdeA*, *hdeB* and *hdeD* genes that code for periplasmic chaperones and an inner membrane protein respectively (Gajiwala and Burley, 2000; Kern *et al.*, 2007; Mates *et al.*, 2007). The removal of *gadE* from an EvgSc strain will effectively abolish promoter activity of the GAD system and the *hde* genes found on the AFI (Burton *et al.*, 2010). Considering that the main acid resistance structural genes regulated by EvgAS will be inactive, we hypothesise that the

survival of an EvgSc Δ*gadE* strain will be significantly reduced. To test this hypothesis, a Δ*gadE* deletion strain was constructed using Datsenko and Wanner mutagenesis (Datsenko and Wanner, 2000). The survival of the MGA Δ*gadE* strain with and without induction and the MGA Δ*gadE* +p*evgAS*cN573L strain were tested using the standard acid resistance assay described in Chapter 2.

The results of this assay are shown in figure 5.3.1. The average levels of survival of MGA with and without induction, and MGA +p*evgAS*cN573L are shown on the figure for comparison. The average survival of the MGA Δ*gadE* strain over 3 biological repeats was below the limit of detection for the assay. These results showed that without induction basal levels of GadE can confer some acid resistance. The average survival of MGA Δ*gadE* under inducing conditions over 3 biological repeats was 0.075%. This is a significant reduction in survival compared to the induced MGA (t test, p value <0.001). This indicates that all of the survival conferred by normal acid induction is dependent on GadE. The average survival of MGA Δ*gadE* +p*evgAS*cN573L over 3 biological repeats was 39.67%. This is lower than the survival of MGA +p*evgAS*cN573L, which was 55.39%, but the difference was not significant (t test, p value 0.331). This result shows that in an EvgSc mutant, mechanisms independent of GadE activation are capable of conferring acid resistance.

In summary, the Δ*gadE* deletion had significant effects on survival of the MGA strain with and without induction. This is consistent with previous work regarding the role of GadE in acid resistance (Burton *et al.*, 2010; Hommais *et al.*, 2004). However, the EvgSc mutation N574L was capable of conferring an acid resistant phenotype in the absence of *gadE*.

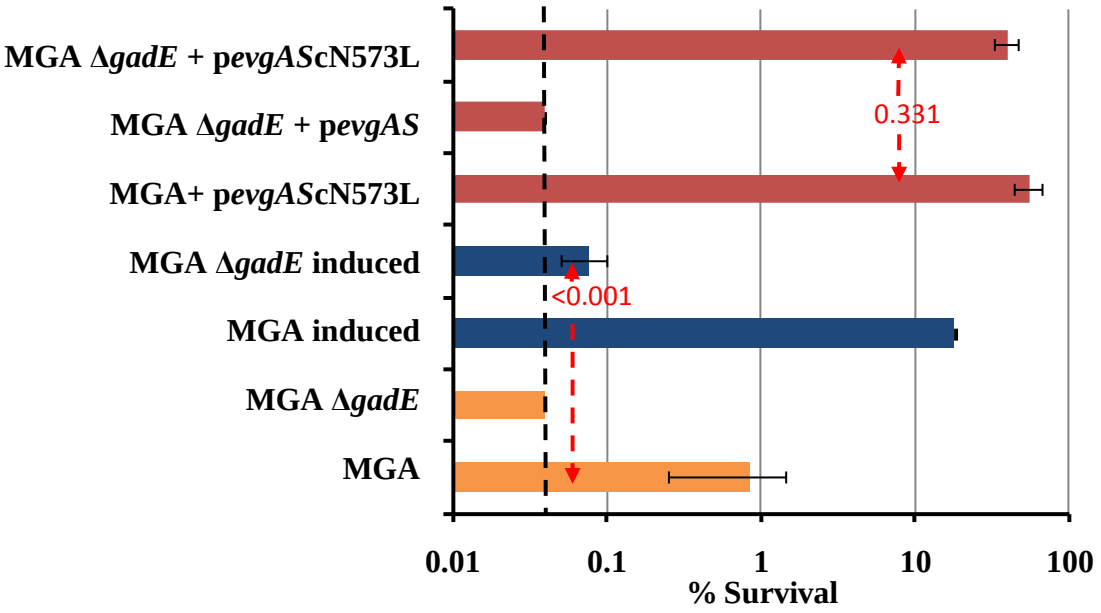


Figure 5.3.1. Survival of *E. coli* containing an EvgSc mutation in a *gadE* background after 2 hours at pH 2.5. The survival of MGA Δ*gadE* containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA Δ*gadE*, with and without induction, and MGA containing plasmid p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively. Comparisons made by T-test are indicated by red arrows, p value from the test is indicated in red.

This could mean that other acid resistance genes, independent of GadE activation, are capable of conferring high levels of acid resistance at pH 2.5. However, considering the importance of the genes regulated by GadE in the response to acid resistance, an alternative hypothesis is that these genes are being activated by independently of GadE in the EvgSc mutant background.

To test this hypothesis, the activities of a subset of AR2 promoters were assayed. The MGA, MGA +p*evgAS*, MGA +p*evgAS*cN573L and MGA +*evgAS*cN573L Δ*gadE* strains were transformed with the pLUX*acp*p, pLLUX*gadBp*, pLUX*gadEp*, pLUX*hdeAp* and pLUX*slp*p promoter probe plasmids. The promoter activity was measured using the single time point promoter probe assay, during exponential phase growth and without induction. The single time point promoter probe assays are explained in detail in Chapter 2

Figure 5.3.2 shows the promoter activity of the *acp*, *gadB*, *gadE*, *hdeA* and *slp* genes in MGA, MGA +p*evgAS*, MGA +p*evgAS*cN573L, MGA +p*evgAS*cN573L Δ*gadE* and MGA +p*evgAS*cN573L Δ*ydeO*. The results from the MGA +p*evgAS*cN573L Δ*ydeO* strain will be discussed in section 5.3.2. The activity of the *acp* promoter, which is a control promoter not involved in AR2 or acid resistance, was the same for all of the genetic backgrounds. The activities of the remaining promoters in the MGA and MGA +p*evgAS* strains were very low (purple and blue bars respectively). All other promoters were more active in the MGA +p*evgAS*cN573L strain compared to MGA +p*evgAS*. This is consistent with previous results regarding the activity of AR2 promoters in the MGA +p*evgAS*cN573L strain, see Chapter 4. The activity of the *gadB* and *hdeA* genes is abolished in the MGA +p*evgAS*cN573L Δ*gadE* strain.

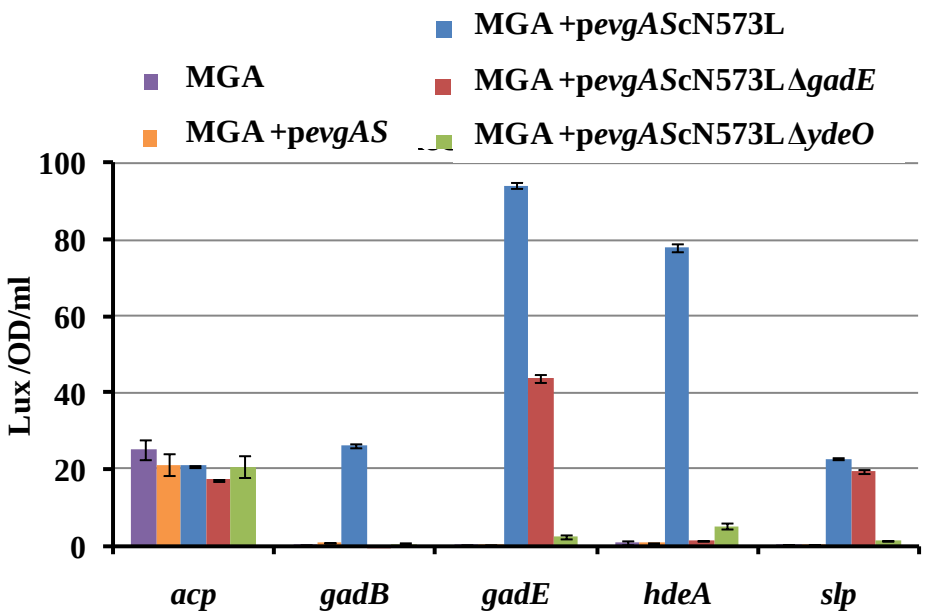


Figure 5.3.2. Affect of a *gadE* and *ydeO* deletion mutant on the activity of the *acp, gadB, gadE, hdeA* and *slp* promoters in an *evgAS*cN573L constitutive mutant. The promoter activity was measured during exponential phase growth at pH 7. The luciferase values were corrected for the OD_{600} and volume of culture. The activity of each promoter in MGA (purple), MGA +p*evgAS*cN573L (blue), MGA +p*evgAS*cN573L Δ*gadE* (red) and MGA +p*evgAS*cN573L Δ*ydeO* strains is shown. Each data point and standard deviation represents the average and standard deviation of three biological repeats respectively.

The activity of the *gadE* promoter was significantly reduced in the MGA Δ*gadE* +p*evgAS*cN573L strain. Finally, the *ydeO* activated promoter, *slp,* was not affected by the Δ*gadE* deletion. These results show that the GadE regulated genes are not being activated by another regulator in the EvgSc strain. Therefore, the resistance observed in MGA +*evgAS*cN573L Δ*gadE* must be due to GadE independent mechanisms.

In summary, the EvgSc conferred acid resistance is independent from GadE regulated mechanisms. The promoter activities in an EvgS constitutive Δ*gadE* background confirmed that the GadE regulated mechanisms were not active in a Δ*gadE* background. Other mechanisms activated by the EvgAS TCS include *ydeP* and the *ydeO* regulated acid resistance genes. This chapter will continue to assay the contributions of each of these components.

5.3.2. Role of YdeO in EvgSc acid resistance

The main contribution of YdeO to acid resistance is through the activation of GadE (Burton *et al.*, 2010). In addition to GadE activation, YdeO activates the *slp* and *dctR*. Overexpression of *ydeO* has been shown to confer acid resistance (Masuda and Church, 2003). As the results above showed that EvgSc conferred acid resistance was unaffected by a *gadE* deletion, a second possible explanation of the acid resistance phenotype is that it results from induction of these YdeO-dependent genes. To test this hypothesis, a Δ*ydeO* deletion strain was constructed using Datsenko and Wanner mutagenesis (Datsenko and Wanner, 2000). The effect of a Δ*ydeO* mutation on the survival of an MGA Δ*ydeO* +p*evgAS*cN573L strain the acid resistance phenotype was measured. The standard acid resistance assay was used. The results of the acid resistance assays are shown in figure 5.3.3.

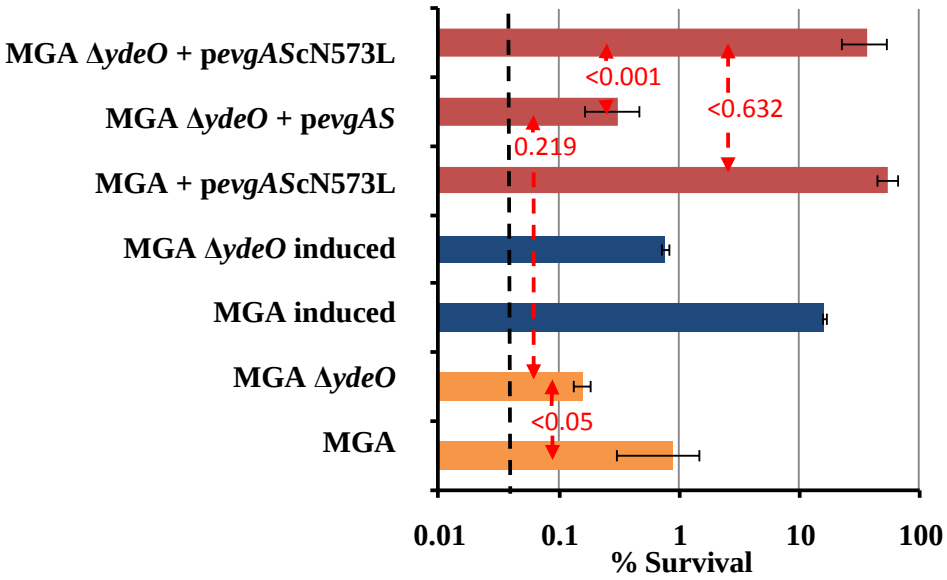


Figure 5.3.3. Survival of *E. coli* containing an EvgSc mutation in a *ydeO* background after 2 hours at pH 2.5. The survival of MGA Δ*ydeO* strain containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA Δ*ydeO*, with and without induction, and the MGA containing plasmid p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively. Comparisons made by T-test are indicated by red arrows, the p values are indicated in red.

The average survival levels of MGA with and without induction, and MGA +p*evgAS*cN573L are also presented on the graph.

The average survival of the MGA Δ*ydeO* strain with and without induction over 3 biological repeats was 0.15 and 0.76% respectively. The survival of the Δ*ydeO* deletion strain is significantly lower than that of MGA, which was 0.88% (t test, p value <0.05). These results show that the YdeO has an important role in the induction of acid resistance in MGA. The average survival of the MGA Δ*ydeO* +p*evgAS*cN573L strain over 3 biological repeats was 37.04%. This level of survival is not significantly different to that of the MGA +p*evgAS*cN573L strain (t test, p value 0.632).

To summarise, these results show that YdeO is important in the MGA strain for both uninduced and induced acid resistance. This is shown by the significant reduction in survival caused by a Δ*ydeO* deletion. However, the presence of an EvgSc mutation confers an acid resistant phenotype even in a Δ*ydeO* deletion mutant. This result confirms that EvgSc resistance is not due to the activation of YdeO regulated genes. Another explanation for this resistance is that the YdeO-dependent acid resistance genes are being activated independently from YdeO in an EvgSc mutant strain. The activity of *slp*, which is a YdeO-dependent promoter, is affected by a *gadE* mutant. To test this hypothesis, the promoter activity of the *acp*, *gadB*, *gadE*, *hdeA* and *slp* genes was tested using the single time point promoter probe assays. The promoter activity was assayed in the MGA, MGA +p*evgAS*, MGA +p*evgAS*cN573L and MGA +p*evgAS*cN573L Δ*ydeO* strains using the standard single time point promoter probe assay.

The promoter activities are shown in figure 5.3.2. The activity of the *acp* promoter, which is a control promoter that is not involved in the AR2 network, is the same in all backgrounds. The activities of the *gadB*, *gadE* and *hdeA* promoters, which are regulated by GadE, are all reduced in the MGA +p*evgAS*cN573L Δ*ydeO* deletion mutant (green bars). This suggests that the activation of *gadE* and all of the promoters it regulates is dependent on YdeO. This result indicates, that in these conditions, *ydeO* independent *gadE* regulation cannot occur via a SafA-PhoPQ-IraM-RpoS cascade or through direct activation by EvgA. However, this is consistent with the observations made by Burton *et al.* in 2010, which showed that a Δ*ydeO* knock out significantly reduced the activity of *gadE* promoter activity.

5.3.3 Effect of a Δ*ydeO* Δ*gadE* double deletion mutant on EvgSc acid resistance

The results presented in this chapter have revealed that the deletion of either *gadE* or *ydeO* has no significant effect on the survival of an EvgSc mutant strain. One hypothesis to explain the acid resistance of the Δ*gadE* +p*evgAS*cN573L and Δ*ydeO* +p*evgAS*cN573L strains is that, in a Δ*gadE* mutant, YdeO regulated genes are able to confer acid resistance, and in a Δ*ydeO* mutant, GadE regulated genes are capable of conferring acid resistance. To test this hypothesis, a Δ*ydeO* Δ*gadE* double deletion strain was made. The survival of this double deletion strain was tested with and without induction and containing plasmids p*evgAS* and p*evgAS*cN573L. Survival was tested using the standard acid resistance assay described in Chapter 2.

Figure 5.3.4 shows the results of the acid resistance assays. The survival of MGA in uninduced and induced conditions was also assayed for comparison.

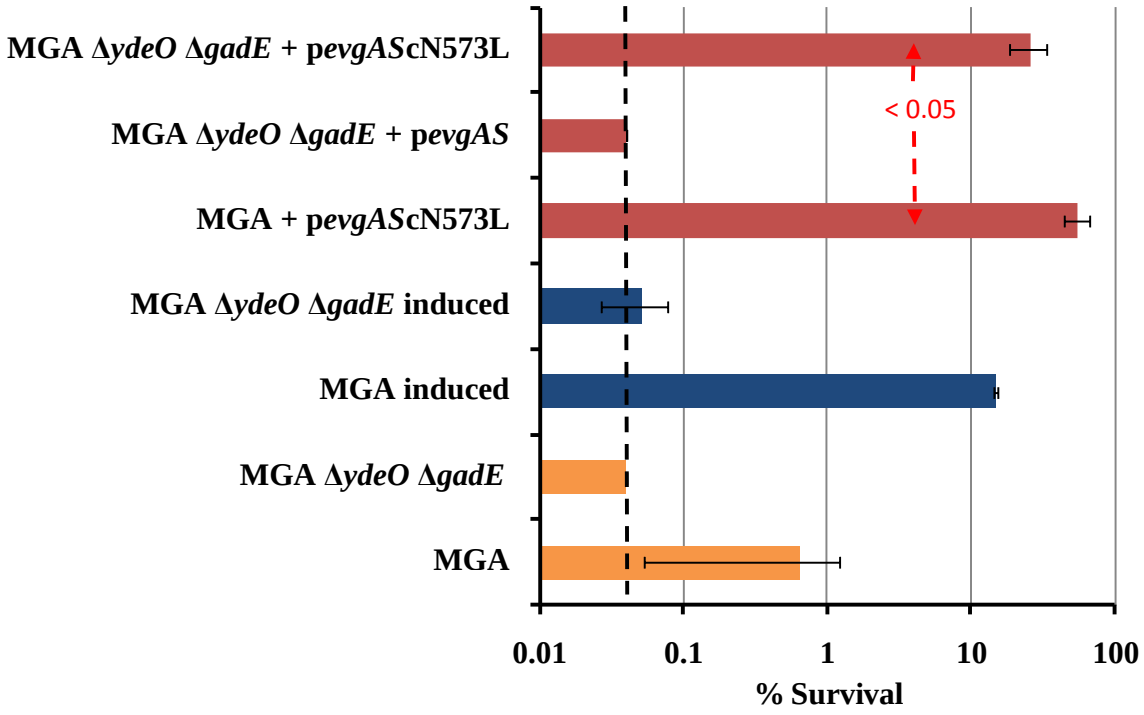


Figure 5.3.4 Survival of *E. coli* containing an EvgSc mutation in a *ydeO gadE* double knockout background after 2 hours at pH 2.5. The survival of MGA Δ*ydeO* Δ*gadE* containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA Δ*ydeO* Δ*gadE*, with and without induction, and MGA containing plasmid p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively.

The survival of MGA +p*evgAS*cN573L is also shown as a reference. The average survival of the MGA Δ*ydeO* Δ*gadE* strain in uninduced and induced conditions over 3 biological repeats was below 0.04% and 0.057% respectively. Firstly, these results show that a double deletion of these genes causes the survival of MGA to drop to below the limit of detection in uninduced conditions. Secondly, these results show that in induced conditions the survival of the MGA Δ*ydeO* Δ*gadE* strain was almost undetectable. The average survival of MGA Δ*ydeO* Δ*gadE* +p*evgAS*cN573L over 3 biological repeats was 26.11%. This is a significant reduction in survival compared to the survival of the MGA +p*evgAS*cN573L strain (t test, p value <0.05). However, this level of survival is still significantly higher than the survival of an induced MGA strain, which had an average survival over 3 biological repeats of 14.82%.

In summary, these results demonstrate that the resistance of the parent strain conferred by acid induction is completely dependent on GadE and YdeO. The Δ*gadE* Δ*ydeO* double deletion significantly reduced the survival conferred by an EvgSc mutation. However, in an EvgSc strain there are acid resistance mechanisms, which are independent of *ydeO* and *gadE* that can confer more resistance than an induced MGA strain.

5.4. Role of YdeP in EvgSc acid resistance

The results described above have shown that the survival of EvgSc strains is not completely due to acid resistance genes regulated by *ydeO* and *gadE*. The only other known acid resistance gene that is regulated by EvgAS is *ydeP* (Itou *et al.*, 2009; Masuda and Church, 2002). Annotated as a putative oxidoreducatse, the role of YdeP in acid resistance is not understood. However, studies have shown that over-expression of *ydeP* can confer acid resistance (Masuda and Church, 2002). Considering these results, we hypothesise that *ydeP* is

conferring acid resistance in an EvgSc strain. To test this hypothesis, a Δ*ydeP* deletion was made in the MGA strain. The MGA Δ*ydeP* strain was tested using the standard acid resistance assay in both induced and uninduced conditions. The survival of MGA Δ*ydeP* +*evgAS*cN573L was also assayed to test the affect on EvgSc conferred acid resistance. The strains were tested using the standard acid resistance assay see Chapter 2.

The average survival of the MGA Δ*ydeP* strain with and without induction over 3 biological repeats was 0.43 and 6.77% respectively (figure 5.4). There was no significant difference in survival between the MGA and MGA Δ*ydeP* strains in uninducing conditions (t test, p value 0.257). There was a significant reduction in survival between the MGA Δ*ydeP* and MGA strains under inducing conditions (t test, p value <0.05). This result shows that YdeP is involved in the inducible acid stress response. The average survival of the MGA Δ*ydeP* +p*evgAS*cN573L strain over 3 biological repeats was 59.47% (figure 5.4). This level of survival is not significantly different from the MGA +p*evgAS*cN573L strain, which was determined as 55.39%. Thus, there is no effect of a Δ*ydeP* mutation on the survival of a EvgSc strain. However, the affect of a YdeP mutation may be compensated by the acid resistance genes activated by the GadE and YdeO.

5.5. Dependence of EvgSc acid resistance on *ydeP, ydeO* and *gadE*

This study has demonstrated that the presence of an EvgSc mutation confers high levels of acid resistance. The results above show that a double deletion of *ydeO* and *gadE* in an EvgSc strain was shown to have reduced survival but only partly. The deletion of *ydeP* had no significant effect on survival.

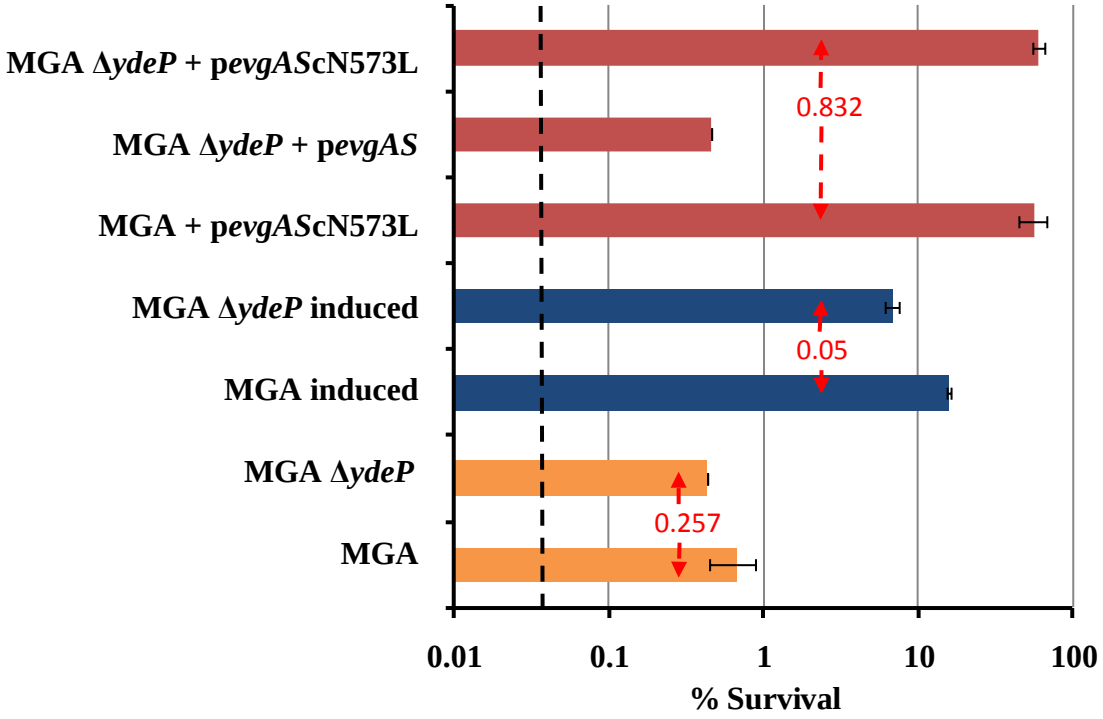


Figure 5.4. Survival of *E. coli* containing an EvgSc mutation in a *ydeP* deletion background after 2 hours at pH 2.5. The survival of a Δ*ydeP* strain containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA Δ*ydeP*, with and without induction, and the MGA containing plasmid p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively. Comparisons made by T-test are indicated using red arrows, values in red represent the p values of these comparisons.

A possible interpretation of these results is that in the *ydeO gadE* double deletion the *ydeP* gene product is conferring acid resistance and when *ydeP* is deleted the *ydeO gadE* regulated genes are conferring resistance. To test this hypothesis, a Δ*ydeP* Δ*ydeO* Δ*gadE* triple mutant strain was constructed. Based on our current knowledge of AR2 regulation, this triple mutant will be sensitive. However, if other mechanisms are involved in the acid resistance of the EvgSc strain then this triple mutant will be resistant. The survival of the triple deletion strain was tested with and without induction and containing plasmids p*evgAS* and p*evgAS*cN573L using the standard acid resistance assay (see Chapter 2).

Figure 5.5 shows the results of the acid resistance assays. The average survival of the MGA Δ*ydeP* Δ*ydeO* Δ*gadE* strain in uninduced and induced conditions over 3 biological repeats was below the limit of detection. These results show that all detectable acid resistance at pH 2.5 is dependent on these genes. The average survival of the MGA Δ*ydeP* Δ*ydeO* Δ*gadE* +p*evgAS*cN573L strain that contains a constitutive on EvgS mutation was below the limit of detection (predetermined to be 0.04%, (Burton *et al.*, 2010)). This result confirms that the acid resistance conferred by an EvgSc strain is dependent on the known EvgAS acid resistance genes and is consistent with the hypothesis that expression of either *ydeP* alone, or the *ydeO* regulated genes, is sufficient for acid resistance.

To summarise, the deletion of all known acid resistance mechanisms regulated by EvgAS causes an acid sensitive phenotype. The presence of either GadE and YdeO or YdeP is enough to confer acid resistance levels higher than those achieved by acid induction. By combining the results of these deletions we can start to understand how each of these components of the AR2 network can confer acid resistance in an EvgSc mutant.

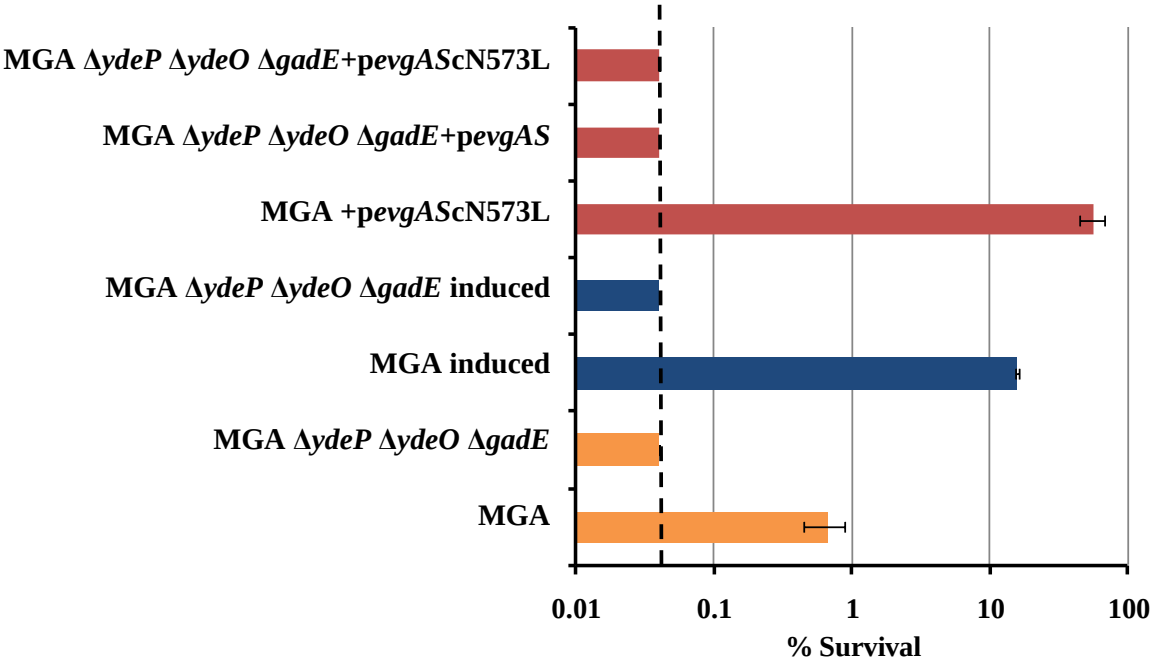


Figure 5.5. Survival of *E. coli* containing an EvgSc mutation in a *ydeP ydeO gadE* background after 2 hours at pH 2.5. The survival of MGA Δ*ydeP* Δ*ydeO* Δ*gadE* containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of the MGA and MGA Δ*ydeP* Δ*ydeO* Δ*gadE*, with and without induction, and the MGA containing plasmid p*evgAS*cN573L. Induced strains were grown for 70 minutes in pH 5.7 prior to acidification at pH 2.5. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively.

5.6. The AR2 network and acid resistance in an EvgSc strain

The AR2 network controls a subset of acid resistance genes that are crucial for acid resistance in *E. coli* K-12 (Burton *et al.*, 2010; Itou *et al.*, 2009; Masuda and Church, 2003). The AR2 network responds to mild acid shock at pH 5.7 (Burton *et al.*, 2010). This response is initiated by EvgAS and results in the activation of the AFI, YdeP and the GAD mechanism. In an EvgSc strain this essentially results in 3 separate routes to acid resistance during pH 2.5 acid shock. The diagram shown in figure 5.6 represents what can be deduced about the AR2 network in an EvgSc mutant, and summarises the conclusions from the results outlined above. The results above show that the individual routes can confer high levels of acid resistance in isolation. Surprisingly, the activation of GadE by the recently characterised strain SafA-PhoPQ-IraM-RssB-RpoS was not essential to EvgSc acid resistance. This was revealed by the high survival seen by a Δ*phoP* Δ*rpoS* double deletion mutant.

5.7 Role of GadE in EvgSc resistance at pH 2.1

We have used chromosomal mutations to dissect the contributions of all AR2 components in the acid resistance phenotype of EvgSc strains. These assays tested the strains at pH 2.5. This pH was used as it is the same pH that was used during the evolution experiment (described in chapter 3). In addition, pH 2.5 is the level of acidity that is most commonly used in the literature for testing acid resistance (Castanie-Cornet *et al.*, 1999). Interestingly, the results presented in Chapter 3 show that strains containing the EvgSc mutations are capable of conferring acid resistance at pH 2. The experiments described above show that the EvgSc mutations can confer GadE independent acid resistance. This result is surprising as the GAD system is the most robust system for survival at pH 2- 2.5.

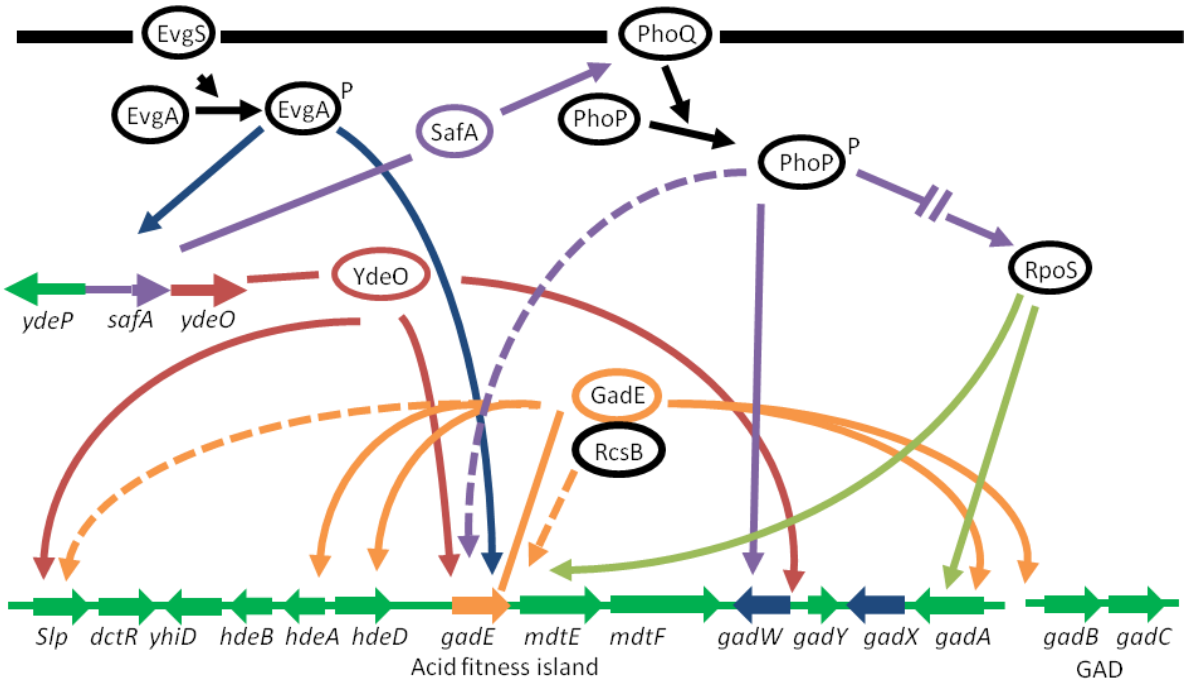


Figure 5.6. Activation of the AFI and GAD genes by the AR2 network. The main constituents of the AR2 network that control acid resistance are shown in this figure. The regulator proteins are represented by open circles. Genes are shown by wide arrows, structural acid resistance genes are shown in dark green. Activation of these genes by regulators is shown by thin arrows. A thick black line represents the inner membrane of the cell and space below that line represents the cytoplasm. A broken line of activation between PhoP and RpoS represents an indirect interaction. The IraM-RssB components of this activation were left out of the figure for simplification.

Based on the results at pH 2.5 in an EvgSc Δ*gadE* mutant strain, we hypothesise that the GadE independent acid resistance systems that confer acid resistance at pH 2.5 can also do so at pH 2.1. To test this hypothesis, the strains with deletions in *rpoS-phoP*, *gadE*, *ydeO* and *ydeP* were transformed with p*evgASc*N573L. The standard acid resistance assay was used to test each strain. The only adjustment was that each strain was shocked in M9supp at pH 2.1.

The results of the acid resistance assays are shown in figure 5.7. The average survival of MGA and MGA +p*evgAS* over 3 biological repeats was 0.57 and 0.49% respectively. There was no statistical difference between the MGA and the MGA +p*evgAS* strains (t test, p value <0.005). The average survival of the MGA +p*evgAS*cN573L over 3 biological repeats was 82.82%. The average survival of the MGA Δ*phoP* Δ*rpoS* +p*evgAS*cN573L and MGA Δ*ydeP* +p*evgAS*cN573L strains over 3 biological repeats was 56.12% and 75.43% respectively. These results show that there was no significant difference in survival between the Δ*phoP* Δ*rpoS* and Δ*ydeP* deletion strains compared to the MGA +p*evgAS*cN573L strain (t test, pvalues 0.342 and 0.699 respectively). The average survival of the MGA Δ*ydeO* +p*evgAS*cN573L strain over 3 biological repeats was 11.24%. This was a significant drop in resistance compared to MGA +p*evgAS*cN573L (t test, p value <0.005). Finally, the average survival of the MGA Δ*gadE* +p*evgAS*cN573L strain over 3 biological repeats was below the limit of detection. These results show that acid resistance at pH 2.1 requires the GadE regulated acid resistance mechanisms. Similar levels of acid resistance cannot be conferred by the other YdeO regulated genes or by YdeP. In fact, removal of YdeP has no effect on survival. In addition, these results support previous observations that the main activator of GadE expression is YdeO and not PhoP and RpoS as the survival of a Δ*ydeO* mutant was lower than that of a Δ*phoP* Δ*rpoS* double mutant at pH 2.1.

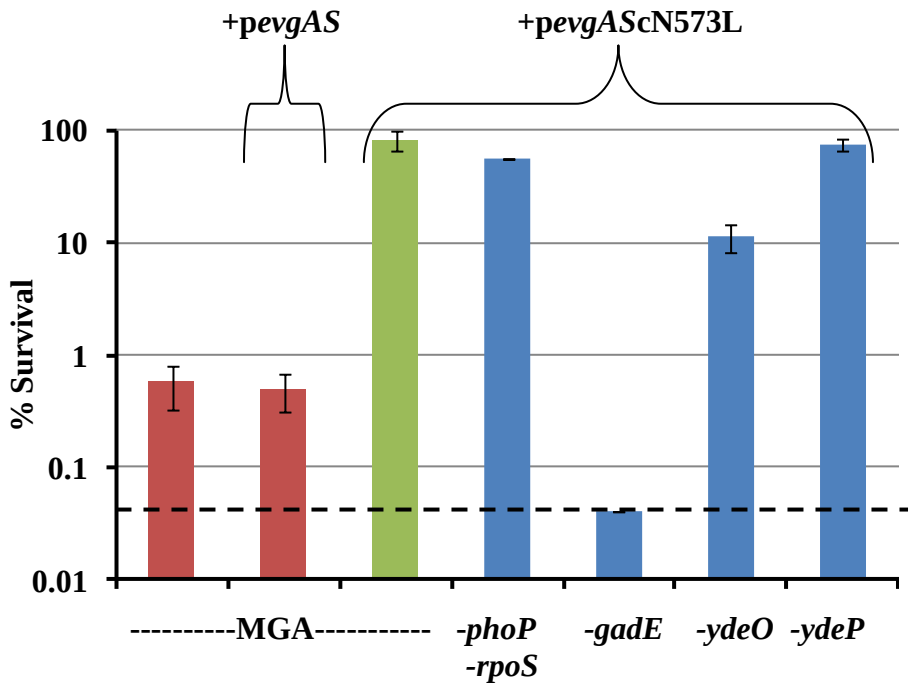


Figure 5.7. Survival of MGA and the MGA +p*evgAS*cN573L in various mutant backgrounds at pH 2.1. The survival of MGA, Δ*rpoS* Δ*phoP*, Δ*gadE*, Δ*ydeO* and Δ*ydeP* strains containing the plasmid p*evgAS*cN573L that contains a constitutive copy of *evgS*, was compared to the survival of MGA and MGA +p*evgAS*. Strains were tested by acidification at pH 2.1 for 2 hours. The plasmid p*evgAS*, which contains a wild type copy of the *evgAS* operon was included as a control. The dotted line represents the limit of detection for this assay, which was predetermined to be 0.04%. All of the data points and error bars represent the average and standard deviation of at least 3 independent biological repeats respectively.

This is shown by the significantly reduced survival of the Δ*ydeO* strain, which is presumably due to a reduction in GadE induction. No such reduction was observed with the Δ*rpoS* Δ*phoP* double deletion strain.

5.8. Discussion

5.8.1 EvgSc acid resistance and the AR2 network

The presence of an EvgSc mutation has been shown by this study to have two consequences. The first is the acid resistant phenotype. The second is the activation of the AR2 network and the acid resistance genes that the AR2 network regulates. In this chapter we have analysed the link between the activation of AR2 components and the acid resistance phenotype of the EvgSc strains. The deletion of *evgA*, the response regulator of the EvgAS TCS, conferred an acid sensitive phenotype in acid induced and EvgSc strains. These results revealed that all inducible acid resistance is EvgA dependent, which is consistent with results from other groups (Burton *et al.*, 2010; Itou *et al.*, 2009; Masuda and Church, 2003). These results also showed that all EvgSc acid resistance was EvgA dependent. As a consequence, the role of all other acid resistance mechanisms of *E. coli* in the survival of an EvgSc mutant can be ruled out. Finally, the reduced survival of an EvgSc *evgA* mutant confirms that the action of the EvgSc mutation is dependent on EvgA. The *evgA* phenotype was also successfully complemented using p*evgA*.

5.8.2 The individual components of AR2 and the EvgSc phenotype

Based on the current understanding of the AR2 network, there are many routes to acid resistance from EvgAS activation. However, a study by Eguchi *et al.* in 2011 highlighted a new activation pathway that was crucial for the survival of an EvgSc mutant. The EvgSc

mutant, EvgS1, was shown to activate *gadE* via a SafA-PhoPQ-IraM-RssB-RpoS activation cascade. The PhoP and RpoS proteins are pivotal to this activation and to the EvgS1 acid resistance phenotype. However, this study has shown that the RpoS PhoP pathway is not essential, and that with both *rpoS* and *phoP* knocked out, an EvgSc strain can survive extreme acid shock at pH 2.1 using our acid resistance assay. In fact, subsequent analysis of the GadE, YdeO and YdeP components of the AR2 network showed that none of these components are essential to acid resistance at pH 2.5. However, it should be noted that acid resistance varies greatly between assay conditions. The studies by Eguchi *et al* use a different acid resistance assay to our own. Key differences include the medium, Eguchi *et al.* used LB; the phase of growth, acid resistance assays were started when cultures reached OD_{600} 0.6 and the method of acidification, Eguchi *et al.* added a predetermined volume of HCl to the culture whereas our technique diluted cells into large volumes of pH 2.5 medium. A comparison of strains used by Eguchi *et al.* with the strains used in the study was made using the acid resistance assay described in Chapter 2. Published survival levels for EvgS1 were much lower than those measured by the acid resistance assay used in this assay. In addition, when p*evgAS*cN573L was introduced into the Δ*rpoS* mutant strain used by Eguchi *et al.* the level of survival was unaffected. These results suggest that the main difference between the results observed by this study is due to the assay method (see Appendix 1).

Surprisingly, YdeP alone was capable of conferring acid resistance as a *gadE ydeO* double mutant was still resistant. However, this resistance was only present in EvgSc strains. Induction of *ydeP* by mild acid shock didn't confer resistance. This was shown by the acid sensitive phenotype of an MGA *ydeO gadE* double mutant strain after induction. When the acid resistance phenotypes are compared in various mutant backgrounds, it has become clear

that the main effect of the EvgSc mutations was to introduce redundancy into *E. coli* acid resistance that is not present in an acid induced strain. Survival of the MGA after induction is completely dependent on the *ydeO gadE* double mutant. Presumably the only difference between the EvgSc strains and an induced MGA strain is the level of expression of the AR2 regulated genes. Specifically, the additional protection conferred by YdeP in the EvgSc strains must due to over-expression. Acid resistance due to YdeP has been shown by other groups (Masuda and Church, 2002). Considering the fact that increased expression can confer high levels of acid resistance, it is surprising that *E. coli* fails to achieve such high levels when induced. One possible explanation is that expression of YdeP to such high levels could confer a considerable fitness cost. Alternatively, the increased expression or translation of YdeP could also be dependent on other environmental factors that are not present in the inducing medium. It is noted that in order to confirm these phenotypes the following complementations must be done. The *gadE*, *ydeO* and *ydeP* loci must be put back into the Δ*gadE* Δ*ydeO* Δ*ydeP* triple mutant strain individually, to confirm that alone, at wild type expression levels, they can confer high levels of acid resistance.

In summary, this chapter has linked the acid resistance phenotype with the activation of the AR2 network. Moreover, the roles of individual AR2 components in EvgSc and induced acid resistance have been characterised. The perturbation of the AR2 network by the EvgSc mutations has conferred a redundancy in the AR2 network. This redundancy is caused by the enhanced protection by YdeP caused by over-expression and is not present in the MGA strain under acid induced conditions.

5.8.3 Summary

1) The acid resistance of EvgSc mutant strains is completely dependent on the response regulator EvgA. The EvgA deletion phenotype can be complemented by the addition of a plasmid encoded EvgA.
2) The activation of the AR2 network is the cause of the acid resistance phenotype conferred by the EvgSc mutations.
3) Activation of *gadE* by the SafA-PhoPQ-IraM-RssB-RpoS cascade is not essential to the EvgSc phenotype. This result contradicts the findings of Eguchi *et al.,* 2010. The main activator of *gadE* was shown to be YdeO.
4) A *gadE ydeO* double deletion caused all inducible resistance to be abolished. However, despite a significant reduction in survival, the MGA *ydeO gadE* +p*evgA*ScN573L strain was still more acid resistant than the induced MGA strain.
5) YdeP was shown to confer acid resistance in the absence of *ydeO* and *gadE*. This resistance was only present in an EvgSc strain. As a result the EvgSc mutations have conferred redundancy in the AR2 network that is not present in the MGA strain.
6) Survival of the MGA +p*evgA*ScN573L strain in pH 2.1 medium was completely dependent on GadE.

CHAPTER 6: RESULTS (4): Interactions between two-component systems in the AR2 network

6.1. Additional interactions of the PhoPQ system in the AR2 network

6.1.1. High-temporal resolution promoter probe assays and the AR2 network

The role of the EvgSc mutations in acid resistance and the AR2 network has been described in previous chapters. The analysis of these mutations has built on work by Neil Burton. The study by Burton *et al.* in 2010 accurately assayed promoter induction in response to mild acid shock (pH 5.7). Novel dynamics of promoter induction were unveiled using a luciferase reporter system to measure promoter activity at high-temporal resolution. The study by Burton *et al*., among others, has shown that investigating the timing of interactions can reveal subtle differences required for adaption to stress conditions (Burton *et al.*, 2010; Kalir *et al.*, 2001; Shin *et al.*, 2006; Temme *et al.*, 2008; Zaslaver *et al.*, 2004). The work described in section 6.1 includes the contributions made to Burton *et al.* 2010.

6.1.2. The affect of Δ*phoP* on *safA-ydeO* promoter activity after induction by mild acid shock

Previous research by Neil Burton used the luciferase promoter probe library to screen most of the AR2 promoters in a variety of backgrounds. This approach aimed to find novel interactions within this network by monitoring promoter activity, in different mutant backgrounds, after induction by mild acid shock. Various promoters of genes involved in the AR2 network were cloned into the luciferase reporter plasmid pLUX (Burton *et al.*, 2010). These promoter probe plasmids were then transformed into different mutant backgrounds. The promoter activity was measured with and without induction using the high-temporal resolution promoter probe assays. For details of the mutants and pLUX plasmids constructed see Burton *et al.*, 2010. The high-temporal resolution promoter probe assays are described in detail in Chapter 2. Briefly, cultures were grown into exponential phase before each was

separated into 6 individual wells of a microtitre plate. Measurements of luciferase activity were made every 70 seconds. After 24 measurements at pH 7, three of the wells were acidified to pH 5.7, thus inducing the cultures. Measurements were then taken for a further 84 time points. Many novel aspects of the AR2 network were revealed during this analysis. In particular, the Δ*phoP* mutant, which is the response regulator of the PhoPQ TCS, was shown to negatively affect promoter activity of EvgA regulated promoters (Burton *et al.*, 2010; Groisman *et al.*, 1992). Of these promoters, the *safA-ydeO* is most important to this section. Figure 6.1 shows the promoter activity of the *safA-ydeO* promoter in both MGA and Δ*phoP* backgrounds, with and without induction. The induction dynamics indicated that the *safA-ydeO* promoter was activated by induction. The use of high-temporal resolution assays has revealed the complexity of induction at the *safA-ydeO* promoter. The profile of activation over time shows two separate increases in activity (1^{st} at t = 10 to 30 minutes, 2^{nd} at t = 45 to 120 minutes) broken by a reduction in activity (t = 30 to 45 minutes). This reduction in activity in the MGA strain was hypothesised to be due to a repressor binding as part of a negative feedback loop. The activation of the *safA-ydeO* promoter was increased in a Δ*phoP* background, and the repression, which was observed in the MGA background, was not present. These results show that the PhoPQ TCS is making additional interactions with the pathway other than those previously described (Eguchi *et al.*, 2007; Eguchi *et al.*, 2004; Zwir *et al.*, 2005). This interpretation of the data led to further investigation into the interaction between the PhoPQ TCS and the *safA-ydeO* promoter.

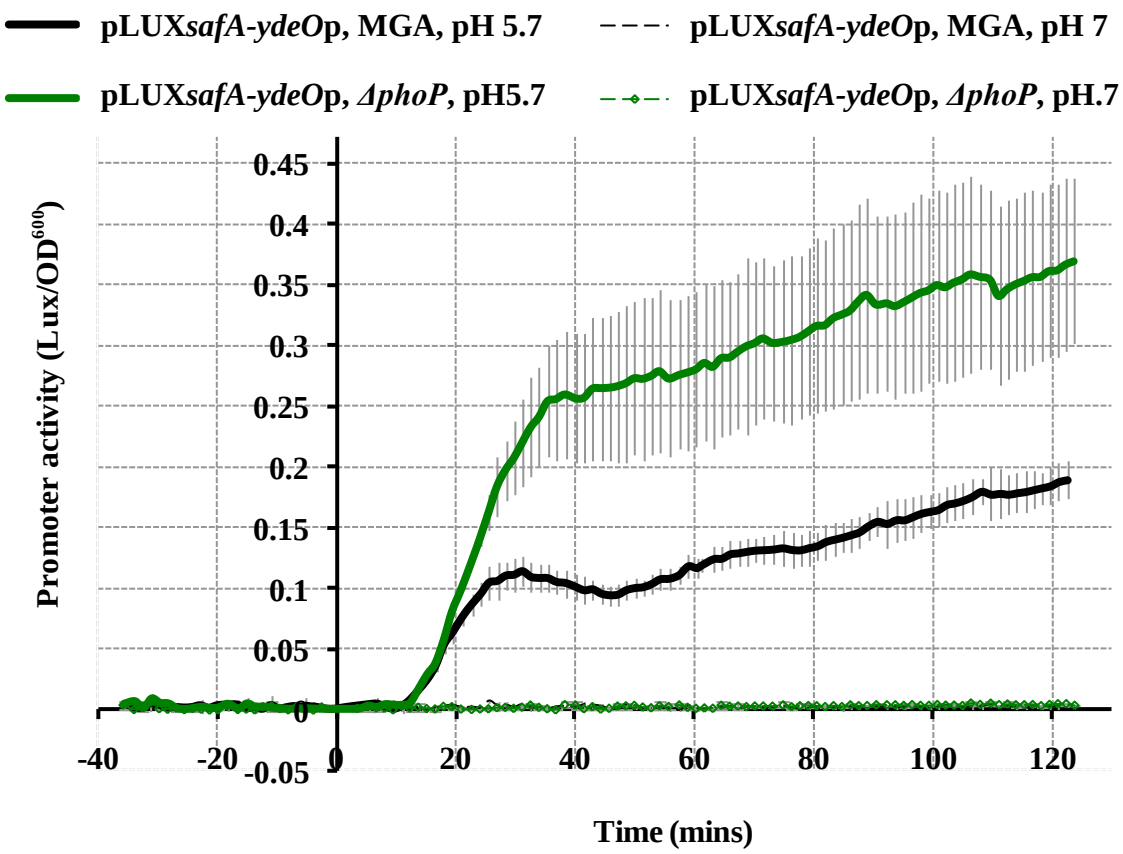


Figure 6.1. Induction of the *safA-ydeO* promoter by mild acid shock in MGA and *phoP* deletion backgrounds. Black lines show the activity of the *safA-ydeO* promoter in an MGA background with (solid line) and without (dashed line) induction. Green lines show the activity of the *safA-ydeO* promoter in a Δ*phoP* background with (solid line) and without (dashed line) induction. Time zero indicates the point of acidification to pH 5.7. Values and error bars represent the average and standard deviation of three biological repeats respectively (adapted from Burton *et al.*, 2010).

6.1.3. The affect of *ΔsafA* on *safA-ydeO* and *mgtA* promoter activities after induction by mild acid shock

The results described in 6.1.2 show a difference in induction of the *safA-ydeO* promoter in a Δ*phoP* mutant compared to the MGA strain. One explanation of the observed difference is that activated PhoP is repressing the *safA-ydeO* promoter. SafA has already been shown to activate the PhoPQ TCS (Eguchi *et al.*, 2007). Considering this we hypothesise that activated PhoP is repressing the *safA-ydeO* promoter, forming a simple negative feedback loop. Alternatively, unphosphorylated PhoP could be repressing the promoter. To test this hypothesis, a Δ*safA* deletion mutant was constructed using Datsenko and Wanner mutagenesis (Datsenko and Wanner, 2000). If phosphorylated PhoP caused the repression, the Δ*safA* deletion will abolish the cross talk between the EvgAS and PhoPQ TCSs and therefore remove the repression at the *safA-ydeO* promoter. The Δ*safA* mutant was transformed with the pLUX*safA-ydeO* and pLUX*mgtA* plasmids. The *safA-ydeO* promoter activity in Δ*safA* background will test the aforementioned hypothesis. The *mgtA* gene is directly activated by phosphorylated PhoP (Zwir *et al.*, 2005). The activity of the *mgtA* promoter was used as a control to indicate whether PhoPQ were still active in a Δ*safA* mutant. Promoter activity of the *safA-ydeO* and *mgtA* promoters was measured using the high-temporal resolution promoter probe assay with the LUX plasmids in both MGA and Δ*safA* backgrounds.

Figure 6.2 shows the affect of a Δ*safA* deletion on *safA-ydeO* (red line) and *mgtA* (blue dotted line). The *safA-ydeO* and *mgtA* promoter activities were also assayed in MGA backgrounds, shown by black solid and green dotted lines respectively. The *safA* mutation caused increased induction of the *safA-ydeO* promoter.

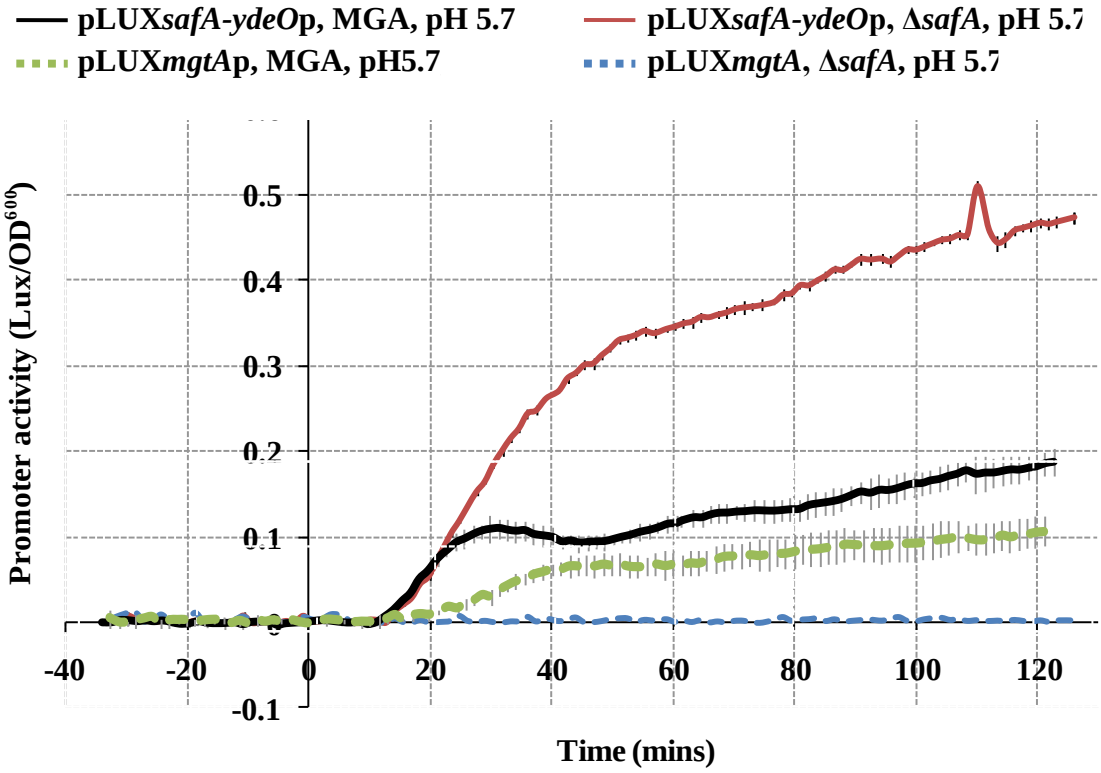


Figure 6.2. Activity of the *safA-ydeO* and *mgtA* promoters in MGA and *safA* deletion backgrounds. Solid lines represent *safA-ydeO* promoter activity in MGA (black line) and *safA* mutant (red line) backgrounds. Dotted lines represent *mgtA* promoter activity in MGA (green) and *safA* mutant (blue) backgrounds. Time zero indicates the point of acidification to pH 5.7. Values and error bars represent the average and standard deviation of three biological repeats respectively.

This result suggests that when PhoP is not activated by SafA there is no repression of the *safA-ydeO* promoter. The activity of the *mgtA* promoter was completely abolished in a Δ*safA* mutant. This result shows that the deletion of *safA* removes all crosstalk between the EvgAS and PhoPQ TCSs, which is consistent with previously published work (Eguchi *et al.*, 2007; Eguchi *et al.*, 2004).

To summarise the results above, when activated, the PhoPQ system is involved in the repression of the *safA-ydeO* promoter. These results reveal a novel negative feedback loop between in the PhoPQ TCS and the *safA-ydeO* promoter. However, it is not know whether the interaction between PhoP and the *safA-ydeO* promoters is direct or indirect. Recent experiments by Ashley Robinson, which support a model for direct binding, are described in the discussion below.

6.2. Novel interactions of RcsB in the AR2 network

6.2.1. The role of RcsB in the *E. coli* acid resistance phenotype

Although originally characterised for its role in the regulation of capsule synthesis, the *rcs* phosphorelay has been implicated in the regulation of acid resistance genes in *E. coli*. RcsB was first implicated in acid resistance by Castanie-Cornet *et al.* in 2007, they showed that stationary phase acid resistance is dependent on RcsB. By assaying the acid resistance of *rcsA, rcsB, rcsC* and *rcsD* knockout strains, Castanie-Cornet *et al.* showed that the other components of the *rcs* phosphorelay system were not required. Finally, the over-expression of RcsB was shown to reduce the level of survival in extreme acid conditions (Castanie-Cornet *et al.*, 2007). In summary, this work indicated that basal levels of unphosphorylated RcsB were required for acid resistance during stationary phase. Recently, the role of RcsB in the

regulation of acid resistance genes has been thoroughly characterised. The binding of GadE to all promoters of genes involved in acid resistance has been shown to be completely dependent on RcsB (Castanie-Cornet *et al.*, 2010; Krin *et al.*, 2010a). Anaylsis of promoter binding *in vitro* indicated that RcsB and GadE form a heterodimer (Castanie-Cornet *et al.*, 2010). Throughout this chapter this will be termed the RcsB-GadE heterodimer model. The heterodimer model suggests that the role of RcsB would be limited to this interaction, and only GadE dependent promoters would be affected by an *rcsB* deletion. However, the authors did not test the promoter activity of GadE independent promoters.

6.2.1.1. Role of RcsB in exponential phase acid resistance

All previously published phenotypic analysis of *rcsB* mutants have been done during stationary phase (Castanie-Cornet *et al.*, 2010; Castanie-Cornet *et al.*, 2007; Krin *et al.*, 2010a). As previously described, the activation of acid resistance by entry in to stationary phase requires different regulatory interaction. However, the activation of acid resistance in both stationary phase and expeonential phase induction conditions results in the activation of GadE. Considering this, the interaction between GadE and RcsB is predicted to be independent from activation by either condition. We hypothesise that the role of RcsB is essential to both stationary phase resistance cultures and exponential phase cultures with and without induction. We wanted to assay the contribution of *rcsB* to *E. coli* acid resistance during exponential phase and under inducing conditions. To test the contribution of RcsB to acid resistance in an Δ*rcsB* deletion mutant was created. The survival of the Δ*rcsB* mutant was assayed using the standard acid resistance assay, during exponential phase growth, with and without induction.

The results of the acid resistance assays are shown in figure 6.3 A. This section is only be concerned with the MGA and Δ*rcsB* strains, the complementation data will be discussed in the next section. The average survival of the MGA strain, with and without induction, over 3 biological repeats was 0.8% and 24.% respectively. These levels of survival are consistent with those previously reported by this study. The average survival of the Δ*rcsB* strain over three biological repeats, with and without induction, was below the level of detection (previously determined as 0.04%, Burton *et al.*, 2010). This results shows that all exponential phase resistance, in both induced and uninduced conditions, is completely dependent on *rcsB.*

To summarise, these results show that acid resistance during exponential phase, without induction, is completely dependent on *rcsB.* This level of uninduced survival is the same as that observed in a *gadE* mutant (see Chapter 5). However, the level of survival of the Δ*gadE* mutant is higher than that of an Δ*rcsB* mutant under inducing conditions (*gadE* survival is shown in figure 5.3.1.). Considering the current understanding of the AR2 network, the Δ*rcsB* mutant should show the same survival as a Δ*gadE* mutant. The reduced survival shown by an Δ*rcsB* strain could be due to additional interactions made by RcsB independently from GadE.

6.2.1.2 Dependence of EvgSc acid resistance on RcsB

The results described above show that when induced, the Δ*rcsB* mutant has a more severe affect on survival compared to the Δ*gadE* mutant. This suggests an additional role of RcsB in the regulation of acid resistance. In Chapter 5 we characterised the acid resistance of four EvgSc mutants (constitutively active EvgS histidine kinases). There was no affect of a *gadE* deletion mutant on the survival of an EvgSc mutant strain.

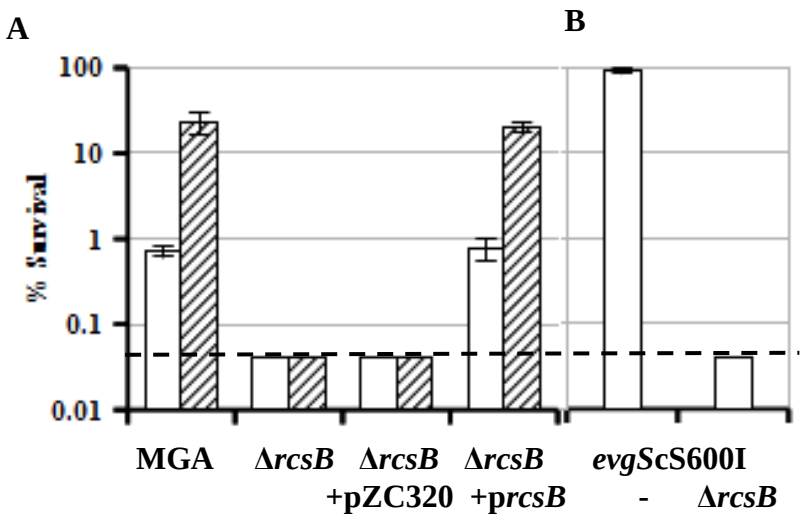


Figure 6.3. Acid inducible resistance is dependent on *rcsB*. Survival represents the scored cells after 2 hours at pH 2.5 in M9supp expressed as a % of total cells at time zero. **A** Shows MGA and Δ*rcsB* mutant strains with and without complementation plasmid p*rcsB* (pZC320 is a control), after induction by acid shift (striped bars), or uninduced (white bars). **B** Uninduced *evgSc*S600I strain, which contains a constitutive EvgS (EvgSc) compared to uninduced Δ*rcsB* mutant of the same strain. Dotted line represents the limit of detectable survival determined as 0.04 %. Error bars represent the standard deviation of 3 repeats.

Considering the current RcsB-GadE heterodimer model, the survival of an EvgSc mutant should be unaffected by an Δ*rcsB* mutant. To test this hypothesis, an *evgS*cS600I Δ*rcsB* strain was created using Datsenko and Wanner mutagenesis (Datsenko and Wanner, 2000). Details of the mutagenesis can be found in Chapter 2. The *evgS*cS600I strain contains an *evgS* mutant that encodes a constitutive on EvgS (EvgSc). The acid resistance phenotype of this strain was tested using a standard acid resistance assay.

The survival of the *evgS*cS600I Δ*rcsB* strain is shown in figure 6.3 B. The average survival over 3 biological repeats was below the limit of detection. This results shows that RcsB is essential to all EvgSc conferred acid resistance. Also, this result indicates that the RcsB-GadE heterodimer model is, at the very least, incomplete as the impact of a *gadE* mutation is not as severe.

6.2.1.3. Complementation of the RcsB phenotype using a plasmid encoded copy of *rcsB*

The experiments described above show that the deletion of *rcsB* confers an acid sensitive phenotype in an acid induced strain and in an EvgSc strain. This is a surprising result as the role of RcsB in the regulation of acid resistance is thought to be limited to its interactions with GadE. However, it is also possible that other genes may have been affected when *rcsB* was mutated, and a secondary site mutation could be the cause of the observed sensitivity. The introduction of a resistance cassette into the chromosome requires a homologous recombination event. To increase the frequency of a homologous recombination lambda red genes are expressed in the strain. The frequency of other recombination events is also increased and this can lead to secondary mutations (Hobman *et al.*, 2007). To confirm that the acid sensitive phenotype is due solely due to the loss of *rcsB*, and not any other genetic

mutations, the Δ*rcsB* mutant was complemented by the addition of a plasmid encoded *rcsB*. The construction of p*rcsB* complementation plasmid is described in detail in Chapter 2. Briefly, the plasmid contains the *rcsB* gene under the control of its native promoter (-1144 to +674, relative to the RcsB translation start site). It is important that the *rcsB* gene is expressed at wild type levels, as the over-expression of *rcsB* has already been shown to confer an acid sensitive phenotype. The survival of the MGA +p*rcsB* and the Δ*rcsB* +p*rcsB* strains was assayed with and without induction. Strains were tested using the standard acid resistance assay, which is described in detail in Chapter 2.

Figure 6.3 A, shows the results of these assays. The survival of the Δ*rcsB* +p*rcsB* with and without induction was 21.5 and 0.9 % respectively. The survival of these strains was equivalent to the MGA strain with and without induction. These results show that the introduction of the plasmid encoded *rcsB* fully complemented the Δ*rcsB* phenotype. The survival of the *evgS*cS600I Δ*rcsB* +p*rcsB* strain was 97.8%, which is actually higher than the *evgS*cS600I strain (31.8%, see 4.3).

The results described above confirm that the deletion of the *rcsB* gene was the cause of the acid sensitive phenotype. These results verify that Δ*rcsB* phenotype is more sensitive to acid stress compared to the Δ*gadE* phenotype. If the current model is correct, a Δ*gadE* mutant phenotype should be the same as the Δ*rcsB* mutant phenotype. The affect of an Δ*rcsB* mutant on both the survival of MGA with induction, and the survival of an EvgSc strain was more severe than that of a Δ*gadE* mutant. The latter result showed that EvgSc strains could survive extreme acid without *gadE*. In fact, a triple mutant of *ydeO ydeP* and *gadE* needed to be made in an EvgSc strain to confer sensitivity equivalent to that of an EvgSc Δ*rcsB* mutant. In

addition, these results also validate the p*rcsB* plasmid as a suitable complementation plasmid. In summary, RcsB must be making additional interactions with acid resistance genes other than those previously identified.

6.2.2. Analysis of Δ*rcsB* and Δ*gadE* mutant strains using high-temporal resolution promoter probe assays

Mutagenesis of the *rcsB* gene has revealed that the contribution of RcsB to the acid resistance phenotype is more than that of GadE. This is a surprising result as the current role of RcsB in the regulation of acid resistance genes is the formation of a heterodimer with GadE (Castanie-Cornet *et al.*, 2010; Krin *et al.*, 2010a). In previous chapters the promoter probe plasmids have been used to assay the promoter activity of genes involved in the AR2 network. Here we aim to uncover additional interactions of RcsB in the AR2 network by comparing the induction dynamics of an Δ*rcsB* mutant and the MGA strain. The induction dynamics of each of these promoters was compared to the promoter activity in a Δ*gadE* strain (Burton *et al.*, 2010). To test the promoter activities of AR2 genes in each genetic background, the promoter probe plasmids were each transformed into each strain. The following plasmids were used: pLUX*evgA*p, pLUX*safA-ydeO*p, pLUX*ydeP*p, pLUX*gadE*p, pLUX*slp*p, pLUX*mgtA*p, pLUX*gadB*p, pLUX*hdeA*p, pLUX*gadA*p, pLUX*gadX*p, pLUX*gadY*p and pLUX*gadW*p. The promoter activity was measured using the standard high-temporal resolution promoter probe assay, which is described in detail in Chapter 2. The cultures were induced by growth in pH 5.7 M9supp medium. To compare the induction of each promoter in MGA and Δ*rcsB* backgrounds, and to measure the significance of the difference, a single time point was chosen and compared. The induction of each promoter was compared 40 minutes after induction as this was the highest point of initial induction. These comparisons are shown in

Table 6.1. This section is only be concerned with the exponential phase comparisons, stationary phase comparisons will be discussed later in this chapter.

The induction of each promoter, in each background, by mild acid shock, is shown in figure 6.4 A. The promoter activity in a Δ*gadE* strain is also shown on each graph for comparison. The Δ*gadE* data was generated and published by Neil Burton in Burton *et al* in 2010. All plotted lines on each graph are induced cultures. The *evgA* promoter in an Δ*rcsB* background shows no significant difference in promoter activity compared to either the Δ*gadE* or the MGA strain. This result demonstrates that any effect of a Δ*rcsB* mutation is not due to an affect the expression level of EvgAS. It is therefore likely to be due to an effect on the ability of the EvgAS TCS to signal to the AR2 network.

The promoters in the AR2 network can be divided in to three groups: GadE independent promoters, GadE partially dependent promoters, and GadE dependent promoters. The GadE dependent promoters include: *gadB*p, *hdeA*p, and *gadA*p; the activities of each of these promoters were abolished in a Δ*rcsB* background. This result is consistent with the GadE-RcsB heterodimer model. The GadE independent promoters include: *safA-ydeO*p, *ydeP*p, *mgtA*p and *gadX*p; the activities of these promoters were not affected by a Δ*gadE* deletion. However, the activities of these promoters were completely removed by an Δ*rcsB* mutation. The *safA-ydeO* and *ydeP* promoters are directly regulated by EvgA. This result shows that the activation of all EvgA dependent AR2 promoters is completely dependent on *rcsB*. The *mgtA* promoter is activated indirectly by SafA, the loss of activation at this promoter, in an Δ*rcsB* mutant, is consistent with the loss of activity at the *safA-ydeO* promoter. Finally, the partially GadE dependent promoters include: *gadE*p, *slp*p, *gadW*p and *gadY*p.

Table 6.1 Relative activity of AR2 promoters

	Promoter dependency[a]		Relative activity	
Promoter	GadE	EvgA	Exp[b]	Stat[c]
gadA	d	d	0.050*	0.005*
gadB	d	d	0.014**	-
hdeA	d	d	0.002***	0.027**
hdeD	d	d	0.100*	0.160*
gadE	pd	d	0.013***	0.086**
gadW	i	d	0.208*	0.534ns
gadY	i	d	0.034**	0.950ns
slp	i	d	0.039***	-
safA	i	d	0.030**	nd
mgtA	i	d	0.166***	0.915ns
ydeP	i	d	0.002**	nd
evgA	i	i	0.900ns	-
gadX	i	i	0.283*	0.256**

[a] d, dependent; i, independent; pd, partially dependent; as determined in (1)

[b] Promoter activity in Δ*rcsB* strain relative to wild type after 40 minutes of induction at pH 5.7

[c] Promoter activity in Δ*rcsB* strain relative to wild type in stationary phase

nd: Not determined, as expression levels too low to measure accurately

-: not done

*, ** and *** indicate P values > 0.95, > 0.99 and > 0.999 respectively.

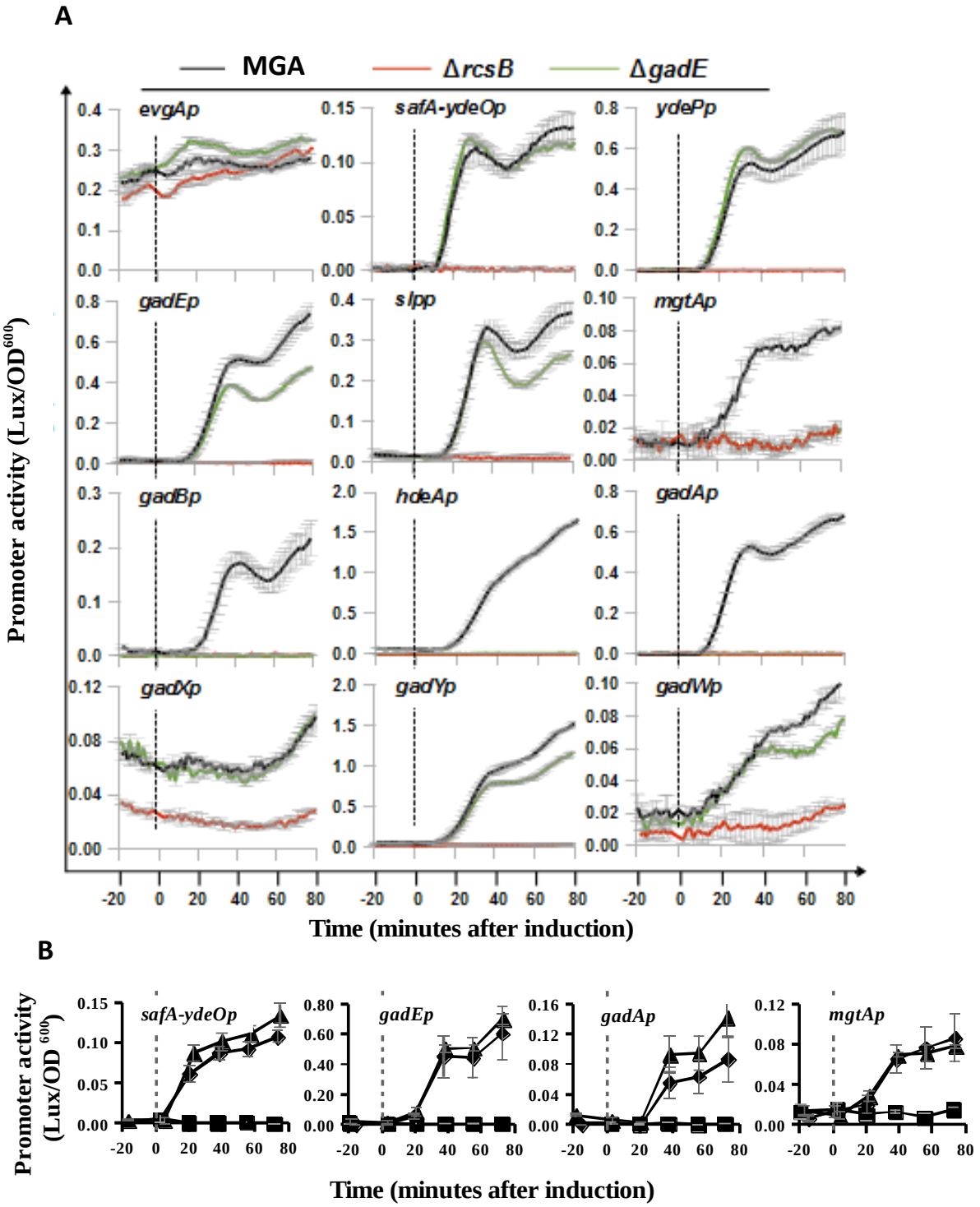

A
MGA
ΔrcsB
ΔgadE
evgAp
safA-ydeOp
ydePp
gadEp
slpp
mgtAp
gadBp
hdeAp
gadAp
gadXp
gadYp
gadWp
Promoter activity (Lux/OD600)
Time (minutes after induction)
B
safA-ydeOp
gadEp
gadAp
mgtAp
Promoter activity (Lux/OD 600)
Time (minutes after induction)

Figure 6.4. Induction dynamics of the AR2 promoters in MGA, *rcsB* deletion and *gadE* deletion backgrounds. A Induction of each promoter by pH 5.7 in MGA (Black lines), ΔrcsB (Red lines) and ΔgadE (Green lines), measured by the high temporal resolution promoter probe assay. Promoter name is included in each graph. Time zero represents the point of acidification, also indicated by a vertical dotted line. Luciferase values were corrected by division of OD^{600}. Each value and error bar represents the average and standard deviations of either 3 biological repeats (MGA and Δ*rcsB*) or 3 technical replicates (Δ*gadE*). **B** Complementation of promoter activity due to induction in an Δ*rcsB* mutant (square data points) by p*rcsB* (diamond data points) compared to MGA promoter activity (triangular data points). All properties are the same as described in **A** except the following. The inclusion of p*rcsB* to supply a plasmid encoded copy of *rcsB* under the control of its native promoter. Each value and error bar represents the average and standard deviation of 3 biological repeats respectively.

There is a minor reduction in activation of these promoters in a Δ*gadE* background. There is no activity at any of these promoters in an Δ*rcsB* background. The *slp* and *gadE* promoters are activated by *ydeO*, which is also inactive in a Δ*rcsB* mutant. The *gadY* and *gadW* promoters are activated by *gadX*, the loss of activity at these promoters in an Δ*rcsB* mutant is consistent with the loss of activity at the *gadX* promoter.

The loss of activity of the *safA-ydeO, gadE, gadA* and *mgtA* promoters was complemented by introducing a plasmid encoded copy of *rcsB*. Figure 6.4 B shows that activity of all promoters in the MGA (triangle data points), Δ*rcsB* (square data points) and Δ*rcsB* +p*rcsB* strains (diamond data points). The activity of each promoter was restored to MGA levels in the presence of p*rcsB*.

In summary, the activation of the AR2 network by mild acid shock during exponential phase is completely dependent on *rcsB*. These results show that the RcsB-GadE heterodimer model is not in itself sufficient to explain the role of RcsB in acid resistance. Consequently, these results do show that RcsB is making other interactions in the AR2 network. The activation of the AR2 network in response to acid stress is completely dependent on the EvgAS TCS.

6.2.3. Role of RcsB in acid resistance and the AR2 network during stationary phase

The results described above demonstrate the affect of an Δ*rcsB* mutant on the induction of the AR2 network, by mild acid shock, during exponential phase growth. Acid resistance in *E. coli* can also be induced by entry into stationary phase (De Biase *et al.*, 1999). To test the role of RcsB in stationary phase cells we assayed the acid resistance and the activity of the AR2 network, during stationary phase, without induction. The standard acid resistance assay was

used to measure the survival of the MGA and Δ*rcsB* strains during stationary phase. The results of these assays are shown in figure 6.5 A. The average survival of the MGA strain was 95.34 %. The high level of survival of *E. coli* K-12 in stationary phase is consistent with previously reported stationary phase levels of resistance (De Biase *et al.*, 1999; Giangrossi *et al.*, 2005; Ma *et al.*, 2003b). The survival of the Δ*rcsB* strain was below the level of detection. This result shows that even stationary phase acid resistance is completely dependent on RcsB. Finally, the sensitive phenotype could be complemented by the addition of the p*rcsB* plasmid figure 6.5 B. This result confirms the previous work of Castanie-Cornet *et al.*

The single time point promoter probe assay was used to measure the activity of all promoters described in 6.2.2, in both MGA and Δ*rcsB* strains. This assay was adjusted from the standard single time point assay, which is described in Chapter 2, by growing the cells into stationary phase, diluting 10-fold and shaking the cultures for 10 minutes before assaying for promoter activity.

The activity of each promoter in the Δ*rcsB* strain relative to the MGA strain is shown in table 6.1. There is no significant reduction in expression at the *gadW*, *gadY* and *mgtA* promoters. Relative expression levels for EvgA regulated promoters could not be determined due to the low levels of expression. These results show that the EvgAS system is not activated by slow growth and requires exposure to mild acid shock. The *gadA*, *hdeA*, *hdeD* and *gadE* promoters all showed a significant decrease in expression in the *rcsB* mutant. These results are consistent with the RcsB-GadE heterodimer model and demonstrate that RcsB is required for induction of GadE regulated promoters during entry into stationary phase. The activity of the *gadX* promoter was also reduced in an Δ*rcsB* background during entry into stationary phase.

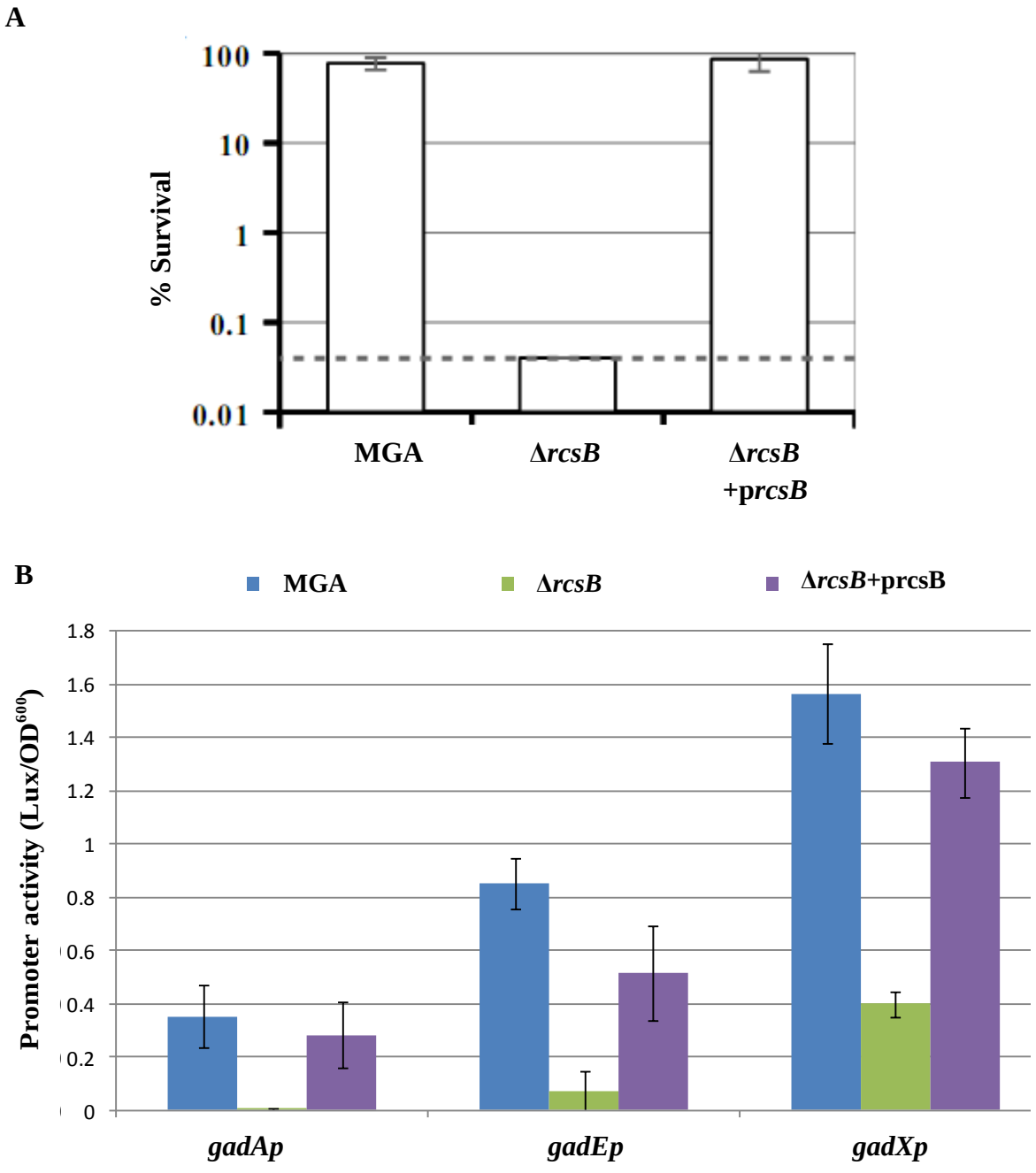


Figure 6.5. Impact of an *rcsB* deletion mutant on acid resistance and AR2 promoter activity during stationary phase. A The survival of MGA, Δ*rcsB* and Δ*rcsB* +p*rcsB* strains after entry into stationary phase. Plasmid p*rscB* is a complementation plasmid that encodes a copy of *rcsB* under the control of the *rcsB* native promoter. Survival was measured by exposing the strains to pH 2.5 for 2 hours, after which the % survival was calculated. **B** Promoter activity of the *gadA* (*gadAp*), *gadE* (*gadEp*) and *gadX* (*gadXp*) genes in MGA (Blue bars) and Δ*rcsB* (Green bars) backgrounds with (Purple bars) and without complementation plasmid p*rcsB* during stationary phase. Promoter activity was measured during stationary phase growth using a single time point promoter probe assay. Luciferase values were corrected for cell number by division by the OD_{600}. In **A** and **B** Each value and error bar represents the average and standard deviation of 3 biological repeats respectively.

This result is not consistent with the GadE-RcsB heterodimer model as the *gadX* promoter is activated independently from GadE in stationary phase cells.

The loss of activity of the *gadA*, *gadE* and *gadX* promoters was complemented by introducing a plasmid encoded copy of *rcsB*, figure 6.5 B. The plasmid pZC320 (the empty plasmid backbone) was also introduced into the strains as a control. The activity of all promoters was restored to MGA levels by the addition of p*rcsB*. Addition of the control promoter pZC320 had no affect on promoter activity.

In summary, the stationary phase results show that RcsB is essential to stationary phase acid resistance and the activation of promoters. These results while consistent with the RcsB-GadE heterodimer model, show that this model cannot fully explain the role of RcsB in acid resistance. RcsB is making other connections within the AR2 network in both stationary and exponential phases.

6.3 Discussion

6.3.1. Role of the PhoPQ TCS in the regulation of acid resistance

The analysis of induction dynamics of a promoter can reveal subtle but important differences in regulation between different mutant backgrounds. The response regulator PhoP was implicated in the regulation of this promoter by assaying the level of activity of the *safA-ydeO* promoter in a Δ*phoP* mutant background. A Δ*phoP* mutant caused an increase in *safA-ydeO* promoter activity (Burton *et al.*, 2010). As the SafA protein has already been shown to activate the PhoPQ TCS, it was suspected that activated PhoP was negatively regulating the *safA-ydeO* promoter. The results described above showed that a Δ*safA* mutant stopped all

activation of PhoPQ in response to mild acid shock and relieved all repression of the *safA-ydeO* promoter. These results do not confirm the direct interaction of PhoP with the *safA-ydeO* promoter. However, work by Ashley Robinson which was published in Burton *et al.* in 2010, revealed two putative PhoP binding sites at the *safA-ydeO* promoter. Single nucleotide mutations of conserved residues in these sites conferred similar levels of de-repression as a Δ*phoP* or *safA* deletion mutant, indicative of reduced binding of PhoP to these sites (Appendix 2) (Burton *et al.*, 2010). This strongly suggests that PhoP does bind the promoter directly but *in vitro* assays such as DNase I footprint assays or EMSAs are required to conclusively prove direct binding. In the context of biological networks, this interaction would represent an indirect negative feedback loop. Negative feedback loops in biological systems enable a faster response to stimuli (Rosenfeld *et al.*, 2002). The dynamics of induction observed at the *safA-ydeO* promoter is suggestive of an activation overshoot, which is then compensated by negative feedback loops. In the context of the *safA-ydeO* promoter this involves two negative feedback loops: direct autoregulation by YdeO, and indirectly by SafA via PhoP (Burton *et al.*, 2010). Presumably, these interactions are temporally separated as the indirect feedback by SafA-PhoP requires additional steps. Unfortunately, the biological significance of these interactions is currently unknown. However, it should be noted that the discovery of such regulatory interactions can only be made using time course experiments. An accurate picture of a regulatory network requires understanding of the temporal component; this takes representations of biological networks beyond 2D pictures and herein lays the necessity for systems biology and mathematical models of biological networks.

6.3.2. RcsB, the AR2 network and acid resistance

The response regulator RcsB was originally characterised as a regulator of capsule synthesis. However, recently RcsB has been implicated in acid resistance in *E. coli* and the AR2 network (Castanie-Cornet *et al.*, 2007). RcsB was found to form a heterodimer with GadE, this has been referred to as the RcsB-GadE heterodimer model (Castanie-Cornet *et al.*, 2010; Krin *et al.*, 2010a). These results showed that the activation of acid resistance by RpoS requires RcsB. They also show that this is because the binding of all GadE-dependent promoters requires RcsB. However, the results described in this study show that an Δ*rcsB* mutant has a more sensitive phenotype compared to a Δ*gadE* mutant in an EvgSc strain (MGA containing a constitutive on EvgS). The survival of an EvgSc mutant strain has been shown to be independent of GadE and high levels of survival have been observed in a Δ*ydeO* Δ*gadE* double knock out. These results showed that RcsB is essential to all EvgSc conferred acid resistance.

Comparison of AR2 promoter activity between Δ*rcsB* and MGA strains revealed that all induction by mild acid shock requires RcsB. This included promoters such as *safA-ydeO* and *mgtA*, which are not regulated by GadE. Comparison of both acid resistance and promoter activity between Δ*rcsB* and MGA strains in stationary phase confirmed that RcsB is essential for acid resistance and promoter activation. These results, while consistent with the original GadE-RcsB heterodimer model, show that it is not sufficient to explain all the actions of RcsB in acid resistance. Based on the results above it is likely that RcsB is interacting with the EvgAS TCS. There are three possible interactions that RcsB could be making with EvgAS. Each interaction is shown in figure 6.6. RcsB could be required either for EvgA binding to promoters, the phosphorylation of EvgA by EvgS or for the activation of EvgS.

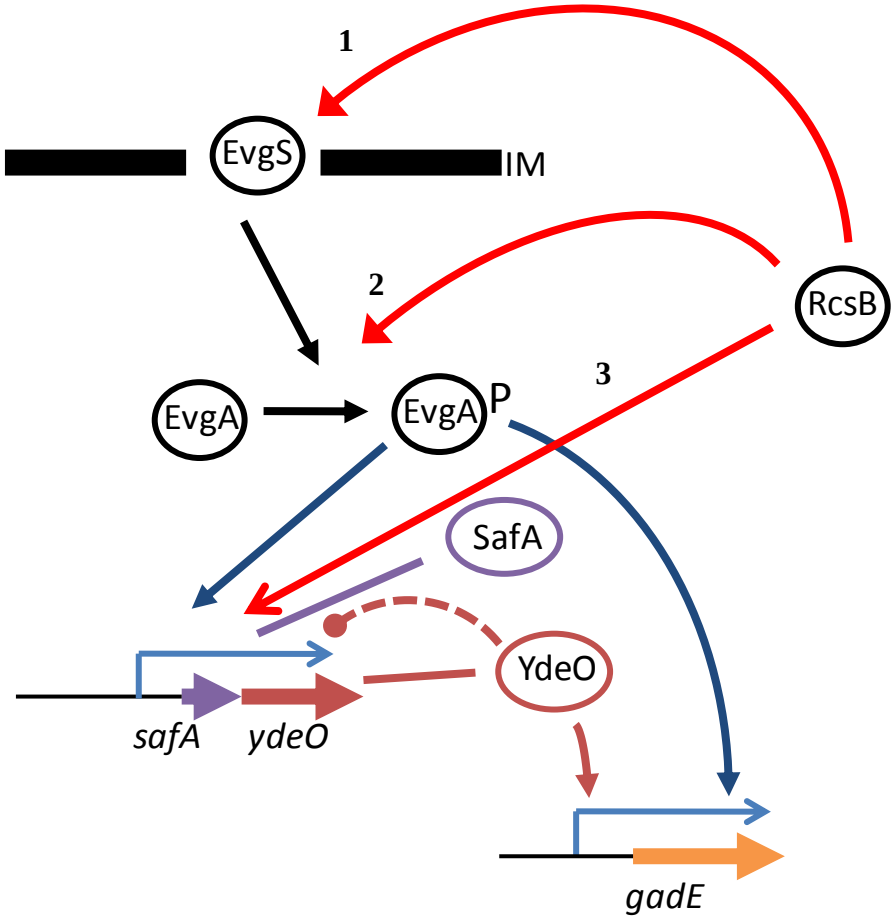


Figure 6.6. Diagrammatic representation of possible interactions of RcsB with the EvgAS TCS. The *safA* (Purple), *ydeO* (Dark red) and *gadE* (Orange) genes are shown by thick arrows. Thin coloured arrows show interactions between proteins and promoters. Red arrows labelled 1, 2 and 3 represent the possible interactions between RcsB and the EvgAS TCS. 1 represents the interaction required for EvgS activation, 2 represents the interaction for activation of EvgA and 3 represents the interaction for EvgA binding to DNA. IM represents the inner membrane of the *E. coli* cell.

The data presented in this chapter can rule out the latter of these interactions as the Δ*rcsB* mutant phenotype was dominant over the EvgSc mutation. This results shows that RcsB cannot be required for EvgS activation. Considering that heterodimer formation is common for RcsB, which also makes such complexes with RcsA, TviA and BglJ to regulate colanic acid synthesis, antigen VI expression and sugar transport, respectively (Venkatesh *et al.*, 2010; Wehland and Bernhard, 2000; Winter *et al.*, 2009). It is reasonable to suggest that RcsB is required for activation of promoters by EvgA. One hypothesis to explain this is that RcsB is forming a heterodimer with EvgA and that the formation of this heterodimer is essential for EvgA promoter binding.

6.3.3. Summary

1) Deletion of *safA* causes PhoPQ to be inactive under inducing conditions. The removal of PhoP or inactivation of PhoP caused increased activation of the SafA promoter. Mutation of conserved residues in a putative PhoP binding site caused a similar increase in activation, indicating that these sites may be bound by PhoP
2) The survival of cells induced by entry into stationary phase, induction by mild acid shock during exponential phase or by an EvgSc mutation is completely dependent on RcsB. However, a Δ*gadE* mutant does not have the same level of sensitivity.
3) Activation of all EvgA regulated promoters is completely dependent on RcsB. As a result the induced AR2 network is completely inactive in an Δ*rcsB* mutant strain.
4) RcsB must be making additional interactions with the AR2 network. The AR2 profile in an RcsB background suggests that RcsB is interacting with the EvgAS system.
5) The fact that the *rcsB* mutation is not suppressed by the EvgSc mutation shows that the role of RcsB is not likely to be activation of EvgS.

6) Stationary phase results are consistent with the RcsB-GadE heterodimer model. However, the Δ*rcsB* mutation also affected the *gadX* promoter. This suggests that RpoS activated promoters are also affected by an Δ*rcsB* knock out.
7) RcsB has been shown to form heterodimers with other regulators. On this basis, a reasonable hypothesis to explain the above results is that it can also form heterodimers with EvgA, and that these are necessary for EvgA-mediumted activation of AR2 promoters. If this interaction occurred it would also be required for EvgA promoter binding.

CHAPTER 7: Final discussion

7.1 Chapter discussions and future work

The results presented in this thesis impact on many areas of prokaryotic gene regulation. Lab-based evolution to select for mutations that conferred increased resistance to extreme acid challenge. High-throughput sequencing, which is a method still under development, was used to sequence the whole genome of 4 out of 5 evolved strains. The concept of combining these methods is still relatively new. The mutations characterised by whole genome re-sequencing included *evgS* mutations that were found in all 5 evolved strains. EvgS is the HK of the EvgAS TCS that has been shown to respond to mild acid shock and activate acid resistance genes. The *evgS* mutations were all missense mutations that were found to be located in the PAS domain of the EvgS protein. The *evgS* mutations were characterised in isolation; when tested, the *evgS* mutations conferred high levels of acid resistance and AR2 promoter activity. This suggested that the *evgS* mutations have altered the structure and function of the EvgS protein and had caused it to become constitutively-on. The AR2 network, which is activated by EvgS, was dissected by analysing promoter activity and acid resistance in an EvgSc strain in different mutant backgrounds. As a result of this analysis the EvgSc mutations were shown to confer redundancy in the AR2 network. In addition, these results reiterated the role of YdeP in acid resistance. Finally, the EvgSc mutations were used to investigate a novel role of RcsB in the regulation of acid resistance. This work is an ideal starting point for investigations into many areas of microbiology and genetics. In particular, this work impacts on lab-based evolution, the mechanisms of TCSs and the role of RcsB in AR2. This chapter will discuss the key findings of each results chapter relevant to these areas and outline future experiments.

7.2. Lab-based evolution

The results in Chapter 3 show that the evolved strains have high levels of acid resistance and all five strains have a mutation in *evgS*. More specifically, three strains (Aa-Ea) had different *evgS* mutations and two strains (Ha and Ga) had evolved the same mutation separately. The evolution experiment, done by P.B. Russell, used extreme acid challenge to evolve an acid resistant phenotype. The cultures were non-clonal during the evolution experiment. Clonal isolates were only isolated at the end of the experiment. Cultures were stored at stages during the process. Consequently, the strains isolated at the end of the evolution experiment may not represent the population as a whole. Moreover, there may be strains carrying other *evgS* mutations or alternative acid resistance mutations altogether. Additional, experiments that could investigate this evolution experiment and the design of additional evolutionary experiments are discussed below.

7.2.1. What is the frequency of the *evgS* mutations in the evolving population?

As described above, this study concentrated on clonal isolates from a population. The population may contain cells with other mutations that contribute to acid resistance. High-throughput sequencing has many applications. This study used high-throughput sequencing to re-sequence the whole genomes of four evolved strains and capture all of the genetic information. Another application of high-throughput sequencing is deep sequencing. Deep sequencing is the sequencing of a particular loci or set of loci using sensitive PCR based methods. The aim of such an approach is to quantify the proportions of different alleles in a mixed population. The data from this analysis would reveal the frequency of the *evgS* mutations in the final population. This analysis can also be applied to the intermediumte

populations, which would reveal the origin and progression of the *evgS* mutations through the populations during the course of the evolution experiment.

7.2.2. What mutations would be selected if the evolution experiment was repeated without EvgAS?

The evolution experiment selected mutations in *evgS* in five individually evolving cell lines. Subsequent analysis of other sequenced strains of *E. coli* revealed that the EAEC strain 042 does not have an *evgAS* locus. Repeating the evolution experiment using *E. coli* 042 could reveal other interesting mutations that confer acid resistance in the absence of *evgAS*. Alternatively, the absence of an *evgAS* locus could reduce the evolvability of the strain preventing the development of acid resistance altogether.

7.2.3. What is the relative fitness of the evolved strains and EvgSc strains compared to the MGA strain?

In Chapter 3 the growth rate of the evolved strains and the MGA strain were analysed. The evolved strains were shown to grow slower in pH 7 M9supp medium. However, competition experiments are a more accurate assessment of the fitness of a strain. It would be particularly interesting to compete the evolved strains with the EvgSc strains. It is possible that the other mutations in the evolved strains could be compensating for the impact of the *evgS* mutation. Thus, explaining the presence of other mutations found in the evolved strains.

7.3. EvgS constitutively active mutants

The results in Chapter 4 confirmed that the EvgSc mutations found in the evolved strains are responsible for the acid resistance phenotype. In addition, results in Chapter 4 indicated that

the EvgSc mutations are causing the AR2 network to be active in conditions when it is normally inactive. The affect of acid induction on strains containing an EvgSc mutation was assayed using high temporal resolution promoter probe assay. The results indicated that three of the EvgSc mutants (G685A, N573L and S584F) were actually repressed by acid induction and one EvgSc mutant (S600I) was induced by acid induction. These results show that although the EvgSc mutations confer an acid resistant phenotype and activation of the AR2 network, they are not affecting the EvgS protein in the same way. In addition, the promoter activity in the AR2 network varies between EvgSc mutants suggesting the some EvgSc mutations are on more than others. The mutations characterised by this study could provide insights into the function of the EvgS protein. In addition, the EvgSc mutations are found in the PAS domain of the protein. The functions of PAS domains vary and little is understood about the role of the PAS domain in EvgS. Experiments using these mutants are described below.

7.3.1. Is it possible to use SMALP to solubilise EvgS?

The biggest obstacle to overcome when working with membrane proteins is their solubility. Until recently, the main methods used to study inner membrane bound proteins have required either detergent, or researchers have simply used the cytoplasmic component (Bock and Gross, 2002). Each method has its caveats. Detergents make the biophysical analysis of proteins difficult as they compromise many techniques, and HKs quite often require membrane binding to function properly. Recent advances in nanodisc technology have made the solubilising of whole membrane proteins possible without detergent. One such technology is being developed at the University of Birmingham called SMALP (styrene maleic acid lipid particle) (Knowles *et al.*, 2009). SMALP can be used to provide a hydrophobic interaction

interface for transmembrane domains. The use of SMALP will enable the study of EvgS and EvgSc mutants by an array of biophysical techniques such as Isothemal calorimetry, surface Plasmon resonance, analytical ultracentifugation, phosphotransfer assays, and X-ray crystallography. The use of such techniques may answer some of the questions described below.

7.3.2. What structural changes occur as a result of the EvgSc mutations?

The use of x-ray crystallography and AUC could reveal the effects of the EvgSc mutations on structure. Solved structures would reveal the interactions and conformational changes that the protein goes through during activation. We would hypothesise that the EvgSc mutants are in an active confirmation as opposed to the wild type protein. The action of TCS often centres on a dimerisation or monomerisation event. AUC is an excellent technique for measuring dimerisation. AUC would also allow the comparison of monomer-dimer equilibriums between EvgSc mutants and the wild type. In addition to AUC, surface plasmon resonance has been used in the past to investigate the interactions between TCSs including EvgAS.

7.3.3 What is the phosphotransfer rate of the EvgSc mutants *in vitro*?

This study has used the promoter activity of the EvgA regulated genes in order to assay the activity of the EvgAS TCS. It is possible to assay the activity by monitoring phosphotransfer between purified EvgS and EvgA *in vitro*. These techniques require that these proteins are solubilised and that EvgS is tethered to the membrane. We hypothesise that the EvgSc mutants would transfer phosphate at a higher rate than that of the wild type. However, it is possible the EvgSc mutants require additional factors. Alternatively, the mutations could be

preventing EvgS from dephosphorylating EvgA. The only way to analyse these interactions in isolation is by using *in vitro* techniques.

7.4. Interactions of RcsB in the AR2 network

The results described in Chapter 6 show that the RR RcsB is forming novel interactions in the AR2 network. This work has shown that the activation of all EvgA regulated genes is completely dependent on RcsB. However, the molecular details of this interaction have not been identified. Future experiments regarding the role of RcsB in the AR2 network are described below.

7.4.1. Does RcsB form a heterodimer with EvgA?

The dependence of EvgA regulated promoters on RcsB suggests that RcsB is either required for the activation of EvgA or the binding of EvgA to promoters. RcsB has been shown to form interactions with other regulators. Considering this, we hypothesise that RcsB is forming a heterodimer with EvgA. To test this hypothesis, the purified EvgA and RcsB can be used in an EMSA with the *ydeP* or *safA-ydeO* promoters. This can be combined with a nested deletion of the *ydeP* promoter, point mutations and a DNaseI footprint to confirm the binding site. Using these techniques we should be able to determine the relationship between EvgA and RcsB. We also hypothesise that EvgA and RcsB form a heterodimer independently of DNA. In order to test this hypothesis EvgA could be co-purified with RcsB or used in a "protein pull down" experiment. Co-purification is often used as evidence of protein-protein interactions. It is also noted there are few occasions where regulators have been shown to form heterodimers and even fewer that form heterdimers with so many different regulators of

different processes. The molecular interactions that RcsB makes with other proteins would be a particularly interesting subject of research.

7.5. General Discussion

The investigations into the regulation of acid stress response in *E. coli* have uncovered a very complicated regulatory network, to which, this study has added. However, the process of sensing and responding to a stress is theoretically far simpler than the AR2 network. One of the main overriding questions, when investigating AR2 is; what is the reason for this complexity? One possible answer to this question is that the complexity of this network is there simply to provide robustness to the system, which in turn will increase the evolvability of the strain. This is a well established concept in evolutionary biology and several examples of redundancy have been shown in biological systems (Masel and Trotter, 2010; Raman and Wagner, 2011; Woods *et al.*, 2011). A particularly striking example of robustness was published by Isalan *et al*. in 2008. This study rewired the some of the major transcription factors by using alternative promoters to control their expression. The major finding of this work was that the growth rates of strains containing new/rewired networks were indifferent from the wild type in optimum growth conditions. Moreover, a subset of strains containing rewired networks, were able to survive certain selection pressures better than the wild type (Isalan *et al.*, 2008). Robustness would enable the network to incur alterations, such as mutations, gene duplications and horizontal transfer, without dramatically impacting on the fitness of the strain. This is similar to the evolutionary theory of survival of the flattest, which argues that biological systems, which are robust, can accumulate a larger proportion of neutral mutations (Wilke *et al.*, 2001). This would allow the system greater freedom to move around the evolutionary space and increase the chance that the system will find a new fitness

peak. However, this concept does assume that the complexity of the system is not functional, and that the regulatory network is inefficient.

An alternative explanation is that the interactions found within complex networks, such as AR2, are required by the organism in order to respond to stimuli appropriately. A notable example of this is combinatorial sensing. Combinatorial sensing is the temporal or spatial summation of signals from multiple sensors. Recently it has been suggested that TCSs in bacteria are using combinatorial sensing to respond accurately to a stimulus or stress (Clarke and Voigt, 2011). By cross-talking to each other TCSs, and the components they regulate, can increase the complexity of the response. There have been many examples of TCSs cross-talking. These interactions could be enabling the cell to respond to a complex environment accurately and efficiently, as opposed to uneconomic features that enable robustness and evolvability. Recent advances in the understanding of signalling by PhoPQ have shown that although PhoPQ activates the transcription of *mgtA* in response to acid shock and low periplasmic Mg^{2+}, the *mgtA* transcript is only processed in to protein in low Mg^{2+} conditions (Choi *et al.*, 2009; Park *et al.*, 2010). As a result a combination of signals is interpreted by the PhoPQ system at the post-transcriptional level and the MgtA protein is only made under the correct conditions. Furthermore, because of this post-transcriptional regulation, the cell can regulate MgtA production based on periplasmic and cytoplasmic conditions.

In summary, the interactions identified by this study, and by preceeding studies, have shown that at a transcriptional level, the regulatory factors controlling acid stress response form a complex regulatory network. The biological significance of this network is not currently

appreciated. Similar levels of investigation of post-transcriptional regulation are required to understand the importance of these interactions, and to untangle these networks.

APPENDIX 1

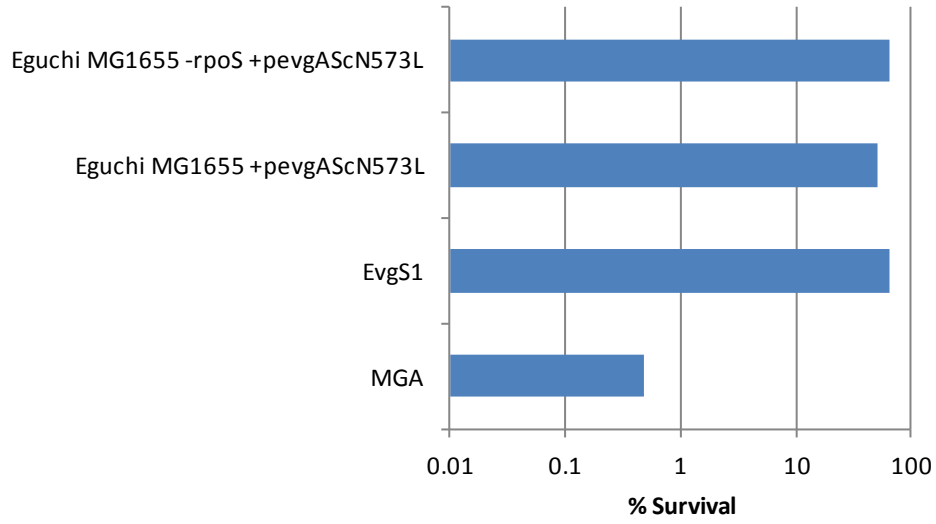


Figure A1. Comparison of the Eguchi *et al.* strains using the acid resistance assay described in Chapter 2. The survival of the MGA strain was compared to the EvgS1 mutant, which is a constitutive-on EvgS mutant; the *E. coli* MG1655 strain used by Eguchi *et al.* (Eguchi MG1655) with p*evgAS*cN573L and with the *rpoS* mutant strain used by Eguchi *et al.* with p*evgAS*cN573L. Each data point represents 1 biological repeat.

APPENDIX 2

(a)

```
                          A      mut3
                          AG     mut2
                           G     mut1
TAAGCATACTGATTAACGATTTTTAACGTTATCCGCTAAATAAACATTTTGAAATG
```

(b)

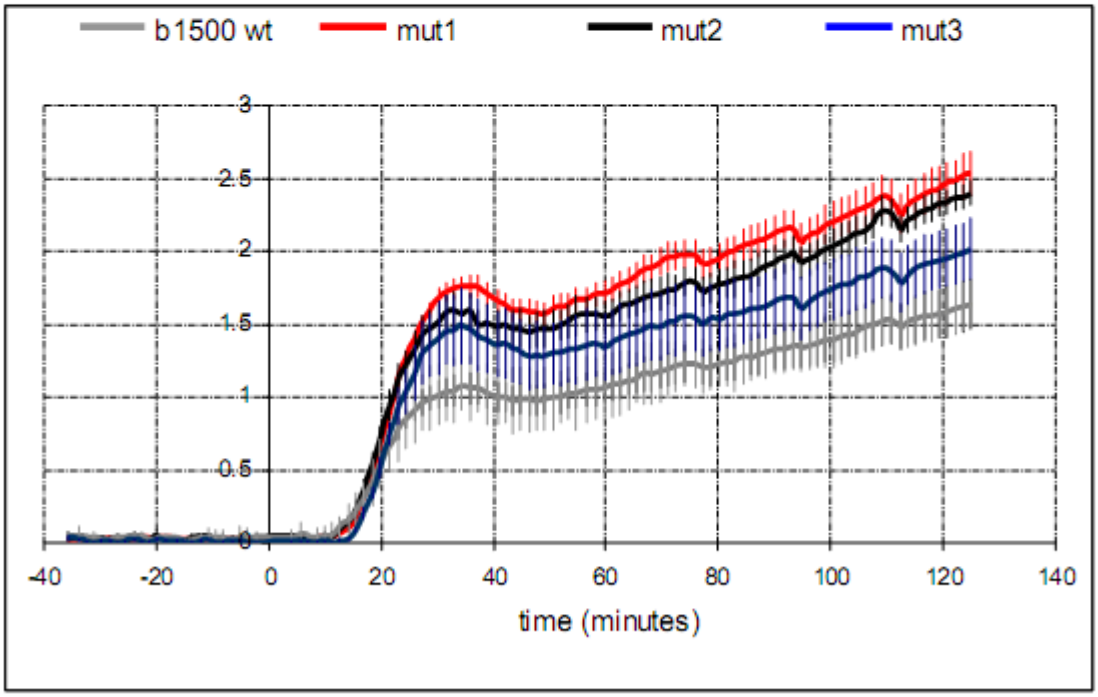


Figure A2. Affect of point mutations in putative PhoP binding sites on *safA-ydeO* promoter activity. The *safA-ydeO* (known as B1500) promoter activity was measured using the same high-temporal resolution promoter probe assay as described in Chapter 2 (b). The promoter activity of three point mutants, shown in (a) was compared using the same assay. Values and error bars represent the average and standard deviation of three biological repeats respectively. (Adapted from Burton *et al.*, 2010)

BIBLIOGRAPHY

Aiso, T., Murata, M. and Gamou, S. (2011) Transcription of an antisense RNA of a *gadE* mRNA is regulated by GadE, the central activator of the acid resistance system in *Escherichia coli*. *Genes Cells* **16:** 670-80.

Alexander, D. M. and St John, A. C. (1994) Characterization of the carbon starvation-inducible and stationary phase-inducible gene *slp* encoding an outer membrane lipoprotein in *Escherichia coli*. *Mol Microbiol* **11:** 1059-71.

Andrell, J., Hicks, M. G., Palmer, T., Carpenter, E. P., Iwata, S. and Maher, M. J. (2009) Crystal structure of the acid-induced arginine decarboxylase from *Escherichia coli*: reversible decamer assembly controls enzyme activity. *Biochemistry* **48:** 3915-27.

Auger, E. A., Redding, K. E., Plumb, T., Childs, L. C., Meng, S. Y. and Bennett, G. N. (1989) Construction of *lac* fusions to the inducible arginine- and lysine decarboxylase genes of *Escherichia coli* K12. *Mol Microbiol* **3:** 609-20.

Bantscheff, M., Perraud, A. L., Bock, A., Rippe, K., Weiss, V., Glocker, M. and Gross, R. (2000) Structure-function relationships in the Bvg and Evg two-component phosphorelay systems. *Int J Med Microbiol* **290:** 317-23.

Barrick, J. E., Kauth, M. R., Strelioff, C. C. and Lenski, R. E. (2010) *Escherichia coli* rpoB mutants have increased evolvability in proportion to their fitness defects. *Mol Biol Evol* **27:** 1338-47.

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E. and Kim, J. F. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461:** 1243-7.

Bearson, B. L., Lee, I. S. and Casey, T. A. (2009) *Escherichia coli* O157 : H7 glutamate- and arginine-dependent acid-resistance systems protect against oxidative stress during extreme acid challenge. *Microbiology* **155:** 805-12.

Benjamin, M. M. and Datta, A. R. (1995) Acid tolerance of enterohemorrhagic *Escherichia coli*. *Appl Environ Microbiol* **61:** 1669-72.

Bentley, R. and Meganathan, R. (1982). Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev* **46**: 241-80.

Blaser, M. J. and Newman, L. S. (1982) A review of human salmonellosis: I. Infective dose. *Rev Infect Dis* **4:** 1096-106.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453-62.

Blethen, S. L., Boeker, E. A. and Snell, E. E. (1968) Argenine decarboxylase from *Escherichia coli*. I. Purification and specificity for substrates and coenzyme. *J Biol Chem* **243:** 1671-7.

Bock, A. and Gross, R. (2002) The unorthodox histidine kinases BvgS and EvgS are responsive to the oxidation status of a quinone electron carrier. *Eur J Biochem* **269:** 3479-84.

Buelow, D. R. and Raivio, T. L. (2010) Three (and more) component regulatory systems - auxiliary regulators of bacterial histidine kinases. *Mol Microbiol* **75:** 547-66.

Burton, N. A., Johnson, M. D., Antczak, P., Robinson, A. and Lund, P. A. (2010) Novel Aspects of the Acid Response Network of E. coli K-12 Are Revealed by a Study of Transcriptional Dynamics. *J Mol Biol* **401:** 726-42

Busby, S. and Ebright, R. H. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293:** 199-213.

Cabedo, H., Macian, F., Villarroya, M., Escudero, J. C., Martinez-Vicente, M., Knecht, E. and Armengod, M. E. (1999) The *Escherichia coli trmE* (*mnmE*) gene, involved in tRNA modification, codes for an evolutionarily conserved GTPase with unusual biochemical properties. *EMBO J* **18:** 7063-76.

Campbell, A. J., Watts, K. J., Johnson, M. S. and Taylor, B. L. (2010) Gain-of-function mutations cluster in distinct regions associated with the signalling pathway in the PAS domain of the aerotaxis receptor, Aer. *Mol Microbiol* **77:** 575-86.

Capitani, G., De Biase, D., Aurizi, C., Gut, H., Bossa, F. and Grutter, M. G. (2003) Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *EMBO J* **22:** 4027-37.

Castanie-Cornet, M. P., Cam, K., Bastiat, B., Cros, A., Bordes, P. and Gutierrez, C. (2010) Acid stress response in *Escherichia coli*: mechanism of regulation of *gadA* transcription by RcsB and GadE. *Nucleic Acids Res* **38:** 3546-54.

Castanie-Cornet, M. P. and Foster, J. W. (2001) *Escherichia coli* acid resistance: cAMP receptor protein and a 20 bp cis-acting sequence control pH and stationary phase expression of the *gadA* and *gadBC* glutamate decarboxylase genes. *Microbiology* **147:** 709-15.

Castanie-Cornet, M. P., Penfound, T. A., Smith, D., Elliott, J. F. and Foster, J. W. (1999) Control of acid resistance in *Escherichia coli*. *J Bacteriol* **181:** 3525-35.

Castanie-Cornet, M. P., Treffandier, H., Francez-Charlot, A., Gutierrez, C. and Cam, K. (2007) The glutamate-dependent acid resistance system in *Escherichia coli*: essential and dual role of the His-Asp phosphorelay RcsCDB/AF. *Microbiology* **153:** 238-46.

Chaudhuri, R. R., Sebaihia, M., Hobman, J. L., Webber, M. A., Leyton, D. L., Goldberg, M. D., Cunningham, A. F., Scott-Tucker, A., Ferguson, P. R., Thomas, C. M., Frankel, G., Tang, C. M., Dudley, E. G., Roberts, I. S., Rasko, D. A., Pallen, M. J., Parkhill, J., Nataro, J. P., Thomson, N. R. and Henderson, I. R. (2010) Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS One* **5:** e8801.

Chenchik, A. A., Bibilashvili, R., Mirzabekov, A. D. and Shik, V. V. (1982) [Contacts of *Escherichia coli* RNA polymerase subunits with nucleotides of *lac*UV5 promoter]. *Mol Biol (Mosk)* **16:** 35-46.

Cherepanov, P. P. and Wackernagel, W. (1995) Gene disruption in *Echerichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158:** 3880-5.

Choi, E., Groisman, E. A. and Shin, D. (2009) Activated by different signals, the PhoP/PhoQ two-component system differentially regulates metal uptake. *J Bacteriol* **191:** 7174-81.

Clarke, D. J. (2010) The Rcs phosphorelay: more than just a two-component pathway. *Future Microbiol* **5:** 1173-84.

Clarke, E. J. and Voigt, C. A. (2011) Characterization of combinatorial patterns generated by multiple two-component sensors in *E. coli* that respond to many stimuli. *Biotechnol Bioeng* **108:** 666-75.

Coldewey, S. M., Hartmann, M., Schmidt, D. S., Engelking, U., Ukena, S. N. and Gunzer, F. (2007) Impact of the *rpoS* genotype for acid resistance patterns of pathogenic and probiotic *Escherichia coli*. *BMC Microbiol* **7:** 21.

Conrad, T. M., Frazier, M., Joyce, A. R., Cho, B. K., Knight, E. M., Lewis, N. E., Landick, R. and Palsson, B. O. (2010) RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal medium. *Proc Natl Acad Sci U S A* **107:** 20500-5.

Crozat, E., Hindre, T., Kuhn, L., Garin, J., Lenski, R. E. and Schneider, D. (2011) Altered regulation of the OmpF porin by Fis in *Escherichia coli* during an evolution experiment and between B and K-12 strains. *J Bacteriol* **193:** 429-40.

Daley, D. O., Rapp, M., Granseth, E., Melen, K., Drew, D. and von Heijne, G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308:** 1321-3.

Dame, R. T., Wyman, C. and Goosen, N. (2000) H-NS mediumted compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* **28:** 3504-10.

Datsenko, K. A. and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97:** 6640-5.

De Biase, D., Tramonti, A., Bossa, F. and Visca, P. (1999) The response to stationary-phase stress conditions in *Escherichia coli*: role and regulation of the glutamic acid decarboxylase system. *Mol Microbiol* **32:** 1198-211.

De Biase, D., Tramonti, A., John, R. A. and Bossa, F. (1996) Isolation, overexpression, and biochemical characterization of the two isoforms of glutamic acid decarboxylase from *Escherichia coli*. *Protein Expr Purif* **8:** 430-8.

Diez-Gonzalez, F. and Karaibrahimoglu, Y. (2004) Comparison of the glutamate-, arginine- and lysine-dependent acid resistance systems in *Escherichia coli* O157:H7. *J Appl Microbiol* **96:** 1237-44.

Dorman, C. J. (2007) H-NS, the genome sentinel. *Nat Rev Microbiol* **5:** 157-61.

DuPont, H. L., Formal, S. B., Hornick, R. B., Snyder, M. J., Libonati, J. P., Sheahan, D. G., LaBrec, E. H. and Kalas, J. P. (1971) Pathogenesis of *Escherichia coli* diarrhea. *N Engl J Med* **285:** 1-9.

Eguchi, Y., Ishii, E., Hata, K. and Utsumi, R. (2011) Regulation of Acid Resistance by Connectors of Two-Component Signal Transduction Systems in *Escherichia coli*. *J Bacteriol* **193:** 1222-8.

Eguchi, Y., Itou, J., Yamane, M., Demizu, R., Yamato, F., Okada, A., Mori, H., Kato, A. and Utsumi, R. (2007) 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* **104:** 18712-7.

Eguchi, Y., Okada, T., Minagawa, S., Oshima, T., Mori, H., Yamamoto, K., Ishihama, A. and Utsumi, R. (2004) Signal transduction cascade between EvgA/EvgS and PhoP/PhoQ two-component systems of *Escherichia coli*. *J Bacteriol* **186:** 3006-14.

Eguchi, Y. and Utsumi, R. (2005) A novel mechanism for connecting bacterial two-component signal-transduction systems. *Trends Biochem Sci* **30:** 70-2.

Elena, S. F. and Lenski, R. E. (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* **4:** 457-69.

Etzkorn, M., Kneuper, H., Dunnwald, P., Vijayan, V., Kramer, J., Griesinger, C., Becker, S., Unden, G. and Baldus, M. (2008) Plasticity of the PAS domain and a potential role for signal transduction in the histidine kinase DcuS. *Nat Struct Mol Biol* **15:** 1031-9.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res* **38:** D211-22.

Foster, J. W. (2004) *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat Rev Microbiol* **2:** 898-907.

Freeman, J. A., Lilley, B. N. and Bassler, B. L. (2000) A genetic analysis of the functions of LuxN: a two-component hybrid sensor kinase that regulates quorum sensing in *Vibrio harveyi*. *Mol Microbiol* **35:** 139-49.

Gajiwala, K. S. and Burley, S. K. (2000) HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *J Mol Biol* **295:** 605-12.

Gale, E. F. and Epps, H. M. (1942) The effect of the pH of the medium during growth on the enzymatic activites of bacteria (*Escherichia coli* and *Micrococcus lysodeiktius*) and the biological significance of the changes produced. *Biochem J.* **36**: 600-18.

Gallegos, M. T., Michan, C. and Ramos, J. L. (1993) The XylS/AraC family of regulators. *Nucleic Acids Res* **21:** 807-10.

Giangrossi, M., Zattoni, S., Tramonti, A., De Biase, D. and Falconi, M. (2005) Antagonistic role of H-NS and GadX in the regulation of the glutamate decarboxylase-dependent acid resistance system in *Escherichia coli*. *J Biol Chem* **280:** 21498-505.

Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R. and Kiley, P. J. (2006) IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in *Escherichia coli*. *Mol Microbiol* **60:** 1058-75.

Gilles-Gonzalez, M. A., Gonzalez, G., Perutz, M. F., Kiger, L., Marden, M. C. and Poyart, C. (1994) Heme-based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation. *Biochemistry* **33:** 8067-73.

Gong, S., Ma, Z. and Foster, J. W. (2004) The Era-like GTPase TrmE conditionally activates *gadE* and glutamate-dependent acid resistance in *Escherichia coli*. *Mol Microbiol* **54:** 948-61.

Gong, S., Richard, H. and Foster, J. W. (2003) YjdE (AdiC) is the arginine:agmatine antiporter essential for arginine-dependent acid resistance in *Escherichia coli*. *J Bacteriol* **185:** 4402-9.

Gorden, J. and Small, P. L. (1993) Acid resistance in enteric bacteria. *Infect Immun* **61:** 364-7.

Grainger, D. C., Hurd, D., Goldberg, M. D. and Busby, S. J. (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* **34:** 4642-52.

Groisman, E. A. (2001) The pleiotropic two-component regulatory system PhoP-PhoQ. *J Bacteriol* **183:** 1835-42.

Groisman, E. A., Heffron, F. and Solomon, F. (1992) Molecular genetic analysis of the *Escherichia coli* phoP locus. *J Bacteriol* **174:** 486-91.

Gunsalus, R. P. and Park, S. J. (1994) Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. *Res Microbiol* **145:** 437-50.

Hengge-Aronis, R. (2002) Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev* **66:** 373-95, table of contents.

Herring, C. D., Glasner, J. D. and Blattner, F. R. (2003) Gene replacement without selection: regulated suppression of amber mutations in *Escherichia coli*. *Gene* **311:** 153-63.

Hersh, B. M., Farooq, F. T., Barstad, D. N., Blankenhorn, D. L. and Slonczewski, J. L. (1996) A glutamate-dependent acid resistance gene in *Escherichia coli*. *J Bacteriol* **178:** 3978-81.

Hert, D. G., Fredlake, C. P. and Barron, A. E. (2008) Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. *Electrophoresis* **29:** 4618-26.

Heuveling, J., Possling, A. and Hengge, R. (2008) A role for Lon protease in the control of the acid resistance genes of *Escherichia coli*. *Mol Microbiol* **69:** 534-47.

Hirakawa, H., Inazumi, Y., Senda, Y., Kobayashi, A., Hirata, T., Nishino, K. and Yamaguchi, A. (2006) N-acetyl-d-glucosamine induces the expression of multidrug exporter genes, mdtEF, via catabolite activation in *Escherichia coli*. *J Bacteriol* **188:** 5851-8.

Hobman, J. L., Patel, M. D., Hidalgo-Arroyo, G. A., Cariss, S. J., Avison, M. B., Penn, C. W. and Constantinidou, C. (2007) Comparative genomic hybridization detects secondary chromosomal deletions in *Escherichia coli* K-12 MG1655 mutants and highlights instability in the flhDC region. *J Bacteriol* **189:** 8786-92.

Holm, L., Ouzounis, C., Sander, C., Tuparev, G. and Vriend, G. (1992) A database of protein structure families with common folding motifs. *Protein Sci* **1:** 1691-8.

Hommais, F., Krin, E., Coppee, J. Y., Lacroix, C., Yeramian, E., Danchin, A. and Bertin, P. (2004) GadE (YhiE): a novel activator involved in the response to acid environment in *Escherichia coli*. *Microbiology* **150:** 61-72.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A. and Bertin, P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40:** 20-36.

Huang, Z. J., Edery, I. and Rosbash, M. (1993) PAS is a dimerization domain common to Drosophila period and several transcription factors. *Nature* **364:** 259-62.

Isalan, M., Lemerle, C., Michalodimitrakis, K., Horn, C., Beltrao, P., Raineri, E., Garriga-Canut, M. and Serrano, L. (2008) Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452:** 840-5.

Itou, J., Eguchi, Y. and Utsumi, R. (2009) Molecular mechanism of transcriptional cascade initiated by the EvgS/EvgA system in *Escherichia coli* K-12. *Biosci Biotechnol Biochem* **73:** 870-8.

Iuchi, S. and Lin, E. C. (1988) *arcA* (dye), a global regulatory gene in *Escherichia coli* mediumting repression of enzymes in aerobic pathways. *Proc Natl Acad Sci U S A* **85:** 1888-92.

Iuchi, S. and Lin, E. C. (1993) Adaptation of *Escherichia coli* to redox environments by gene expression. *Mol Microbiol* **9:** 9-15.

Iyer, R., Williams, C. and Miller, C. (2003) Arginine-agmatine antiporter in extreme acid resistance in *Escherichia coli*. *J Bacteriol* **185:** 6556-61.

Johnson, M. D., Burton, N. A., Gutierrez, B., Painter, K. and Lund, P. A. (2011) RcsB Is Required for Inducible Acid Resistance in *Escherichia coli* and Acts at gadE-Dependent and -Independent Promoters. *J Bacteriol* **193:** 3653-6.

Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M. G. and Alon, U. (2001) Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292:** 2080-3.

Kannan, G., Wilks, J. C., Fitzgerald, D. M., Jones, B. D., Bondurant, S. S. and Slonczewski, J. L. (2008) Rapid acid treatment of *Escherichia coli*: transcriptomic response and recovery. *BMC Microbiol* **8:** 37.

Kato, A., Ohnishi, H., Yamamoto, K., Furuta, E., Tanabe, H. and Utsumi, R. (2000) Transcription of *emrKY* is regulated by the EvgA-EvgS two-component system in *Escherichia coli* K-12. *Biosci Biotechnol Biochem* **64:** 1203-9.

Kern, R., Malki, A., Abdallah, J., Tagourti, J. and Richarme, G. (2007) *Escherichia coli* HdeB is an acid stress chaperone. *J Bacteriol* **189:** 603-10.

Kieboom, J. and Abee, T. (2006) Arginine-dependent acid resistance in Salmonella enterica serovar Typhimurium. *J Bacteriol* **188:** 5650-3.

Knowles, T. J., Finka, R., Smith, C., Lin, Y. P., Dafforn, T. and Overduin, M. (2009) Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc* **131:** 7484-5.

Krin, E., Danchin, A. and Soutourina, O. (2010a) Decrypting the H-NS-dependent regulatory cascade of acid stress resistance in *Escherichia coli*. *BMC Microbiol* **10:** 273.

Krin, E., Danchin, A. and Soutourina, O. (2010b) RcsB plays a central role in H-NS-dependent regulation of motility and acid stress resistance in *Escherichia coli*. *Res Microbiol*.

Kuper, C. and Jung, K. (2005) CadC-mediumted activation of the *cadBA* promoter in *Escherichia coli*. *J Mol Microbiol Biotechnol* **10:** 26-39.

Lee, D. J., Bingle, L. E., Heurlier, K., Pallen, M. J., Penn, C. W., Busby, S. J. and Hobman, J. L. (2009) Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiol* **9:** 252.

Lin, E. C. and Iuchi, S. (1991) Regulation of gene expression in fermentative and respiratory systems in *Escherichia coli* and related bacteria. *Annu Rev Genet* **25:** 361-87.

Lin, J., Lee, I. S., Frey, J., Slonczewski, J. L. and Foster, J. W. (1995) Comparative analysis of extreme acid survival in *Salmonella typhimurium*, *Shigella flexneri*, and *Escherichia coli*. *J Bacteriol* **177:** 4097-104.

Lin, J., Smith, M. P., Chapin, K. C., Baik, H. S., Bennett, G. N. and Foster, J. W. (1996) Mechanisms of acid resistance in enterohemorrhagic *Escherichia coli*. *Appl Environ Microbiol* **62:** 3094-100.

Ling, J., Sharma, M. and Bhagwat, A. A. (2008) Role of RNA polymerase sigma-factor (RpoS) in induction of glutamate-dependent acid-resistance of *Escherichia albertii* under anaerobic conditions. *FEMS Microbiol Lett* **283:** 75-82.

Ma, Z., Gong, S., Richard, H., Tucker, D. L., Conway, T. and Foster, J. W. (2003a) GadE (YhiE) activates glutamate decarboxylase-dependent acid resistance in *Escherichia coli* K-12. *Mol Microbiol* **49:** 1309-20.

Ma, Z., Masuda, N. and Foster, J. W. (2004) Characterization of EvgAS-YdeO-GadE branched regulatory circuit governing glutamate-dependent acid resistance in *Escherichia coli*. *J Bacteriol* **186:** 7378-89.

Ma, Z., Richard, H. and Foster, J. W. (2003b) pH-Dependent modulation of cyclic AMP levels and GadW-dependent repression of RpoS affect synthesis of the GadX regulator and *Escherichia coli* acid resistance. *J Bacteriol* **185:** 6852-9.

Ma, Z., Richard, H., Tucker, D. L., Conway, T. and Foster, J. W. (2002) Collaborative regulation of *Escherichia coli* glutamate-dependent acid resistance by two AraC-like regulators, GadX and GadW (YhiW). *J Bacteriol* **184:** 7001-12.

Malpica, R., Franco, B., Rodriguez, C., Kwon, O. and Georgellis, D. (2004) Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proc Natl Acad Sci U S A* **101:** 13318-23.

Mascher, T., Helmann, J. D. and Unden, G. (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* **70:** 910-38.

Masel, J. and Trotter, M. V. (2010) Robustness and evolvability. *Trends Genet* **26:** 406-14.

Masuda, N. and Church, G. M. (2002) *Escherichia coli* gene expression responsive to levels of the response regulator EvgA. *J Bacteriol* **184:** 6225-34.

Masuda, N. and Church, G. M. (2003) Regulatory network of acid resistance genes in *Escherichia coli*. *Mol Microbiol* **48:** 699-712.

Mates, A. K., Sayed, A. K. and Foster, J. W. (2007) Products of the *Escherichia coli* acid fitness island attenuate metabolite stress at extremely low pH and mediumte a cell density-dependent acid resistance. *J Bacteriol* **189:** 2759-68.

Meng, S. Y. and Bennett, G. N. (1992a) Nucleotide sequence of the *Escherichia coli* cad operon: a system for neutralization of low extracellular pH. *J Bacteriol* **174:** 2659-69.

Meng, S. Y. and Bennett, G. N. (1992b) Regulation of the *Escherichia coli* cad operon: location of a site required for acid induction. *J Bacteriol* **174:** 2670-8.

Mika, F. and Hengge, R. (2005) A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in E. coli. *Genes Dev* **19:** 2770-81.

Miyatake, H., Mukai, M., Park, S. Y., Adachi, S., Tamura, K., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T. and Shiro, Y. (2000) Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: crystallographic, mutagenesis and resonance Raman spectroscopic studies. *J Mol Biol* **301:** 415-31.

Moreau, P. L. (2007) The lysine decarboxylase CadA protects *Escherichia coli* starved of phosphate against fermentation acids. *J Bacteriol* **189:** 2249-61.

Naryshkina, T., Mustaev, A., Darst, S. A. and Severinov, K. (2001) The beta ' subunit of *Escherichia coli* RNA polymerase is not required for interaction with initiating nucleotide but is necessary for interaction with rifampicin. *J Biol Chem* **276:** 13308-13.

Neely, M. N., Dell, C. L. and Olson, E. R. (1994) Roles of LysP and CadC in mediumting the lysine requirement for acid induction of the *Escherichia coli cad* operon. *J Bacteriol* **176:** 3278-85.

Nishino, K. and Yamaguchi, A. (2001) Overexpression of the response regulator evgA of the two-component signal transduction system modulates multidrug resistance conferred by multidrug resistance transporters. *J Bacteriol* **183:** 1455-8.

Nishino, K. and Yamaguchi, A. (2002) EvgA of the two-component signal transduction system modulates production of the *yhiUV* multidrug transporter in *Escherichia coli*. *J Bacteriol* **184:** 2319-23.

Opdyke, J. A., Kang, J. G. and Storz, G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* **186:** 6698-705.

Park, S. Y., Cromie, M. J., Lee, E. J. and Groisman, E. A. (2010) A bacterial mRNA leader that employs different mechanisms to sense disparate intracellular signals. *Cell* **142:** 737-48.

Park, Y. K., Bearson, B., Bang, S. H., Bang, I. S. and Foster, J. W. (1996) Internal pH crisis, lysine decarboxylase and the acid tolerance response of *Salmonella typhimurium*. *Mol Microbiol* **20:** 605-11.

Pelosi, L., Kuhn, L., Guetta, D., Garin, J., Geiselmann, J., Lenski, R. E. and Schneider, D. (2006) Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* **173:** 1851-69.

Qi, Y., Rao, F., Luo, Z. and Liang, Z. X. (2009) A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from Acetobacter xylinum. *Biochemistry* **48:** 10275-85.

Raman, K. and Wagner, A. (2011) Evolvability and robustness in a complex signalling circuit. *Mol Biosyst* **7:** 1081-92.

Reinelt, S., Hofmann, E., Gerharz, T., Bott, M. and Madden, D. R. (2003) The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. *J Biol Chem* **278:** 39189-96.

Richard, H. and Foster, J. W. (2004) *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J Bacteriol* **186:** 6032-41.

Richard, H. and Foster, J. W. (2007) Sodium regulates *Escherichia coli* acid resistance, and influences GadX- and GadW-dependent activation of gadE. *Microbiology* **153:** 3154-61.

Richard, H. T. and Foster, J. W. (2003) Acid resistance in *Escherichia coli*. *Adv Appl Microbiol* **52:** 167-86.

Rosenfeld, N., Elowitz, M. B. and Alon, U. (2002) Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* **323:** 785-93.

Sayed, A. K. and Foster, J. W. (2009) A 750 bp sensory integration region directs global control of the *Escherichia coli* GadE acid resistance regulator. *Mol Microbiol* **71:** 1435-50.

Sayed, A. K., Odom, C. and Foster, J. W. (2007) The *Escherichia coli* AraC-family regulators GadX and GadW activate *gadE*, the central activator of glutamate-dependent acid resistance. *Microbiology* **153:** 2584-92.

Schell, M. A. (1993) Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47:** 597-626.

Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H. and Kiley, P. J. (2001) IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. *Proc Natl Acad Sci U S A* **98:** 14895-900.

Senda, T. and Ogawa, N. (2005) [Structure function relationship of the LysR type transcriptional regulator]. *Tanpakushitsu Kakusan Koso* **50:** 1535-40.

Shi, J. and Biek, D. P. (1995) A versatile low-copy-number cloning vector derived from plasmid F. *Gene* **164:** 55-8.

Shi, X., Waasdorp, B. C. and Bennett, G. N. (1993) Modulation of acid-induced amino acid decarboxylase gene expression by hns in *Escherichia coli*. *J Bacteriol* **175:** 1182-6.

Shimada, T., Makinoshima, H., Ogawa, Y., Miki, T., Maeda, M. and Ishihama, A. (2004) Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. *J Bacteriol* **186:** 7112-22.

Shin, D., Lee, E. J., Huang, H. and Groisman, E. A. (2006) A positive feedback loop promotes transcription surge that jump-starts *Salmonella* virulence circuit. *Science* **314:** 1607-9.

Small, P., Blankenhorn, D., Welty, D., Zinser, E. and Slonczewski, J. L. (1994) Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: role of rpoS and growth pH. *J Bacteriol* **176:** 1729-37.

Small, P. L. and Waterman, S. R. (1998) Acid stress, anaerobiosis and *gadCB*: lessons from Lactococcus lactis and *Escherichia coli*. *Trends Microbiol* **6:** 214-6.

Smith, J. L. (2003) The role of gastric acid in preventing foodborne disease and how bacteria overcome acid conditions. *J Food Prot* **66:** 1292-303.

Soderback, E., Reyes-Ramirez, F., Eydmann, T., Austin, S., Hill, S. and Dixon, R. (1998) The redox- and fixed nitrogen-responsive regulatory protein NIFL from Azotobacter vinelandii comprises discrete flavin and nucleotide-binding domains. *Mol Microbiol* **28:** 179-92.

Stanek, M. T., Cooper, T. F. and Lenski, R. E. (2009) Identification and dynamics of a beneficial mutation in a long-term evolution experiment with *Escherichia coli*. *BMC Evol Biol* **9:** 302.

Stim, K. P. and Bennett, G. N. (1993) Nucleotide sequence of the adi gene, which encodes the biodegradative acid-induced arginine decarboxylase of *Escherichia coli*. *J Bacteriol* **175:** 1221-34.

Stock, A. M. and Mowbray, S. L. (1995) Bacterial chemotaxis: a field in motion. *Curr Opin Struct Biol* **5:** 744-51.

Studemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D. and Hengge, R. (2003) Sequential recognition of two distinct sites in sigma(S) by the proteolytic targeting factor RssB and ClpX. *EMBO J* **22:** 4111-20.

Tabor, H., Hafner, E. W. and Tabor, C. W. (1980) Construction of an *Escherichia coli* strain unable to synthesize putrescine, spermidine, or cadaverine: characterization of two genes controlling lysine decarboxylase. *J Bacteriol* **144:** 952-6.

Tapley, T. L., Franzmann, T. M., Chakraborty, S., Jakob, U. and Bardwell, J. C. (2010) Protein refolding by pH-triggered chaperone binding and release. *Proc Natl Acad Sci U S A* **107:** 1071-6.

Tapley, T. L., Korner, J. L., Barge, M. T., Hupfeld, J., Schauerte, J. A., Gafni, A., Jakob, U. and Bardwell, J. C. (2009) Structural plasticity of an acid-activated chaperone allows promiscuous substrate binding. *Proc Natl Acad Sci U S A* **106:** 5557-62.

Taylor, B. L. and Zhulin, I. B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63:** 479-506.

Temme, K., Salis, H., Tullman-Ercek, D., Levskaya, A., Hong, S. H. and Voigt, C. A. (2008) Induction and relaxation dynamics of the regulatory network controlling the type III secretion system encoded within Salmonella pathogenicity island 1. *J Mol Biol* **377:** 47-61.

Tramonti, A., De Canio, M. and De Biase, D. (2008) GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the gadY-gadW divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. *Mol Microbiol* **70:** 965-82.

Tramonti, A., De Canio, M., Delany, I., Scarlato, V. and De Biase, D. (2006) Mechanisms of transcription activation exerted by GadX and GadW at the *gadA* and *gadBC* gene promoters of the glutamate-based acid resistance system in *Escherichia coli*. *J Bacteriol* **188:** 8118-27.

Tramonti, A., Visca, P., De Canio, M., Falconi, M. and De Biase, D. (2002) Functional characterization and regulation of *gadX*, a gene encoding an AraC/XylS-like transcriptional activator of the *Escherichia coli* glutamic acid decarboxylase system. *J Bacteriol* **184:** 2603-13.

Tucker, D. L., Tucker, N. and Conway, T. (2002) Gene expression profiling of the pH response in *Escherichia coli*. *J Bacteriol* **184:** 6551-8.

Tucker, D. L., Tucker, N., Ma, Z., Foster, J. W., Miranda, R. L., Cohen, P. S. and Conway, T. (2003) Genes of the GadX-GadW regulon in *Escherichia coli*. *J Bacteriol* **185:** 3190-201.

Typas, A., Becker, G. and Hengge, R. (2007) The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Mol Microbiol* **63:** 1296-306.

Venkatesh, G. R., Kembou Koungni, F. C., Paukner, A., Stratmann, T., Blissenbach, B. and Schnetz, K. (2010) BglJ-RcsB heterodimers relieve repression of the *Escherichia coli* bgl operon by H-NS. *J Bacteriol* **192:** 6456-64.

Vescovi, E. G., Ayala, Y. M., Di Cera, E. and Groisman, E. A. (1997) Characterization of the bacterial sensor protein PhoQ. Evidence for distinct binding sites for Mg^{2+} and Ca^{2+}. *J Biol Chem* **272:** 1440-3.

Vogt, G., Etzold, T. and Argos, P. (1995) An assessment of amino acid exchange matrices in aligning protein sequences: the twilight zone revisited. *J Mol Biol* **249:** 816-31.

Waterman, S. R. and Small, P. L. (1996) Identification of sigma S-dependent genes associated with the stationary-phase acid-resistance phenotype of *Shigella flexneri*. *Mol Microbiol* **21:** 925-40.

Waterman, S. R. and Small, P. L. (2003a) Identification of the promoter regions and sigma(s)-dependent regulation of the *gadA* and *gadBC* genes associated with glutamate-dependent acid resistance in *Shigella flexneri*. *FEMS Microbiol Lett* **225:** 155-60.

Waterman, S. R. and Small, P. L. (2003b) Transcriptional expression of *Escherichia coli* glutamate-dependent acid resistance genes *gadA* and *gadBC* in an hns *rpoS* mutant. *J Bacteriol* **185:** 4644-7.

Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L. and Olson, E. R. (1992) Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J Bacteriol* **174:** 530-40.

Watts, K. J., Sommer, K., Fry, S. L., Johnson, M. S. and Taylor, B. L. (2006) Function of the N-terminal cap of the PAS domain in signaling by the aerotaxis receptor Aer. *J Bacteriol* **188:** 2154-62.

Wehland, M. and Bernhard, F. (2000) The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. *J Biol Chem* **275:** 7013-20.

Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E. and Adami, C. (2001) Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412:** 331-3.

Winter, S. E., Winter, M. G., Thiennimitr, P., Gerriets, V. A., Nuccio, S. P., Russmann, H. and Baumler, A. J. (2009) The TviA auxiliary protein renders the *Salmonella enterica* serotype Typhi RcsB regulon responsive to changes in osmolarity. *Mol Microbiol* **74:** 175-93.

Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. and Lenski, R. E. (2006) Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci U S A* **103:** 9107-12.

Woods, R. J., Barrick, J. E., Cooper, T. F., Shrestha, U., Kauth, M. R. and Lenski, R. E. (2011) Second-order selection for evolvability in a large *Escherichia coli* population. *Science* **331:** 1433-6.

Wu, Y. and Outten, F. W. (2009) IscR controls iron-dependent biofilm formation in *Escherichia coli* by regulating type I fimbria expression. *J Bacteriol* **191:** 1248-57.

Young, B. A., Gruber, T. M. and Gross, C. A. (2004) Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science* **303:** 1382-4.

Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M. G. and Alon, U. (2004) Just-in-time transcription program in metabolic pathways. *Nat Genet* **36:** 486-91.

Zwir, I., Shin, D., Kato, A., Nishino, K., Latifi, T., Solomon, F., Hare, J. M., Huang, H. and Groisman, E. A. (2005) Dissecting the PhoP regulatory network of *Escherichia coli* and Salmonella enterica. *Proc Natl Acad Sci U S A* **102:** 2862-7.