Extracytoplasmic Stress Responses Induced by a Model Secretin

A dissertation presented in partial fulfilment of the

requirements for the degree of

Doctor of Philosophy

in

Biochemistry

At Massey University, Manawatū,

New Zealand

Julian Spagnuolo

2015

Abstract

Pathogenic bacteria export large proteins and protein complexes, including virulence factors, using dedicated transenvelope multiprotein machinery, collectively called secretion systems. Four of these protein export machines found in Gram-negative bacteria, type 2/3 secretion systems, filamentous phage assembly-secretion system and the type 4 pilus assembly system contain large homologous gated channels, called secretins, in the outer membrane. Secretins are radially symmetrical homomultimers (luminal diameter 6-8 nm) interrupted by an internal septum or gate. Expression of these channels imposes a fitness cost to bacteria. While stress induced by model secretin pIV has been previously investigated using microarrays, this thesis is the first RNA-seq characterisation of secretin stress responses. Furthermore, this is the first comparison of stress imposed by a closed-gate secretin (wildtype pIV), vs. an isogenic leaky-gate variant, the latter serving as a model of an open-gate substrate-secreting channel. The high sensitivity to changes in gene expression and low background noise of the RNA-seq approach have greatly expanded the known secretin stress responses to include the SoxS, CpxR and RcsB/RcsAB regulons, in addition to the known involvement of the Psp response. A synthetic lethality analysis of candidate genes in these pathways suggested that the leaky-gate secretins, besides rendering the Psp response essential for survival, also stimulate the SoxS and RcsB/RcsAB regulons for protection of the cells. Knowledge of the secretin stress expanded by this work helped identify potential targets for development of much-needed antibiotics against toxin-secreting Gram-negative bacteria.

Foreword & Acknowledgements

Bristling with anticipation, excitement and unbridled passion for science, a young man, longboard in hand, walked into an office, the first of many meetings to come. The seasoned scientist sitting across the table rebuffed his notion to save the world with ‘super-phage’ – bacteria evolve resistance too quickly. So began my journey, the start of my Masters degree which, would later weave into the, at times, perilous tale of my Doctorate.

As the months lapsed into years my ears dried, the salt hardened and increasingly I was left to my own devices, given such a free reign that, in my haste and lust, I quickly became lost in a myriad of leads, methodology and gene-regulation. I was rather lucky to have a mentor to pull me back on the path. In hindsight, I don’t think it was easy to supervise me – Jasna, Murray, my vocabulary just isn’t sufficient to express how important your support was.

Standing on the precipice of completion, I wonder just who I have done this for. Was this journey started and completed in the name of some Muse? Or was it for my own greater glory? Having stood on stage in various capacities I can safely say I do not care for glory, only acknowledgement of quality and substance. The Muses, well they disappear and reappear like the mist on the wind, I wish I could say I doubt they were important. Alas, they are attributable for waking parts of my mind that drove my journey and allow me to do the things I can. I won’t name these people, but anytime I pick up a pipette or my pen I remember each and every one. To these people – this thesis is just as much yours as mine. Live long.

Abbreviations

ABC – ATP binding cassette

ATP – Adenosine triphosphate

BH – Benjamini-Hochberg multiple testing correction

BSA – Bovine serum albumin

cAMP – Cyclic adenosine monophosphate

cDNA – Complementary DNA

Cm – Chloramphenicol

CPS – Capsular polysaccharide

CV – Coefficient of variation

DE – Differential expression

DEPC – Diethyl pyrocarbonate

DNA – Deoxyribonucleic acid

EDTA – Ethylenediaminetetraacetic acid

ELISA – Enzyme linked immunosorbent assay

EM – Electron microscopy

EPS – Extracytoplasmic polysaccharide

FDR – False discovery rate

FFSS – Filamentous phage assembly/secretion system

G3P – Glucose-3-phosphate

Km - Kanamycin

IHF – Integration host factor

IM – Inner membrane

IPTG – Isopropyl β-D-1-thiogalactopyranoside

LPS – Lipopolysaccharide

NAD – Nicotinamide adenine dinucleotide

NADP – Nicotinamide adenine dinucleotide phosphate

OM – Outer membrane

OMP – Outer membrane protein

ONPG – Ortho-Nitrophenyl-β-galactoside

PCA – Principle component analysis

PEG – Polyethylene glycol

PGA – poly-β-1,6-N-acetylglucosamine

PMF – Proton motive force

POTRA – Polypeptide transport associated

PSP – Phage shock protein

$PulD_{PBS}$ – Pilotin binding domain of PulD

RCF – Relative centrifugal force

RIB – Reiterative *ihf* bacterial interspersed mosaic element

RIN – RNA integrity number

ROS – Reactive oxygen species

RNA-seq – RNA sequencing

RNA – Ribonucleic acid

rRNA – ribosomal RNA

SDS – Sodium dodecyl sulfate

SPA – Single particle analysis

sRNA – Small RNA

T2SS – Type 2 secretion system

T3SS – Type 3 secretion system

T4PS – Type 4 pilus assembly system

TCA – Tricarboxylic acid cycle

Tet – Tetracycline

UAS – Universal activation sequence

UDP – Uridine diphosphate

List of Tables

List of Figures

Table of Contents

Chapter 1

1 Gram-negative Membrane Biology

The cell wall of Gram-negative bacteria such as *Escherichia coli*, *Klebsiella pneumonia*, *Yersinia enterocolitica* and *Salmonella typhimurium* is composed of 3 elements; an asymmetrical outer-membrane (OM) composed of lipopolysaccharide (LPS) on the outer surface and phospholipid on the inner surface; a periplasm containing peptidoglycan, rigidifying the cell; and a symmetrical cytoplasmic, or inner-membrane (IM), composed of phospholipid (Figure 1).

At the most fundamental level, the cell wall serves to protect the vital cellular functions occurring within the cytoplasm by maintaining conditions favourable for DNA replication and cell division. Additionally, it buffers against environmental changes such as nutrient availability, pH and osmolarity. The dual membrane structure of Gram-negative's cell wall makes them extremely recalcitrant to environmental change and the noxious compounds they encounter within their respective habitats. As such, the cell wall (and the molecular machinery it contains) is the structure that senses and communicates environmental conditions to the bacterium's genetic machinery, which then mounts an appropriate response.

Underlying the differences in IM and OM structures are their disparate functions. Whereas the OM acts as a selective barrier to undesirable and often deleterious compounds, the IM is host to cellular machinery critical for survival of the bacterium in diverse environments. As such, the important differences between the IM and OM are not only in their lipid composition, but also in the

proteins they host. Arising from inner/outer membrane asymmetry is the problem of protein targeting to a specific membrane, a functionally important aspect of bacterial physiology. Localisations of some IM and OM proteins have been experimentally verified, although prediction algorithms can make informed guesses where they may be located based upon the presence/absence of specific targeting motifs, such as, signal peptides, lipoboxes and transmembrane domains. Among the 4000, or so, protein coding genes in *E. coli* K-12 only 981 have had their cellular location experimentally verified (Daley *et al.*, 2005). Predictions have placed 46.1% of *E*. *coli* K-12 MG1655 proteins in the cytoplasm, 25.9% in the IM and just 1.9% and 3.8% in the OM and periplasm, respectively, while 1.1% are predicted to be extracellular and 21.2% of protein's still have no known localisation (Rey *et al.*, 2005, Yu *et al.*, 2011).

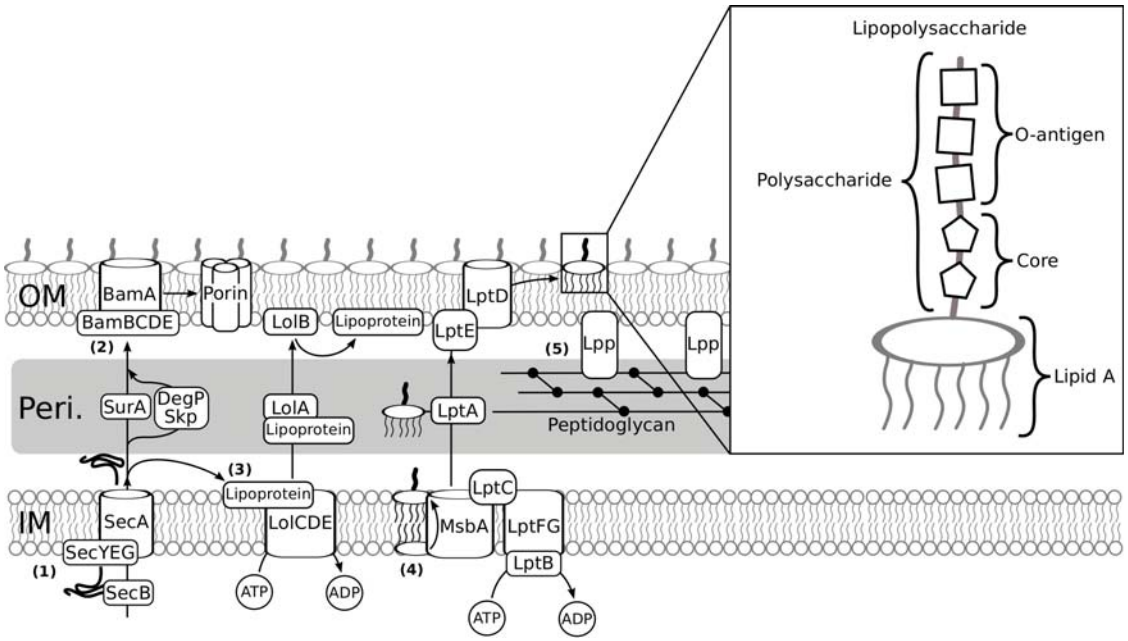

Figure 1: Structure of the Gram-negative cell wall and biogenesis machinery.

(1) Nascent polypeptides with an N-terminal signal sequence are recognised by SecB, which targets the polypeptide to the SecYEGA transport complex for export to the periplasm (Peri.). **(2)** OM beta-barrel proteins are trafficked by SurA to the BamABCDE complex for assembly into the OM. **(3)** Lipoproteins are first associated with the IM and are shuttled to the periplasmic face of the OM by the Lol system. **(4)** Lipid A-core is first synthesised at the cytoplasmic face of the IM and the O-antigen is ligated to the core polysaccharide after MsbA transitions the lipid A-core molecule across the IM. Once in the periplasm, LPS is trafficked and assembled into the OM by the Lpt machinery. **(5)** Cross-linked peptidoglycan stretches around the cell within the periplasm, in parallel with the OM, to which it is pinned by the Lpp, lipoprotein. The LPS (inset) comprises 3 main components: hydrophobic Lipid A, Core polysaccharide, and variable O−antigen polysaccharides.

1.1 The Outer Membrane

The OM bilayer is asymmetrical – the lipid composition of it's two layers is different. The outer-leaflet is composed, predominantly of a tripartite molecule, lipopolysaccharide (LPS) – containing the following components: a lipid A anchor, the endotoxin responsible for eliciting a strong immune response; core oligosaccharide (containing heptose, glucose and galactose) and the variable O-antigen (Figure 1, inset), a repetitive glycan polymer central to the barrier function of the OM (Nikaido, 2003). Conversely, the inner leaflet of the OM is composed of phospholipids, which can cause the OM to become leaky to hydrophobic molecules (including detergents) if they are forced into the outer-leaflet by environmental stress or disruption of the OM biogenesis machinery (Nikaido, 2003, Nikaido, 2005). However, the periplasm and OM contain machinery to ameliorate the effects of phospholipids in the outer leaflet (Malinverni & Silhavy, 2009).

1.1.1 Function of Outer Membrane Proteins

Outer membrane proteins (OMPs) fall into 3 main classes: channels, lipoproteins, and components of surface organelles. OM channels are the transporters of the Gram-negative cell wall, responsible for the selective nature of the OM barrier. The smaller channels, known as porins, allow entry of nutrients up to 100 Daltons but prevent diffusion of larger molecules. General porins, such as OmpC and OmpF, are non-specific transporters (Benson *et al.*, 1988, Cowan *et al.*, 1992, Misra & Benson, 1988b). Some specialised porin homologues, like LamB, are transporters of specific molecules and may act as

receptors for phage infection (Freundlieb *et al.*, 1988, Randall-Hazelbauer & Schwartz, 1973).

Larger channels, such as the secretins of the multicomponent protein-secretion machinery - the Type 2 and 3 secretion systems - are vital to the transport of larger complexes and contain structural elements preventing unregulated passage of solutes or proteins (Burghout *et al.*, 2004, Frye *et al.*, 2006, Marciano *et al.*, 2001, Nouwen *et al.*, 1999, Russel & Kazmierczak, 1993). Other large OM complexes include complexes that assemble various pili, as well as the flagellar motility machinery, responsible for surface attachment and cellular movement, respectively (Berg, 2003, Craig *et al.*, 2003, Craig *et al.*, 2006). Other OM channels have functions associated with the biogenesis of the OM, that is, assembling other integral proteins into the OM (the Bam complex, Figure 1, (Wu *et al.*, 2005)) and maintaining the asymmetry of the OM bilayer (PagP, PldA; Malinverni and Silhavy (2009)).

Integral OM proteins are predominantly barrels built from transmembrane β−strands (β−barrels; (Koebnik *et al.*, 2000)). They are either multimers – such as the trimeric multidrug exporter TolC (Koronakis *et al.*, 1997) and the OmpF and OmpC porins of *E. coli*. β−barrels can also be monomeric channels such as FhuA or BtuB (Baslé *et al.*, 2006, Weiss & Schulz, 1992). However, alpha-barrel structures have also been identified, as in the *E. coli* capsular polysaccharide translocon, lipoprotein Wza (Dong *et al.*, 2006).

Lipoproteins are transported across the periplasm by the lipoprotein sorting system encoded by *lolABCDE* and are anchored into the inner leaflet of the OM via the covalently attached lipid moieties through an obligatory N-terminal cysteine residue (Figure 1). Lipoproteins have multiple roles including assembly and localisation of proteins into the OM (Koo *et al.*, 2012); mitigation of stress (Onufryk *et al.*, 2005); and signal transduction (Castanie-Cornet *et al.*, 2006). The presence of a conserved lipobox sequence around the signal-sequence cleavage site signals that the mature polypeptide is a lipoprotein (Hayashi & Wu, 1990). The N-terminal cysteine residue (of the mature lipoprotein) is lipidated in a three-step reaction by Lgt, LspA and Lnt. This reaction is essential for anchoring the lipoproteins in the membranes (Gupta & Wu, 1991, Sankaran & Wu, 1994, Tokunaga *et al.*, 1984a, Tokunaga *et al.*, 1984b). Retention in the periplasmic face of the IM is dependent on the presence of an aspartate residue, i.e. a negative charge, at position 2 of the mature polypeptide (Yamaguchi *et al.*, 1988). Positively charged residues at position 3 can suppress IM retention of lipoproteins (Gennity & Inouye, 1991). It is thought that the electrostatic complementarity between the negatively charged N-terminal Asp-residue of IM lipoproteins and phosphatidylethanolamine in the IM is strong enough to prevent interaction with the IM LolCDE complex that transfers lipoproteins to the OM and therefore, constitutes a Lol-avoidance signal (Hara *et al.*, 2003, Tokuda & Matsuyama, 2004). LolCDE is a non-canonical ABC-transporter. In contrast to canonical ABC-transporters, LolCDE does not participate in the transport of molecules across the membrane. Typically, ABC-transporters do participate in this process. LolC and LolE span the membrane, while LolD contains an ATP-binding cassette (Yakushi *et al.*, 2000). LolCDE

recognise lipoproteins that do not contain the avoidance signal at position 2, and links them to LolA in an ATP-dependent manner. LolA shuttles lipoproteins across the periplasm to LolB, which releases them into the periplasmic leaflet of the OM (Yokota *et al.*, 1999).

1.1.2 Biogenesis of the Outer Membrane

1.1.2.1 Synthesis of Lipopolysaccharides

The synthesis of LPS begins at the cytoplasmic leaflet of the IM, utilizing a constitutively expressed pathway of nine enzymes. LpxACDHBK catalyse the 6-step production of Lipid IV-A by the fatty acylation of two UDP-N-acetylglucosamine molecules (Babinski *et al.*, 2002, Crowell *et al.*, 1986, Galloway & Raetz, 1990, Garrett *et al.*, 1997, Kelly *et al.*, 1993, Ray & Raetz, 1987, Young *et al.*, 1995). WaaA and LpxLM attach two CMP-Kdo residues and perform an acylation reaction to form hexaacylated Kdo-Lipid A (Belunis & Raetz, 1992, Clementz *et al.*, 1997, Six *et al.*, 2008). Three glycosyl transferases, WaaPYQ, attach additional sugars to Lipid A, to complete the core oligosaccharide – the lipid A-core molecule is referred to as LPS (Yethon *et al.*, 1998). An ABC-transporter, MsbA, then flips the Lipid A-core molecule to the periplasmic face of the IM (Doerrler *et al.*, 2004). The O-antigen repeats are separately synthesised in the cytoplasm by enzymes encoded by the *rfb* gene cluster, carried across the IM by the undecaprenyl diphosphate-carrier molecule and ligated onto the core oligosaccharide by WaaL (Klena *et al.*, 1992). *E. coli* K-12 strains do not produce O-antigen to attach to the LPS owing to defects in the *rfb* gene cluster. In addition, *E. coli* K-12 may contain two pathways for LPS synthesis: one creates core oligosaccharides suitable for O-antigen attachment;

the other produces lipooligosaccharides with altered core structures that prevent O-antigen attachment (Klena *et al.*, 1992, Stevenson *et al.*, 1996).

1.1.2.2 Shuttling Lipopolysaccharide to the OM

The Lpt (*L*ipo*p*olysaccharide *t*ransport) machinery in the IM, periplasm and OM (Figure 1) shuttles the mature LPS molecule from the periplasmic surface of the IM and assembles it into the outer leaflet of the OM. All components of the Lpt machinery are essential in Gram-negative bacteria (The exception is *Neisseria meningitidis*, which can survive without an OM). In all other Gram-negative bacteria, depletion of any Lpt-component causes accumulation of LPS within the periplasm, aberrant WaaL-mediated modification of the LPS core-oligosaccharide with colanic acid, increased OM permeability, and cessation of growth. It should be noted that in all cases, depletion of Lpt components does not affect OMP assembly or trafficking (Ruiz *et al.*, 2008, Sperandeo *et al.*, 2007, Sperandeo *et al.*, 2008, Wu *et al.*, 2006). However, defects in LPS transport or synthesis result in the movement of phospholipid from the inner leaflet to the outer leaflet of the OM – this activates expression of two proteins, PagP and PldA which eliminate phospholipids in the outer leaflet and converts the phospholipid to lyso-phosopholipid, respectively (Malinverni & Silhavy, 2009) – these will be described in greater detail in section 1.3.1.

The Lpt complex responsible for shuttling LPS intermediates to the OM is composed of seven proteins, LptBFG-LptC-LptA-LptDE. The IM components, LptBFG form an ABC-transporter, with LptFG forming the transmembrane component and the IM-associated cytoplasmic protein, LptB, providing the ATP-

binging cassette to drive the transport of LPS embedded in the periplasmic face of the IM, after it has been flipped across the IM by MsbA (Ruiz *et al.*, 2008, Sperandeo *et al.*, 2007).

To span the periplasm and bridge the IM and OM Lpt complexes, two proteins are required, LptC and LptA. The N-terminal domain of the former, LptC, is anchored into the IM while its C-terminal domain interacts with the second periplasmic bridging protein, LptA. It is, as yet, unknown whether LptC plays a deeper role in LPS transport (Narita & Tokuda, 2009). Completing the periplasmic bridge, the C-terminal domain of LptA interacts with the N-terminal domain of the OM β-barrel protein, LptD. Additionally, LptA may play a sensor role in LPS transport due to the observation that LPS transport/assembly defects result in the degradation of LptA, effectively breaking the periplasmic bridge (Freinkman *et al.*, 2012, Sperandeo *et al.*, 2011). The final OM components, OM β-barrel, LptD and the OM lipoprotein, LptE are responsible for flipping LPS from the inner leaflet to the outer leaflet of the OM (Wu *et al.*, 2006).

1.1.2.3 Targeting and assembly of outer membrane proteins

A number of OMPs, like the porins LamB, OmpF and OmpC, require the β-barrel assembly machinery, Bam (Figure 1). This is an OM protein complex composed of an essential OM β-barrel, BamA, and four semi-redundant lipoproteins BamBCDE (Jain & Goldberg, 2007, Wu *et al.*, 2005). Structural analysis of BamA revealed the presence of five N-terminal polypeptide transport-associated (POTRA) domains (POTRA1-5) conserved in Gram-

negative bacteria, mitochondria and chloroplasts. These domains project into the periplasmic space and act as scaffolds with which the BamA-associated lipoproteins, BamBCDE, interact. Of the five domains, only POTRA5 is essential for the interaction of all associated lipoproteins, although the folding catalyst, BamB, requires POTRA2-5. Domain deletions of BamA were expressed in *E. coli* to investigate the requirement of domains 1 to 4 for the assembly of LamB, a trimeric porin-family OM β-barrel protein. It was shown that domains 3 and 4 are essential for the integration and assembly of LamB into the OM, whereas domains 1 and 2 were only partially required, since BamA lacking these domains was still able to assemble LamB in the OM albeit at reduced levels (Kim *et al.*, 2007). Other investigators observed the recognition, by the POTRA domains, of a C-terminal 10 amino acid signature sequence identified in most OM proteins (Robert *et al.*, 2006). This signature sequence consisted of a C-terminal aromatic residue, usually phenylalanine or tryptophan, and hydrophobic residues at positions 3, 5, 7, and 9 from the C-terminus. After export into the periplasm *via* the Sec or Tat pathways (For reviews see: Economou, 2002, Palmer & Berks, 2012), mature polypeptides are guided by a periplasmic chaperone, SurA, to the periplasmic POTRA domains of BamA (Bennion *et al.*, 2010). Two other periplasmic chaperones, Skp and DegP, are responsible for processing aggregated or off-pathway proteins by either degrading them, or channel them to SurA for refolding and eventual targeting to the Bam complex (Rizzitello *et al.*, 2001). Depletion of any component of the Bam complex leads to the reduction in the number of assembled OMPs in the OM (Doerrler & Raetz, 2005). The mechanism of OMP folding and membrane insertion performed by Bam is still unclear.

1.2 The Periplasm

Lying between the IM and OM is an aqueous layer known as the periplasm. This space contains a thin peptidoglycan layer – a polymer of N-acetyl glucosamine-N-acetyl muramic acid – pinned to the OM by the highly abundant lipoprotein, Lpp (Braun & Wolff, 1970, Schleifer & Kandler, 1972). This structure rigidifies the cell wall, protecting against mechanical shear forces and preventing the internal turgor pressure from exploding the cell. The importance of the peptidoglycan layer is exemplified by the action of beta-lactams that block peptidoglycan synthesis, resulting in production of OM vesicles (blebbing), loss of cell shape and protoplast formation (Lederberg, 1956).

Aside from being home to the peptidoglycan layer, the periplasm is host to a large number of proteins with a diverse range of functions. It is here that the seemingly schizophrenic stressful stimuli and environmental signals are interpreted, integrated and passed on to the respective signal transduction pathways embedded in the IM, ultimately regulating the activity of transcription factors. Additionally, the periplasm contains a range of molecular chaperones such as Skp, DegP and SurA that are responsible for protein quality within this compartment together with disulphide bridge formation and isomerization pathway proteins, DsbAB and DsbCD (Missiakas *et al.*, 1996, Sklar *et al.*, 2007).

The molecular chaperone, SurA, when overexpressed, was observed to suppress the induction of the sigma-E response and prevent the periplasmic

accumulation of misfolded proteins (Missiakas *et al.*, 1996). Furthermore, it was found to compensate for the depletion of any DsbABCD protein, which normally results in periplasmic and OM protein misfolding. Cross-linking experiments indicated that SurA interacted directly with BamA, and its up-regulation in Δ*degP*/Δ*skp* double mutants indicated these two proteases have a role in shepherding off-pathway and misfolded proteins back to Bam (Jain & Goldberg, 2007, Sklar *et al.*, 2007).

In addition to chaperones, binding proteins (eg. MalE – maltose binding protein) are present to shuttle compounds that cross the OM (sugars, iron-complexes) from their respective OM channels to IM transporters that ferry them into the cytoplasm (Felder *et al.*, 1999, Tam & Saier Jr, 1993). The biosynthetic machinery required for the formation of peptidoglycan is also found within the periplasm. As many of these processes are susceptible to disruption by toxic compounds, including antibiotics (β-lactams), detoxifying enzymes (e.g. β-lactamases) are also present in the periplasm.

1.3 The Inner Membrane

1.3.1 Phospholipid Synthesis

The phospholipids present in both leaflets of the IM and the inner leaflet of the OM are synthesised by an integral IM glycerol-3-phosphate acyltransferase, PlsB, that catalyses the first committed step in the production of phospholipid (Kessels *et al.*, 1983). The phospholipid backbone and substrate of PlsB, glycerol-3-phosphate (G3P), is derived either as a product of glycolysis (Cronan & Bell, 1974, Edgar & Bell, 1979), or obtained from the environment in the form

of glycerol imported by GlpF and converted into G3P by GlpK (Lin, 1976, Voegele *et al.*, 1993, Weissenborn *et al.*, 1992). The type 2 fatty acid cycle provides the long-chain acyl groups tethered to a carrier protein, acyl-ACP (Rock & Jackowski, 2002, Zhang & Rock, 2008). The acyl-ACP molecule is, in turn, passed onto integral IM acyltransferases of the Pls pathway (**P**hospho**l**ipid **s**ynthesis), leading to the formation of the phospholipid molecule and integration into the IM. A flippase is required to 'flip' phospholipids from the cytoplasmic leaflet of the IM to the periplasmic leaflet. It is thought this step is mediated by the potassium transporter, KcsA, and the leader peptidase, the protein that normally removes the signal sequence from exported and/or membrane proteins (Kol *et al.*, 2003). From there, the process by which phospholipids move from the IM to the inner leaflet of the OM is unclear. What is known, is that at least two integral OM enzymes (PagP, PldA) and one pathway (Mla – **m**aintenance of **l**ipid **a**symmetry) exist in *E. coli* that maintains the asymmetrical nature of the OM (Malinverni & Silhavy, 2009). An enzyme, PagP, whose expression is regulated by the transcription factor, PhoP, eliminates phospholipid in the outer leaflet by flipping them back into the inner leaflet, while PldA converts the phospholipid to lyso-phosopholipid by cleaving one palmitate moiety and transferring it to Lipid A, creating hepta-acylated LPS. While this still leaves lyso-phospholipid in the OM, creating a steric disturbance in the OM, the production of hepta-acylated LPS is thought to improve OM stability.

Phospholipids are also transported retroactively away from the OM to be reinserted into the IM by the dedicated Mla pathway comprised of a retrograde

IM ABC-transporter and periplasmic binding proteins (Malinverni & Silhavy, 2009).

1.3.2 Inner Membrane Proteins

The IM is a site of much activity. It is the location of the cytochrome complexes of the electron-transport chain and oxidative phosphorylation. The IM is also the starting point for secretion of large protein complexes from the cytoplasm to the external milieu. As described in previous sections the IM is also the site of synthesis of phospholipids and LPS. In addition, synthesis of extracytoplasmic polysaccharides (EPS) and capsular polysaccharides (CPS) also begins at the IM. Functional annotation clustering of the 25.9% of all *E*. *coli* MG1655 proteins predicted to be IM localised indicates that these proteins are predominantly involved in 2-component signal transduction, transport, or respiration (anaerobic or aerobic), with several basal complexes of bacterial secretion systems embedded into the IM (Huang da *et al.*, 2009a, Huang da *et al.*, 2009b).

Another critical function of the IM is to integrate environmental signals and convey them to the transcriptional regulation machinery within the cytoplasm. Two main strategies exist to accomplish this task: (1) 2-component phosphorelay systems, in which activity of a cytoplasmic response regulator is dependent on its phosphorylation by an integral IM sensor kinase; (2) Proteolytic cascades, where the activities of specific regulatory proteases is modulated by extracellular or periplasmic stimuli and this activity controls the turnover of IM-associated response regulators, or the cleavage of inhibitory factors/domains. Both of these strategies are central to the processing,

interpretation and response to stressful environmental stimuli, and will be discussed later.

1.4 Type 2 and 3 secretion systems

Secretion of folded proteins is a keystone to many bacterial lifestyles. However, the double-membrane envelope of Gram-negative bacteria poses a significant barrier to the secretion of such substrates. Evolution has provided these bacteria with several related secretion systems or pathways for the secretion of proteins and macromolecular complexes through both the IM and OM while minimising the disturbances to the cell envelope (Figure 2, Figure 4). There are more than half-a-dozen such secretion systems that differ in complexity, mode of function, and substrates they secrete. Some secretion systems assemble/secrete filamentous phage, exchange genetic material with other organisms or assemble specialised surface structures (Gerlach & Hensel, 2007, Hueck, 1998, Russel, 1998).

Among these secretion systems four have a homologous large OM channel, called secretin: type 2 secretion system (T2SS), type 3 secretion system (T3SS), type IV pilus assembly system (T4PS) and filamentous phage assembly/secretion system (FFSS). T3SS is essential to virulence processes in pathogenic bacteria such as *Salmonella enterica* Typhimurium, where perturbing the secretion of effector proteins abolishes the ability to invade host cells and as a result prevents cytotoxicity and dramatically reduces persistence in cell culture and murine models (Miki *et al.*, 2004). T3SSs are also required to

inhibit activation of phagocytes, thereby preventing destruction of *Salmonella* within phagosomes and lysosomes (Galan & Collmer, 1999). A T2SS secretes pentameric cholera toxin protein, the key virulence factor of *Vibrio cholerae* (Davis *et al.*, 2000). The same system is also required for the secretion of lipase (Sikora *et al.*, 2007). T2/3SS share similarities with the FFSS. Indeed, the cholera toxin encoding phage CTXϕ of *V. cholerae* utilises the secretin of the host cell T2SS for its assembly and release (Davis *et al.*, 2000). There is also some similarity between Type 4 pili assembly (T4PS) and phage assembly, in that the major structural subunits of both pilus and phage are integrated in the IM, prior to the assembly into a filamentous structure *via* hydrophobic interactions of the transmembrane helices of the major subunits (Bayan *et al.*, 2006, Craig & Li, 2008).

The assembly machinery of filamentous phage can be thought of as a simplified version of T2/3SS, requiring fewer accessory proteins but utilising homologous and structurally similar gated OM-channels of the secretin family (Figure 2).

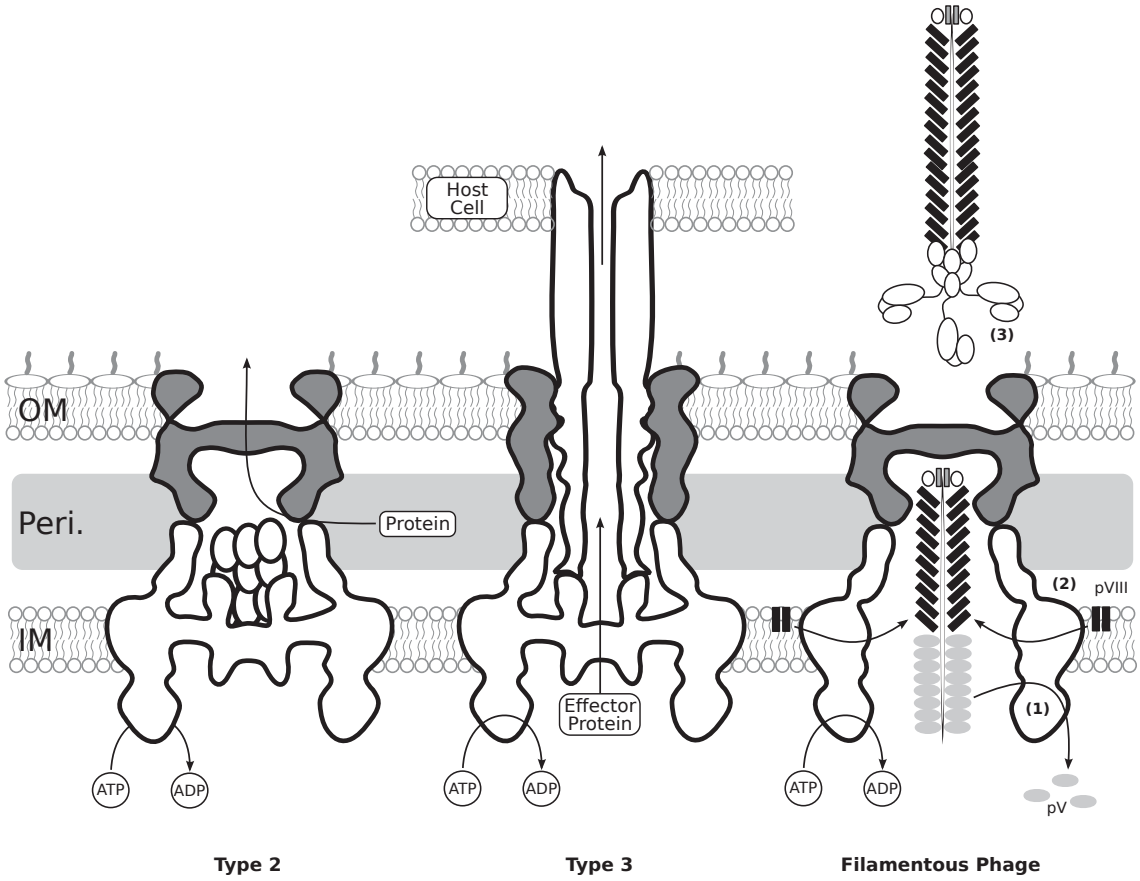

Figure 2: Overview of Type 2 and 3 Secretion Systems in contrast with the filamentous phage assembly system

IM components vary between systems and species. However, some generalities exist: export is driven by ATP hydrolysis; and each utilise a homologous OM component - the secretin channel (dark grey). **Type 2** secretion substrates enter the secretion apparatus in the periplasm for export to the extracellular melieu. **Type 3** secretion substrates enter a molecular needle threaded through the centre of the secretion apparatus and are injected directly into target cells. The assembly of **Filamentous phage** initiated when the packaging signal targets the ssDNA phage genome (in complex with pV) to the IM pI/pXI complex, which strips pV away **(1)**. Major coat protein pVIII is processively assembled onto the growing phage filament **(2)** as it is transported through an open pIV secretin channel (not shown). When assembly is terminated by the addition of pIII, the phage is released into the environment **(3)**. Source of silhouettes: (Marlovits *et al.*, 2004, Opalka *et al.*, 2003).

1.4.1 pIV – a model secretin

The OM component of the filamentous phage assembly machine, pIV, is a member of the secretin family of homologous proteins, which are large multimeric gated channels, embedded in the OM of Gram-negative bacteria (Kazmierczak *et al.*, 1994, Linderoth *et al.*, 1997). Cryo-EM and gold-labelling experiments have revealed that the pIV multimer has 14-fold radial symmetry (Opalka *et al.*, 2003) and is composed of 14 identical subunits encoded by the genome of filamentous bacteriophage f1.

Cryo-electron microscopy (EM) and single particle analysis (SPA) of pIV has revealed the three-tiered ring structure. pIV has large internal chambers at the channel entrance and exit, ranging from 6.0 nm in diameter at the N-ring to 8.8 nm in diameter at the C-ring (Opalka *et al.*, 2003) – large enough to allow passage of bulky proteins and antibiotics, which cannot pass through the LPS leaflet of the OM. However, *E. coli* expressing pIV is not sensitive to these compounds. The impermeability of pIV to these compounds is attributable to a septum that blocks the lumen of the channel at the level of the M-ring (Figure 3). Structures of the T3SS secretins InvG and YscC of *Salmonella typhimurium* (Marlovits *et al.*, 2004, Oliver & Marlovits, 2011) and *Yersinia enterocolitica* (Burghout *et al.*, 2004), respectively, and the T2SS secretin PulD of *Klebsiella oxytoca* (Chami *et al.*, 2005, Nouwen *et al.*, 1999, Tosi *et al.*, 2014) have also been obtained by Cryo-EM and SPA revealing very similar structures (Figure 3).

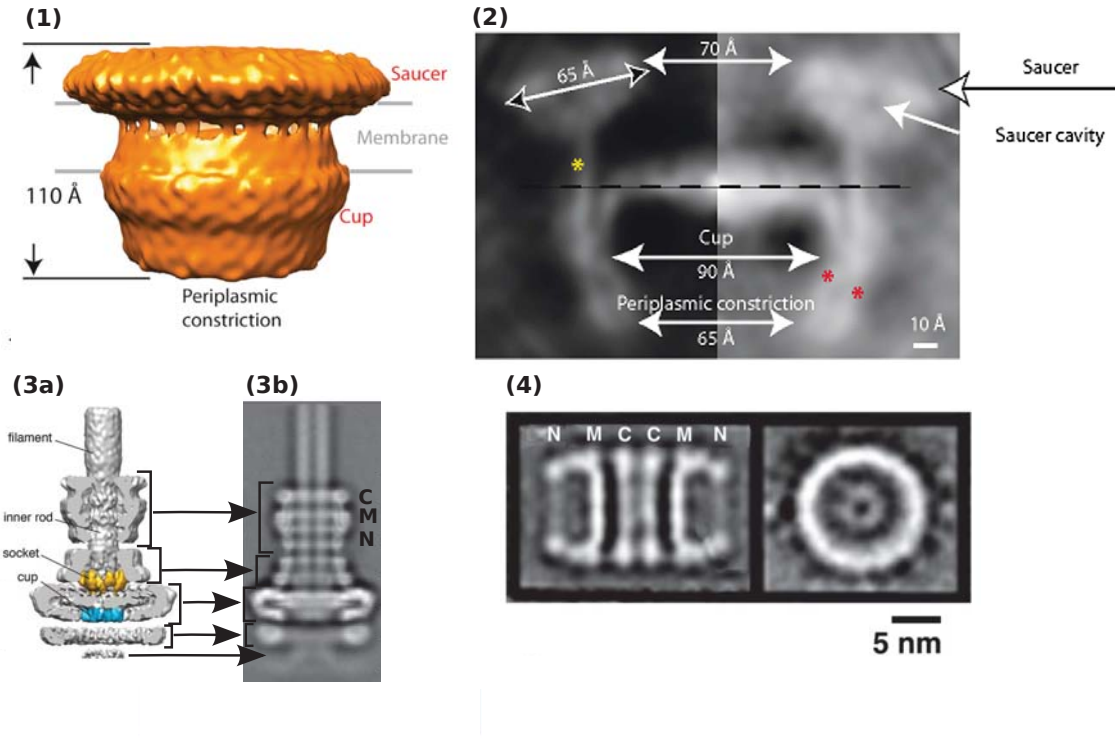

Figure 3: Cryo-EM and 3D reconstructions of PulD, InvG and pIV.

(1) Side view of *Klebsiella oxytoca* secretin PulD 3D reconstruction viewed at 1.3 sigma. **(2)** Sagittal density slice taken from 3D reconstruction using SPIDER, contrast enhanced on right-most section to aid visualisation. Red asterisk indicates the two 10 Å walls that form the cup. A yellow asterisk indicates the arm linking the cup to the saucer (Tosi *et al.*, 2014). **(3a)** 3D reconstruction of the Inv needle complex, including basal body, from *Salmonella* Typhimurium. Arrows indicate the equivalent structure within the class average Cryo-EM image **(3b)** (Oliver & Marlovits, 2011). **(4)** Single particle reconstruction of pIV, side (left) and top (right) view. N, the periplasmic N-terminal domains; M, middle ring forming the septum; C, the C-terminal and membrane embedded ring. Equivalent structures have been labelled similarly within the InvG single particle reconstruction **(3b)**. When purified, pIV multimers form dimers that interact via their C rings (Opalka *et al.*, 2003).

Sequence alignments of secretins have identified a conserved C-terminal domain (Figure 4), which spans the outer membrane and forms the middle ring, including the septum (or gate) and the cup. The N-terminal domain is less well conserved and must be disordered, as it does not show in the cryo-EM and SPA analysis of T2SS and T3SS secretins. Domain swapping experiments between pIV homologs from two related filamentous phages f1 and Ike have suggested that this region may contain substrate specificity determinants (Daefler *et al.*, 1997a, Daefler *et al.*, 1997b, Linderoth *et al.*, 1997). 3D SPA revealed an interesting structural feature of pIV common to other secretins; the lack of electron dense regions between the C and M rings, punctuated by connecting regions (Figure 3; (Opalka *et al.*, 2003, Tosi *et al.*, 2014). Density mapping of contoured longitudinal sections of the InvG and PulD secretin reveal low-density regions correlating to the C-M connecting regions seen in pIV 3D reconstructions (Marlovits *et al.*, 2004, Tosi *et al.*, 2014). Importantly, differences in the shape and density of InvG secretin rings can be seen by comparing Cryo-EM images in the presence and absence of the cognate substrate, the needle complex. When the needle complex is assembled the periplasmic N ring, and OM embedded C ring expands in diameter to accommodate the needle. The septum disappears completely, suggesting a complex gate opening/closing mechanism (Marlovits *et al.*, 2004).

In addition to the septum revealed by SPA (Figure 3), it was also found in electrophysiology experiments that the application of high voltage (minimum of 200 mV) was required to open the pIV channel (Marciano *et al.*, 2001). This was

later corroborated by similar electrophysiology experiments with wildtype Yersinia secretin, YscC (Burghout *et al.*, 2004). Together, these structural and physiological findings suggest the presence of a gate structure. Electrophysiology experiments of purified leaky mutant pIV-S324G (Marciano *et al.*, 1999) revealed a markedly-decreased voltage threshold required for opening the channel (80 mV), suggesting a lower resistance to electrical force and/or differences in protein conformation.

Earlier attempts to understand the pIV gate have created leaky-gate mutants defined as permissivity to maltopentaose in a Δ*lamB* background, or sensitivity to bile salts (deoxycholate) and/or the antibiotics vancomycin and bacitracin. These agents do not normally affect Gram-negative bacteria unless the integrity of the OM has been perturbed. Antibiotic and detergent sensitivity tests have been historically used to characterise leaky-mutations in OM porins (Benson & Decloux, 1985, Benson *et al.*, 1988, Misra & Benson, 1988a, Misra & Benson, 1988b, Wandersman *et al.*, 1979). Mutation of the highly conserved pIV residue G355 resulted in a leaky gate (Russel, 1994). This residue lies in a highly conserved segment (PFAM family alignment consensus TLVLGGLT). More recently, random mutagenesis of pIV with selection for maltopentaose permissivity in a Δ*lamB* background identified two regions, called GATE1 and GATE2, within the C-terminal homology domain that have a role in the integrity of the secretin gate. Substitution mutation of a number of residues within GATE1 or GATE2 conferred a leaky gate phenotype – mucoidy and sensitivity to bile salts, vancomycin and bacitracin. Four other previously identified leaky mutants (N295S, S324G, A329T, N335I/S, and G367S) were mapped to the

GATE1 and GATE2 regions (Spagnuolo, 2010, Spagnuolo *et al.*, 2010) and one (G367S) downstream of the highly conserved TLVLGGLT motif. Another group of mutations characterised in pIV affect stability of the multimer and the turnover of pIV in the periplasm, indicating folding and/or multimerisation defects. Several such mutants were found interspersed with the leaky mutants in the primary sequence of pIV (Linderoth *et al.*, 1996, Linderoth *et al.*, 1997, Marciano *et al.*, 1999, Marciano *et al.*, 2001, Russel, 1994, Spagnuolo *et al.*, 2010). Leaky mutants variably affect the function of pIV in filamentous phage assembly, whereas destabilizing mutants all render pIV inactive because they are unable to form channels in the OM. Similar mutagenesis experiments with the pIV homolog, PulD, have indicated GATE1 and GATE2 exist in other secretins, located in similar regions of the C-terminal homology domain (Whitaker, 2012).

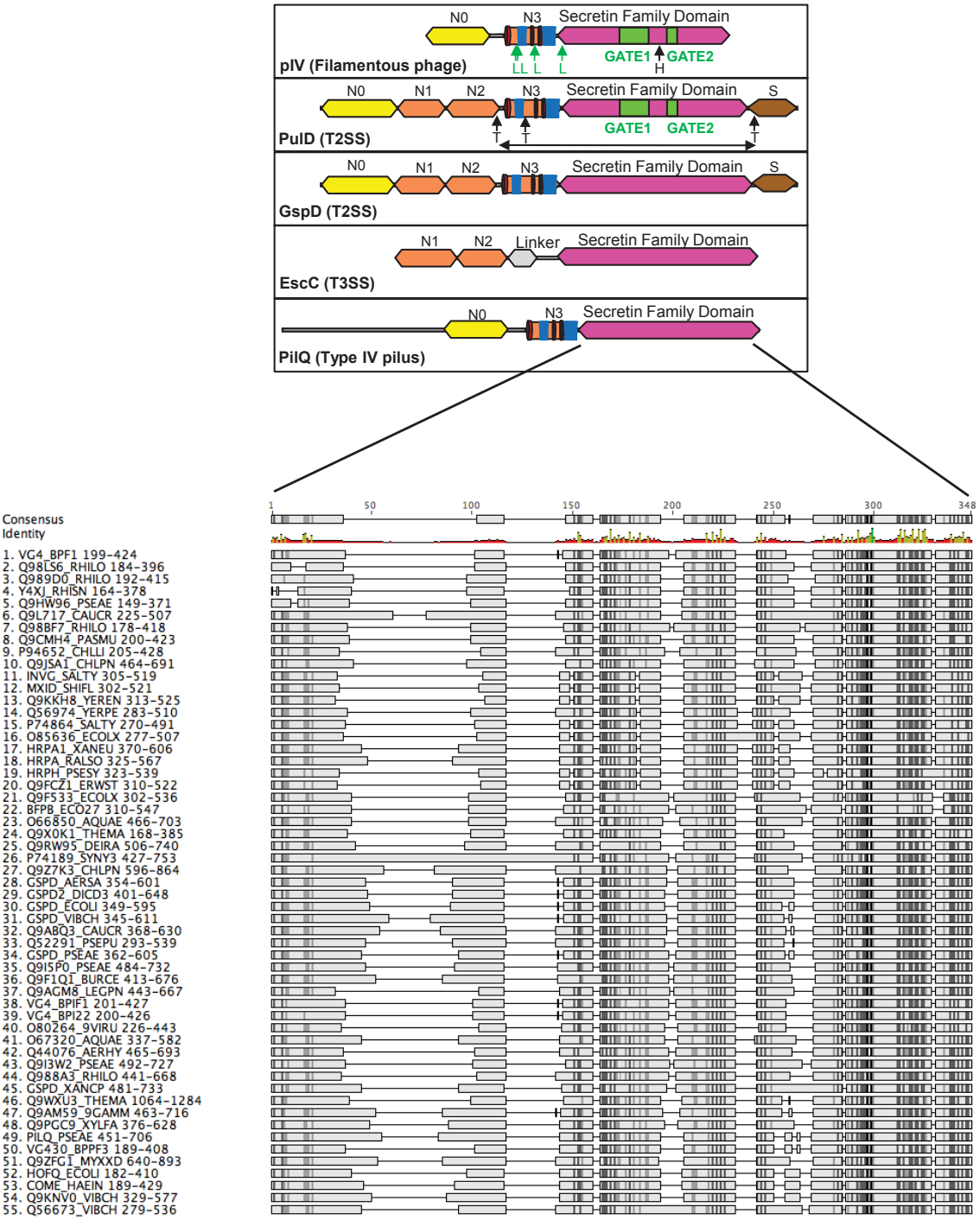

Figure 4: Domain architecture of 4 secretins and the secretin homology domain PFAM seed alignment.

(TOP) pIV domain organisation compared to that of PulD (T2SS of *K. oxytoca*; (Chami *et al.*, 2005)), GspD (EPEC T2SS; (Korotkov *et al.*, 2009)), and EscC (EPEC T3SS; (Spreter *et al.*, 2009)). Yellow diamond, N0 or FpvA-like domain; orange diamond, N1, N2, and N3, KH-fold domains; grey diamond, T3SS-specific N-terminal subdomain; pink diamond, secretin family domain, brown diamond, pilotin-binding domain; H, histidine-tag, T, trypsin cut sites; Linker, variable linker. Green boxes labelled GATE1 and GATE2 within the pIV and PulD secretin homology domain indicate the two regions responsible for gate integrity, Green arrows labelled L indicate point mutations with the pIV N-terminal that also confer leaky gate phenotypes (Spagnuolo, 2010, Spagnuolo *et al.*, 2010). GATE1 and GATE2 regions have been identified in equivalent locations within PulD ((Whitaker, 2012)). Double-headed arrow under PulD schematic indicated the trypsin resistant core of the multimeric channel. Shapes in the $N3_{pIV}$ and $N3_{PulD/GspD}$ schematics: secondary structures modelled based on structures of $N1_{GspD}$ (Korotkov *et al.*, 2009) and $N1/N2_{EscC}$ (Spreter *et al.*, 2009): red ovals, beta-strands; blue shapes, alpha-helices. **(BOTTOM)** PFAM seed alignment of the secretin family domain pfam number PF00263). Filamentous phage f1 secretin, pIV, is listed at the top (VG4_BPF1 199-424), showing sequence homology of residues 199 – 424 of the nascent propeptide. Boxes indicate regions of homology; darker shading indicates a higher level of conservation. Lines indicate breaks in homology (Kearse *et al.*, 2012, Finn *et al.*, 2014).

1.4.2 The secretin C-terminal homology domain

Secretins are grouped as a family based on a large C-terminal "secretin" homology domain (Figure 4). Several research groups have conducted truncation and protease digestion experiments to structurally and functionally characterise this region. Circular dichroism spectra of the trypsin-resistant PulD C-terminus (comprising part of the N3 sub-domain, and a complete C-terminal domain) indicated only 27% of the polypeptide is in β-strand conformation. This is a low β-strand content compared to most other integral OMPs (Chami *et al.*, 2005). Although most OM proteins contain transmembrane β-strands, it is possible that trans-OM segments are α-helices as in the Wza capsular polysaccharide export channel and VirB10 type 4 secretion (conjugation) system OM channel (Beis *et al.*, 2004, Chandran *et al.*, 2009, Dong *et al.*, 2006). Alternatively, the apparent 1:3 ratio of β-strands to α-helices could reflect the ratio of the membrane spanning vs. periplasmic portions in the analysed multimer, as was reported for TolC, an OM channel containing a large periplasmic domain (Koronakis *et al.*, 1997).

Truncation of PulD to express only the C-terminal secretin homology domain (PulD-CS) revealed that the N-terminal domain in this secretin is not involved in membrane insertion, multimerisation or localisation. Furthermore, negatively stained EM and single particle reconstruction showed that PulD-CS was still able to form 3-tiered barrels, showing that the C-terminal domain is structurally involved in gating and forms the major part of the N ring. The lack of significant difference in the transmembrane electrochemical potential of strains expressing

wildtype PulD or PulD-CS is due to its accurate targeting by a cognate chaperone, PulS (Guilvout *et al.*, 2006). If erroneously inserted into the IM, due to the absence of signal sequence, PulD is toxic to *E. coli* (Guilvout *et al.*, 2011).

Attempts to His-tag the pIV secretin at the N- or C-termini prohibited multimerisation and/or destabilised the multimer. pIV His-tagged at positions A308 or D347 formed functional multimers, however stability was compromised in the absence of the stabilising spontaneous mutation S318I, which increased stability of the multimer in comparison to the untagged wildtype pIV (Linderoth *et al.*, 1996). Transmembrane β-strand (TMBETA) prediction (Gromiha *et al.*, 2005) suggested that the least interfering his-tag site, A308, was situated between two predicted transmembrane β-strands (Spagnuolo, 2010, Spagnuolo *et al.*, 2010). Secretin family PFAM alignments (Figure 4) further showed that the A308 His-tag was located between two conserved regions, which correspond to the TMBETA-predicted β-strands. From this we can infer that sequence elements within the C-terminal domain of secretins are important for inter-subunit interactions.

Mutagenesis of the secretin PulD followed by selection for leaky gate mutants yielded a number of leaky mutations, all mapping to similar locations in the C-terminal homology domain as GATE1 and GATE2 in pIV (Whitaker, 2012). However, in another report a unique screen to isolate PulD assembly mutants, as those overcoming the lethality of PulD multimers expressed without a signal sequence that inserted into the IM from the cytoplasmic side to select for multimerisation-incompetent mutants, identified a GATE1 region mutant (G458S

in PulD, G358S in pIV) identical to that found in the leaky gate screen (Guilvout *et al.*, 2011, Guilvout *et al.*, 2014, Huysmans *et al.*, 2013). This was perplexing since leaky gate mutants must be stable enough to permit maltopentaose through, yet multiple experiments indicate these residues confer a stability defect. It is possible that GATE1 has roles in both multimer stability and gate integrity. This finding, however, may reflect different mechanisms of protein folding/quality control in the cytoplasm, where mutants in this screen have folded, to those mechanisms in the periplasm, where the mutants in the leaky gate screen folded.

1.4.3 Secretin OM targeting, folding and assembly

Secretin monomers are translocated to the periplasm via the SecBA-SecYEG system. However, the subsequent folding and targeting pathway is not well understood and varies between secretins. Some secretins are localised to the OM with the help of an OM lipoprotein, or pilotin, which interacts with a dedicated pilotin-binding domain at the C-terminus of the secretin. Representative secretin-pilotin pairs from T2SS and T3SS, whose targeting has been studied in some detail, are PulD-PulS and InvG-InvH (Crago & Koronakis, 1998, Guilvout *et al.*, 2006). Some secretins, like HxcQ (T2SS), BfpB and TcpQ of the type 4 pilus assembly system, contain a lipoprotein-like domain and they are targeted to the OM by fatty acylation – therefore they contain an "intramolecular" pilotin. However, secretins BfpB and TcpQ require additional partner proteins for OM localisation (Viarre *et al.*, 2009). Removal of the pilotin, pilotin-binding-domain, or lipid-attachment motifs, all result in the mislocalisation

of secretins to the IM. In contrast to other pilotin-dependent secretins, PulD is degraded in the absence of its cognate pilotin, PulS. Altogether, published work on several secretins from different secretion systems points to a basic common domain architecture and varied requirements for additional folding, multimerisation, membrane insertion and OM targeting domains and accessory factors.

Pilotin-dependent secretins are probably shuttled to the OM via the Lol lipoprotein-sorting pathway (Collin *et al.*, 2011, Okon *et al.*, 2008). Given that multimerisation of InvG and HxcQ is not affected by the absence of the pilotin or fatty-acylation, these localisation factors are likely to have no role in the formation of the multimer. However, it has been suggested for *N. meningitidis* PilQ secretin that the Bam complex and additional cognate lipoproteins (besides pilotins) encoded by the respective secretion system gene clusters, may have a function in multimerisation or stabilisation of the multimer. Yet, this is the only known case of a secretin requiring the Bam complex (Carbonnelle *et al.*, 2005, Collin *et al.*, 2007, Volokhina *et al.*, 2009, Voulhoux *et al.*, 2003). In contrast to the secretins discussed above, that require accessory factors or domains for OM targeting and/or multimerisation, the pIV secretin of the FFSS has no known localisation factors and membrane fractionation shows the localisation of pIV to both the IM and OM (Russel & Kazmierczak, 1993). Similarly, *V. cholerae* T2SS secretin, EpsD, appeared not to require additional localisation proteins. However, a chromosomally-encoded pilotin, AspS, has recently been identified (Dunstan *et al.*, 2013).

Even though pilotins efficiently target their cognate secretins to the OM, depending on the secretin under investigation, there is always a variable IM-associated pool of secretin multimers (Crago & Koronakis, 1998, Daefler *et al.*, 1997a), suggestive that pilotins sequester secretins, decreasing their IM association, rather than acting as chaperones. The exception is PulS, which protects the PulD secretin from proteolysis in the periplasm. In the absence of cognate pilotins, secretins that are dependent upon them are mislocalised entirely to the IM (Crago & Koronakis, 1998, Daefler & Russel, 1998, Guilvout *et al.*, 2006). Fusing the pilotin-binding domain of PulD ($PulD_{PBD}$) to a self-targeted secretin, pIV, and concurrently expressing the pilotin, PulS, together with pIV–$PulD_{PBD}$ allows efficient targeting to the OM. Furthermore, proteolysis is mediated by the $PulD_{PBD}$ domain, and interaction of PulS with pIV-$PulD_{PBD}$ prevents degradation, independent of PulS lipidation and of OM targeting, as long as the interaction is maintained (Carbonnelle *et al.*, 2005, Daefler *et al.*, 1997a, Koo *et al.*, 2008). Together, these findings suggest a model where secretin monomers are sequestered by their lipopeptide pilotin, or localisation factor, before they insert into the IM. Proteolysis of unsequestered secretins secures the OM targeting pathway and at the same time protects the host from the formation of potentially toxic secretin channels in the IM.

Whether secretins are transported to the OM as monomers or, multimers, ready for insertion remains to be revealed. The sheer size of secretin multimers (up to 1500 kDa) would make it exceedingly difficult to pass through the murein layer, even if a cognate transport protein exists. Substitution of pIV residue P375

(Russel, 1994) prevents efficient multimerisation of pIV. However, these substitution mutants were still found associated with both the IM and OM fractions, suggesting that secretins cross the periplasm as monomers and association of monomers with membranes could precede multimerisation and insertion (Russel, 1994). This observation poses the question of causality between multimerisation and membrane insertion of self-targeting secretins.

From these observations a model can be proposed, whereby OM biogenesis of secretins occurs by co-targeting the secretin monomers associated with a lipoprotein to the OM, followed by multimerisation and insertion into the membrane. Alternatively, lipoproteins may not act as pilots; rather they only prevent premature multimerisation/insertion by titrating monomers away from the IM. OM targeting could be a consequence of the OM lipoprotein *itself* being targeted to the OM by the Lol pathway. The involvement of the Lol lipoprotein-sorting pathway in targeting the secretin to the OM has yet to be confirmed.

1.5 Extracytoplasmic Stress Responses

Bacteria have evolved a plethora of interacting pathways to cope with stochastic environmental stressors, which can cause catastrophic damage to metabolism and cellular integrity if not circumvented by stress responses. *E. coli* K-12 has six documented extracytoplasmic stress response pathways, of which three have been characterised: the RpoE/σ^E, Cpx, and Bae pathways. These three pathways are induced by changes to OM homeostasis and misfolded proteins in the periplasm (Duguay & Silhavy, 2004, Raffa & Raivio, 2002, Rowley *et al.*,

2006, Ruiz & Silhavy, 2005). Another stress pathway, the phage shock protein (Psp) response, is most likely induced by agents and conditions that interfere with the homeostasis of the IM: filamentous phage assembly, extreme heat-shock, high salt concentrations, ethanol and perturbation of the proton-motive force (PMF) across the IM, amongst other signals. Stressful stimuli are sensed through a signal-transduction pathway involving multiple components, of which at least one is IM-embedded, to control induction. The IM sensor protein controls the release/activity of alternative sigma factors (RpoE/σ^E) or transcriptional-regulators, as in the Cpx, Bae and Psp responses (Rowley *et al.*, 2006, Ruiz & Silhavy, 2005). Transcription factors, such as CpxR and PspF can act to regulate gene expression in one of three ways; positive regulation, to increase gene expression; negative regulation, to decrease gene expression; and finally, dual regulation, where gene expression can be either increased or decreased, activity of the transcription factor is allosterically regulated.

In addition to the above stress responses regulating protein composition and quality within the membranes, other responses, such as the *r*egulation of *c*apsular *s*ynthesis (Rcs) regulon, protect against osmotic shock. This response is controlled by a 2-component sensor kinase phosporelay, which in turn regulates expression of colanic acid capsules and effectors that mitigate the effects of imbalanced intracellular osmolytes (Na^+/K^+) (Sledjeski & Gottesman, 1996).

1.5.1 The Phage Shock Protein Response

Expression of secretins, in the absence of other components of the secretion system, imposes stress onto the host cell (secretin-stress), which somehow induces the Psp response. For secretins, induction correllates with the mislocalisation of a significant portion of the secretin multimer into the IM: with pilotin-guided secretins, which are precisely targeted to the OM (like PulD) being non-inducers, and secretins that insert into both IM and OM (like pIV) being inducers. Microarray analyses of global transcription during secretin-stress in *Y. enterocolitica, E. coli* K-12, and *S. typhimurium* (Becker *et al.*, 2005, Bury-Mone *et al.*, 2009, Jovanovic *et al.*, 2006, Lloyd *et al.*, 2004, Seo *et al.*, 2007, Seo *et al.*, 2009) suggested the response to secretin-stress is limited to the *psp* operon; no other stress operons seemed to be induced.

Originally identified in filamentous phage f1-infected *E. coli* (Brissette *et al.*, 1990), the Psp regulon has been found to encode some 7 proteins (PspA-G). The main *psp* operon corresponds to *pspA-E*; *pspF* is located just upstream and is divergently transcribed from a constitutive sigma-70 dependent promoter. A separate gene, *pspG*, however, is not linked to the rest of the *psp* genes but is co-regulated (Figure 5). The Psp response is induced by a range of stressors that are also known to induce other stress responses (RpoE, and Cpx). The Psp response has also been shown to play a role in maintaining the PMF (therefore protecting ATP-production) under PMF-dissipating conditions, such as SecYEG translocon overloading (Kleerebezem & Tommassen, 1993, Kleerebezem *et al.*, 1996), or secretin mislocalisation to the IM. However, dissipation of PMF alone

was found to be insufficient to induce *psp*, indicating PMF dissipation could be a symptom of a broader *psp*-inducing stressor (Engl *et al*., 2011). Inhibition of phospholipid biosynthesis was also shown to induce the Psp response (Bergler *et al*., 1994). Furthermore, defects in Lol, LppX, and LPS all result in incorporation of phospholipids into the outer leaflet and general disturbance of the OM (Ruiz & Silhavy, 2005), resulting in induction of the Psp response. This latter set of defects may indirectly increase the level of IM stress by disrupting phospholipid homeostasis, leading to induction of *psp*.

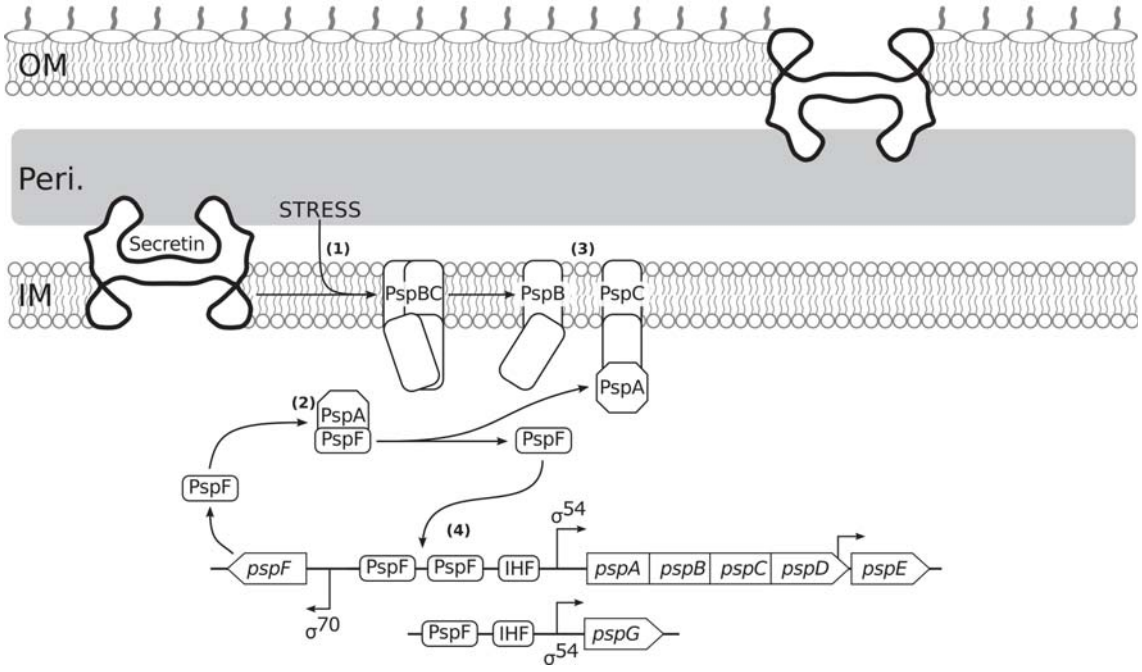

Figure 5: Organisation of the phage shock protein regulon and induction of the Psp response.

In the uninduced state **(1)**, the cytoplasmic domain of PspB is covered by that of PspC and cannot inhibit the negative regulator, PspA, which binds and inactivates the response activator, PspF **(2)**. Inducing signals, such as mislocalised secretins, causes a shift in PspC binding partners, disassociating from PspB and sequestering PspA **(3)**, freeing PspF to bind to the upstream activating sequences (UAS). The *pspA*/*pspG* promoters are sigma-54 dependent **(4)** (Flores-Kim & Darwin, 2012b, Flores-Kim & Darwin, 2015, Jovanovic *et al.*, 1997, Yamaguchi *et al.*, 2013).

Induction of the *psp*-regulon requires activation of the divergent IHF-enhanced σ^{54}-dependent *pspA* promoter by the binding of the specific transcriptional activator, PspF, and general DNA-bending protein, IHF, to the respective target sequences, UAS1/2 and IHF boxes (Figure 5). The *psp*-specific activation by PspF requires detection of an inducing signal by the membrane-embedded domain of the polytopic integral IM protein PspC in the PspBC complex (Flores-Kim & Darwin, 2012b). An as yet unknown signal causes PspC to change binding partners, dissociating from the PspBC complex and sequestering PspA (Flores-Kim & Darwin, 2015). Now liberated of PspA-mediated inactivation, PspF can interact with the UAS sequences upstream of the *pspA-E*/*pspG* -35/-10 sequences, and in the presence of IHF, that bends the DNA, can interact with RNA polymerase (σ^{54}-dependent) and stimulate transcription of the *pspA-E* and *pspG* operons. Analysis of *psp* transcripts revealed that expression from the *pspF* promoter is inhibited by the expression of the *pspA-E* operon. However, PspF continues to be expressed owing to a stable pool of *pspF* mRNA. Two inverted repeat elements, named RIB, lie downstream of *pspF* and stabilise the transcript against RNAse degradation (Jovanovic & Model, 1997). The analysis of *pspA-E* mRNA transcripts and promoters revealed that the last gene in the *psp* operon, *pspE*, is independently transcribed from a temperature-sensitive σ^{70}–dependent promoter such that *pspE*-mRNA levels independent of the sigma-54-dependent promoter upstream of *pspA*. That is to say, expression of *pspE* is independent of the *pspF*-induced sigma-54 promoter upstream of *pspA* (Brissette *et al.*, 1990, Brissette *et al.*, 1991, Jovanovic *et al.*, 1996, Jovanovic *et al.*, 1997).

The effector proteins of the *psp* response are PspA, PspD, E, and G. The involvement of PspBC is dependent on the nature of the inducing stimuli, as they are not required for *psp*-induction by heat-shock, but are essential for induction by secretin mislocalisation (the most severe inducer of the *psp* response). Adding to the confusion, PspBC are partially required for induction during overloading of the SecYEG IM translocon, known to dissipate the PMF (Kleerebezem & Tommassen, 1993, Kleerebezem *et al.*, 1996). In *Y. enterocolitica* PspA is dispensable, while PspB and PspC have been shown to be essential for survival of secretin stress, establishing a dual effector/regulator role for these two proteins (Horstman & Darwin, 2012, Maxson & Darwin, 2006). PspD is a cytoplasmic protein associated with the IM, but is not involved in the regulation of *psp*-induction. PspE localises to the periplasm where it is thought to act like a rhodanese, reconstituting Fe-S clusters required for respiration (Adams *et al.*, 2002, Adams *et al.*, 2003). The *pspG* gene, which is unlinked to the *psp*-operon, is co-regulated with the operon due to its PspF-σ^{54}-dependent IHF-enhanced *pspA* promoter. Localisation experiments indicated it coats the inner surface of the IM and additionally, it is only found in organisms with *a psp* operon. Expression analysis showed that PspG had a role in regulating the motility of *E. coli* (Engl *et al.*, 2009, Green & Darwin, 2004, Lloyd *et al.*, 2004, Jovanovic *et al.*, 2006).

Investigation into the protein-protein interactions of the Psp proteins revealed changes in the behaviour and localisation of regulator and effector proteins

upon induction. In the 'off' state, PspBC and PspAF complexes are highly mobile, moving within the IM and cytoplasm, respectively. Upon detecting a *psp*-inducing signal, PspA switches binding partners and binds PspC, which localises to the poles of induced cells (Flores-Kim & Darwin, 2015). Fluorescently tagged secretin, YscC, mislocalised to the IM has also been shown to migrate to the cell poles. Furthermore, aggregates of the secretin and the separation of IM from OM at the cell poles have been observed in *psp*-null cells expressing mislocalised secretins (Adams *et al.*, 2003, Diepold *et al.*, 2010, Horstman & Darwin, 2012, Yamaguchi *et al.*, 2013).

A wide range of stressors induces the Psp response implying a diverse range of intermediary events that transduce stress signals to the Psp regulon. The identity and mechanism of these events are not clear. Microarray analysis of transcription during secretin-stress showed the conditional involvement of the ArcAB, a two-component metabolic regulation system in inducing Psp under microaerobic conditions. The ArcAB system is thought to regulate the switch between aerobic and anaerobic metabolism (Jovanovic *et al.*, 2006, Jovanovic *et al.*, 2009).

Several studies have implicated the RpoE/σ^E-regulon in Psp responses. However, the outcomes of these studies are different in three enteric bacteria, *Y. enterocolitica*, *E. coli*, and *S. typhimurium*. A transposon-insertion library screened for synthetically lethal phenotypes caused by YsaC secretin expression in *Y. enterocolitica* identified an RpoE/ σ^E-null mutant containing an

uncharacterised suppressor mutation (RpoE/σ^E is essential in *Y. enterocolitica*). Other mutants isolated included DegS-null and TrkA-null mutants (Seo *et al.*, 2007). DegS is a periplasmic protease essential for activating the RpoE pathway. TrkA is a potassium transporter, probably implicated due to the involvement of potassium in regulation of osmotic pressure. The mutations in *rpoE*, *degS*, and *trkA* obtained in this screen did not affect Psp induction in the absence of the YsaC (Seo *et al.*, 2007). Therefore, these proteins have a buffering role in the protection of *Y. enterocolitica* from secretin stress in the absence of the induction of the Psp regulon. In *E. coli*, microarray analysis of *rpoE* knock-down mutants showed that they displayed little effect on *psp* induction in response to secretin stress, suggesting similarity to *Yersinia* (Becker *et al.*, 2005, Seo *et al.*, 2007, Rhodius *et al.*, 2006). In contrast, analysis of *S. typhimurium rpoE* mutants (*sans* secretin) revealed strong up-regulation of PspA protein due to PMF dissipation. Thus, depending on the organism, the RpoE/σ^E response may be implicated in maintaining PMF either directly or indirectly through PspA induction.

Interestingly, *P. aeruginosa*, a bacterium lacking a *psp* operon, is not sensitive to expression of native secretins, and this secretin resistance does not require an AlgU response (an ortholog of RpoE) (Seo *et al.*, 2009). Transcriptome analysis revealed no regulatory changes of any gene due to secretin expression in wildtype strains. However, random transposon mutagenesis isolated a secretin sensitive mutant with an insertion at locus PA0943. This PA0943-null mutant was unable to export substrates of the Hxc T2SS, and displayed

drastically reduced production of T4PS secretin XcpQ and its mislocalisation. Global OMP localisation, however, was not affected and the toxicity of secretin mislocalisation could be suppressed by the expression of plasmid-borne *Y. enterocolitica* PspBC. PA0943 showed sequence similarity to LppX and Lol proteins involved in periplasmic trafficking, and fractionations showed that it was a periplasmic protein.

In summary, although various stressors affecting the IM are shown to induce the *psp* operon, the inducing signal of secretin-stress has not been identified as yet. The data from *E. coli*, *Y. enterocolitica*, and *S. typhimurium* have not identified any proteins in the secretin-stress induction pathway other than those encoded by the *psp* operon itself (Jovanovic *et al.*, 2006, Jovanovic *et al.*, 2009, Rowley *et al.*, 2006, Ruiz & Silhavy, 2005, Seo *et al.*, 2007).

1.5.2 The Regulation of capsular synthesis regulon

Mutations of the cytoplasmic regulatory protease, Lon, often cause a mucoid phenotype due to the overproduction of exopolysaccharide capsules. This observation allowed identification of the regulatory components of the Rcs regulon. Individual *rcs* genes, *rcsA*, *rcsB* and *rcsC* were identified using lac fusions of the capsular polysaccharides synthesis genes, *cps*/*man*, responsible for synthesis of colonic acid capsules, in Lon-positive vs Lon-negative backgrounds (Gottesman *et al.*, 1985). Sequence homology to known 2-component regulators suggested the sensor kinase and effector roles for RcsC and RcsB, respectively, and later identified RcsD as a phosphorelay protein

transmitting the activating phosphate from the C-terminal domain of RcsC to the N-terminal domain of RcsB (Stout & Gottesman, 1990, Takeda *et al.*, 2001). Upon activation by phosphorylation, RcsB forms a homodimer ($RcsB_2$) and heterodimer with RcsA (RcsAB) (Figure 6) (Stout *et al.*, 1991). The presence of C-terminal helix-turn-helix motifs in both RcsA and RcsB indicated these proteins bind DNA. The accumulation of RcsA, but not RcsB, in *lon*$^-$ cells revealed RcsA as the target for degradation by the Lon protease (Torrescabassa & Gottesman, 1987). Further characterisation of RcsA and RcsB determined that RcsB was the primary DNA-binding dual transcriptional regulator of the regulon, with RcsA acting as an ancillary protein, amplifying the potency of transcriptional regulation by RcsB at RcsA-dependent promoters, which are characterised by a conserved RcsAB box 50-100 nt upstream of the -35 promoter region. Conversely, RcsA-independent promoters, require binding of $RcsB_2$ to a conserved box immediately upstream of the -35 promoter region for expression (Brill *et al.*, 1988, Francez-Charlot *et al.*, 2003, Wehland & Bernhard, 2000).

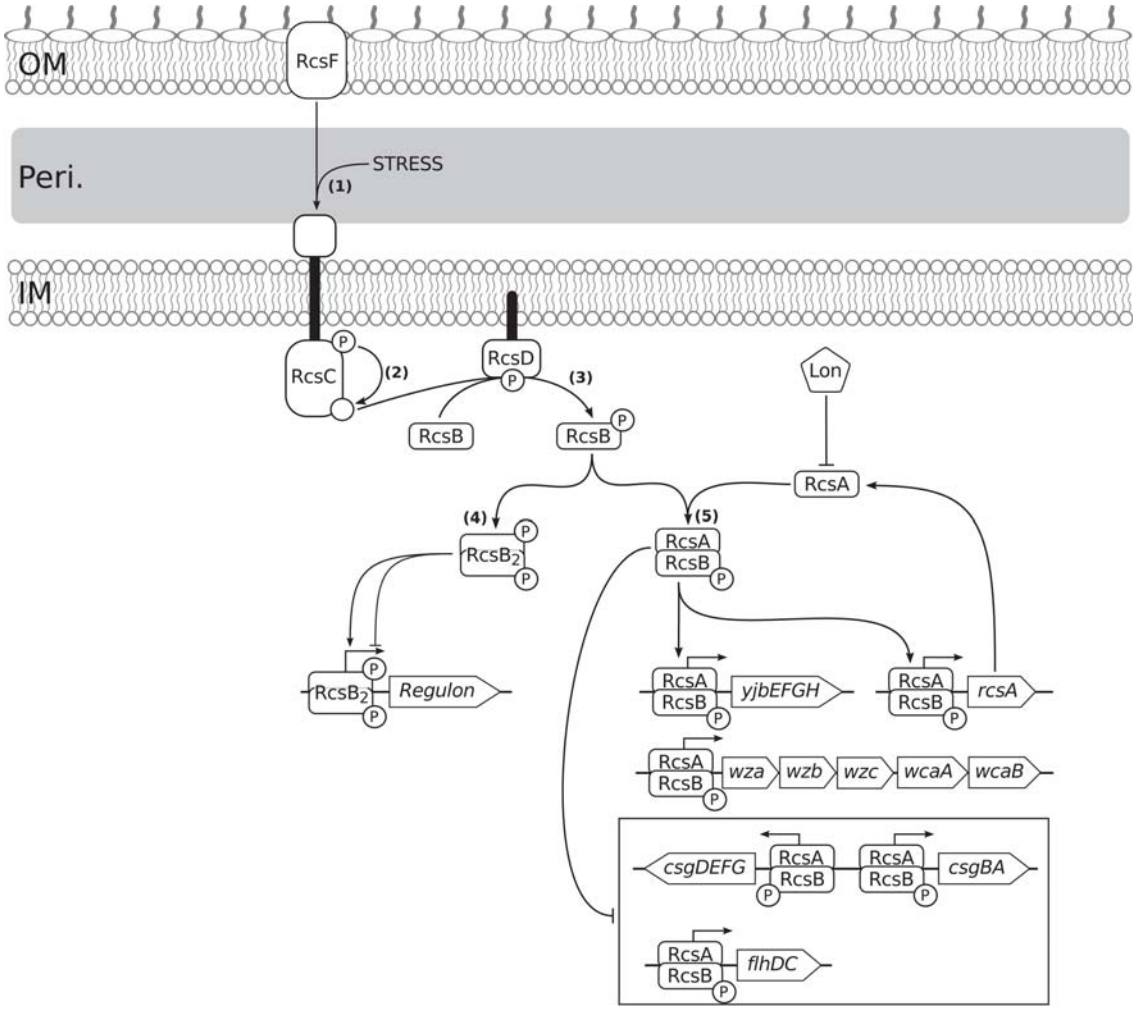

Figure 6: Regulation of the Rcs phosphorelay.

(1) OM protein RcsF transmits inducing signals from the OM to the RcsC sensor kinase in the IM. Peptidoglycan damage may also activate the pathway by directly stimulating RcsC. **(2)** RcsC transfers phosphate to intermediate protein RcsD which activates RcsB by phosphorylation **(3)**. RcsB may then form a homodimer to activate the RcsA-independent **(4)** branch of the regulon, or complex with RcsA to activate expression of *rcsA*, and regulate the RcsA-dependent branch of the regulon **(5)**. Both $RcsB_2$ and RcsAB act as dual transcriptional regulators, able to either repress or activate expression of regulon members.

Identifying suppressors of the *ftsZ84* division mutation lead to the discovery of the OM protein, RcsF, that induces a mucoid phenotype in an RcsB-dependent manner, by transducing the Rcs-inducing signals detected at the OM to RcsC (Gervais & Drapeau, 1992). However, RcsF is dispensable for induction of capsule synthesis and the Rcs regulons can be induced by peptidoglycan damage, thus the Rcs-phosphorelay is sensitive to both OM and periplasmic stressors (Laubacher & Ades, 2008).

The Rcs-regulon can be described as having two branches, defined by the dependence on RcsA for full expression from the Rcs-regulated promoters (Figure 6). Typically, interaction of RcsAB at RcsA-dependent promoters induce the expression of capsular polysaccharide synthesis operons (Ferrieres *et al.*, 2007, Wehland & Bernhard, 2000), but RcsAB has also been shown to autoregulate the expression of *rcsA* (Ebel & Trempy, 1999, Wehland & Bernhard, 2000) and repress motility by inhibiting the expression of the motility master regulator pseudogene *flhDC*. The expression of curli pilin is similarly repressed by the RcsAB dimer (Francez-Charlot *et al.*, 2003, Vianney *et al.*, 2005). Conversely, $RcsB_2$ binds RcsA-independent promoters and controls the expression of a surprisingly wide variety of genes, ranging from effectors of osmotic shock and acid resistance, to regulators of cell division, and an alternative sigma factor, RpoS/σ^{38}. It has been suggested the Rcs regulons directly and indirectly affect the expression of as much as 2.5% of the *E. coli* K−12 genome (Ferrieres & Clarke, 2003).

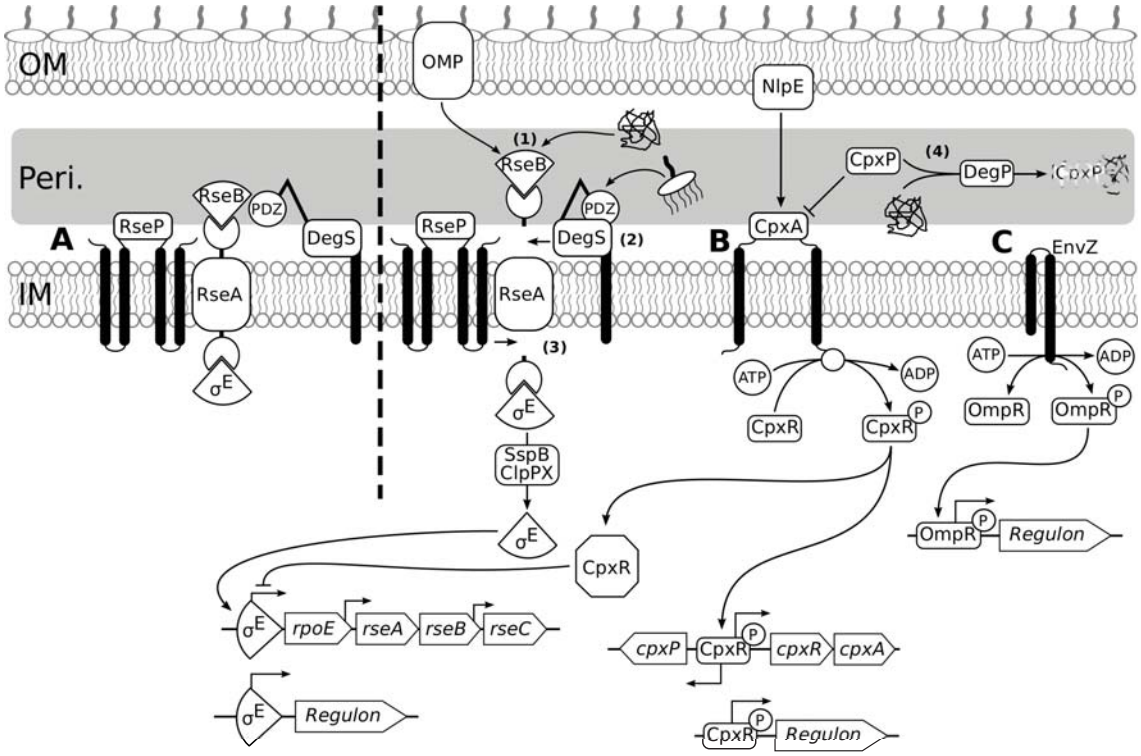

Figure 7: Regulation of the sigma-E, Cpx and EnvZ/OmpR responses.

(A) In the absence of stress, the IM-embedded anti-sigma factor, RseA, sequesters sigma-E (σ^E), preventing expression from sigma-E dependent promoters. **(1)** Overproduction of OM-proteins (OMP), formation of aggregates and/or misfolded proteins, stimulate a conformational shift in RseB, allowing binding of protease, DegS, to RseA. **(2)** The PDZ domain of DegS blocks access to RseA, preventing cleavage of the RseAB complex. In addition to overproduced, misfolded or aggregated OMPs, the accumulation of LPS intermediates within the periplasm promotes conformational shift in the PDZ domain of DegS, allowing access to RseA, and eventual release of sigma-E. The removal of the RseAB complex causes a conformational shift in the cytoplasmic portion of RseA, allowing an IM protease, RseP, to cleave the RseA-sigma-E complex, releasing it from the membrane. **(3)** This complex is then further processed by cytoplasmic proteases (Ssp/ClpPX), which degrade the cytoplasmic domain of RseA, freeing sigma-E to allow expression of sigma-E dependent promoters. **(B)** Similar to the sigma-E response, Cpx response is induced by aggregated and/or misfolded proteins, and is specifically induced by the over-production of lipoprotein NlpE. The signal activates the CpxA cytoplasmic ATPase domain, activating the Cpx response regulator (CpxR) by phosphorylation. Phosphorylated CpxR activates the expression of the *cpx* operon, and regulates the expression of regulon members. **(4)** CpxP can interact with misfolded/aggregated proteins and consequently becomes targeted for destruction by DegP, freeing CpxA kinase-activity from CpxP-mediated inhibition **(C)** The EnvZ/OmpR 2-component signal transduction pathway reacts to changes in osmolarity. Conformational changes in EnvZ result in the activation of the response regulator, OmpR by phosphorylation. OmpR-P can then regulate the expression of OmpR-dependent promoters. In low osmolarity environments, the phosphatase activity of EnvZ dominates, dephosphorylating OmpR and shutting down the response, a process also requiring ATP.

1.5.3 The sigma-E stress response

The alternative sigma factor, sigma-E, encoded by *rpoE*, regulates the expression of the best-understood extracytoplasmic stress response in Gram-negative bacteria. It is essential for viability, however, one suppressor of Δ*rpoE* essentiality was identified – Δ*hicB*, which encodes the antitoxin of the HicAB toxin-antitoxin pair. Interestingly, overexpression of the toxin, HicA, promoters the cleavage of mRNA and decreases growth rate but does not cause cell death (Button *et al.*, 2007, Jorgensen *et al.*, 2009). It is thought that bacterial toxin-antitoxin pairs play a deeper role in regulation of stress responses because RNAse activity in several toxins has been observed. Moreover, they show preference for degrading sigma-70 (i.e. housekeeping genes) transcripts, promoting translation of stress-related transcripts dependent on alternative sigma factors, such as sigma-E or sigma-54 (Bertram & Schuster, 2014, Wang *et al.*, 2011).

Accumulation of misfolded and overproduced OM proteins in the periplasm induces the sigma-E response by way of a proteolytic cascade involving 3 IM proteins; the proteases DegS and RseP; an anti-sigma factor, RseA; and a periplasmic sensor protein RseB (Figure 7A) (Collinet *et al.*, 2000, De Las Penas *et al.*, 1997, Kanehara *et al.*, 2002, Missiakas *et al.*, 1997, Walsh *et al.*, 2003). Additionally, the accumulation of off-pathway LPS intermediates in the periplasm (originating from the LPS transport and assembly machinery) cause the release of Sigma-E from RseA, by interacting with and stimulating a conformational shift in the PDZ domain of DegS (Figure 7B) (Lima *et al.*, 2013).

The sigma-E response regulates the expression of 82 transcription units (genes and operons) encoding proteins with a wide variety of functions (Keseler *et al.*, 2013). Notably these are periplasmic proteases (DegP), chaperones (SurA), as well as the OMP and LPS assembly complexes (Bam and Lpt, respectively) (Dartigalongue *et al.*, 2001, Rhodius *et al.*, 2006). It also regulates the production of OMPs by way of regulatory sRNAs, and the expression of DNA-repair pathways and cell division proteins (Al Mamun *et al.*, 2012, Johansen *et al.*, 2006).

1.5.4 The Cpx stress response

First shown to be required for conjugal DNA transfer (Silverman, 1985), the *cpxRA* operon (named for *c*onjugative *p*lasmid *e*xpression) encodes a 2-component sensor kinase system controlling expression of 38 transcriptional units (Keseler *et al.*, 2013). Upon stimulation, the IM sensor kinase, CpxA, is autophosphorylated; it further transfers this phosphate to the DNA-binding response regulator CpxR, thereby activating the response (Figure 7B). The periplasmic inhibitor of the response, CpxP, is encoded by a divergently oriented *orf* located just upstream of the *cpxRA* operon and is induced by binding of phosphorylated CpxR to its operator. CpxP shows structural similarity (and shares 26.5% sequence identity) to cryptic periplasmic chaperone, Spy, and inhibits autophosphorylation of CpxA. It is thought this action limits the length of the response to a short time period, preventing runaway expression of the regulon (Danese & Silhavy, 1998, DiGiuseppe & Silhavy, 2003, Raivio *et al.*, 1999, Thede *et al.*, 2011).

The members of the Cpx regulon overlap significantly with the other extracytoplasmic stress responses (notably sigma-E), but the response is induced by different signals – attachment to surfaces, over-production of lipoproteins and misfolding of pilus subunits. The latter signal may also be detected by the sigma-E response (Miyadai *et al.*, 2004, Otto & Silhavy, 2002, Snyder *et al.*, 1995). This response activates expression of genes involved in multidrug resistance (*acr*, *mdt*), OMPs (*ompC*), and periplasmic protein quality factors (*dsbA*, *ppiA*, *degP*), while repressing adherence, taxis and motility factors, through inhibition of pili assembly (*csg*, *motB*). Additionally, CpxR inhibits expression of the *rpoE-rseABC* operon, blocking concomitant activation of the sigma-E stress response. This has a consequence of altering gene expression to suit a biofilm lifestyle, increasing resistance to heavy metals, antibiotics and other phenotypes typical of cells growing in biofilms or switching from planktonic to biofilm lifestyles (Dorel *et al.*, 2006).

1.5.5 Osmotic shock responses

Osmotic stress induces a number of overlapping extracytoplasmic stress responses, making it difficult to discern whether the response to osmotic shock is mediated by one or more regulons. It is likely that several participate, but expression of individual genes within the regulons is highly tuned by sRNAs and binding of transcriptional regulators to operators. Some of these interactions may be competitive and differential binding affinity has been shown to tune gene expression in at least three overlapping regulons; SoxS, MarA, and Rob (Martin *et al.*, 2000).

The EnvZ/OmpR 2-component sensor kinase system is sensitive to changes in osmolarity, which are detected by the IM sensor protein, EnvZ. This causes phosphorylation and activation of the DNA-binding transcriptional regulator, OmpR, allowing regulation of a number of OMPs (Figure 7C), including the differentially regulated porins OmpC and OmpF, induced by OmpR in response to high and low osmolarity, respectively (Mizuno & Mizushima, 1990, Lan & Igo, 1998). Kinase and phosphatase activities of the cytoplasmic domain of EnvZ are controlled by the strength of the osmotic signal (Forst *et al.*, 1989, Yoshida *et al.*, 2002).

Biofilm formation is a common response to osmotic shock, the first step in which is to halt the expression of flagella. OmpR-P (phosphorylated OmpR) is able to inhibit expression of the motility master regulator pseudogene *flhDC*, a role also shared by the RcsAB branch of the Rcs regulon, another regulon induced in response to osmotic shock (Francez-Charlot *et al.*, 2003, Shin & Park, 1995). The second step in biofilm formation is to promote surface adhesion by expression of curli pilus, encoded by the *csgDEFG* operon. This operon is positively regulated by OmpR-P, promoting the expression of the curli assembly machinery (CsgEFG) and the DNA binding dual-regulator, CsgD, which activates expression of the curli subunits from the *csgBA* operon (Reviewed by Barnhart & Chapman, 2006).

Interestingly, the curli assembly machinery and subunit-encoding operons are negatively regulated by the accumulation of subunits (curlin) in the periplasm *via* the induction of the Cpx response (and expression of the *cpxRA* operon).

CpxR inhibits the expression of the *csgDEFG* and *csgBA* operons, preventing runaway production of curli in the face of long-lasting osmotic shock (Prigent-Combaret *et al.*, 2001).

The regulator of capsular exopolysaccharide synthesis also inhibits the csgDEFG and csgBA operons. This curli down-regulation coincides with up-regulation of exopolysaccharide production (Figure 6) (Ferrieres *et al.*, 2007, Jubelin *et al.*, 2005, Vianney *et al.*, 2005, Wehland & Bernhard, 2000). Thus, cross-talk between Cpx, OmpR/EnvZ and Rcs not only controls the cessation of planktonic living and the transition to early biofilms *via* inhibition of flagellar genes and induction of curli, but also from early to mature biofilms by repression of curli assembly and induction of capsular exopolysaccharide production. It should be noted that the *csg* operons are co-regulated by a large number of other factors that are important for adaptation to, for example, the host environment (Keseler *et al.*, 2013, Salgado *et al.*, 2013).

1.6 Oxidative Stress Responses

Oxidative stress caused by reactive oxygen species (ROS) is highly detrimental to all cellular functions. However, production of ROS is a natural consequence of aerobic respiration, and bacteria have evolved regulatory pathways to deal with specific types of oxidative stress. ROS are depleted primarily by the expression of superoxide dismutase (SodA), which catalyses the formation of hydrogen peroxide from ROS, and catalase (KatE), which converts hydrogen peroxide to water and oxygen. The SoxS and OxyR regulons are responsible for prevention and mitigation of oxidative stress in Gram-negative bacteria. The

activity of the corresponding DNA-binding transcriptional master regulators is regulated by ROS and the redox state of the cell. Both of these regulons overlap with two other regulons involved in resistance to antibiotics, heavy metals and organic solvents; MarA and Rob. SoxS, MarA and Rob regulate a common set of genes that overlaps with those regulated by OxyR. In addition to the common set of genes, each master regulator controls the expression of a set of specific genes. SoxS, MarA and Rob recognise a common consensus sequence in the promoters of regulated genes. In contrast, OxyR recognises repeated ATAG motifs separated by 10 nt, and binds operators in the form of two distinct homocomplexes, depending on its redox state.

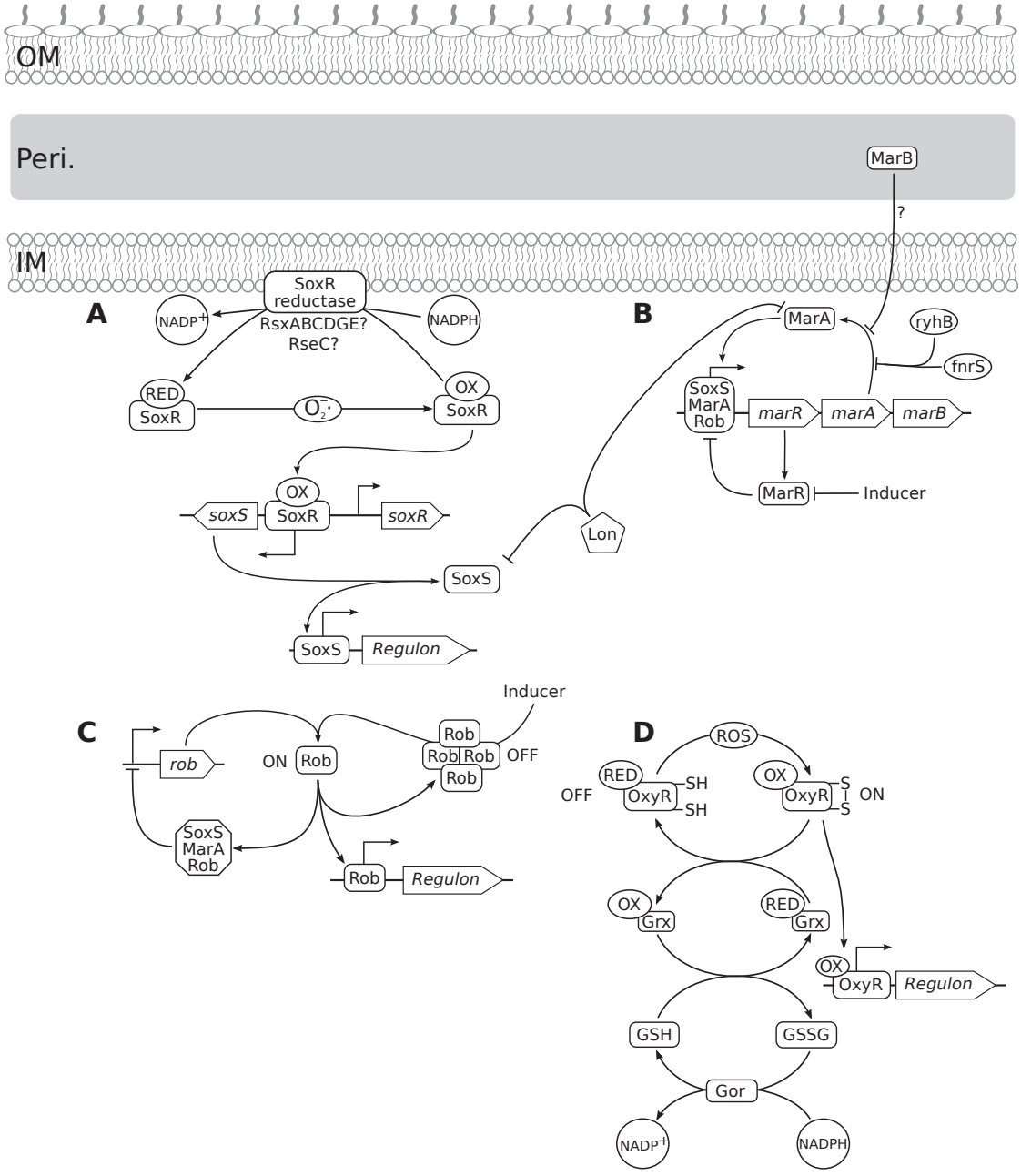

Figure 8: Oxidative stress response pathways of Gram-negative bacteria.

(A) The Superoxide stress response (Sox) is regulated by two proteins, SoxR and SoxS, the former of which is oxidised (OX) by reactive oxygen species (O_2^-•/ROS) allowing it to activate transcription of *soxS*, a DNA-binding transcriptional regulator and global regulator of the response. Integral IM proteins, RsxABCDGE and RseC are thought to cycle oxidised SoxR back to its reduced form (RED). Following dissipation of ROS, SoxR is reduced, halting transcription of *soxS*, and remaining SoxS is degraded by Lon protease. **(B)** Transcription of the *marRAB* operon encoding the response activator (MarA), repressor (MarR) and periplasmic response modulator (MarB) can be activated by three activators; SoxS, MarA and Rob, all binding to the same sequence in the promoter region. Small regulatory RNAs, ryhB, and fnrS, inhibit translation of the *marA* transcript. It is throught that the periplasmic protein MarB indirectly inhibits translation of *marA*. No other function for MarB has been identified. MarR-mediated constutive inhibition of *marRAB* transcription can be inactivated by an inducing signal, resulting in activation by SoxS, MarA or Rob. **(C)** In contrast to SoxS and MarA, Rob transcription-regulation activity is controlled somewhat differently to the other three responses described here. The C-terminal domain of Rob binds and sequesters other Rob proteins in the absence of an inducing signal (dipyridyl, bile salts). Upon binding of an inducer to the C-terminal domain, Rob-aggregates disperse, allowing Rob monomers to interact with Rob-regulated promoters. **(D)** The fourth oxidative stress response is very similar to the SoxRS response in that the master regulator (OxyR) is sensitive to oxidation (and activation) by ROS and that the oxidised form (OX) is able to activate transcription from OxyR-regulated promoters. Two proteins are responsible for redox cycling of OxyR; Glutaredoxin 1 (Grx) that reduces oxidised OxyR, and is itself reduced by Glutathione (GSH). Grx reduction results in oxidation of GSH to Glutathione disulphide (GSSG), which is then reduced by Glutathione reductase (Gor) in an NADPH-dependent reaction. Both *grx* and *gor* are members of the OxyR regulon, thus their expression is tied to the activation of the OxyR response and the redox state of the cell.

1.6.1 The Superoxide stress response

The superoxide stress response (Sox) is regulated by two cytoplasmic proteins; the redox-sensing transcriptional activator, SoxR; and the DNA-binding dual transcriptional regulator, SoxS (Figure 8A). The homodimeric regulator, SoxR, contains two iron-sulfur clusters sensitive to oxidation by ROS and nitric oxide (Fujikawa *et al.*, 2012). Redox-cycling drugs, like paraquat, have also been shown to directly oxidise SoxR and the response can be activated in the absence of ROS by deregulation of cellular NADPH homeostasis (Gu & Imlay, 2011, Krapp *et al.*, 2011). In either reduced or oxidised states, SoxR binds two known operators; that of *soxS*, inducing expression of the SOS regulon; and under iron-limiting conditions, SoxR may also bind the operator of *fumC* (Fuentes *et al.*, 2001). However, only the oxidised SoxR dimer is able to activate transcription from these promoters. Furthermore, due to the close proximity of the divergently transcribed *soxS* and *soxR orfs* and overlapping promoter regions, binding of SoxR between the *soxS* -10 and -35 promoter regions inhibits expression of *soxR* (Hidalgo *et al.*, 1998), thereby autoregulating its own expression. The half-life of SoxS is limited, and without a pool of oxidised SoxR to activate its transcription, it is quickly degraded by Lon protease, shutting off the SoxRS regulon in the absence of oxidative stress (Griffith *et al.*, 2004).

It is thought that the same enzyme responsible for the reduction of cytochrome *c* is responsible for reducing SoxR (specifically the SoxR iron-sulfur clusters) in a NADPH-dependent manner (Kobayashi & Tagawa, 1999), returning it to its 'off' state, thereby linking the Sox regulon to electron transport. Using a *soxS*

promoter transcriptional reporter and insertional inactivation libraries, two loci were identified as having a role in reducing SoxR; *rsxABCDGE* and *rseC* (Gaudu *et al.*, 1997, Koo *et al.*, 2003). The *rseC* gene is a member of the rpoE/σ^E regulon and is co-transcribed from the same promoter as *rpoE* and *rseABC*. However, unlike RseA and RseB, it is not a regulator of the σ^E regulon. Furthermore, primer extension experiments suggested *rseC* may be transcribed independently from a σ^E–independent promoter situated upstream of *rseC*, within the *rseB orf* (De Las Penas *et al.*, 1997) (Figure 7). Located in the IM, RseC was suggested to have a role in thiamine synthesis independent of σ^E in *Salmonella typhimurium* because a RseC-null mutant was unable to grow in the absence of thiamine, and insertional inactivation of *rpoE* did not confer a growth defect in the absence of thiamine (Beck *et al.*, 1997). RseC shows homology to *rnfF* of the *fdxC-fdxN* gene cluster from *Rhodobacter capsulatus*. This *R. capsulatus* gene cluster is located upstream of the *rnfABCDGE* operon and is transcribed divergently from it. This operon is homologous to the *rsxABCDGE* cluster present in *E. coli* K-12.

The SoxR-reducing locus was identified through genetic-linkage with RseC in *Rhodobacter capsulatus.* Independent inactivation of each of the members of the *rsxABCDGE* operon caused similarly increased expression from the *soxS* promoter, indicating that, at the least, they cooperate to reduce SoxR, or act as a larger complex (Koo *et al.*, 2003). The Rsx homologs RnfA, RnfD and RnfE are predicted integral IM proteins with RnfB and RnfC located at the cytoplasmic periphery of the IM, stabilising the complex. Further sequence comparison of the *R. capsulatus rnf* orfs with homologs from *H. pylori*, *E. coli*

and *V. alginolyticus* revealed RnfA is similar to two transmembrane subunits (NqrA and NqrE) of the Na^+ translocating NADH:quinone oxidoreductase complex from *V. alginolyticus* and also to the C-terminal domain of RnfE, encoded at the 5' end of the operon. RnfB and RnfC were reported to contain 2[4Fe-4S]-type ferredoxin-like domains, with RnfC also containing putative NADH and FMN binding sites, highly similar to those found in complex 1 of the electron transport chain of bacteria (Kumagai *et al.*, 1997).

SoxS shares homology with two other DNA-binding transcriptional regulators, MarA and Rob. Each of these proteins binds a shared but poorly conserved operator sequence called the *mar*/*sox*/*rob* box. While these proteins regulate a common, core set of promoters, the precise number and identity of MarA/SoxS/Rob regulon members is subject to debate and each protein independently regulates a number of other genes that are not members of the core MarA/SoxS/Rob regulon. Furthermore, MarA, SoxS and Rob confer differing levels of regulatory power on the common set of promoters in the MarA/SoxS/Rob regulon, due to differences in binding affinity of each regulator to the particular *mar*/*sox*/*rob* boxes (Martin *et al.*, 2000). At present time, the *E. coli* transcriptional regulation database, regulonDB, has 26 promoters (controlling expression of 40 genes) annotated as being regulated by SoxS, MarA or Rob, including that of the *marRAB* and *rob* operons, the latter of which is repressed by both SoxS and MarA (Alekshun & Levy, 1999, Jair *et al.*, 1996, Martin *et al.*, 1999, Martin & Rosner, 2002, Salgado *et al.*, 2013). However, transcriptional cross-talk only occurs between Rob and MarA regulons, both clearly downstream of SoxS. Therefore, regulation of SoxS expression and by

association that of the SoxS regulon, is not controlled by any gene within the MarA/SoxS/Rob regulon (Chubiz *et al.*, 2012). A more recent study, that used biological relationship analysis of similarly expressed gene clusters, estimated the number of genes in the SoxRS regulon to be 119, including those annotated in RegulonDB. Functional characterisation placed these genes into three broad groups; (1) NADPH regeneration; (2) removal of xenobiotics and recycling of damaged macromolecules; (3) damage prevention (Blanchard *et al.*, 2007). The major issue in determining the precise identity of promoters of the genes, encoding effector proteins, directly regulated by the SoxRS regulon is that some of these control expression of DNA-binding transcriptional regulators – creating a transcriptional cascade effect with multiple levels of regulation.

1.6.2 The MarA Regulon

The *m*ultiple *a*ntibiotic *r*esistance (Mar) response is regulated by three proteins encoded by the *marRAB* operon (Figure 8B; Cohen *et al.* (1993)). The response activator, MarA, binds the shared MarA/SoxS/Rob box in the promoter region, activating expression of the *marRAB* operon, and other members of the MarA regulon. Free MarA is sensitive to proteolysis by Lon, which quickly shuts down the response in the absence of inducing signals. However, when bound to DNA, MarA is protected from Lon-mediated degradation (Griffith *et al.*, 2004). Hfq-dependent small RNAs *ryhB* and *fnrS* are also able to post-transcriptionally inhibit translation of the *marA* transcript, preventing activation of the MarA regulon (Boysen *et al.*, 2010, Geissmann & Touati, 2004, Wright *et al.*, 2013). Deletion of *marB* and quantification of *marA* expression by qRT-PCR indicated MarB reduces the rate of MarA translation. However, it has a mechanism

different from destabilising the *marA* transcript, given that it is a periplasmic protein. It is unknown how MarB exerts its inhibitory effect on MarA expression from the periplasm. No other functions were identified for MarB (Vinue *et al.*, 2013). Together, the actions of sRNA's *ryhB* and *fnrS* and the protein MarB modulate the expression of the response activator MarA.

MarR transcriptionally represses expression of *marRAB* by binding, as a dimer, two repeats in the *mar* operator. Salicylate, phenolic-compounds and multiple antibiotics have been shown to inactivate the repressor activity of MarR – allowing expression of the *marRAB* operon and activation of the MarA regulon (Martin & Rosner, 1995, Martin & Rosner, 2004).

1.6.3 The Rob Regulon

The *r*ight *o*riC *b*inding (Rob) protein was found as one of the proteins that bind the directional origin of replication, *oriC*. Rob was shown to have homology to known DNA-binding transcriptional regulators. Expression analysis suggested that transcription of the *rob* gene is dependent on sigma-S and is highly and constitutively expressed with about 5000-10000 monomers present per cell (Kakeda *et al.*, 1995, Skarstad *et al.*, 1993). Its overexpression from a plasmid resulted in increased resistance to organic solvents, heavy metals and antibiotics (Nakajima *et al.*, 1995). Transcriptional cross-talk between Rob- and SoxS/MarA-regulated genes was later shown and the common SoxS/MarA/Rob binding box was identified (Jair *et al.*, 1996, Martin *et al.*, 1999).

The most notable difference between the Rob, SoxRS and MarRAB regulons lies in their regulation. While SoxRS and MarRAB are regulated by redox cycling and repressor activity, respectively (Figure 8A & B), the transcriptional-regulator function of Rob is regulated by a sequestration-dispersion mechanism. The C-terminal domain of Rob mediates sequestration of Rob monomers into an aggregate in the absence of inducing compounds (notably bile salts and dipyridyl). When these xeno-molecules bind to the Rob C-terminal domain, this interaction causes the aggregate to disperse, freeing the Rob monomers to bind the promoters of regulon members (Figure 8C). The Rob C-terminal domain also protects its N-terminal domain from Lon-mediated proteolysis (Griffith *et al.*, 2009).

Besides the mode of activation, the other difference from SoxRS and MarRAB lies in the amount of protein per cell; Rob is constitutively expressed to 10000 molecules/cell and is not sensitive to proteolysis (Ali Azam *et al.*, 1999, Skarstad *et al.*, 1993). In contrast, both SoxS and MarA pools are swiftly increased upon induction and, depleted following removal of the inducing stressor. SoxR is produced at very low levels, fewer than 100 molecules/cell, SoxS is only detectable when SoxR is oxidised and is produced at around 2500 copies/cell. Similarly, the *marRAB* operon can only be expressed when repressor MarR is inhibited, and the regulon maximally activated when sufficient levels of the activator, MarA are present – at least 10 000 molecules/cell (Martin *et al.*, 2008).

1.6.4 The OxyR Regulon

The *o*xidative stress *r*egulator, OxyR, is a DNA-binding transcriptional regulator governing expression of about 40 genes in response to ROS (Christman *et al.*, 1985, Keseler *et al.*, 2013, Salgado *et al.*, 2013). Its expression is positively regulated by cAMP-activated Crp during exponential phase of growth in culture, and negatively regulated by the alternative sigma-factor sigma-S during stationary phase (Gonzalez-Flecha & Demple, 1997). Like SoxR, both oxidised and reduced forms of OxyR can bind the operator of recognised promoters, but only the oxidised form can activate transcription (Figure 8D; Toledano *et al.* (1994)). Adjacent to the *oxyR* gene, but divergently transcribed, is the Hfq-dependent small regulatory RNA *oxyS*, whose transcription is also activated by oxidised OxyR (Altuvia *et al.*, 1997, Zhang *et al.*, 1998). Due to the overlap of these two promoters, binding of OxyR inhibits its own expression, while its redox state regulates expression of *oxyS* (Christman et al., 1989, Gonzalez-Flecha & Demple, 1997). While OxyR is readily activated by hydrogen peroxide, two proteins whose expression is directly regulated by OxyR; Glutaredoxin 1 (Grx) and Glutathione reductase (Gor), are responsible for reducing OxyR and switching off the response in the absence of inducing signals (Figure 8D; Zheng *et al.* (1998)).

1.7 Advantages of current generation RNA sequencing over microarray technologies

There are several approaches to investigate transcriptional events in model organisms ranging from more specific, target-based tools such as quantitative RT-PCR and northern blotting to global approaches like microarrays and RNA-

sequencing. Global approaches are able to capture snap-shots of all biological processes occurring within cells, thus are well suited for hypothesis discovery and lead-finding projects.

The principal underpinning microarrays is DNA hybridisation; probes to all known *orfs* of an organism are covalently attached to a slide in an ordered grid – the array. Total RNA from the organism of interest is reverse transcribed to cDNA, subsequently labelled with a dye and allowed to hybridise to the array, transcript abundance is measured as fluorescence.

By comparison, RNA-seq does not rely on hybridisation. Instead, mRNA, depleted of rRNA is fragmented; sample-specific primers are ligated to the 5' and 3' ends of the fragments and then converted to cDNA in a process that also attaches 3' sequencing primers. The sample-specific primers also allow the differentiation of sense transcripts from antisense. Rather than relying on probe-based hybridisation to detect specific transcripts, all mRNA fragments are sequenced and sequence, or "read" libraries are returned. These are aligned to reference genomes or gene-models and transcript abundance is measured directly as counts of reads per gene (Levin *et al.*, 2010, Wilhelm & Landry, 2009).

The use of hybridisation and fluorescence to measure transcript abundance in microarrays limits the dynamic range of detectable differential gene expression to a few hundred-fold, compared with several thousand-fold change in gene expression detectable by RNA-sequencing methods. Furthermore, probe

binding efficiency and specificity in microarrays can cause significant background signals, further reducing the reliability and dynamic range of the data (Wang *et al.*, 2009).

1.8 Aims

The primary aim of this work was to increase the understanding of the adaptations made by *E. coli* K-12 in response to the production of large outer membrane channels, secretins, which are common to toxin-, phage-, and type IV pilus-secreting trans-envelope complexes. Previous attempts, utilising microarray technology, to analyse the transcriptomic adaptations of *E. coli* to the production of filamentous phage secretin, pIV, suggested the response was limited to a single genetic circuit, the Psp response (Lloyd *et al.*, 2004, Jovanovic *et al.*, 2006, Seo *et al.*, 2007). Because of the limitations of microarrays, many subtle changes in gene expression may have been missed due to the poor dynamic range and statistical nature of such data. By employing a highly sensitive and robust sequencing strategy using directional, stranded RNA-sequencing, a high-resolution transcriptome analysis was undertaken in order to reveal changes previously missed by microarray analyses. Furthermore, analysis of the transcriptomic response to the production of a leaky pIV mutant (pIV-E292K), compared to the transcriptomic response to wildtype pIV provided insight into the differences in stress casued by non-permissive secretins (like wildtype pIV), that have pore sizes smaller than porins in the closed state, to permissive secretins (like pIV-E292K), that have pore sizes larger than porins in the closed state. Secretin InvG of the type 3 secretion system analysed concurrently with this thesis was shown to be

naturally permissive to maltopentaose and vancomycin (Khanum, 2015). As such, the analysis presented here has provided leads for identification of genes required, not only for survival of stress experienced by production of leaky pIV, but also other secretins that contain pores of the sizes larger than porins in the closed sate, or that open spontaneously, hence are naturally permissive to > 600 Da molecules, such as the secretins of T3SS.

Individual up-regulated genes or master-activators of up-regulated pathways were analysed for synthetic lethality with expression of secretins, to identify stress-relieving factors that allow survival of *E. coli* expressing secretins, findings that could be relevant for cells that secrete filamentous phage or toxins. This analysis resulted in identification of three novel secretin-stress relieving pathways. The regulatory genes and downstream effectors are potential targets that, if inactivated by a drug, would result in killing the secretin-expressing cells in a bacterial population – and by derivation, toxin – or phage-secreting cells in the population. In line with this, parallel work (Khanum, 2015) showed that leads from this thesis proved valid for T3SS secretin InvG.

Chapter 2

2 Materials and Methods

2.1 Bacterial strains, plasmids and growth conditions

All strains used were non-pathogenic laboratory *E. coli* K-12 strains, listed in Table 1. Plasmids used in this project are similarly listed in Table 2.

Unless otherwise stated, all cultures were grown in 2xYT rich media (Sambrook *et al.*, 1989) at 37°C with shaking, supplemented with either 25 µg/mL chloramphenicol (Cm), 50 µg/mL kanamycin (Km), or 5 µg/mL tetracycline (Tet), as dictated by the plasmid's selective marker. Where appropriate, 1 mM IPTG was added when cultures had reached an OD_{600} of 0.2-0.3 to induce protein expression from the plasmids. Protein expression from plasmids was similarly induced by 1 mM IPTG when cells were grown on 2xYT agar plates.

Table 1: Bacterial strains used.

Strain	Parent	Genotype	Source
K1508	MC4100	MC4100, *ΔlamB106*	(Baneyx & Georgiou, 1991)
BW25113		F-,Δ*(araD-araB)567*, Δ*lacZ4787(::rrnB-3)*, λ⁻, *rph-1*, Δ(rhaD-rhaB)568, hsdR514	(Baba *et al.*, 2006)
JW0429-1	BW25113	BW25113, *Δlon::kan*	(Baba *et al.*, 2006)
JW1296-5	BW25113	BW25113, *ΔpspF::kan*	(Baba *et al.*, 2006)
JW1721-1	BW25113	BW25113, *ΔkatE::kan*	(Baba *et al.*, 2006)
JW1732-1	BW25113	BW25113, *Δspy::kan*	(Baba *et al.*, 2006)
JW1935-1	BW25113	BW25113, *ΔrcsA::kan*	(Baba *et al.*, 2006)
JW2205-2	BW25113	BW25113, *ΔrcsB::kan*	(Baba *et al.*, 2006)
JW4023-5	BW25113	BW25113, *ΔsoxS::kan*	(Baba *et al.*, 2006)
JW4024-1	BW25113	BW25113, *ΔsoxR::kan*	(Baba *et al.*, 2006)

Table 2: Plasmids used.

Name	Description	Source
pGZ119EH	Cm^R marker and *colD* origin of replication; protein expression cassette driven by the *tac* promoter.	(Lessl *et al.*, 1992)
pPMR132	Derivative of pGZ119EH; expresses *gIV* wild-type under the control of *tac* promoter	(Russel, 1994)
pPMR132-E292K	Mutant of pPRMR132 containing mutation of *gIV* (G937→A) encoding pIV containing E292→K missense mutation.	(Spagnuolo *et al.*, 2010)
pPMR132-G267D	Mutant of pPRMR132 containing mutation of *gIV* (G863→A) encoding pIV containing G267→D missense mutation.	(Spagnuolo *et al.*, 2010)
pPMR132-V270I	Mutant of pPRMR132 containing mutation of *gIV* (G871→A) encoding pIV containing V270→I missense mutation.	(Spagnuolo *et al.*, 2010)
pPMR132-F288L	Mutant of pPRMR132 containing mutation of *gIV* (T927→G) encoding pIV containing F288→L missense mutation.	(Spagnuolo *et al.*, 2010)
pPMR132-R293C	Mutant of pPRMR132 containing mutation of *gIV* (C940→T) encoding pIV containing R293→C missense mutation.	(Spagnuolo *et al.*, 2010)
pPMR132-I316G	Mutant of pPRMR132 containing mutation of *gIV* encoding pIV containing I316→G missense mutation.	(Spagnuolo *et al.*, 2010)

Name	Description	Source
pPMR132-A329V	Mutant of pPRMR132 containing mutation of *gIV* (C1076→T) encoding pIV containing A329→V missense mutation.	(Spagnuolo *et al.*, 2010)
pBR322	Amp^R and Tet^R markers, *colE*1 (pMB1) origin of replication.	(Bolivar *et al.*, 1977)
pGZt	Tet^R variant of pGZ119EH	This work
pPMR132t	pPMR132 containing the Tet^R marker instead of Cm^R marker.	This work
pPMR132t-E292K	pPMR132E292K plasmid containing the Tet^R marker instead of Cm^R marker.	This work
pCA24N-Lon[a]	pCA24N::*lon, Δgfp*	(Kitagawa *et al.*, 2005)
pCA24N-RcsA[a]	pCA24N::*rcsA, Δgfp*	(Kitagawa *et al.*, 2005)
pCA24N-RcsB[a]	pCA24N::*rcsB, Δgfp*	(Kitagawa *et al.*, 2005)
pCA24N-SoxR[a]	pCA24N::*soxR, Δgfp*	(Kitagawa *et al.*, 2005)
pCA24N-SoxS[a]	pCA24N::*soxS, Δgfp*	(Kitagawa *et al.*, 2005)

[a]pCA24N vector (Kitagawa *et al.*, 2005) contains an expression cassette (encoding an N-terminal his-tag and a C-terminal Gfp tag) under the control of *T5-lac* promoter, Cm^R marker and pUC (high-copy-number) origin of replication. The sequence encoding the C-terminal Gfp tag was deleted from all pCA24N-derived plasmids used in this thesis, as indicated in plasmid description.

2.2 Total Cell Protein Extracts

In each case, 500 μL of culture was mixed with an equal volume of ice-cold 10% trichloroacetic acid (TCA). Precipitated proteins were recovered by centrifugation at 16,060 RCF at 4°C for 10 minutes. Pellets were washed twice with 1 mL ice-cold acetone before resuspending in a volume of 4% SDS such that lysates corresponded to a cell density of 10 OD_{600} units/mL. Lysates were then boiled for 2 min before storing at -80°C.

"Cold competitor" is a Triton X-100 total cell extract of strain K1508 (that does not express pIV). It was used to block pIV polyclonal antiserum from binding to *E. coli* proteins. The cold-competitor extract was prepared in the following manner: 50mL *E. coli* K1508 cultures were grown in 2xYT to an OD_{600} of 0.8 before harvesting by centrifugation. The pellet was resuspended in 1.25 mL of 50 mM Tris-HCl (pH 8) and again pelleted by centrifugation. The pellet was then resuspended in 1.18 mL 50 mM Tris-HCl (pH 8) 5 mM EDTA. To this, 62.5 μg of lysozyme and 100 μL of 100× protease inhibitor (EDTA-free, Roche) was added and incubated at 37°C for 20 minutes. The lysate was then sonicated before a further 100 μL of 100× protease inhibitor (EDTA-free, Roche) was added and incubated for 1 hour at 37°C before storing 50 μL aliquots of cold competitor lysate at -80°C.

2.3 Enzyme Linked Immunosorbant Assay

An ELISA was developed to investigate the time course of PspA expression during secretin stress. TCA extracts of K1508 expressing pIV wildtype (wt),

mutants (G267D, V270I, F288L, E292K, R293C, I316G, A329V), or the vector control (pGZ119EH) were prepared at 0 (before induction), 1, 2, 4 and 8 hrs post-induction as previously described in section 2.2. Samples were diluted into 1x TBS (50 mM Tris-HCl (pH 7.5), 150 mM NaCl) to 0.002 OD_{600} units/100 µL, to reduce SDS concentration below 0.04%; this concentration of SDS does not interfere with ELISA (Lechtzier *et al.*, 2002). A total 100 µL (0.002 OD_{600} units) of diluted TCA extract per sample was subsequently loaded into each well of a 96-well microtitre plate and incubated at 37°C for 1 hr to adsorb proteins. Wells were then washed with 300 µL 1x TBST (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween-20) before adding 200 µL blocking buffer (1x TBS + 2% BSA) and incubating 2 h at room temperature. Post-blocking, wells were washed three times with 300 µL 1x TBST.

To decrease background signal, pIV antiserum was preincubated in 1x TBS containing 0.2% BSA and 50 µL of the cold-competitor lysate (prepared as described in section 2.2) for 1 hour at room temperature prior to dilution. The primary antibodies, α-pIV, α-PspA, and α-Trx, were diluted 1:2000, 1:10000, and 1:2000, respectively, in 1x TBS containing 0.2% BSA. Diluted antibodies (100 µL) were loaded into wells and incubated at room temperature for 1 h. Wells were then washed seven times with 300 µL 1x TBST, followed by addition to each well of 100 µL of the secondary antibody (Horseradish Peroxidase-conjugated anti-Rabbit IgG derived from a Goat (Sigma-Aldrich) in 1x TBS containing 0.2% BSA) at the same dilution as the equivalent primary antisera., The plate was incubated at room temperature for 1 h before washing seven times with 1x TBST and development of the signal.

Aliquots (100 µL) of a commercial developing reagent (1-Step TMB, Thermo-Fisher) containing 3,3',5,5'–Tetramethylbenzidine were loaded into each well and incubated 30 min before stopping the reaction with 25 µL 1M H_2SO_4. Absorbance was measured at 450 nm using a PowerWave XS plate reader (BioTek). Absorbance signals for pIV and PspA were normalised by that of the loading control Trx.

2.4 Cloning and DNA Manipulations

Plasmids pPMR132t and pPMR132t-E292K were derived from pPMR132 and pPMR132-E292K, by replacing the Cm^R marker with Tet^R, to allow selection of double-transformants with ASKA-collection plasmids, which contain Cm^R marker. The source of the Tet^R-encoding gene, plasmid pBR322, was digested by restriction enzyme XmnI (New England Biolabs, USA) to yield 2 fragments; 1932 bp, and 2429 bp with blunt ends. The latter fragment, bearing the Tet^R marker, was purified from a 0.7% agarose gel run in Tris-acetate-EDTA buffer (TAE; 40 mM Tris-HCl, 20 mM Acetic Acid, 1 mM EDTA; (Sambrook *et al.*, 1989)), using the ZymoClean gel DNA recovery kit (Zymo Research, USA). This fragment was ligated to large fragment (4073 bp) of pPMR132-E292K, obtained after ScaI/XmnI digest (New England Biolabs, USA), using T4 DNA Ligase (Roche), to construct pPMR132t-E292K. pPMR132t, was constructed by replacing the sequence encoding C-terminal domain of pIV including the E292K mutation of pPMR132t-E292K with the wild-type sequence. Both pPMR132 and pPMR132t-E292K were cut with XbaI/PstI (New England Biolabs, USA); the small XbaI/PstI fragment of pPMR132 (804 bp), containing the DNA sequence encoding the C-terminal portion of wild-type pIV was ligated to the large (5706

bp) XbaI/PstI pPMR132t-E292K fragment using T4 DNA ligase (Roche), to obtain pPMR132t. The vector control, pGZt, was constructed by removing *gIV* sequence from pPMR132t-E292K by PstI/BglII (New England Biolabs, USA) digest, followed by a blunting reaction using the T4 DNA polymerase-based Quick Blunt kit (New England Biolabs, USA) and subsequent recircularisation of the 5048 bp fragment.

2.5 Total RNA Extraction

Three independent colonies of strain K1508 transformed by each of the three plasmids: pGZ119EH (vector control), pPMR132 (expressing wildtype pIV), or pPMR132-E292K (expressing leaky mutant pIV-E292K) were picked to inoculate separate 5 mL overnight 2xYT cultures (nine cultures in total). The overnight cultures were diluted 1/100 into 20 mL 2xYT medium supplemented with 25 µg/mL Cm. Once an OD_{600} of 0.2 - 0.3 was achieved, IPTG was added to a final concentration of 1 mM to induce the *tac* promoter of the plasmids' expression cassette.

The RNA stabilization agent, RNAprotect (Qiagen), was found to be incompatible with this RNA extraction protocol and was not used. The cells were rapidly chilled by placing on ice and pelleted by centrifugation, followed by exposure to cold phenol, which instantly kills the cells and denatures all proteins. All samples were treated identically during these steps, equalising any possible effects due to temperature-shock (cooling) or shear forces (from centrifugation). If any, this approach would likely induce the transcriptional response of genes involved in cold-shock or shear-force stress equally in all

samples. Consistently, no significant differential expression of these genes was identified among the analysed samples (Chapter 4).

In detail, one hour post-induction, 5 mL of culture was harvested by centrifugation (2 min, 12,000 RCF at 4°C). Supernatant was aspirated and the pellet resuspended in 1 mL resuspension buffer. This buffer (20 mM NaOAc pH 5.5, 1 mM EDTA, 1% SDS) was prepared in DEPC-treated Milli-Q water and washed with phenol:chloroform). An equal volume of cold acid phenol:chloroform (1:1 mix) was added to the resuspended cells, and 0.2 mg of a slurry of acid washed glass beads (Sigma-Aldrich) in phenol:chloroform was added to each tube to assist cell disruption by vortexing at high speed for 6 minutes (3 × 2 min pulses) at 4°C. Phases were separated by centrifugation (5 min, 4500 RCF, 4°C) and the aqueous phase removed to a clean tube. An equal volume of phenol:chloroform was added, vortexed briefly and centrifuged again. This step was followed by separate phenol, then chloroform washes, to remove the remaining protein and phenol, respectively. After the final chloroform wash, the aqueous phase was transferred to a clean tube and precipitated by adding 1/10 volume of 3 M NaOAc pH 7 and 2.5 volumes of ethanol (Mol. Bio grade, Merck). Tubes were mixed by inversion and placed at -80°C overnight before centrifugation (30 min, 12,000 RCF, 4°C). The supernatant was aspirated and the pellet air-dried for 20 min before resuspending in 50 µL TE buffer.

Genomic DNA was removed by adding 10 µL of a master mix containing 2 units TURBO DNAse I (Ambion), 2 units Superase-In RNAse inhibitor (Ambion), in 1x

TURBO DNAse I buffer (Ambion). The master mix was allowed to incubate at room temperature for 20 min before adding to the RNA samples, to allow the RNAse inhibitor to bind any contaminating RNases. DNAse I digest proceeded at room temperature for 30 min, before repurification by phenol:chloroform.

Samples were split into two equal-volume aliquots after the final chloroform wash and ethanol-precipitated separately. One sample was resuspended in 50 µL of DEPC-treated MQ, 5 µL of which was analysed by the Agilent Bioanalyser 2100 to determine quality of the RNA extraction at the Massey Genome Centre (Palmerston North, New Zealand) to be run. The second pellet was not resuspended, but was rather preserved as a pellet. Briefly, 1 mL of 95% ethanol (Mol. Bio grade, Merck) was added to the pellet; the tube then was sealed with Parafilm and stored at -80°C before dry-ice shipping to Ambry Genetics (Los Angeles, USA) for RNA sequencing. Samples with an RNA integrity number (RIN) of 8 or above were sent for sequencing; if a sample had a score below this cut-off value, it was discarded and a new sample was prepared.

2.6 RNA sequencing

To sequence the mRNA transcriptome from each total RNA sample, ribosomal RNA was first depleted from the samples using the Ribo-Zero rRNA depletion kit (EpiCentre), avoiding the use of spin columns – such columns do not bind small RNAs which are potentially involved in post-transcriptional regulation of *E. coli* gene expression. This and subsequent library preparation and sequencing steps were performed by Ambry Genetics (Los Angeles, USA). Paired-end 100 bp directional stranded RNA libraries were prepared using the Illumina TruSeq

Stranded sample preparation kit and subsequently run on 2 lanes of the Illumina HiSeq 2000 platform. The small RNAs from the rRNA-depleted samples was also sequenced by preparing 54 bp single-end libraries using the Illumina TruSeq small RNA library preparation kit and running on 1 lane of the Illumina Genome Analyser. Ambry Genetics performed all library preparation, subsequent quality control, sequencing and demultiplexing of reads.

2.7 Bioinformatics and analysis of the sense transcriptome

Overall quality of raw reads in the libraries obtained from Ambry Genetics (Los Angeles, USA) was determined by FastQC (Andrews, 2010). The read trimming program, Trimmomatic v.0.30, was used to remove poor quality regions at the 5' and 3' ends (Bolger *et al.*, 2014) when the average Phred33 quality score fell below 20 over a window of 4 nucleotides. Illumina Clip (a function within Trimmomatic) was then used to identify and trim any Illumina sequencing primers, removing them from the pool of reads. Trimmed reads were then passed through FastQC again to confirm that the read quality had improved and that reads containing Illumina primer sequences had been removed. At the time of data-acquisition, *E. coli* K-12 strain BW2952 was determined to be the closest sequenced relative of our experimental strain K1508. Both, K1508 and BW2952 are derived from MC4100; K1508 contains a *lamB* deletion mutation, while BW2952 contains a *lacZ* reporter fused to *malG*, otherwise they can be considered to be the same (in the absence of a complete K1508 sequence) (Ferenci *et al.*, 2009). Therefore, BW2952 cDNA and ncRNA gene models, obtained from Ensemble Bacteria (Kersey *et al.*, 2014), were used as the reference gene set in the Bowtie2 alignment (Langmead & Salzberg, 2012).

The sequences for plasmid-derived transcripts (*gIV*, *RNAI*, *RNAII* and *lacI*^{q}) were added to the collated cDNA and ncRNA gene models. Sense and antisense reads that aligned to the gene models were separated using the –norc and –nofw options, respectively, and returned by Bowtie2. The number of sense and antisense reads aligning to each reference gene model was then counted using a custom Perl script (available on request). Prior to differential gene expression analysis by DESeq2, the count data was filtered to remove genes with zero counts across all samples. Biological variation within replicate groups was constrained by removing genes whose counts had a coefficient of variation >0.7. The coefficient of variation (CV) is a standardised measure of the dispersion of a frequency distribution and is defined as the ratio of the standard deviation to the mean, thus gives a measure of how much variability exists in a set of samples/replicates. Genes that did not show differences in expression between replicate groups were also removed by filtering genes that had a coefficient of variation <0.3 between replicate groups. Finally, all genes with fewer than 10 counts in any sample replicate were removed. The filtered count data was then used to determine which genes in the wildtype pIV and pIV-E292K producing cultures were differentially expressed relative to those that did not produce pIV (the vector control; K1508 + pGZ119EH). Statistical significance was corrected for multiple testing by the Benjamini-Hochberg method (Benjamini & Hochberg, 1995). This was achieved using the statistics software R and the DESeq2 package (Anders & Huber, 2010, R Development Core Team, 2013).

2.8 Synthetic lethality and complementation

Due to the large number of transformations in this experiment, procedures for plasmid transformation were streamlined to increase the throughput and shorten the time required to complete the experiments. Competent cells of each Keio strain containing each a single chromosomal mutation of interest (Table 1, (Baba *et al.*, 2006)) were prepared by diluting a fresh overnight culture 1/100 into prewarmed (37°C) 2xYT and incubating with aeration for 1 hour before harvesting cells by centrifugation (Heraeus, Labofuge 400R) and washing twice in ice cold 10 mM $CaCl_2$, 10% glycerol. Aliquots of competent cells (50 µL) were transformed by 100 ng of pGZt, pPMR132t or pPMR132t-E292K, only, or as a mixture with an equivalent amount of the appropriate complementing ASKA plasmid (Table 2; (Kitagawa *et al.*, 2005)). The transformants were selected for on plates using the combinations of antibiotics corresponding to the plasmid markers. Three colonies from each transformation plate were picked to inoculate three 5 mL overnight cultures. To determine the effect of pIV production on viability of transformed cells, serial dilutions of the cultures were made. Undiluted cultures, as well as 10^{-2}. 10^{-4} and 10^{-6}- fold dilutions (25 µL each) were plated onto 2xYT agar containing 5 µg/mL Tet, in the presence or absence of 1 mM IPTG. For complementation experiments, the medium also contained 25 µg/mL Cm to maintain the ASKA complementing plasmids. Plating efficiency of cells transformed with pPMR132t and pPMR132t-E292K in the presence or absence of IPTG was calculated relative to the vector control (pGZt).

Chapter 3

3 Choice of the secretin leaky mutant for transcriptome analysis and validation of the secretin stress transcriptome

Expression of secretins with deregulated gates mimics the stress of virulence factor secretion by enteric bacteria during infection (Horstman & Darwin, 2012, Spagnuolo *et al.*, 2010). Moreover, negative selection for virulence factor secretion under laboratory conditions, where selective pressures of the host environment are absent, has been recently demonstrated (Sturm *et al.*, 2011). The stress of virulence factor secretion is likely complex, with multiple factors perturbing the inner and outer membranes, as well as the pressure imposed by the expression of a large amount of secreted substrates, be they phage filaments or virulence factors. For this reason we sought to analyse the stress of “closed” vs “open” or “leaky” secretins, as models of resting and secreting states. Although various secretins analysed to date have variable degrees of permisivity or “leakage” (Burghout *et al.*, 2004, Marciano *et al.*, 1999, Marciano *et al.*, 2001, Nouwen *et al.*, 1999), it is not possible to compare them to each other as they have other unrelated differences, such as general structure or accessory proteins required for outer membrane targeting or different modes and accuracy of targeting, all of which can impose different stresses unrelated to gate opening. Therefore, it was decided to compare the stress imposed on cells by wildtype pIV secretin as a model of a resting-state secretin, whose pore size is smaller than that of the porins, and that of an isogenic severely leaky mutant of the same secretin (pIV-E292K).

This chapter describes the experiments that were carried out to choose the pIV leaky secretin mutant as model for the toxin-secreting secretion systems. It also

details the strategy for transcriptome analysis as well as initial results that monitor the quality of transcriptome data.

3.1 Characterisation of the pIV secretin mutants

3.1.1 Comparison of leaky pIV mutants

The ELISA assays were performed using samples taken from cultures over an 8 hour time-course, post-induction of pIV expression. Furthermore, as a negative control, an engineered folding mutant (pIV-I316G), that was found in low amounts in the cells and did not appreciably induce the Psp response was included for comparison (Table 3; Spagnuolo *et al.* (2010)). A desirable mutant for transcriptome analysis would have a severely leaky phenotype (reflecting a gate that opens at physiological conditions; Marciano *et al.* (1999)), be functional in filamentous phage assembly (Table 3) and is present in a high amount in the cell (indicating low turnover by periplasmic proteases and hence rapid folding), while inducing the Psp response. The most severely leaky mutant in the collection was pIV-E292K, demonstrating the highest sensitivity to Bac, Van and DOC in the laboratory collection. This mutant was also functional in filamentous phage assembly (30% that of the wild-type pIV).

Table 3: Properties of cells expressing pIV mutants.

	Mutation[a]	f1 assembly[b]	Mp[c]	Van MIC[d]	Bac MIC[d]	DOC Mm[e]
WT pIV	- (wildtype pIV)	+++	N	R	R	2±0
Vector	NA[f] (no pIV)	-	N	R	R	0±0
GATE1	G267D	+++	P	R	R	3±0
	V270I	++	P	3±0	21±3	3±0
	F288L	-	P	R	R	2±0
	E292K	+	P	1±0	3±1	12±1
	R293C	+++	P	R	R	2±0
GATE2	A329V	+++	P	R^-	R	2±0
Folding mutant[g]	I316G	-	N	R	R	0±0

[a] All data in this table are taken from Spagnuolo *et al.* (2010) .

[b] To assess f1 assembly function, dilutions of *gIV* deletion phage R484 were spotted on cells that expressed mutant or wildtype (WT) *gIV* and on vector-containing cells that lacked *gIV*. +++, e.o.p. 0.9-1 and normal plaque size compared to wt f1; ++, e.o.p. 0.5-0.8 and reduced plaque size; +, e.o.p. 0.3-0.4 and very small plaques; +/-, e.o.p. 0.1-0.3, barely detectable plaques; -, e.o.p ~10^{-5} (equivalent to plating on cells containing the vector (pGZ119EH) that lacks *gIV*).

[c] Permisivity to maltopentaose (Mw = 829 Da), demonstrated as ability of a pIV variant to allow growth of a maltoporin mutant on maltopentaose as the sole carbon source. P, permissive; allows growth on maltopentaose; N, non-permissive, does not allow growth on maltopentaose.

[d] Minimal inhibitory concentrations of vancomycin (Mw = 1449 Da) and bacitracin (Mw = 1422 Da) as measured by concentration gradient strips (in µg/mL) are indicated. R, completely resistant at the highest concentration tested (256 µg/ml); Each value was obtained from two cultures (three strips per culture) and includes the standard error of the mean.

[e] Mean annular radii (mm) of growth inhibition zones around discs pre-loaded with 50 µL of 10% DOC and dried prior to use. Average micelle size of DOC is 1,200 - 5,000 Da and critical micellar concentration 0.08 – 0.25%. Each value was obtained from two cultures (three discs per culture) and includes the standard error of the mean.

[f] Not applicable.

[g] Constructed by site-directed mutagenesis in a predicted transmembrane segment between the two GATE regions. Confirmed to be a folding mutant (Spagnuolo *et al*, 2010).

The amount of pIV produced was different among mutants (Figure 9). Whereas leaky mutants, including pIV-E292K, were found at equal or higher amounts as the wildtype pIV, the unstable I316G mutant was present in the least amount, validating quantification (Figure 9A). The amount of pIV produced did not correlate with the position of mutation within the C-terminal homology domain or PspA expression. The production of pIV appears to peak 2 – 4 hours post-induction for most mutants, which is consistent with the pIV half-life being around 3 hours (Spagnuolo *et al.*, 2010). Maximal PspA accumulation does not occur until 4 hours post-induction and seems to plateau at 8 hours or later (Figure 9A & B). Eight hours post-induction, the differential ability of mutants to induce PspA appears to split into three groups; high inducers (G267D, R293C); wildtype-like (E292K); and low inducers (V270I, F288L and A329V), while the amount of PspA in cells expressing the unstable mutant pIV-I316G contained the same level as in the cells not expressing any pIV (transformed with empty vector; Figure 9B).

It is worth noting that reported microarray analyses of secretin-stress transcriptomes were sampled 1 hr post-induction of secretin expression (Lloyd *et al.*, 2004), at which time the phage shock protein regulon mRNA is induced by a factor of 10 or more; however, PspA accumulation lags behind production of the *psp* mRNA. Nevertheless, the amount of PspA protein in cells expressing the most severely leaky mutant, pIV-E292K, is highest of all at that time point, albeit other mutants were associated with higher PspA accumulation at later time-points. Consistently with the severely leaky phenotype, cells producing pIV-E292K (Figure 9C) showed decreased growth rate post-induction. This

property was noted for other severely leaky secretins (Horstman & Darwin, 2012, Marciano *et al.*, 1999, Spagnuolo *et al.*, 2010).

Based on the severely leaky phenotype, functionality in phage assembly, stability of protein and induction of the Psp response, mutant pIV-E292K was chosen for the transcriptome analysis.

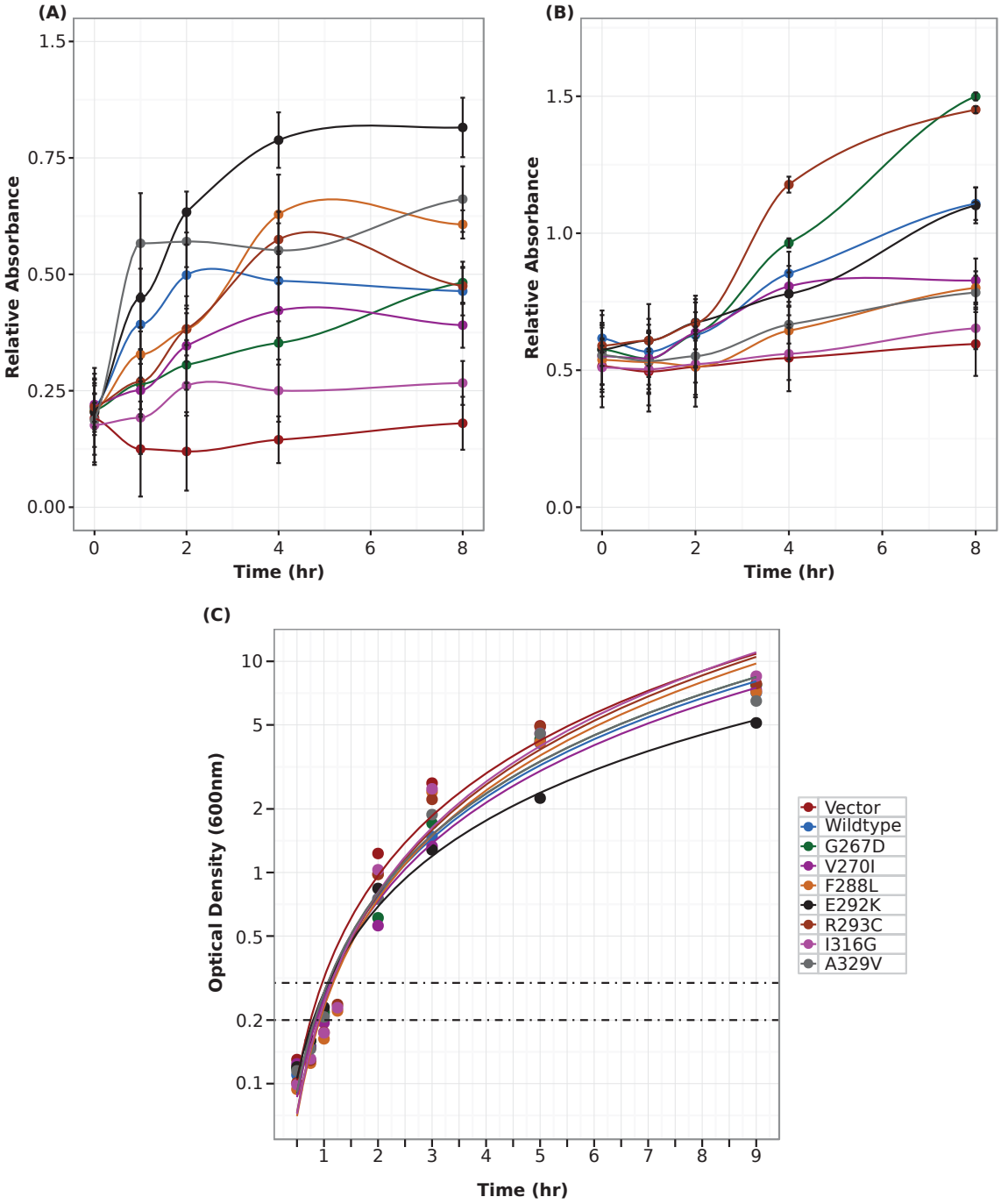

Figure 9: Effect of pIV GATE mutations on the expression of PspA and growth of *E. coli* K1508.

(A) Amount of pIV relative to TrxB (loading control)**. (B)** Amount of PspA relative to TrxB (loading control). Error bars in **(A)** and **(B)** repesent the standard error of the mean for 3 technical replicates; curves were fitted by locally weighted regression; time is relative to the point of induction. **(C)** Growth curves of K1508 expressing a variety of pIV mutants. These cultures were the source of total protein extracts used in the ELISA. pIV expression was induced by 1 mM IPTG once cultures reached OD_{600} 0.2 – 0.3 (dashed lines). Log regression fit the growth curves. Legend applies to all three graphs in this figure.

3.2 Experimental design of the secretin stress transcriptome analysis

Comparison of the global transcriptional response stress imposed by wildtype secretin pIV *vs.* a gate-deregulated mutant (pIV-E292K) was undertaken here as a well-controlled approach to clearly separate the stress of compromised outer and inner membranes from that caused by secretion of the substrates.

To assess the transcriptomic response of *E*. *coli* to the gate-disregulated secretin, rRNA-depleted mRNA, derived from *E*. *coli* K-12 strain K1508 (Table 1) cultures expressing either wildtype pIV or the severely leaky mutant, pIV-E292K, were sequenced by next-generation Illumina technology (RNA-seq), using a stranded 100 bp paired-end strategy that determined from which strand the transcript was derived, hence the "sense" expression was clearly distinguished from the "antisense". The secretin-expressing transcriptomes were compared to a control that did not express a secretin (but contained the cloning vector from which the secretins were expressed). Three biological replicates of each strain were analysed.

3.2.1 Pre- and Post-sequencing Quality Control

In order to obtain a reliable representation of the transcriptome, it is critical that starting total RNA material extracted from cultures be intact. To determine the quality of RNA, the samples were analysed by electrophoresis in the Agilent BioAnalyser (Figure 10; Schroeder *et al.* (2006)). All lanes had integrity scores (RIN) >8.5 (the recommended cut-off is 8). Furthermore, they showed no sign of genomic DNA (gDNA) contamination, usually visible as a diffuse band running above the 23S rRNA (Figure 10). While there are faint bands detectable

between the 16S and 5S rRNA of wildtype-1, wildtype-2, E292K-2 and E292K-3 (Figure 10B), the RIN scores for these samples indicated that RNA was intact and of sufficient quality for preparation of libraries for transcriptome sequencing by RNA-seq.

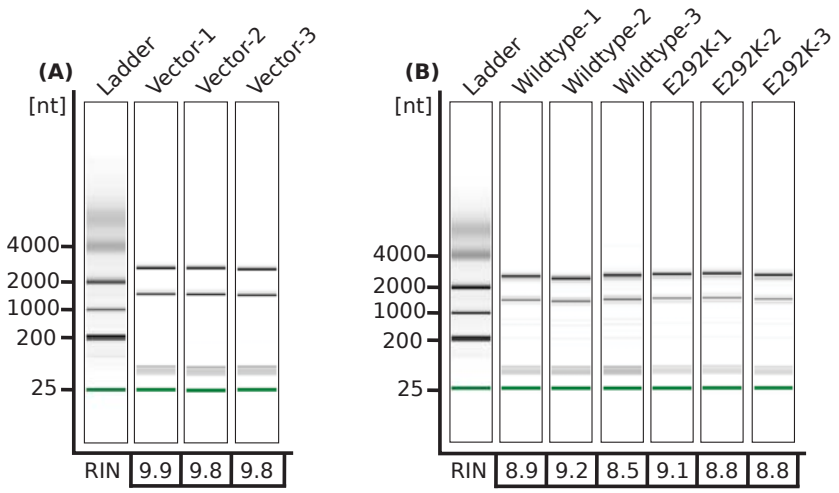

Figure 10: Integrity of total RNA.

The quality of purified total RNA for transcriptome sequencing was analysed by electrophoresis in the Agilent BioAnalyser. The RNA integrity number (RIN) determined by Agilent Bioanalyser 2100 software is shown below the corresponding lane. **(A)** Lanes: Ladder, the RNA standard (sizes indicated along the length of the electrophoregram); Vector–1 to Vector–3; total RNA isolated from three cultures containing the empty vector pGZ119EH. **(B)** Lanes: Ladder (as in **(A)**); wildtype–1 to wildtype-3, and E292K–1 to E292K-3, total RNA isolated from three cultures expressing each the wild-type pIV fom plasmid pPMR132 or mutant pIV-E292K from plasmid pPMR132-E292K, respectively (**Table 2**). The strains were derived by transformation of K1508 with the appropriate plasmid (**Table 2**).

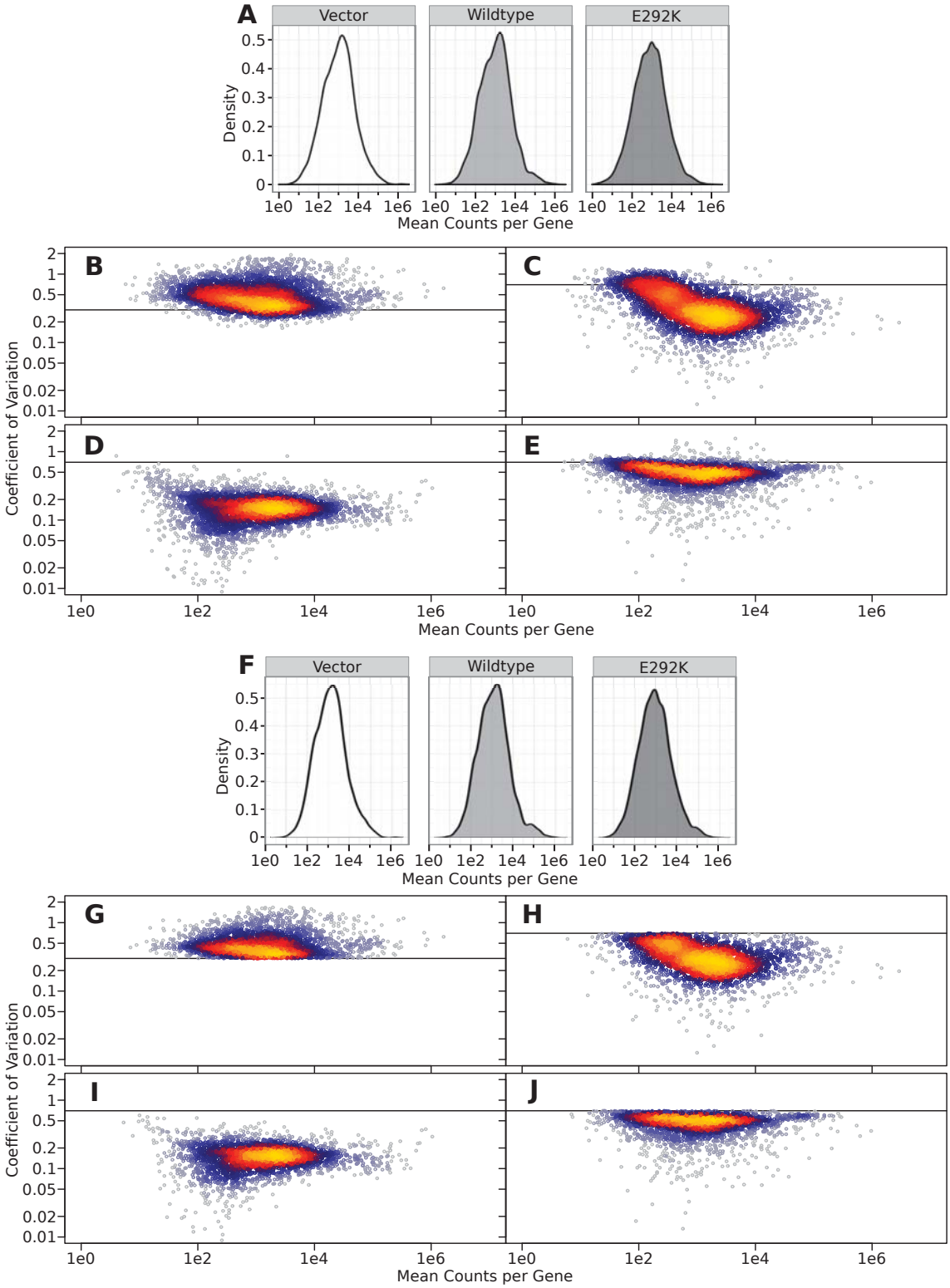

Figure 11: Variability of Counts per gene within and across sample replicates.

(A) Distribution of mean counts per gene prior to removing highly variable genes within sample replicates (e.g. within E292K-1, E292K-2, E292K-3, or Vector-1, Vector-2, Vector-3) and genes that do not show variability across samples (e.g. wildtype *vs.* vector *vs.* pIV-E292K). **(B)** Variability of gene counts across all samples prior to filtering – differentially expressed genes should have a high coefficient of variation. The black horizontal line indicates the 0.3 CV cut off used to filter genes unlikely to be differentially expressed. **(C, D, E)** Variability of gene counts within Vector, Wildtype and pIV-E292K replicates, respectively, prior to filtering. The black horizontal line indicates the 0.7 CV cut off used. **(F)** Distribution of counts after removing highly variable genes within samples (CV>0.7, points below the line in **C**, **D**, and **E**) and those genes which are unlikely to be differentially expressed (CV<0.3, points above the line in **(B)**). **(G)** Variability of counts across all samples after removing genes unlikely to be differentially expressed (CV<0.3, points above the line in **B**). **(H, I, J)** Variability of gene counts within Vector, Wildtype and pIV-E292K replicates, respectively, after removing unreliable counts (CV>0.7, points below the line in **C**, **D**, and **E**). The black horizontal line indicates the 0.7 CV cut off used. Heat coloring of points in **B, C, D, E, G, H, I, & J**, indicates the density of points in that region – hotter colours (reds to yellows) indicate greater density of points, cooler colours (blues to greys) indicate lower density of points.

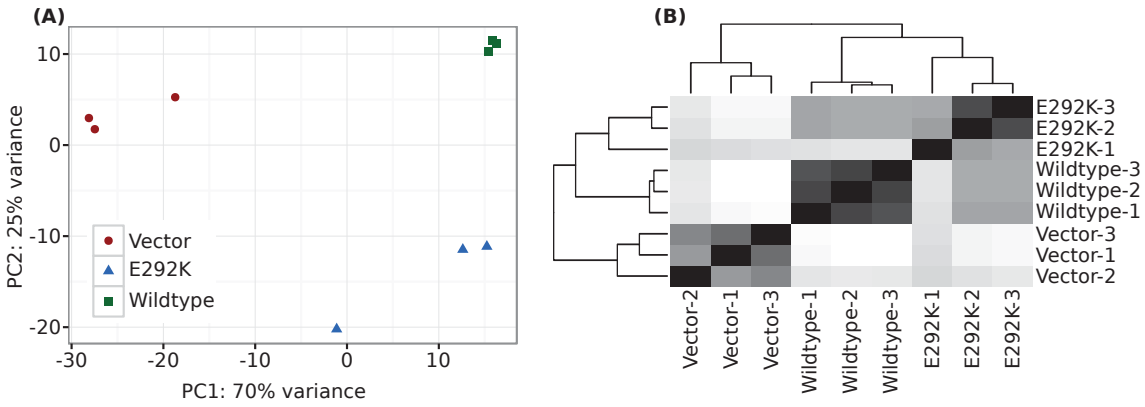

Figure 12: Variation of normalised transcript counts.

(A) Principal component analysis of regularized log-transformed normalised counts. **(B)** Euclidean distance using regularized log-transformed normalised counts. Dendrogram lengths are relative.

The stranded sequence libraries, one for each of the nine sequenced transcriptomes, were prepared and sequenced by the service provider (Ambry Genetics, USA). Each of the nine libraries was distinguished from others by a specific barcode introduced during library preparation. The sequence data from each library was aligned to the 4255 reference cDNA genes (coding and non-coding) from *E. coli* BW2952, obtained from Ensembl Bacteria, using Bowtie2 (Ferenci *et al.*, 2009, Kersey *et al.*, 2014, Langmead & Salzberg, 2012). For each sequence library, the number of sequencing reads aligned to each gene was counted using a custom Perl script (Cox, 2013, http://mpcox.github.io/mapcount/) to create a "count library" for each transcriptome.

To increase the power of statistical testing, a filtering strategy was used to remove unreliable and noisy data. The coefficient of variation (CV), defined as the ratio of the standard deviation to the mean, provided a measure of each gene's count variability among the replicates of the same sample (to remove unreliable gene counts) and across all replicates (to remove genes that are unlikely to be differentially expressed). Genes with highly variable counts (CV>0.7) between replicates of the same sample are unlikely to provide reliable estimates of differential expression. Additionally, genes that do not show any count variation across all samples and replicates (regardless of sample type, CV<0.3) are unlikely to be differentially expressed, thus their removal decreases the rate of false discovery by decreasing the number of tests performed and eases analysis by removing uninteresting data. After removal of all genes with zero counts in any sample replicate (Section 2.7), 4081 out of

4260 genes, including genes encoded by plasmids pGZ119EH (vector), pPMR132 (expressing the wildtype pIV) and pPMR132-E292K (expressing the leaky pIV mutant; Table 2) remained for further filtering using the described strategy.

The number of genes with a high CV was greatest in the vector replicates and lowest in the wildtype replicates, with 437 and 2 genes, respectively, displaying a CV >0.7 (Figure 11C & D). However, the average CV was highest in the pIV-E292K replicates and lowest in the wildtype replicates – 0.49 and 0.16, respectively (Figure 11D & E). The counts within the wildtype replicates appear to be the most consistent, and therefore most reliable (Figure 11C, D, E & H, I, J). Globally, most genes appear to show variation across all sample replicates, with only a small fraction (515 genes with CV < 0.3) showing no or very little variation (Figure 11B, G). Collectively, 1091 genes were removed by the sample CV and global CV filters used. Finally, to ensure the genes being analysed are expressed genes, all genes with fewer than 10 counts per replicate were removed from the dataset, culling a further 29 genes.

The distribution of mean counts per gene for each sample did not change radically (Figure 11A & F) after filtering by the coefficient of variation. The likely reason for this being that, the mean counts of genes that were removed due to low CV across all samples or high variability within the pIV-E292K replicates (Figure 11B & E), were roughly centred around the mode of the distributions for each sample (Figure 11A & F). The exception is the distribution of mean counts per gene within the vector replicates, which did change somewhat (Figure 11A

& F). The mean counts of the genes removed due to high variability within these replicates were skewed to the left of the distribution mode (Figure 11C).

Of the 4255 coding and non-coding genes aligned against, only 2961 genes remained in the count libraries after the filtering steps. These count libraries were analysed using the R package, DESeq2 to normalise the count libraries and perform differential expression testing (Love *et al.*, 2014, R Development Core Team, 2013). To determine whether normalisation is correct and no abnormalities exist within the data - i.e. unusual gene expression or dissimilar replicates – biological replicates were analysed for 'sameness'. The distribution of significantly differentially expressed genes, as a first measure of the transcriptome's sensitivity, was determined in parallel.

A principal component analysis (PCA) and heat mapping the Euclidean transcriptome-to-transcriptome distance was first used to determine the relationships within and among the replicates. Two issues encountered during library preparation are template jumping during RNA-seq library amplification and incorrect demultiplexing of bar-coded samples that are sequenced together. These issues were either not encountered, or negligibly affected the data. If these issues did occur one would expect the PCA to show larger sample-to-sample distances and overlap between replicates of different samples.

If template jumping is low and the demultiplexing is correct, the replicates of the same samples (defined as derived from genetically identical transcriptomes,

obtained from cultures containing the same plasmid in this work) should be statistically comparable to one another and contrasted to replicas of other samples (transcriptomes from cultures containing a different plasmid). Before performing the PCA, a regularised log-transformation (a function built into DESeq2; Equation 1) was applied to the normalised counts to reduce the effects of a few highly expressed genes, which can vary considerably within the sample replicates. The PCA analysis of thus treated data showed that the replicates of the same sample clustered together and did not overlap considerably with replicates from other samples (Figure 12).

The E292K-1 replicate had greater distance from its siblings (E292K -2 and –3). This is correlated with a much smaller number of counts (i.e. the total number of reads) in the count library of E292K-1 in comparison to that of other replicates, requiring stronger normalisation. Inspection of pIV-E292K, wildtype and vector reads aligning to *gIV* revealed a low number of reads from other samples (e.g. 2 pIV-E292K to 106,510 wt pIV for a wildtype sample), indicating template jumping. The larger sizes of E292K -2 and -3 replicate read counts meant that the relative number of reads from other libraries was smaller. The small library size of the E292K-1 replicate probably amplified the effect of mixed reads, resulting in the observed discrepancy from the other two pIV-E292K sample replicates. Thus, while statistical power of some tests will be diminished slightly, the design of differential expression testing is not sensitive the issues seen in the E292K-1 replicate.

Equation 1: Regularised log transformation function of DESeq2

$$\log_2(q_{ij}) = \beta_{i0} + \beta_{ij}$$

q_{ij} is a parameter proportional to the expected true concentration of reads for gene *i* and sample *j*. β_{i0} is an intercept that does not undergo shrinkage (regularisation) and β_{ij} is the sample-specific effect which is shrunk toward zero based on the dispersion-mean trend over the entire dataset. This transformation accounts for differences in sequencing depth (Love *et al.*, 2014).

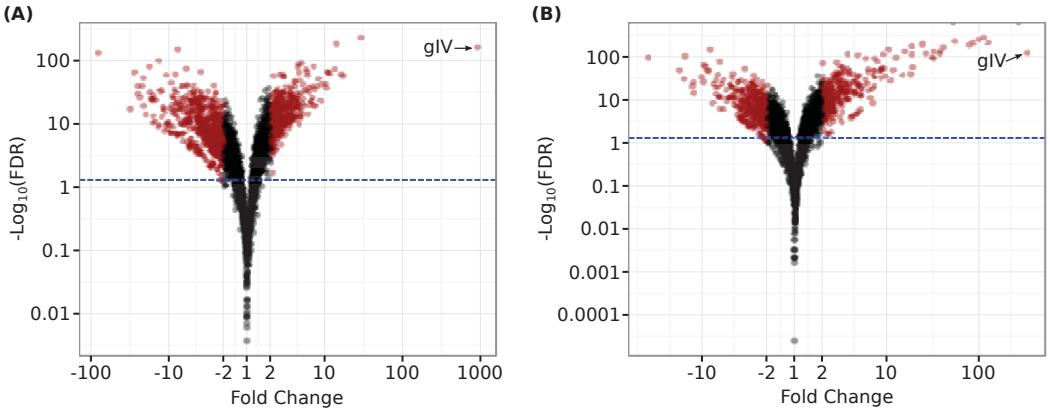

Figure 13: Rate of false discovery within differentially expressed genes.

False discovery rates among the genes differentially expressed in the transcriptomes from wildtype pIV-expressing **(A)** and pIV-E292K mutant-expressing **(B)** cultures**,** relative to that of the vector control cultures. Red colouring of points indicate change of a factor of at least 2. Blue dashed line indicates the false discovery rate (FDR) = 0.05, the cutoff used in analysis. The point corresponding to gene IV transcript (gIV) is labelled in each chart. Y-axis shows the $-\log_{10}$ of FDR (Benjamini-Hochberg adjusted p-values (Benjamini & Hochberg, 1995)). X-axis displays the shrunken fold-changes reported by DESeq2.

The R package DESeq2 uses a negative binomial generalised linear model to estimate expression values for each gene and from this calculates differential expression, applying a Wald test using the $\log_2$ fold-change standard error to determine whether $\log_2$ fold-change is equal to zero (Love *et al.*, 2014). The count libraries obtained from cells experiencing secretin stress (wildtype or the mutant, pIV-E292K) were contrasted against a count library obtained from cells transformed by an empty vector, thus were not experiencing secretin-dependent stress. The lower limits of significant differential gene expression testing appear to be around +/- 2-fold-change for both wildtype and pIV-E292K *vs.* vector tests.

3.3 Discussion

3.3.1 pIV mutant characterisation

The pIV-E292K leaky mutant was chosen for RNA-sequencing and transcriptomic analysis based on its severely leaky phenotype, wildtype-like accumulation of pIV suggesting fast folding, functionality in filamentous phage secretion and wildtype-like induction of PspA. Although some growth retardation (increased generation time) was observed for the pIV-E292K-expressing cells, this was expected based on the severely leaky phenotype of the mutant; in this respect pIV-E292K was similar to severely leaky secretins such as those from T3SS (Horstman & Darwin, 2012).

3.3.2 Sampling time

Analyses of the PspA accumulation time-course showed that at the 1 hour timepoint post-induction of pIV, the amount of PspA is still very low. However,

all published microarray experiments were performed at 1 hour post induction. At that timepoint, the *pspA* transcript is increased 10-fold (Spagnuolo et al., 2010), even though the PspA accumulation is less than 1.5-fold. Nevertheless, the 1 hour timepoint was chosen for transcriptome analysis by RNAseq in this work, to allow comparison with the published microarray data for wildtype pIV (Jovanovic *et al.*, 2006, Lloyd *et al.*, 2004). Another advantage of this timepoint is that it avoids physiological changes caused by entry of the culture into the stationary phase of growth and allows focusing of the transcriptional response to secretin expression.

3.3.3 Transcriptome - quality assessment and preliminary analysis

The quality assessment of the transcriptome by PCA showed that the three biological replicas for each sample clustered together. Minor template jumping (cross-contamination of libraries) has been observed, however the number of cross-contaminating reads was too low (~2 in 100,000) to interfere with the results of analysis.

Chapter 4

4 Characterisation of the secretin stress transcriptome

Previous work that analysed the wildtype pIV stress response using microarrays, only identified up-regulation of the Psp regulon with minor differential expression of genes involved in motility and protein assembly/secretion (Jovanovic *et al.*, 2006, Lloyd *et al.*, 2004). The analysis presented in this thesis identified a number of genes that were significantly up-regulated or down-regulated by both the wildtype pIV and the permissive mutant pIV-E2924K. In the following sections, general analysis of the functions enriched in various significantly changed gene populations will be assessed, followed by overview of envelope stress responses and detailed analysis of affected genes and pathways.

4.1 Differential Gene Expression under secretin stress

Of the 2961 genes in each count library (after pre-filtering), DESeq2 determined that 2309 and 1895 were significantly differentially expressed (DE), reflected by the expected probability that the finding is a false positive (i.e. the false discovery rate, FDR) of less than 0.05 in the wildtype pIV and pIV-E292K *vs*. vector analyses, respectively. Of the 2309 DE genes in the wildtype pIV producing cells *vs.* vector, 388 and 667 genes, respectively, were up-regulated or down-regulated by, at least, a factor of 2 (Figure 13A). Among the 1895 significantly DE genes in the pIV-E292K *vs*. vector analysis, 312 and 367 genes were up-regulated or down-regulated, respectively by, at least, a factor of 2 (Figure 13B). Comparison of the genes within the two transcriptomes revealed only 13 genes up-regulated in response to pIV-E292K that were down-regulated

in response to wildtype pIV and none that are down-regulated in response to pIV-E292K that were upregulated in response to wildtype pIV (Figure 14).

Initial analysis of genes significantly differentially expressed in cells producing both the wildtype pIV and pIV-E292K by heat mapping revealed a common set of differentially expressed genes that are predominantly down-regulated (Figure 15). However, a small subset of genes (13) that are up-regulated in response to production of the leaky pIV-E292K mutant, is down-regulated in response to wildtype pIV production (Figure 14, Figure 15C). This smaller subset of genes appears to specifically respond only to a leaky secretin, used here to model the physiological state of an open secretin during substrate secretion (Figure 15B).

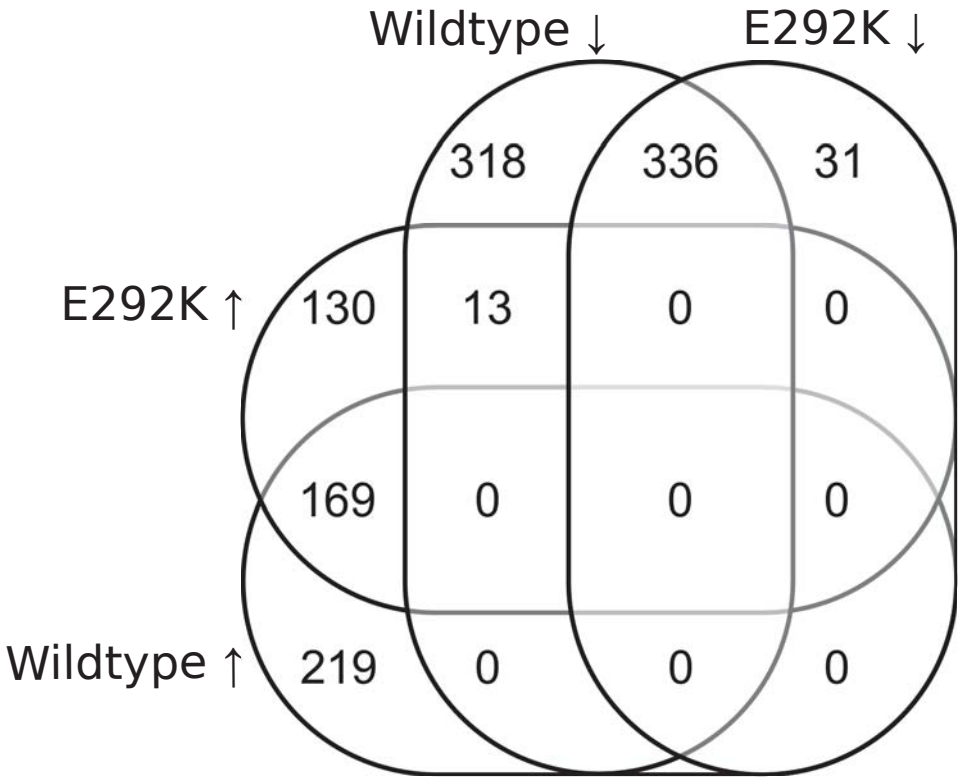

Figure 14: Comparison of significant differential gene expression in response to the production of wildtype pIV or pIV-E292K.

FDR < 0.05, Fold-change > 2 for Wildtype ↑ and E292K ↑; Fold-change < -2 for Wildtype ↓ and E292K ↓. n(Wildtype ↑) = 388; n(Wildtype ↓) = 667; n(E292K ↑) = 312; n(E292K ↓) = 367.

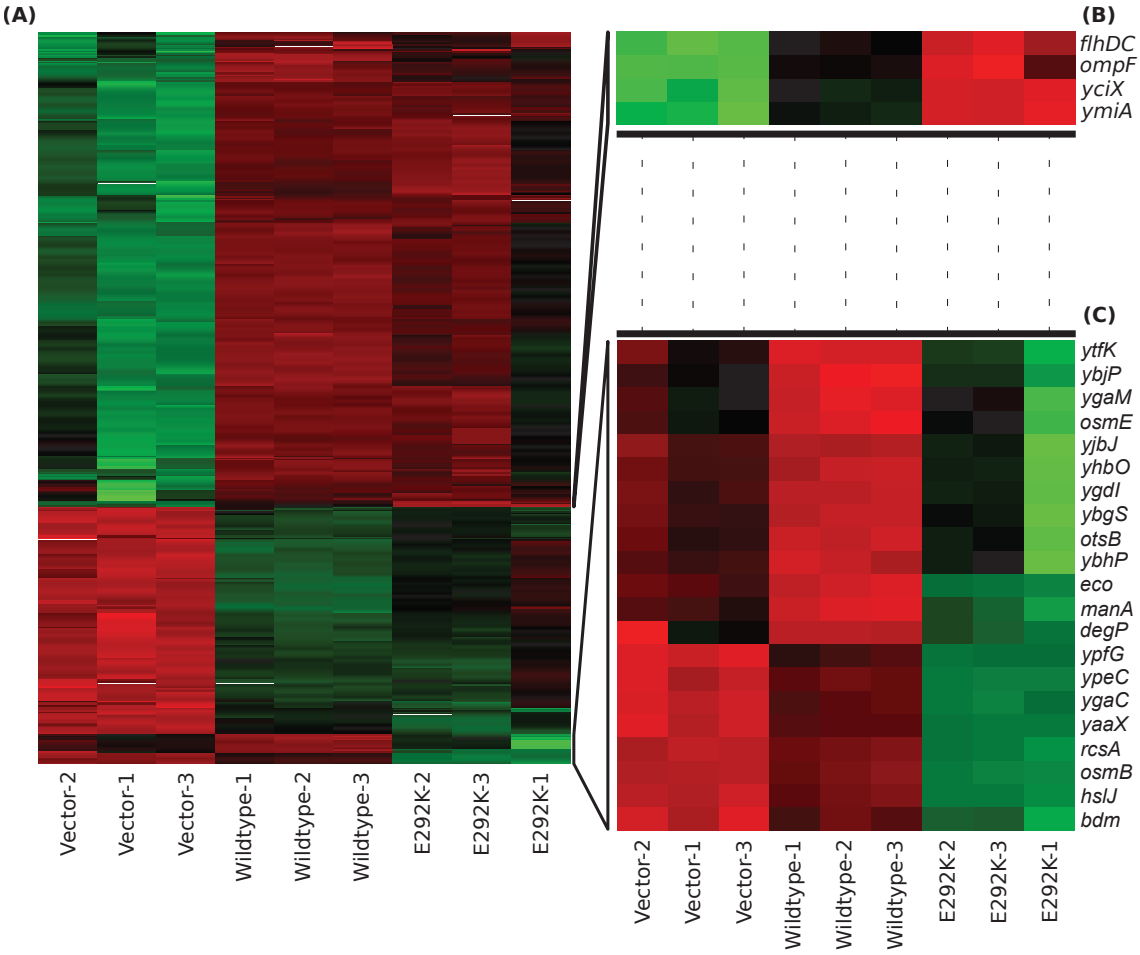

Figure 15: Heat map of significantly differentially expressed genes.

(A) Hierarchical clustering by Pearson correlation performed on the regularised log-transformed normalised counts for genes significantly differentially expressed (FDR < 0.05 & fold-change ≥2 or ≤−2) in *both* the pIV-E292K *and* wildtype *vs*. vector analyses (Number of genes = 518). **(B)** Genes down-regulated in response to production of leaky pIV-E292K. **(C)** Genes up-regulated in response to expression of leaky mutant pIV-E292K, this cluster includes the 13 genes found to be down-regulated in response to wildtype pIV production (**Figure 14**). Plasmid-borne genes were removed prior to heat mapping to avoid skewing the scale in favour of over-expressed plasmid-borne gIV. Heat colouring of cells is based on the global relative counts per gene, red indicates low counts (low expression); green indicates higher counts (higher expression), relative to all samples.

To examine whether differentially regulated genes encode proteins that share common functions or localisation, gene ontology term enrichment was performed using the web service PANTHER on significantly differentially expressed gene lists (fold-change of at least two-fold; FDR<0.05) to identify functional clusters (Mi *et al.*, 2013a, Mi *et al.*, 2013b, Thomas *et al.*, 2006). This analysis identifies statistical enrichment of certain annotated functions within the lists of genes (in this case differentially expressed). In the wildtype pIV producing cells, common PANTHER classifications among the significantly up-regulated genes include transport, transferases and nucleic acid binding, whereas the functions enriched among the down-regulated genes include oxidoreductase-activity, kinase-activity and transport (Figure 16). Within the set of genes up-regulated in pIV-E292K-producing cells, transporters, nucleic acid binding proteins and hydrolases are the most common classifications. Down-regulated genes in the pIV-E292K transcriptome appear to be enriched for reductases, dehydogenases, and transporters. (Figure 16).

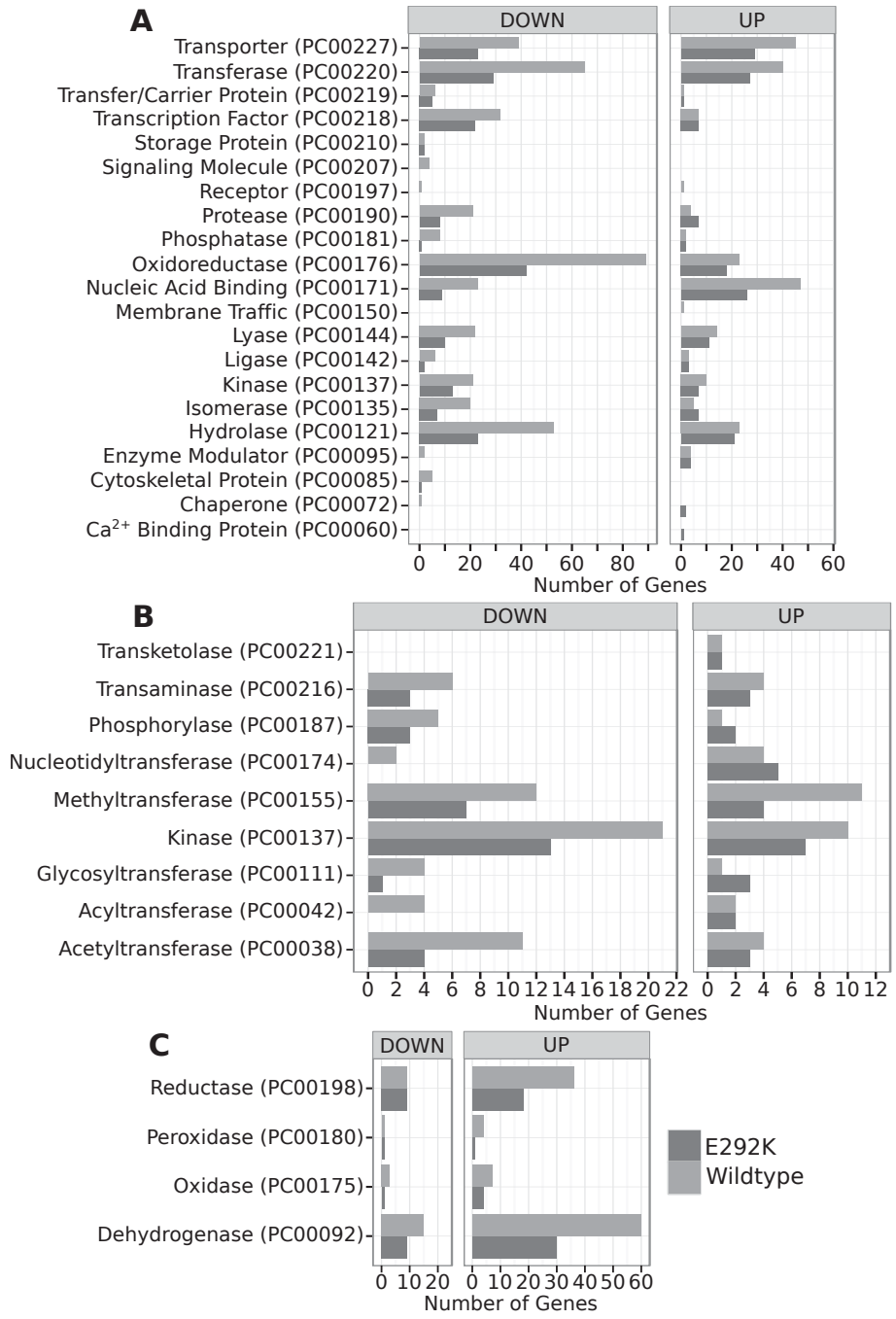

Figure 16: PANTHER protein classifications of genes significantly differentially expressed in response to pIV secretin stress.

(A) All level 1 PANTHER protein classifications for genes significantly down-regulated (DOWN; FDR < 0.05, Fold-Change < -2) or up-regulated (UP; FDR<0.05, Fold-Change > 2) in response to wildtype pIV or pIV-E292K. 603 of 657 unique gene IDs significantly down-regulated in response to wildtype pIV were mapped to the PANTHERdB; 322 of 356 unique gene IDs significantly down-regulated in response to pIV-E292K were mapped to the PANTHERdB. 345 of 378 unique gene IDs significantly up-regulated in response to wildtype pIV were mapped to the PANTHERdB; 286 of 312 unique gene IDs significantly up-regulated in response to pIV-E292K were mapped to the PANTHERdB. **(B)** Level 2 PANTHER protein classifications for those genes classed as “Transferases” in **(A)**. **(C)** Level 2 PANTHER protein classifications for those genes classed as “oxidoreductases” in **(A)**. In all figures, the number of genes annotated with a PANTHER protein class is indicated on the x-axis and the PANTHER protein classification indicated on the y-axis. Some genes, e.g. phage insertase *insH*, are duplicated in the genome and are the source of differences in the numbers of genes between this figure and **Figure 14**. (Mi *et al.*, 2013a, Mi *et al.*, 2013b).

4.2 The Phage Shock Protein Response

As mentioned in the previous sections, the key response to secretin stress in Enterobacteria is the Psp or phage shock protein response. The genes and regulatory circuit of the Psp response are organised as a regulon, composed of one gene cluster, *pspF-A-E*, and an unlinked gene, *pspG* (schematically represented in Figure 17). As the transcriptional response of the Psp regulon to a 'leaky' secretin has not been assessed to date, it was interesting to compare the extent and nature of Psp response to the leaky mutant pIV-E292K relative to the response to wildtype pIV. The expression of the sigma-N (σ^{54}) PspF-dependent secretin-stress induced promoters controlling transcription of *pspA−D* and *pspG* was highly induced, as expected (6- to 15-fold; Figure 17). Interestingly, fold-changes in expression of *pspBCD* and *pspG* genes were comparable between the wildtype- and pIV-E292K-producing cells, suggesting secretin leakiness does not correlate with the extent of *psp* mRNA accumulation (Figure 17). The exception was PspA, which showed higher up-regulation in pIV-E292K-producing than in the wildtype pIV-producing cells. This is in agreement with increased amount of PspA protein in the pIV-E292K-expressing cells in comparison to the wildtype pIV expressing cells 1 hour after induction of pIV expression by IPTG, which is the same timepoint at which the RNA was isolated for the transcriptome analysis (Figure 9 & Spagnuolo *et al.* (2010)). The *pspA* mRNA increases 7.4-fold in the wildtype-expressing cells and 10.8-fold in the pIV-E292K-expressing cells and was much more prominent than the accumulation of the PspA protein after 1 hour of pIV induction (less than 1.5-fold for both the mutant and the wildtype pIV; Figure 9B), hence PspA protein

accumulation lags behind mRNA up-regulation, implying post-transriptional regulation.

In agreement with published observations (Lloyd *et al.*, 2004, Jovanovic *et al.*, 1996, Jovanovic *et al.*, 1997), the *pspF* gene encoding the positive regulator of *pspA-D* and *pspG* promoters is down-regulated, albeit only by two-fold, with the activation of the response. Another gene of the regulon, *pspE,* which is located at the 3' end of the *pspA-E* cluster and controlled independently of the sigma-54 *pspA-D* promoter, was also down-regulated (again by around 2-fold) in response to secretin stress (Brissette *et al.*, 1991, Jovanovic *et al.*, 1997).

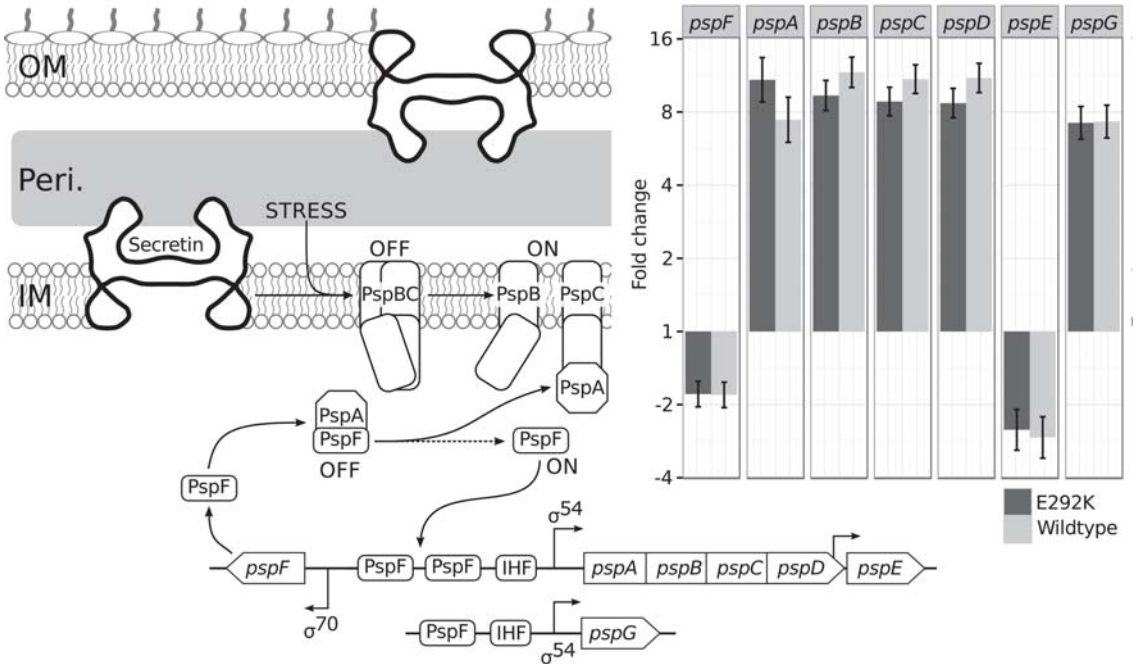

Figure 17: Differential expression of the Psp regulon during secretin stress.

Regulation of the Psp regulon by secretin stress (Left). In the OFF state, PspF is bound to PspA, preventing interaction with the *pspA-E*/*pspG* regulatory regions. Upon induction (ON), PspC dissociates from PspB and sequesters PspA (Flores-Kim & Darwin, 2015), freeing PspF (dashed line) to activate transcription of *pspA-E* and *pspG* together with IHF (Adams *et al.*, 2003, Dworkin *et al.*, 1997, Weiner *et al.*, 1991). Fold changes of all *psp* genes are relative to the vector control (Right). The fold changes determined by DESeq2 are displayed along the y-axis; error bars show the standard error of the fold changes. All fold changes in the graph have a false discovery rate (FDR) of less than 1×10^{-5}.

4.3 Sigma-E and Cpx stress responses

The sigma-E response (Figure 18) protects Gram-negative bacteria from envelope stresses that result in misfolding of OMPs and accumulation of periplasmic LPS intermediates (Lima *et al.*, 2013). This response consists of a number of sigma-E-stimulated genes, including the *rpoE-rseABC* cluster and the sigma-E sigmulon (the complete set of genes regulated, directly or indirectly, by a given sigma factor) composed of a number of genes and operons. While sigma-E was found to be required for secretion of effectors through type 2 and 3 secretion systems in *V. cholerae* and *Y. enterocolitica* (Carlsson *et al.*, 2007, Zielke *et al.*, 2014), activation of the sigma-E response by the secretion system alone (sans secreted effector proteins) or by the secretins alone (without the accompanying type 2 or 3 secretion system machinery) has not been demonstrated. Findings presented here using the RNA-seq approach are consistent with the microarray data (Lloyd *et al.*, 2004, Jovanovic *et al.*, 2006) in that the sigma-E response was not stimulated by the wildtype pIV secretin. Furthermore, a question addressed here was whether a leaky secretin, which is expected to impose an increased envelope stress, like that imposed by type 3 secretin YscC, would also result in activation of the sigma-E response.

The data analysis of differentially expressed genes showed that, regardless of gate phenotype (wildtype or the leaky pIV-E292K mutant) transcription of the *rpoE-rseABC* operon (*rpoE* encodes sigma-E) did not increase (Figure 18A) relative to the vector control. To the contrary, the *rpoE-rseABC* operon is repressed in the transcriptome of cells producing wildtype pIV.

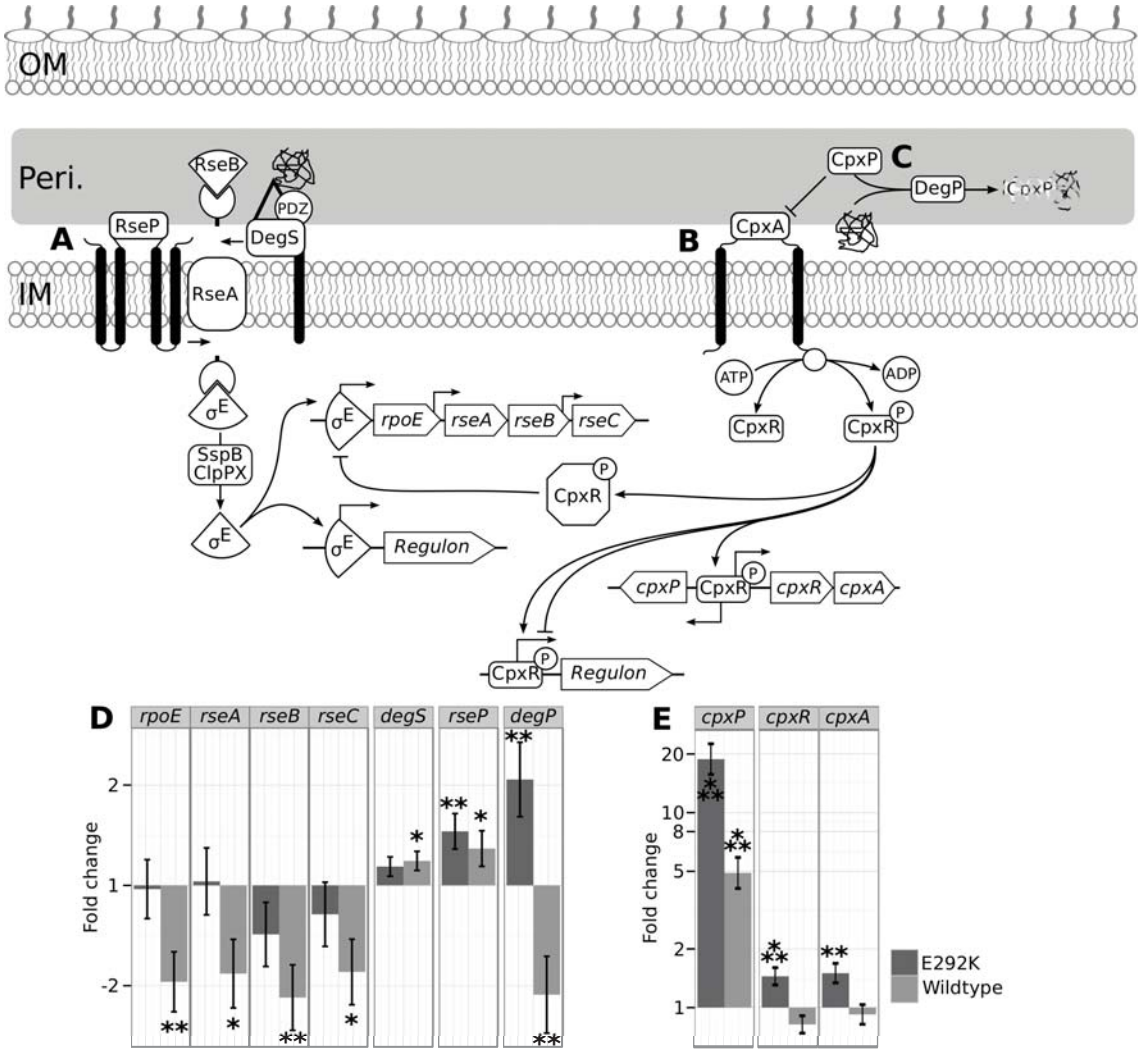

Figure 18: Sigma-E and Cpx response pathways and differential expression of the genes encoding the key regulators of these pathways.

(A) The activation pathway of the sigma-E response to perplasmic stress. **(B)** The Cpx response pathway **(C)** Accumulation of misfolded periplasmic proteins and their consequent degradation by DegP includes degradation of negative regulator CpxP, resulting in the induction of the Cpx response **(D)** Fold change in the *rpoE-rseABC, degS, rseP* and *degP* in the pIV secretin-expressing cells. **(E)** Fold change in the *cpxPRA* genes. Fold changes relative to the vector control, reported by DESeq2 are displayed on the y-axis. Error bars show the standard error of the DESeq2 fold changes. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001.

Phosphorylated transcriptional regulator CpxR negatively regulates the *rpoE* gene encoding sigma-E (De Wulf *et al.*, 2002). Phosphorylation of CpxR is in turn blocked by protein CpxP (Fleischer *et al.*, 2007). However, expression of *cpxP* itself is dependent on phosphorylated CpxR, closing a negative feedback loop. It is therefore interesting that *cpxP* is differentially expressed significantly and highly in cells producing wildtype pIV (5-fold) and pIV-E292K (20-fold) (Figure 18B). The expression of the *cpxRA* operon, under the control of phosphorylated CpxR, was induced marginally but significantly (1.5-fold) in the pIV-E292K producing cells, while it was unchanged in the wildtype pIV producing cells. The up-regulation of *cpxP* expression suggests that CpxR is indeed phosphorylated, and at least in the pIV wildtype producing cells, is likely responsible for the 2-fold repression of the *rpoE-rseABC* operon. It is unclear why the pIV-E292K transcriptome does not show similar or greater repression of the *cpxRA* and *rpoE-rseABC* operons, despite the ~4-fold higher transcription of *cpxP* compared to wildtype, which is expected to reduce phosphorylation of CpxR. It is possible that CpxR-mediated inhibition of the sigma-E response is dose-dependent in a non-linear manner. Otherwise, competition between CpxR-mediated transcriptional inhibition and sigma-E RNA polymerase could cause the lack of differential expression, and allow expression of sigma-E dependent genes, without eliciting a full sigma-E response. Alternatively, pIV-E292K may form misfolded aggregates in the periplasm, which are detected by CpxP and RseP/DegS. This would allow activation of CpxR by CpxA-mediated phosphorylation and release of sigma-E from RseA-inhibition. However, competition between phosphorylated CpxR and sigma-E would prevent any

change in expression of the *rpoE-rseABC* operon as observed in the response to pIV-E292K.

Genes of the CpxR regulon are inconsistently regulated in cells producing wildtype pIV or pIV-E292K (Appendix 1 & 2, Figure 33, Figure 34). Many of these genes are co-regulated by other transcription factors, for example expression of the *marRAB* operon, in addition to CpxR, is activated by SoxS and MarA. Such cross talk between regulons is common and likely essential for fine-tuning stress responses to a specific stressor. Genes with function in protein folding, cell wall repair/biogenesis and regulation of biofilm/motility were among those CpxR-regulon genes differentially expressed in response to the leaky mutant, pIV-E292K. Differential expression and regulation of the CpxR regulon is shown in Appendix 1 & 2 (Figure 33, Figure 34).

Some of the genes significantly differentially expressed in cells producing secretins that belong to the sigma-E sigmulon (FDR < 0.05; Appendix 3 & 4; Figure 35 and 36) have common functions in cell wall repair/biogenesis, or are periplasmic proteases/chaperones and transport. Their expression depends on specific transcription factors in addition to sigma-E, and their transcription in the secretin-producing cells must be induced by these other factors. Regulation and differential expression of these genes is shown in Appendix 3 & 4.

In summary, it could be speculated how the Cpx response is stimulated by secretins in a gate-dependent manner: kinase activity of CpxA is inhibited by CpxP, favouring the phosphatase activity of CpxA (Fleischer *et al.*, 2007).

Misfolded periplasmic proteins (including misfolded secretins) can interact with CpxP, which targets the CpxP-misfolded-protein complex for degradation by DegP (Isaac *et al.*, 2005). This removes the CpxA inhibiting signal, favouring phosphorylation of CpxR. If the folding of pIV-E292K mutant were delayed relative to the wildtype pIV, this would explain the higher transcription of *cpxP* and the lack of change in transcription of *rpoE-rseABC* in response to pIV-E292K compared to wildtype pIV. Low-molecular weight phosphor-donor molecules derived from the Pta-AckA pathway may also activate CpxR. If this is the case, the Cpx regulon expression would be sensitive to changes in core metabolism, manifested by the increase in the phosphor-donor molecules (Wolfe *et al.*, 2008). Both *ackA* and *pta* are up-regulated significantly 3.3 and 2.2-fold (FDR < 0.001), respectively, in response to wildtype pIV and 2.7 and 2.3-fold (FDR < 0.001), respectively, in response to pIV-E292K, hence they may be the source of CpxR activation in response to wildtype pIV.

4.4 EnvZ/OmpR response

The EnvZ/OmpR response is a regulatory circuit that controls expression of different porins in response to the changes in osmolarity and other environmental conditions. The sensor EnvZ activates the transcriptional regulator, OmpR, by phosphorylation. In its phosphorylated state OmpR signals high osmolarity in the environment and ultimately blocks expression of wider porin OmpF while inducing that of the narrow porin OmpC. It is also a positive regulator of *omrA* and *omrB* regulatory sRNA transcription. These sRNAs are repressors of the *ompR-envZ* operon. The expression of *ompR* is down-regulated marginally but significantly in the cells producing wildtype pIV (Figure

19). The genes *omrA* and *omrB* were culled from analysis prior to DESeq2 analysis because < 50 reads mapped to them in any one replicate. Such low counts are indicative of poorly or unexpressed genes; hence the repression of *ompR* is likely caused by another transcription factor. Genes directly regulated by OmpR (Figure 19B) follow the expected pattern of OmpR regulation, suggesting the OmpR regulon is activated.

The porins *ompF* and *ompC* are repressed in response to secretin stress. Expression of these two porins is supposed to be inversely regulated by phosphorylated OmpR in response to osmotic stress, with OmpC up- and OmpF down-regulated in a dose-dependent manner (Lan & Igo, 1998). The RNA-seq data obtained here show that both porins are repressed in response to secretin stress; however, repression of *ompF* is much greater than that of *ompC* in response to the leaky pIV-E292K mutant (Figure 19), whereas both porins are similarly repressed in cells expressing wild-type pIV, whose pore size is smaller than that of the porins. Excess production of OmpC over OmpF is expected in the leaky mutant as this former porin is induced by the osmotic stress or factors that have compromising effect on the outer membrane integrity (see the subsequent sections).

OmpR regulates the *csgDEFG* operon, encoding the control and assembly factors for curli pili in an IHF-dependent fashion (Ogasawara *et al.*, 2010). The transcription of *csgDEFG* was increased about 2-fold in the cells producing wildtype pIV. The first gene in this transcription unit, *csgD*, is the transcriptional

regulator for *csgBA* (curli pilin subunits) and a number of unlinked genes involved primarily in motility, including the flagellar operon.

Master regulator of class II flagellar operons, *flhDC*, is regulated by a number of transcription factors, including OmpR, RcsAB and the sRNAs *oxyS*, *omrA*, *omrB*, and *arcZ*. The *flhDC* operon encoding the motility regulator was strongly down-regulated by pIV production, nearly 20-fold by pIV-E292K and 5-fold by wildtype pIV. This change is consistent with the physiological conditions of the cells. Repression of motility is the rule for high-osmotic environment responses in enteric bacteria; furthermore, expending energy on motility in the time of secretin stress is not conducive to survival.

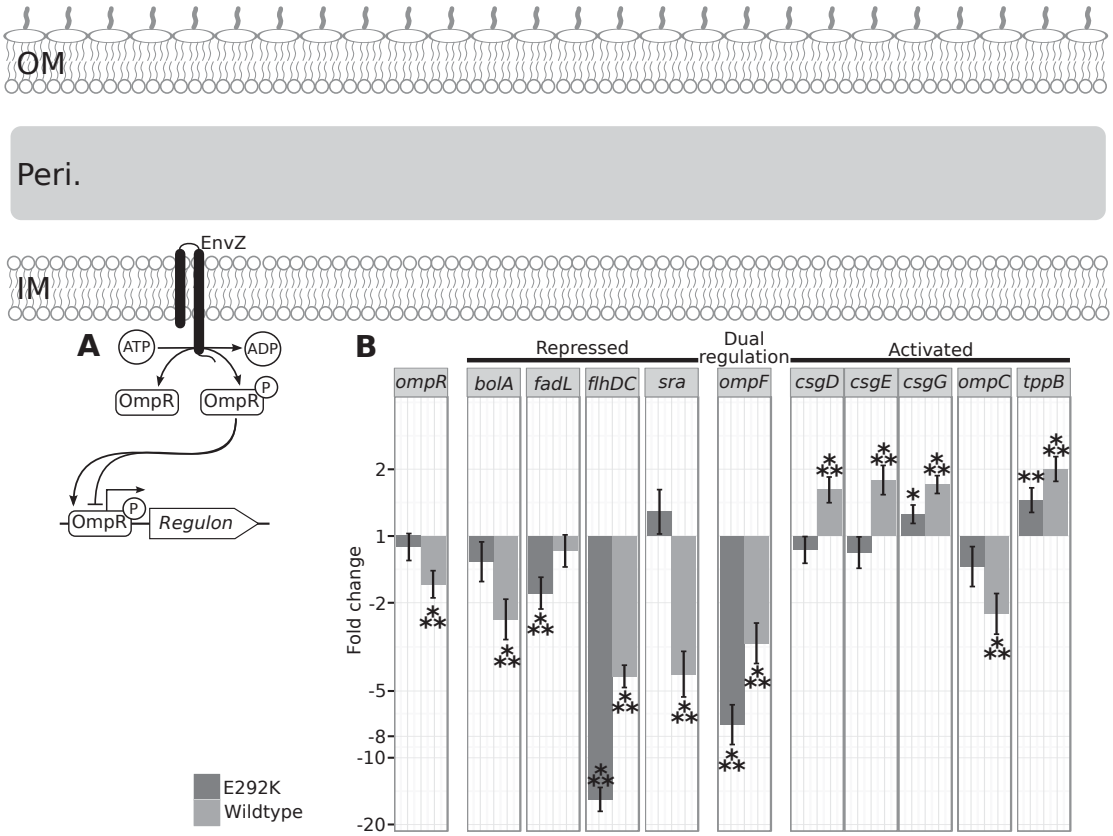

Figure 19: Differential expression of the *ompR* regulon.

(A) Scheme of theEnvz-OmpR two-component system regulation. The sensor kinase, EnvZ, detects and responds to osmotic stress by controlling the rate at which OmpR is phosphorylated (OmpR-P). Under low osmolarity conditions, only low levels of phosphorylated OmpR are present and *ompF* is expressed in favour of *ompC*. Under high-osmolarity conditions, more OmpR-P is present, allowing expression of *ompC* in favour of *ompF* (Lan & Igo, 1998). Thus, *ompF* is considered to be both repressed and activated by OmpR-P, ie is “Dual Regulated”. **(B)** Genes annotated in the EcoCyc database as directly regulated by phosphorylated OmpR were identified in the DESeq2 results, the type of regulation imparted by OmpR-P is labelled above the regulated genes. Genes belonging to the same transcriptional unit are grouped, others are either sole genes in the transcription unit or transcription unit members were removed from the dataset prior to DESeq2 analysis. Fold change relative to the vector control is indicated on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001.

4.5 Oxidative stress responses

SoxR is a redox state sensor, sensitive to the cellular ratio of NADPH/$NADP^+$ and is oxidised by superoxide (Liochev & Fridovich, 2011). Oxidised SoxR activates expression of the *soxS* gene, encoding a transcriptional regulator that induces expression of a regulon involved in the response to oxidative stress. In cells producing wildtype pIV or the leaky mutant pIV-E292K, a ~20-fold increase in *soxS* is detected, indicating that the cells are experiencing oxidative stress, with the induction of the SoxS regulon and likely oxidation of SoxR (Figure 20A). Expression of *soxS* causes the autoinhibition of the *soxR* promoter, thus the lack of differential *soxR* up-regulation is expected. The SoxR reductase (RsxABCDGE) is an inner membrane protein complex that reduces SoxR using NADPH as an electron-donor. Oxidation of other cellular components by superoxide or other reactive oxygen species (ROS) depletes NADPH and inhibits the SoxR reductase. Thus, the ratio of oxidised-SoxR/reduced-SoxR reflects the level of oxidative stress. The expression of *rsxA-E* genes appears to be correlated with the differential expression of *soxS*, which is higher in cells producing wildtype pIV vs. mutant pIV-E292K. This difference probably reflects the variable levels of SoxR oxidation induced by the two pIV variants (Figure 20A). However, whether transcription of the *rsxA-E* gene cluster requires activation by a regulator, such as SoxS, is unknown.

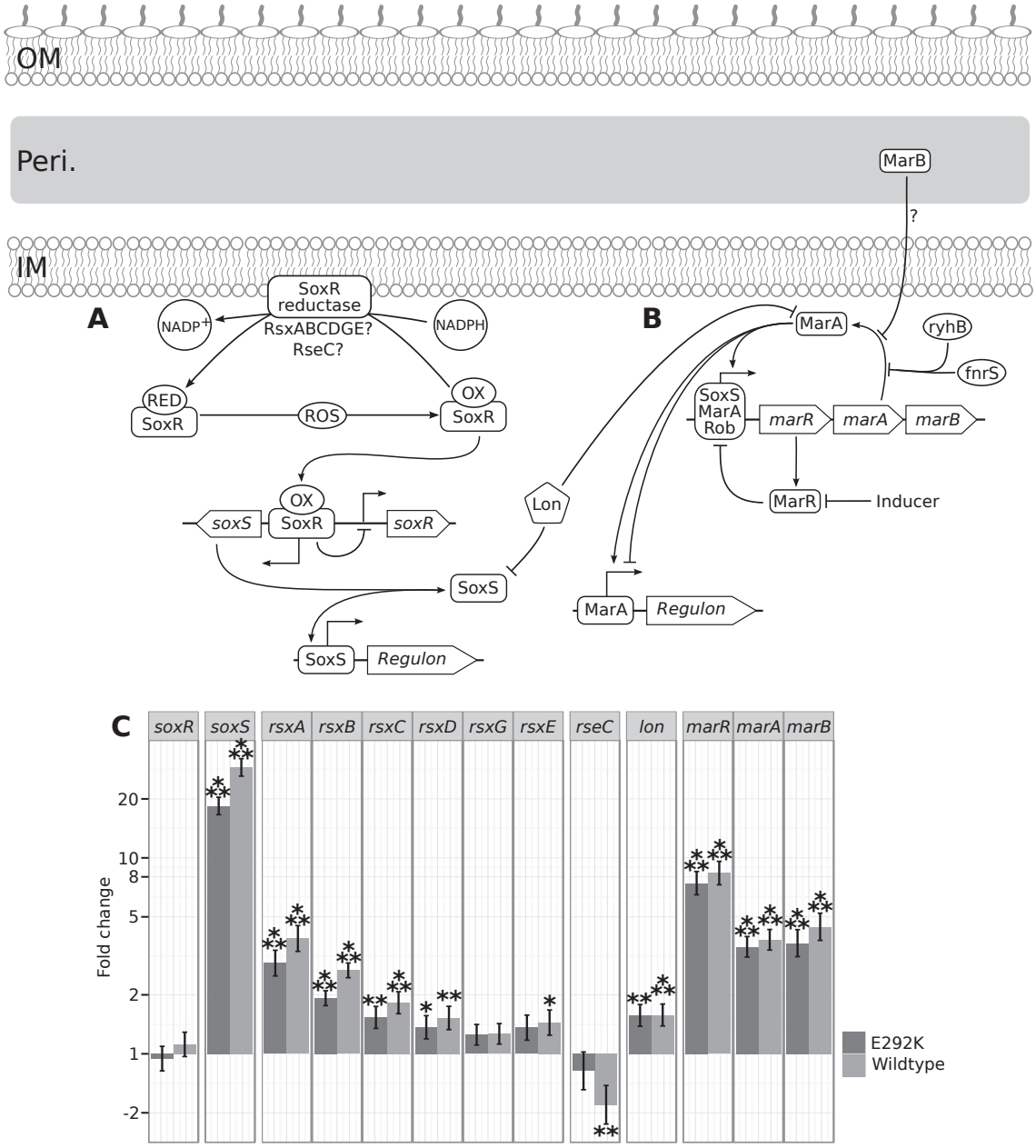

Figure 20: Differential expression of the genes encoding regulators of the *sox* and *mar* regulons.

(A) Regulatory pathways of the superoxide stress (Sox) response. **(B)** Regulatory pathway of the multiple antibiotic resistance (*mar*) regulon. **(C)** Differential gene expression of the *soxRS*, *rsxABCDE, rseC*, *marABR* and *lon* in the secretin-expressing cells. Genes colocalised in the same transcriptional unit have been grouped together. Fold changes relative to the vector control are displayed on the y-axis; * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001.

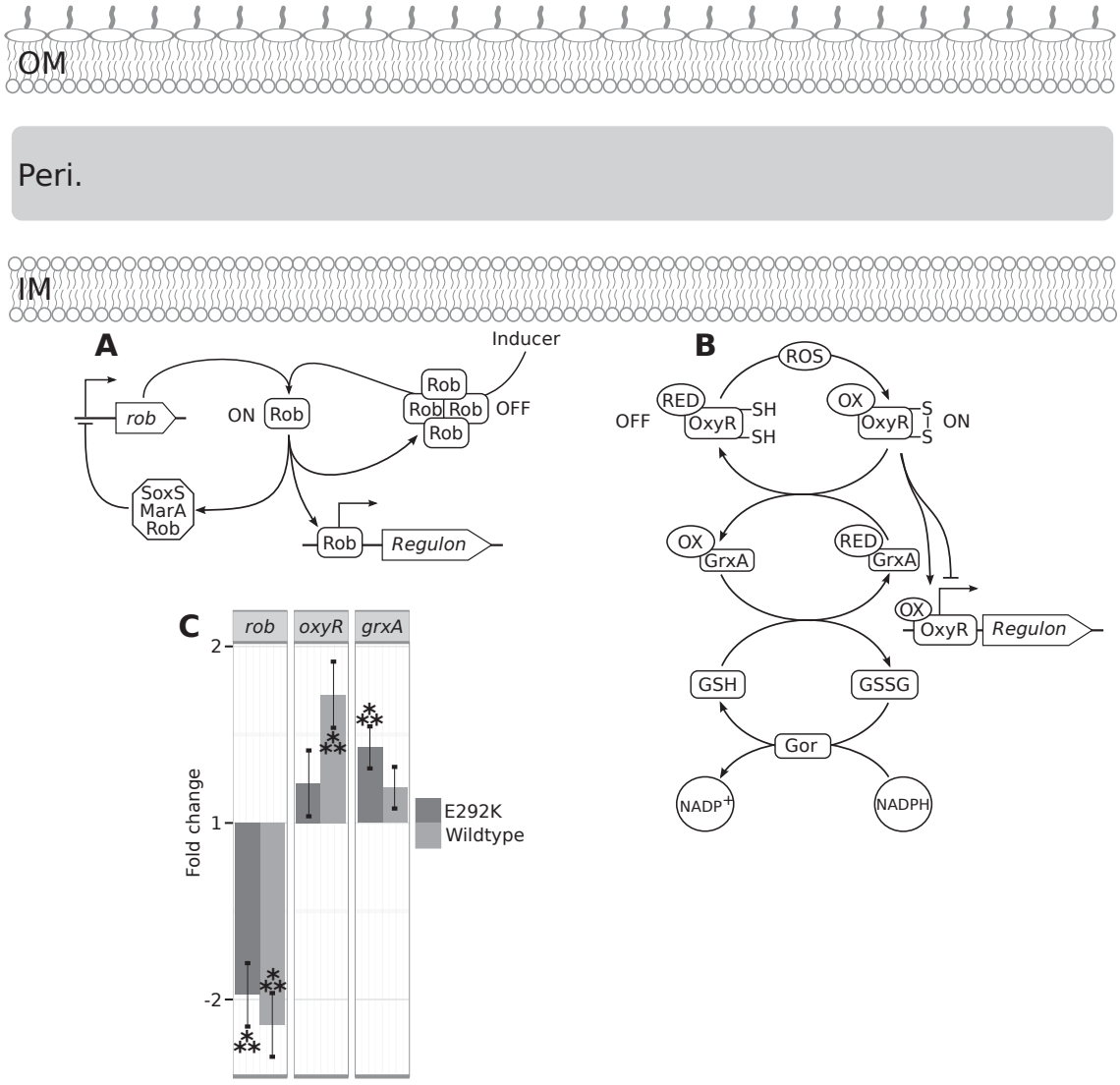

Figure 21: Differential expression of the genes controlling transcription of the *rob* and *oxyR* regulons.

(A) Regulatory pathway of the Rob-dependent regulon. **(B)** Pathways for sensing and removal of hydrogen peroxide. **(C)** Regulation of *rob*, *oxyR* and *grxA* by expression of the wildtype pIV and the leaky muant pIV-E292K. Fold changes relative to the vector control are displayed on the y-axis; * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001.

The *rseC* gene, which is thought to play a role in SoxR redox cycling, is located at the 3' end of the *rpoE-rseA-C* gene cluster. This gene is down-regulated by a factor of 2 in cells producing mutant pIV-E292K, consistent with its regulation by Sigma-E (RpoE). Redox cycling and activity of the SoxR reductase under secretin stress is therefore, likely enabled by the activity of the RsxABCDGE complex (Figure 20A).

Oxidised redox-sensing regulator, OxyR, directly regulates expression of *grxA* and *gor* genes, encoding the proteins required for regulating the redox state of OxyR, Glutaredoxin 1 and glutathione reductase, respectively (Figure 21B). The *grxA* gene is significantly differentially expressed only in cells expressing the leaky mutant, pIV-E292K, however the change is less than 2-fold (Figure 21B). Transcript reads derived from *gor* were removed from the datasets prior to DESeq2 analysis due to low variance across samples, thus was likely not differentially expressed in response to secretin stress. Overall, genes responsible for removal of superoxide were induced strongly, whereas the genes responsible for removal of hydrogen peroxide and regeneration of glutathione remained generally uninduced.

4.6 Effect of secretin production on the SoxS regulatory network

The *marRAB* operon encodes regulators of efflux pumps involved in resistance to multiple antibiotics (Alekshun & Levy, 1999). The inducer of the oxidative stress response, SoxS, is involved in the induction of *marRAB*, together with protein, MarA, encoded by the *marRAB* operon, and Rob encoded elsewhere in the genome. Strong up-regulation of *marR* (8-fold) and *marAB* (4-fold) was

detected in the cells producing wildtype pIV and the mutant pIV-E292K (Figure 20B). In the context of secretin stress, it is likely that SoxS is the regulator responsible for this up-regulation, because of its dominant expression (20-fold) over *marA* (4-fold).

The gene, *rob*, encoding the co-activator of the *marRAB* operon and regulator of the Rob regulon is down-regulated ~2-fold in response to pIV imposed stress (Figure 21A). Inducing compounds cause the non-functional Rob protein aggregates to collapse, allowing Rob to regulate transcription of Rob regulon members and repress its own promoter.

The Lon protease is responsible for turning off both the *sox* and *mar* regulons by degrading transcriptional activators SoxS and MarA. This regulatory protease is marginally (less than 2-fold), but statistically significantly, down-regulated in response to pIV expression (Figure 20C), likely contributing to the increase of SoxS and MarA protein concentration in the cells.

4.7 Differential expression of the regulons controlled by SoxS, MarA, Rob and OxyR

Genes regulated by the above transcription factors can be divided into five groups, based on combinations of the above regulators by which they are controlled.

Genes that are common between the regulons controlled by SoxS, MarA and Rob encode a variety of functions involved in response to stresses such as

antibiotic efflux pumps and oxidoreductases (Randall & Woodward, 2002). Of those, only the genes encoding the efflux pump AcrAB show differential expression greater than 2-fold in the cells producing the pIV secretin variants (Figure 22A). Transcription of these two components of the AcrAB-TolC efflux pump is higher in the cells producing the leaky mutant, pIV-E292K, than wildtype pIV. Several other genes have marginal but statistically significant change in expression. One of these is the gene encoding YbhT (synonym AcrZ) protein that interacts with the AcrAB-TolC transport complex and shows increased transcription in response to production of pIV-E292K, but decreased in response to wildtype pIV (Figure 10A). The same is true of *ygiB*, an uncharacterised gene that is co-transcribed with *tolC*, and may mitigate defects in the TolC protein.

Other genes in this group whose expression is changed significantly (but by less than a factor of two) are *sodA* and *nfsB*. Superoxide dismutase gene, *sodA*, is significantly differentially expressed in response to pIV-E292K, but not in response to wildtype pIV. Conversely, nitroreductase B, *nfsB*, is differentially repressed in response to wildtype pIV, but its expression does not change in response to pIV-E292K.

Three genes regulated by SoxS and MarA but not Rob (Figure 22B) show interesting differential expression in response to pIV production. The LPS core biosynthesis pathway genes, *rfaY* and *rfaZ* are both significantly down-regulated (by ~2-fold) in response to both wildtype pIV and the mutant, pIV-E292K. The gene, *poxB*, encoding pyruvate oxidase is central to aerobic

metabolism and survival in stationary phase (Moreau, 2004), however, is oppositely regulated by the two pIV variants, showing a significant increase and decrease, respectively, in transcript amount in the cells producing pIV-E292K and wildtype pIV.

Only one gene, *ydbK*, amongst those uniquely regulated by SoxS (Figure 22C) shows differential expression of at least 2 fold. *ydbK,* encodes a pyruvate:flavodoxin oxidoreductase induced in response to superoxide. pIV-E292K exhibits a greater inductive effect on *ydbK* transcription than wildtype pIV.

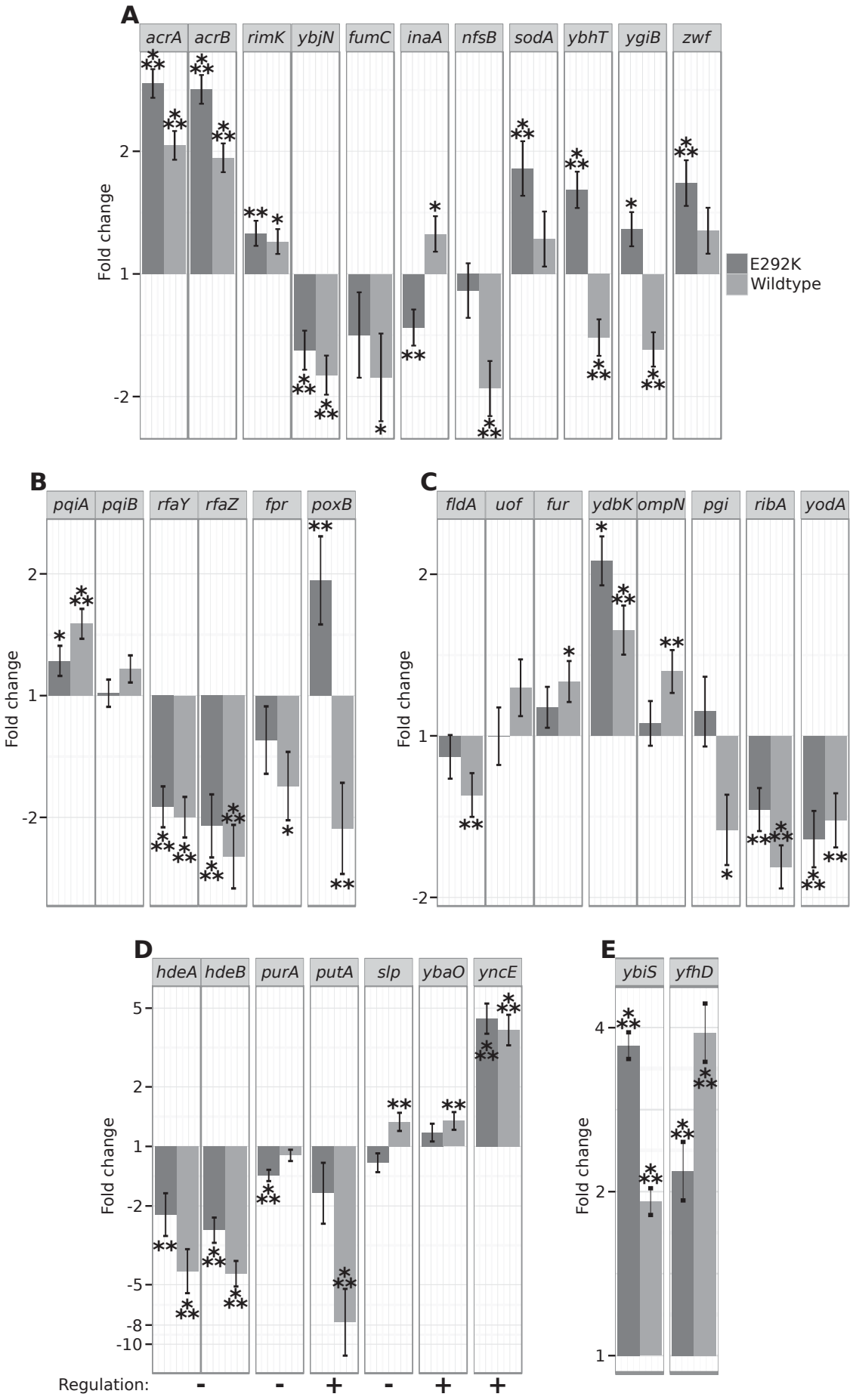
A
acrA
acrB
rimK
ybjN
fumC
inaA
nfsB
sodA
ybhT
ygiB
zwf
Fold change
2
1
-2
E292K
Wildtype
B
pqiA
pqiB
rfaY
rfaZ
fpr
poxB
Fold change
2
1
-2
C
fldA
uof
fur
ydbK
ompN
pgi
ribA
yodA
Fold change
2
1
-2
D
hdeA
hdeB
purA
putA
slp
ybaO
yncE
Fold change
5
2
1
-2
-5
-8
-10
Regulation:
-
-
+
-
+
+
E
ybiS
yfhD
Fold change
4
2
1

Figure 22: Differential expression of the Mar/Sox/Rob-regulated genes

(A) Differential expression of genes commonly regulated by MarA, SoxS and Rob. **(B)** Differential expression of genes commonly regulated by both MarA and SoxS. **(C)** Differential expression of genes unique to the *soxS* regulon. **(D)** Differential expression of genes unique to the *marA* regulon. **(E)** Differential expression of genes unique to the *rob* regulon. Genes belonging to the same transcriptional unit are grouped together. Fold changes relative to the vector control are displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes in **(A)**, **(B)**, **(C)** and **(E)** are all always activated by the respective transcription factor (SoxS, Rob or combinations thereof), whereas the MarA has a reported positive or negative effect, as indicated below each lane representing individual regulated operons and genes (Salgado *et al.*, 2013).

The effect of pIV variants on some genes soley regulated by MarA was much stronger than the effect on promoters regulated by combinations of transcriptional factors (Figure 22D). Among the MarA-regulated genes, the periplasmic chaperones, *hdeA* and *hdeB*, show a reduction in transcript abundance by a factor of at least 2 in response to the production of pIV-E292K and even lower (4-fold) in response to wildtype pIV. A gene, *putA*, which encodes a dual-function protein acting as both proline dehydrogenase and the transcriptional repressor of the *putAP* operon, was repressed nearly 8-fold in response to wildtype pIV but displayed no significant difference in response to pIV-E292K. A periplasmic protein of unknown function, *yncE*, exhibits at least 4-fold increased transcription in response to production of both wildtype pIV and the leaky mutant, pIV-E292K.

Genes uniquely regulated by Rob (Figure 22E), which have altered transcript abundance under pIV secretin stress encode a transpeptidase *ybiS*, linking the major outer membrane lipoprotein, Lpp, to the murein layer and a murein hydrolase, *yfhD*. Both transcripts show significant increases in response to pIV production. However, *ybiS* transcript (increased 4-fold) is twice as abundant in cells producing pIV-E292K as in those producing wildtype pIV. Conversely, the *yfhD* transcript (increased 4-fold) in the cells producing wildtype pIV is twice as abundant as in the cells producing pIV-E292K.

Of the genes solely activated by OxyR (Figure 23, Figure 24) only the *sufABCDSE* operon and *mntH* gene show a significant increase in transcription relative to the vector control (up to 5-fold). The *sufABCD* genes encode an Fe-S

cluster assembly complex, whereas *sufE* encodes a selenocysteine lyase which stimulates a cysteine desulfurase encoded by *sufS*. The *sufABCDSE* operon transcripts are present in slightly higher amounts in the cells producing pIV-E292K relative to the cells producing the wildtype pIV. Divalent metal ion transporter, *mntH*, displays a significant 2-fold increase in transcription in the wildtype pIV-producing cells, but not in cells producing pIV-E292K. The *dps* gene encodes a nucleoid-associated protein that sequesters iron, protects against DNA damage and is required for survival in stationary-phase shows a 5-fold reduction in transcription in response to the production of wildtype pIV but no significant change in response to the production of leaky mutant, pIV-E292K.

Of the genes repressed by OxyR (Figure 23, Figure 24) only *ybjN*, *uxuAB*, and *gntP* exhibit significant decrease in transcript abundance in pIV-producing cells. The putative chaperone, *ybjN*, is co-transcribed with ribosome-associated polyglutamate synthase, *rimK*, yet it shows a decrease in transcription, while *rimK* transcript exhibits a small but significant increase in cells producing wildtype pIV or pIV-E292K. Genes, *uxuA*, encoding an Fe^{2+}-dependent D-mannonate dehydratase, and *uxuB*, encoding D-mannonate oxidoreductase, show significant repression in cells producing wildtype pIV. Transcription of these two genes is repressed in response to production of pIV-E292K, but only *uxuB* shows a significant (FDR < 0.05) but small (less than 2-fold) difference.

The gene encoding a transporter, *gntP*, is significantly repressed in response to both wildtype pIV and pIV-E292K. Repression of *gntP* transcription is greater in response to production of wildtype pIV.

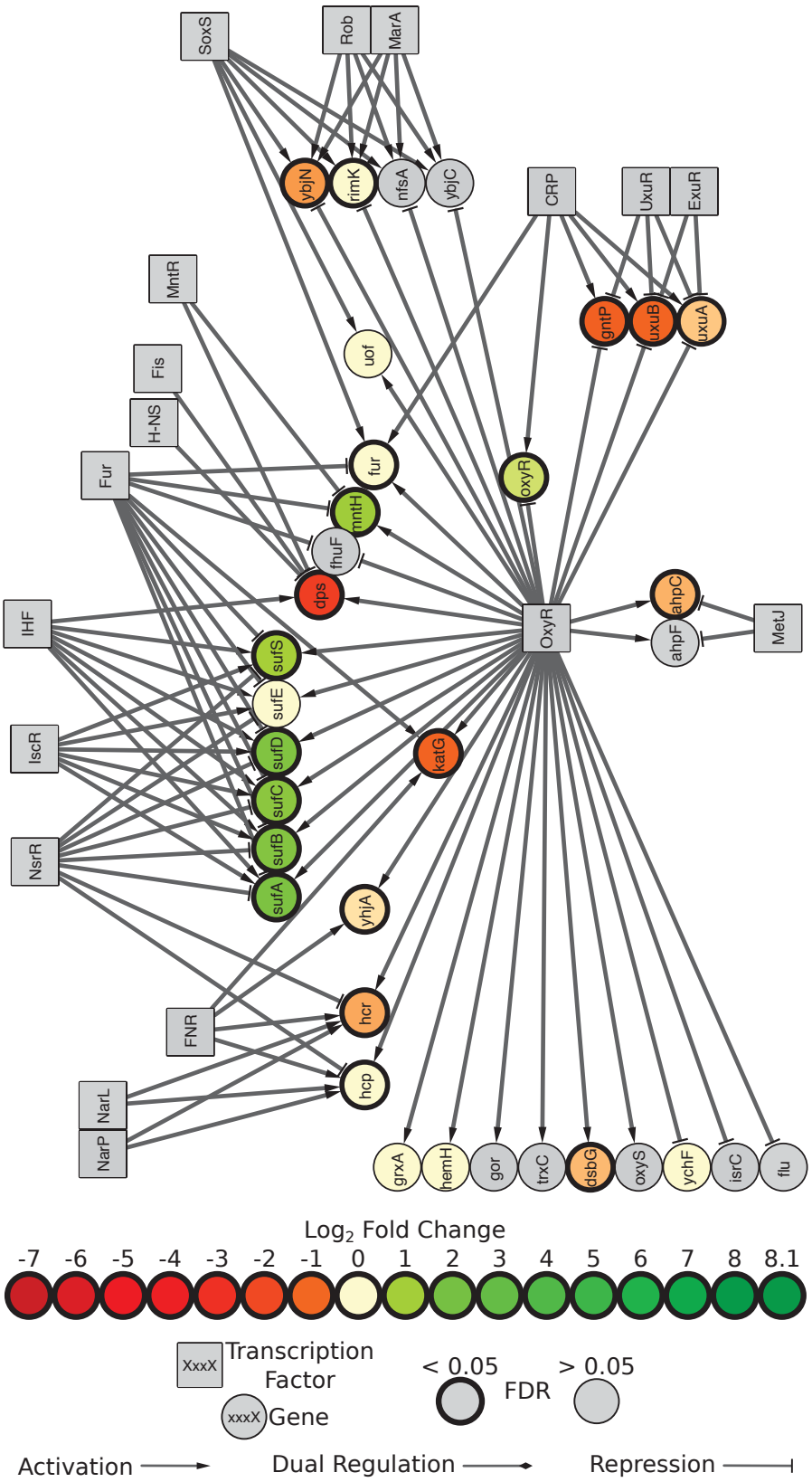

Figure 23: Differential expression of the *oxyR* regulatory network in response to wildtype pIV.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 Fold Changes reported by DESeq2 are relative to the vector control. Colour scale is limited to the minimum and maximum fold-changes of the genes mapped to the entire RegulonDB network.

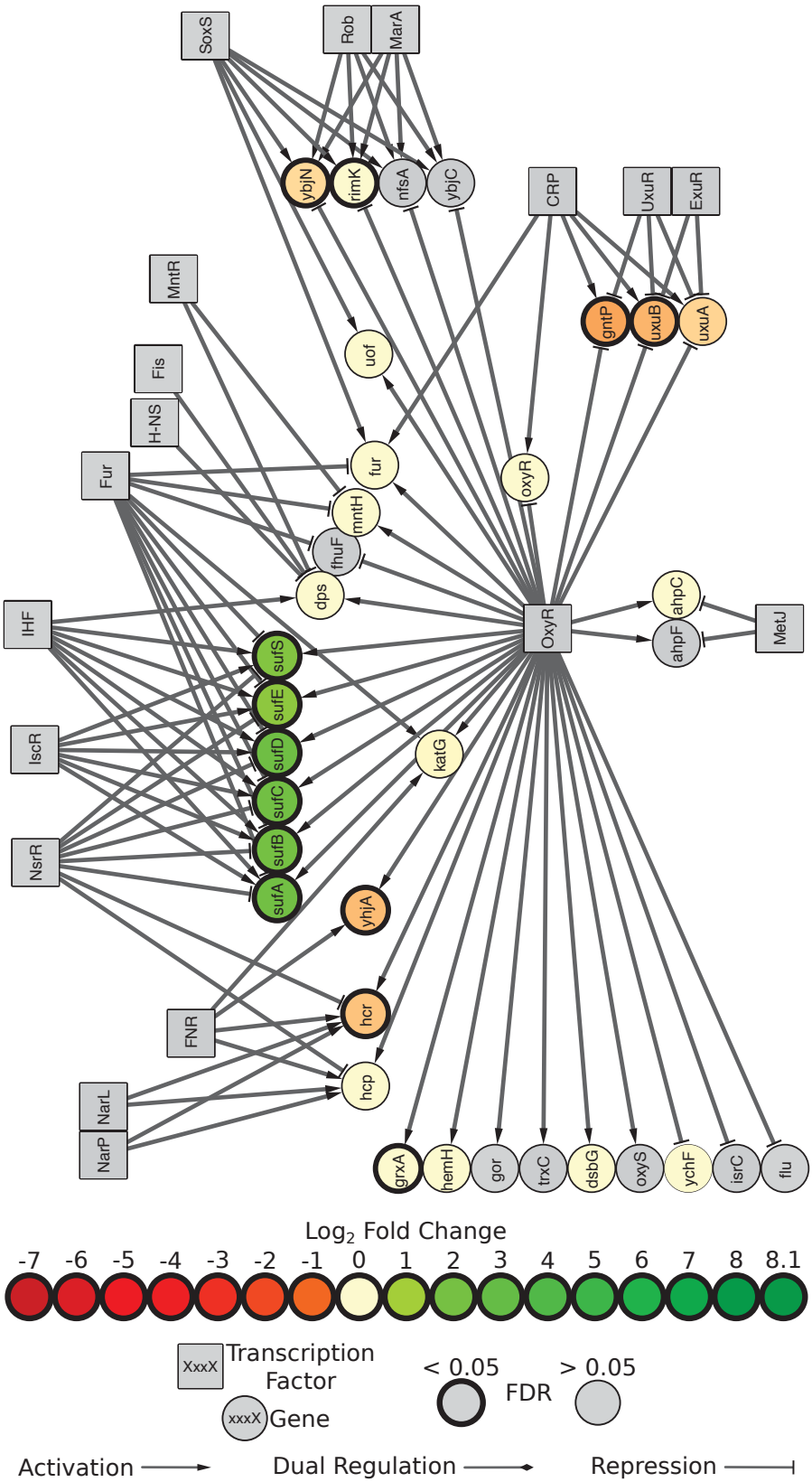

Figure 24: Differential expression of the OxyR regulatory network in response to leaky mutant pIV-E292K.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 fold changes reported by DESeq2 are relative to the vector control.

4.8 Induction of the Rcs stress response

The Rcs response (Figure 25) is induced by a variety of stresses, including peptidoglycan damage and osmotic shock (Laubacher & Ades, 2008, Sledjeski & Gottesman, 1996). It regulates the production of capsular polysaccharides and the expression of up to 2.5% of all *E. coli* genes. The regulatory circuit includes two regulators, RcsA and RcsB, which upon induction bind to the regulatory sequences as heterodimers to induce colanic acid and type 4 capsule polysaccharide export operons, *wza-wzb-wzc-wcaAB* and *yjbEFGH*, respectively, and inhibit expression of motility regulator, *flhDC*, and the curli pilin assembly operon, *csgDEFG* (Wehland & Bernhard, 2000). RcsB can also act alone as a homodimer to regulate a number of genes involved in protein folding/sorting, acid resistance and osmotic shock. RcsB also forms regulatory complexes with BglJ (RcsB-BglJ) and GadE (RcsB-GadE) to regulate a number of transcription factors and IM proteins (Castanie-Cornet *et al.*, 2010, Salscheider *et al.*, 2014). The genes, *rcsB* and *rcsD*, are collocated in the same operon and are transcribed together. The gene, *rcsC*, is located downstream of the *rcsDB* operon but convergently oriented. The remaining *rcs* genes, *rcsA* and *rcsF*, are transcribed from unlinked operons of which they are the sole transcript.

The pIV-E292K mutant displays a highly mucoid phenotype, suggesting a massive production of capsular- and/or exo- polysaccharides. Consistently, the abundance of the *rcsA* transcript increases 77-fold in cells producing the mutant, pIV-E292K, and only 2.7-fold in cells producing wildtype pIV (Figure 25). The *rcsC* and *rcsD* sensor-encoding genes are marginally (2-fold or less)

but significantly up-regulated in cells producing pIV-E292K and repressed in cells producing wildtype pIV. The *rcsF* gene encoding the outer membrane component of the sensor is marginally (less than 2-fold) but statistically significantly down-regulated in cells producing either of the pIV variants.

Expression of *rcsA* is controlled directly by RcsAB (Figure 25); regulation of the remaining Rcs phosphorelay is unclear. A sigma-E dependent promoter has been predicted for *rcsC*, whether this is physiologically relevant is yet to be determined (Salgado *et al.*, 2013).

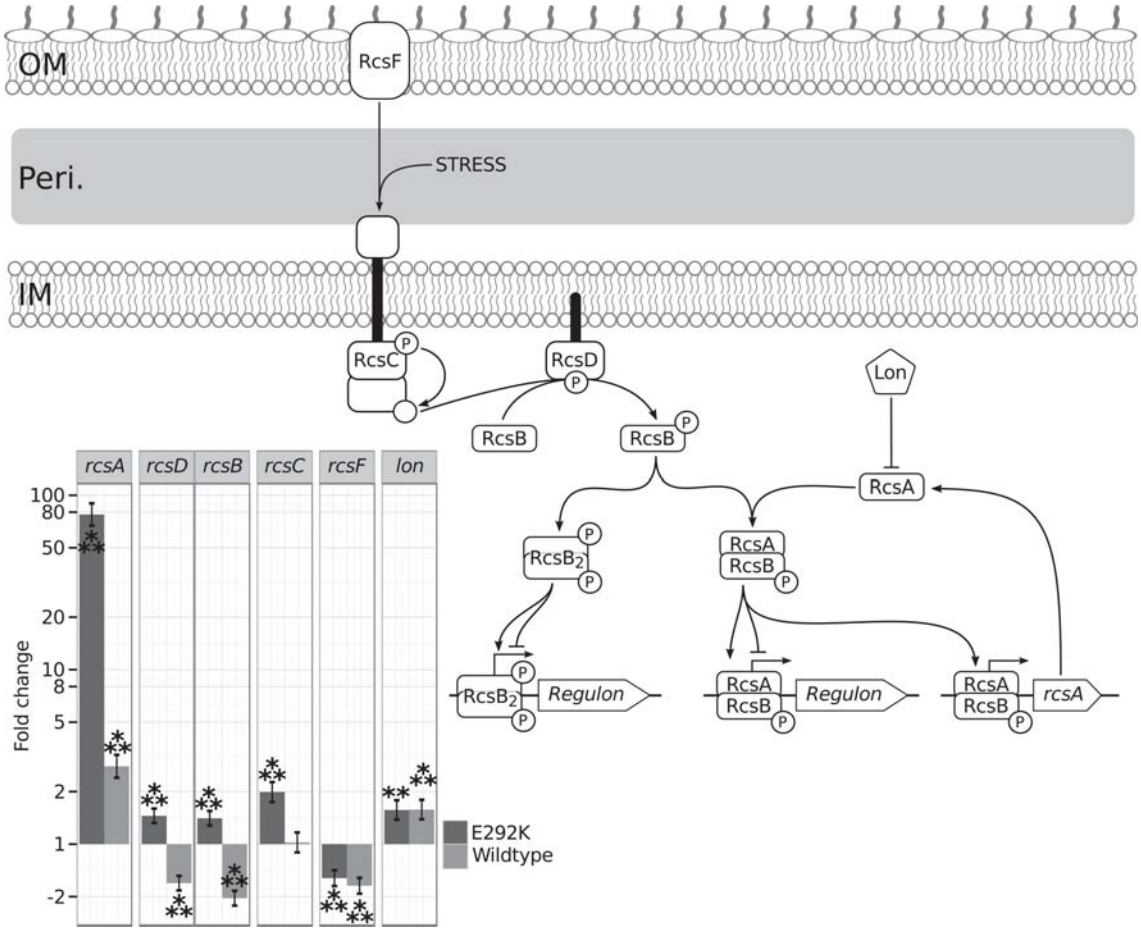

Figure 25: Differential expression of Rcs response regulators.

Top, regulatory pathway of the Rcs regulon. Bottom left, changes in expression. Fold changes relative to the vector control are displayed on the y-axis, * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped.

4.9 Differential expression of the RcsB and RcsAB regulons

There are two branches of genes directly regulated by the Rcs response; the RcsA-independent regulon, requiring only phosphorylated RcsB for regulation; and the Rcs-dependent regulon, requiring RcsB to interact with RcsA to regulate member genes.

Most genes transcriptionally activated by the RcsB homodimer (RcsB) are involved in protein folding and sorting (*lolA*, *hdeAB*), cell division (*ftsA*, *ftsZ*), osmotic shock (*bdm*, *osmB*, *osmC*) and indirectly, acid resistance by regulating transcription factors of the Gad regulon (*gadX*, *gadY*). The RcsB-homodimer regulated genes *bdm* and *osmB*, encoding a poorly characterised modulator of motility and osmotically inducible lipoprotein, respectively, are highly induced (Figure 26A), nearly 20 and 50-fold, respectively, in the cells producing pIV-E292K, whereas they only show a 2-fold increase in transcription in the cells producing wildtype pIV. The *gadAX* operon, encoding glutamate decarboxylase (involved in acid resistance) and a DNA-binding transcriptional regulator, respectively, is repressed by the RcsB homodimer, and activated by RcsB-GadE. Of the two genes, only *gadX* showed a greater than 2-fold decrease in transcription in cells producing wildtype pIV (Figure 26B), whereas no change was detected in cells producing the pIV-E292K mutant. Furthermore, *lolA*, *rarA*, *osmC*, and *ydeP*, respectively encoding; a component of the outer membrane lipoprotein translocation machinery, LolA; a protein involved in replication-associated recombination, RarA; an osmotically inducible peroxiredoxin, OsmC; and an acid resistance protein co-regulated with EvgA, YdeP; show greater than 2-fold differential increases in transcription in cells producing either of the

pIV variants relative to the vector control. Other RcsB-homodimer regulated genes show no significant increase in transcription in the cells producing pIV variants.

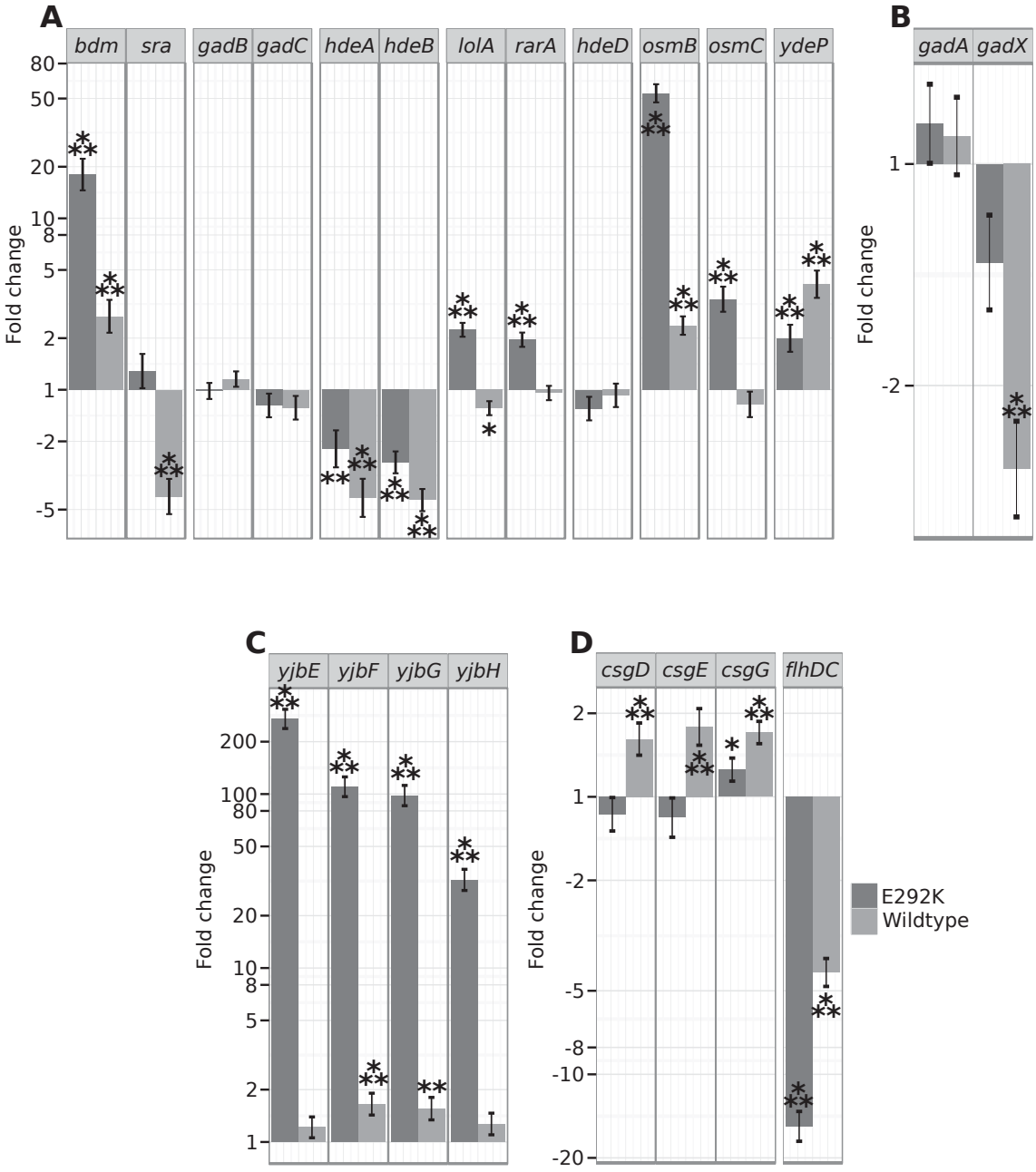

Figure 26: Differential expression of RcsB and RcsAB regulons.

(A) Genes activated by RcsB. **(B)** Genes repressed by RcsB. **(C)** Genes activated by RcsAB. **(D)** Genes repressed by RcsAB. Fold changes relative to the vector control, reported by DESeq2, are displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped.

The transcripts corresponding to capsule synthesis operon *yjbEFGH*, regulated by the RcsAB-heterodimer are highly induced (up to 200-fold) in cells producing leaky pIV-E292K (Figure 26C). In contrast, wildtype pIV whose production does not result in a noticeably mucoid phenotype, accordingly, does not induce this or other type 4 capsule polysaccharide biosynthesis operons. In addition to the *yjbEFGH* operon, RcsAB also positively regulates transcription of the *wza-wzb-wzc-wcaAB* operon. However, these genes were removed from the dataset by manual prefiltering (Figure 11, Section 3.2.1) prior to DESeq2 analysis due to high variance within sample replicates. Independent differential expression testing was performed on these genes (using identical normalisation factors used in the primary analysis), revealing they are highly induced in response to pIV-E292K (but not wildtype), similarly to the increased transcription observed in the *yjbEFGH* operon. The reason for sample-to-sample variability in the expression of this operon remains to be identified.

Genes repressed by RcsAB are not significantly differentially expressed in response to pIV-E292K, except the master regulator of motility, pseudogene *flhDC*, (Figure 26D). The regulator of curli pilin synthesis, *csgD*, and assembly components *csgEFG*, are transcribed in response to wildtype pIV only.

In summary, the observed differential expression of RcsAB-activated genes, indicates that the RcsAB regulon is important for the mitigation of secretin stress when gate function is compromised and when the gate is intact the other stress responses are sufficient to prevent damage.

4.10 Differential expression of transporters

The previous section showed that production of the leaky pIV-E292K mutant results in strong induction of a putative type 4 capsular polysaccharide operon, *yjbEFGH* (Figure 26C, Figure 27B & D), and similar differential expression of a number of osmotically sensitive OMPs and lipoproteins. Differential expression testing of the colanic acid capsule operon, *wza-wzb-wzc-wcaAB*, performed independently of the main transcriptome, indicated a similar transcriptional response to pIV-E292K, but not wildtype pIV (Figure 27A). Given the differential expression of osmotically sensitive lipoproteins, it was prudent to investigate expression of polysaccharide transport-assembly genes. Note that it is not clear which of these are used in the capsular polysaccharide synthesis induced by pIV-E292K.

The *pgaABCD* operon is required by *E. coli* to produce another non-colanic acid capsule exopolysaccharide, poly-beta-1,6-N-acetyl-D-glucosamine (PGA) (Wang *et al.*, 2004). Transcription of this operon is regulated by NhaR, a sodium-sensitive DNA-binding transcriptional regulator whose activity is regulated by intracellular Na^+ and is co-transcribed with the Na^+:H^+ antiporter, NhaA (Figure 28B, C, D & E) (Rahav-Manor *et al.*, 1992). Both genes show up to 5-fold increased transcription in response to production of wildtype pIV. However, while significant increases in transcription of *nhaA* and *pgaD* were observed in cells producing pIV-E292K transcriptome, the change was at least 2-fold lower than that seen in response to wildtype pIV (Figure 28E).

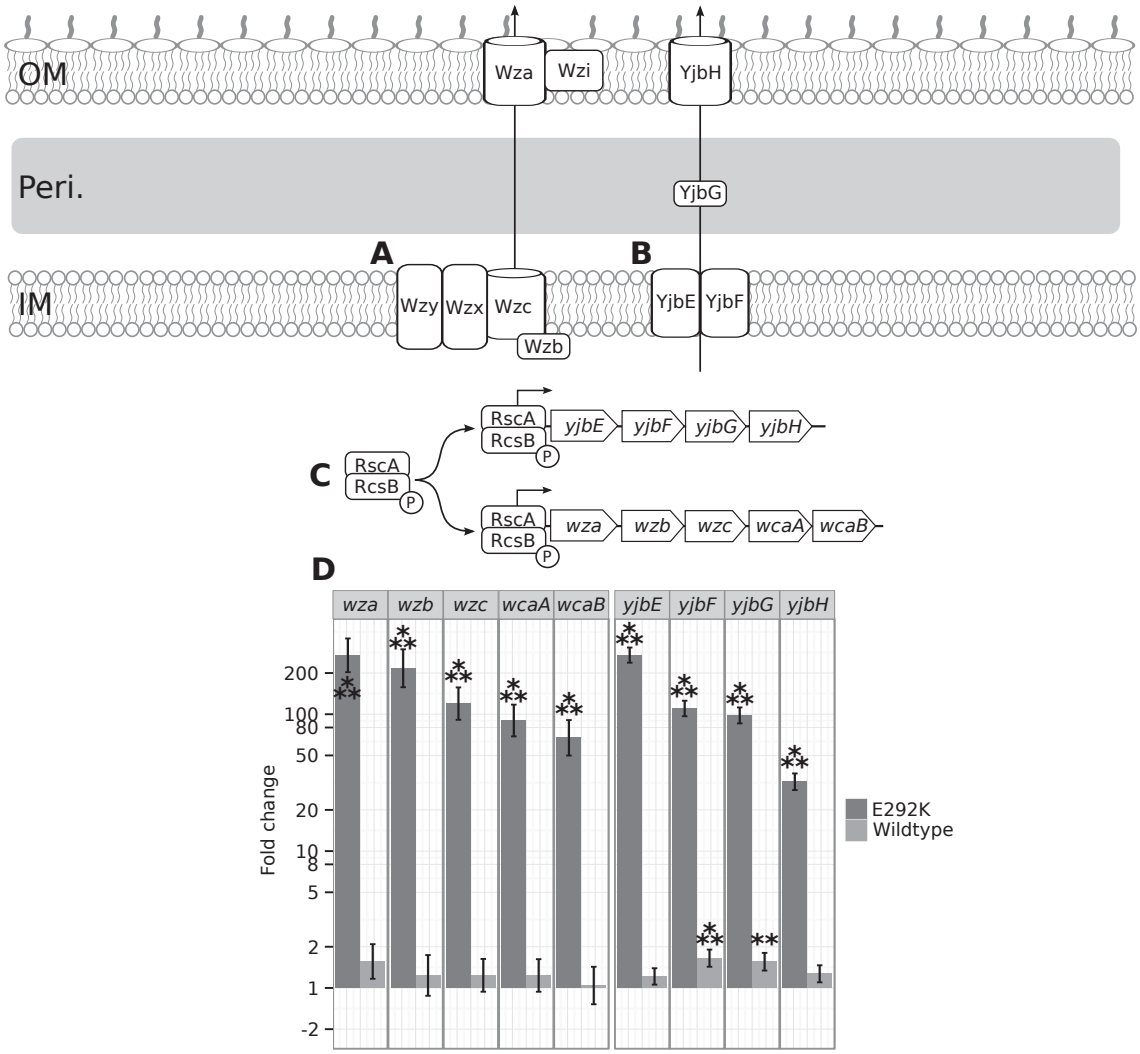

Figure 27: Differential expression of RcsAB-regulated capsule polysaccharide export systems.

(A) Schematic representation of the Wza colanic acid capsule export pathway. **(B)** Representation of the Yjb colanic acid capsule export pathway. **(C)** Schematic representation of the RcsAB-dependent regulation of *wza-wzb-wzc-wcaAB* and *yjbEFGH* operons. **(D)** Differential expression demonstrated in this work. Fold change relative to the vector control reported by DESeq2 is displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped. Differential expression testing of *wza*, *wzb*, *wzc*, *wcaA* and *wcaB* was performed as an independent test (using DESeq2) on that subset of genes only. Normalisation of the raw counts was performed identically to the rest of the data. Subsequent correction for multiple testing was also performed identically using the BH method (Benjamini & Hochberg, 1995).

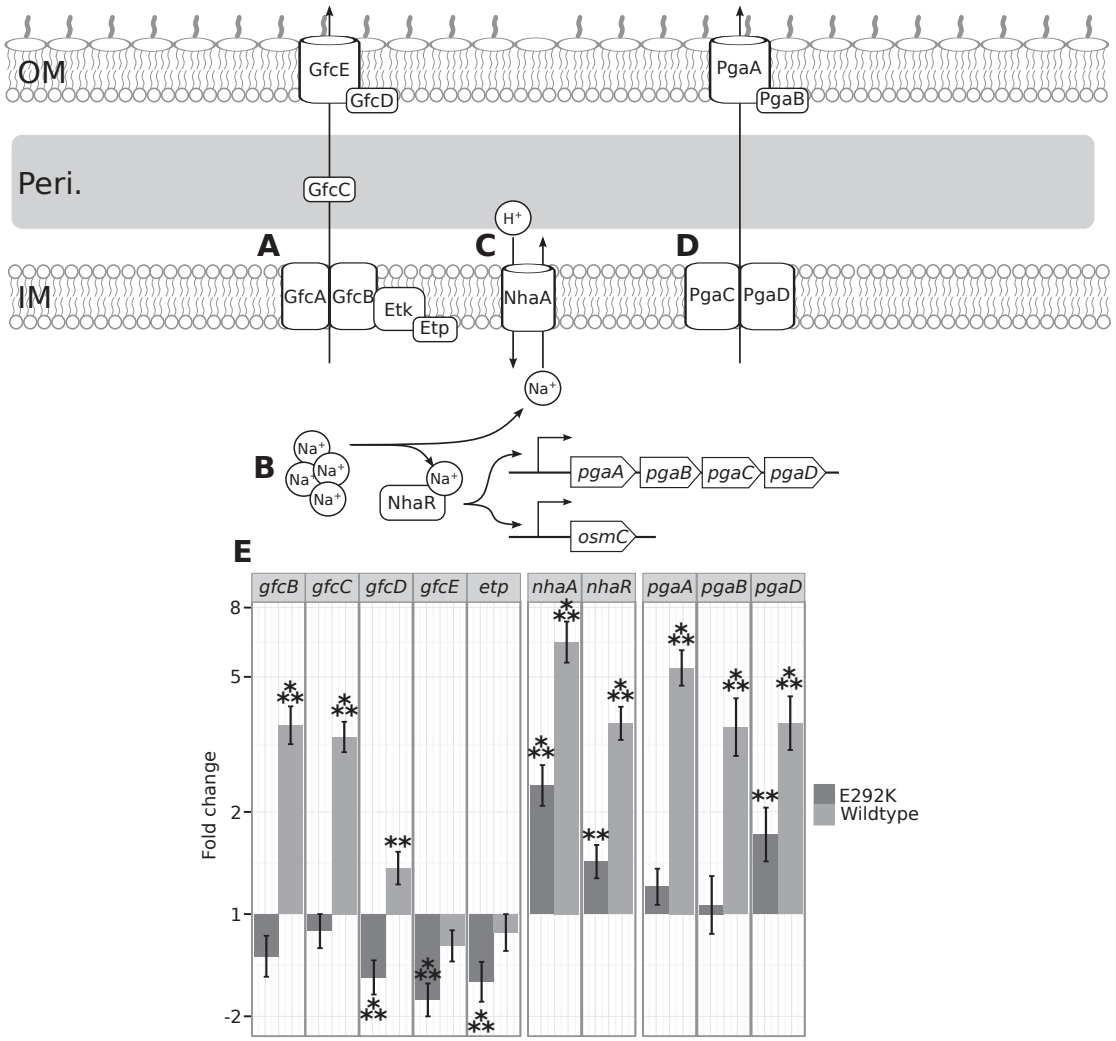

Figure 28: Differential expression of capsule export systems not regulated by RcsAB

(A) Schematic representation of GfcABCDE-Etp-Etk export system for group 4 capsules. **(B)** High cytosolic Na^+ concentrations activate the DNA-binding transcriptional regulator, NhaR. **(C)** Schematic representation of the Na^+:H^+ antiporter, NhaA. **(D)** Schematic representation of the PGA capsule export system. **(E)** Differential expression demonstrated in this work. Fold change relative to the vector control reported by DESeq2 is displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped.

The *gfcABCDE-etp-etk* operon required for type 4 capsule synthesis by uropathogenic *E. coli* O127:H6 is reported to be inactive in *E. coli* K-12 strains owing to a promoter-disrupting insertion (Peleg *et al.*, 2005). However, the production of wildtype pIV increased transcription significantly in three out of the seven genes in this transcription unit. Significantly decreased transcription was also observed in three of the seven genes in response to the production of leaky pIV-E292K; however, this change was not greater than a factor of 2 (Figure 28A & E). It is unknown whether transcription of these genes is biologically significant and no known promoter sequence was predicted in the 5' region preceding *gfcA*.

Aside from producing capsular polysaccharides, Gram-negative bacteria avoid osmotic shock by regulating the transport of osmolytes such as K^+, glycine betaine, proline and arginine. The latter may be imported into the cytoplasm as an osmoprotectant through the import channel ArtPQM (Figure 29A). A mild (less than 2-fold) but statistically significant increase in transcription was observed in the genes encoding the IM arginine import channel, *artPQM*, but not the gene encoding the arginine periplasmic binding protein, *artI*, in response to the leaky mutant, pIV-E292K. Conversely, a statistically significant decrease (up to 2-fold) in transcription of these genes was observed in response to wildtype pIV (Figure 29C). Three transporters that import glycine-betaine or proline into the cell exist in *E. coli*; an ABC-type transporter encoded by *proVWX*; a betaine/proline:H^+ symporter encoded by *proP*; and a proline:H^+ symporter encoded by *proY* (Figure 29B), co-transcribed with *brnQ*, a branched amino acid transporter. All genes showed very strong induction by both variants

of pIV. Interestingly, increase in transcription (up to 10-fold) in response to production of wildtype pIV was more prominent than response to the leaky pIV-E292K mutant (up to 5-fold; Figure 29C).

During conditions of high osmolarity, the trehalose synthesis pathway is activated, using glucose-6-phosphate, derived from glycolysis, to synthesise trehalose in a two-step reaction utilising OtsA and OtsB (Figure 30). Transcription of *otsA* and *otsB* increases up to 3-fold in response to the leaky mutant, pIV-E292K. However, transcription of these genes decreases by up to 3-fold in response to wildtype pIV. The trehalose-6-phosphate sensitive transcriptional regulator, TreR, is transcribed, 2-fold more (relative to the vector control) in response to the leaky mutant, pIV-E292K. A mild (less than 2-fold), but statistically significant increase in transcription of *treR* in response to wildtype pIV was also observed. TreR, in the absence of trehalose-6-phosphate, represses the *treBC* operon. The *treB* gene encodes the IM permease responsible for the import and phosphorylation of α,α-trehalose (Figure 30A), while *treC* encodes a hydrolase converting trehalose-6-phosphate back to glucose in conditions of low osmolarity (Figure 30D). Reads from these genes (and *treF*, *crr*) were originally removed from analysis due to high variation within sample replicates. However, subsequent independent testing revealed an up to 30-fold decrease in transcription of *treBC* in response to leaky pIV-E292K, and up to 170-fold decrease in response to wildtype pIV, relative to the vector control (Figure 30F). This indicates trehalose-6-phosphate is limited and TreR is repressing the *treBC* operon.

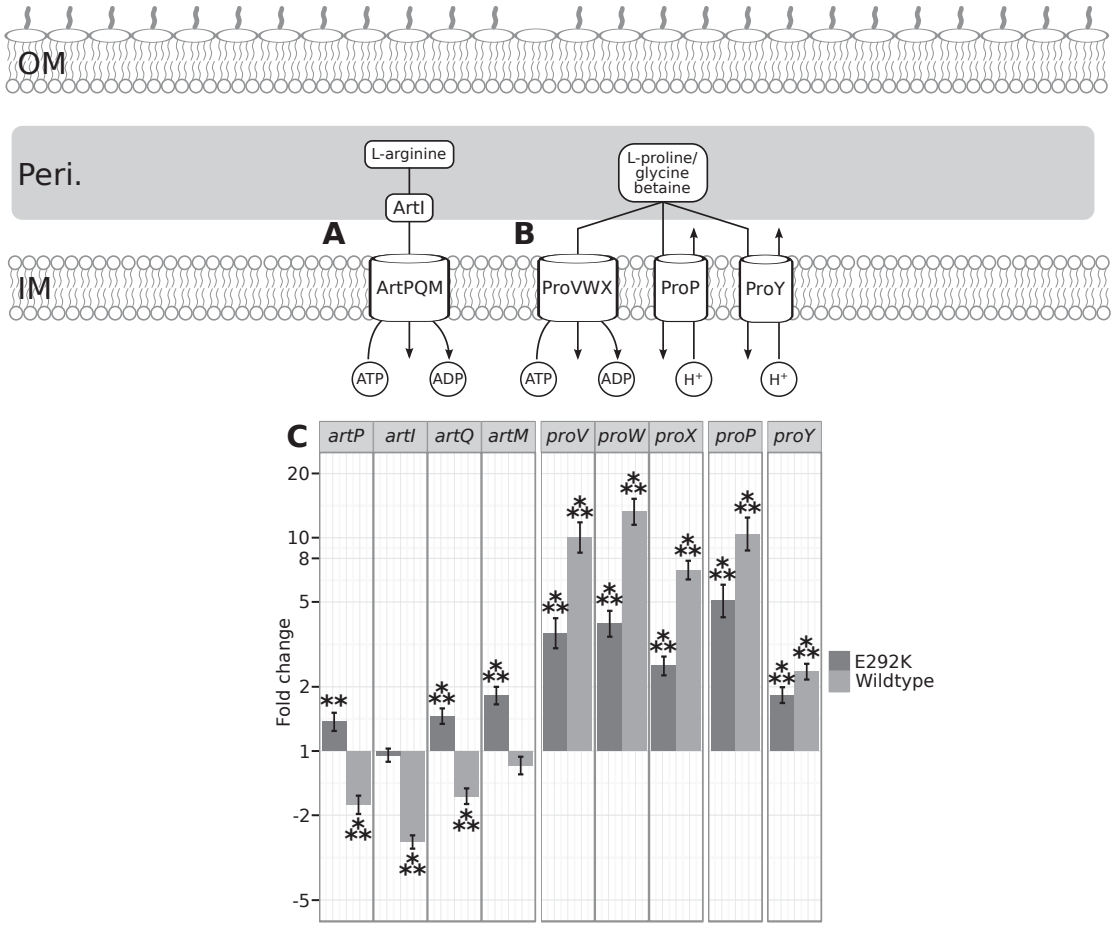

Figure 29: Differential expression of osmoprotectant transporters.

(A) Schematic representation of the L-arginine import system, ArtPIQM. **(B)** Schematic representation of the L-proline/Glycine-betaine import systems encoded by *proVWX*, *prop* and *proY*. **(C)** Differential expression demonstrated in this work. Fold change relative to the vector control reported by DESeq2 is displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped.

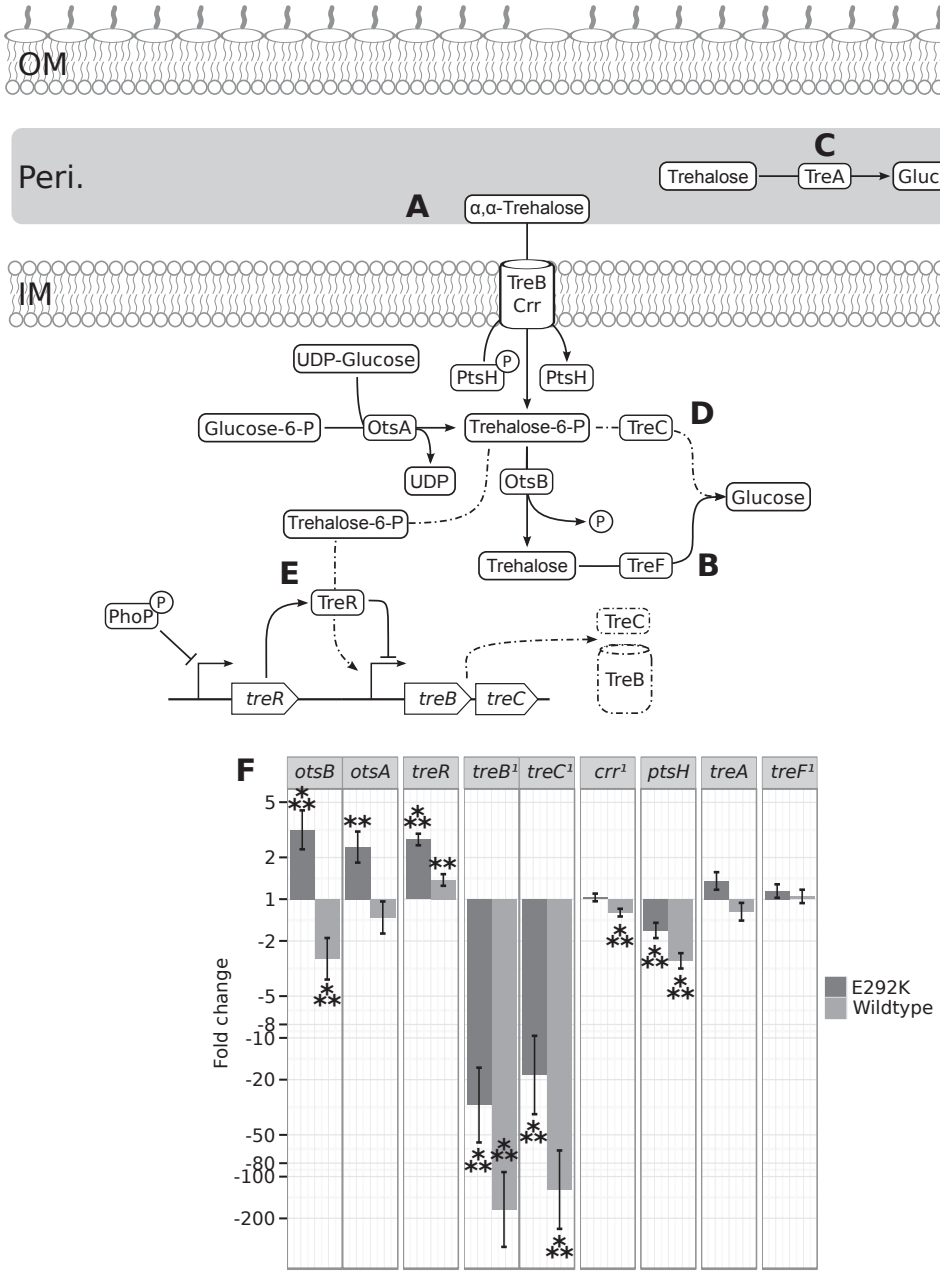

Figure 30: Differential expression of the trehalose synthesis pathway.

(A) Schematic representation of the trehalose synthesis pathway. The precursor, α,α-trehalose is imported and phosphorylated by the phosphotransfer permease, TreB/Crr. Alternatively, Glucose-6-P, derived from glycolysis, is converted into trehalose-6-P and finally trehalose, by OtsA and OtsB, respectively. Under high osmolarity conditions, two trehalases exist to convert trehalose back to glucose, TreF in the cytoplasm **(B)**, and TreA in the periplasm **(C)**. Under low osmolarity conditions (dashed lines), another hydrolase, TreC **(D)**, may also convert trehalose-6-P to glucose. In the absence of trehalose-6-P, the transcriptional regulator, TreR, represses the transcription of the *treBC* operon **(E)**. However, in the presence of trehalose-6-P or during low osmolarity (dashed lines) conditions, TreR derepresses and allows transcription of the *treBC* operon. **(F)** Fold change relative to the vector control reported by DESeq2 is displayed on the y-axis. * FDR 0.01−0.05, ** FDR 0.001−0.01, *** FDR <0.001. Genes within the same transcriptional unit are grouped. [1] Read counts for these genes (*treB*, *treC*, *treF* and *crr*) were removed from primary analysis due to high variability within sample replicates. Differential expression tests were performed independently (using DESeq2) on this subset of genes using identical normalisation and BH correction for multiple testing used in the primary analysis.

4.11 Discussion

The transcriptional response to secretin stress is not confined to the Psp response. The data shown here clearly indicates that in addition to induction of the Psp response, the CpxR, SoxS and Rcs regulons are all involved in the transcriptional response to secretin stress (Figure 17, Figure 18, Figure 20, Figure 25). Furthermore, wildtype pIV, despite not causing phenotypes typical of osmotic shock appears to “prime” cells for osmotic stress, in the event that the secretin gate is disregulated (as in the leaky pIV-E292K mutant). Given the natural permissivity of type 3 secretins, comparison of the transcriptional responses to wildtype pIV and leaky pIV-E292K gives insight into the adaptations by pathogens assembling type 3 secretion systems (Figure 2).

The repression of both osmosensitive porins, *ompC* and *ompF* (Figure 19), combined with the repression of the low-osmolarity *treBC* operon (Figure 30), and increased transcription of Proline importers (Figure 29) caused by production of wildtype pIV indicates production of closed-gate secretins causes some form of osmotic stress. Additionally, in response to the leaky mutant, pIV-E292K, cells increase transcription of *otsAB* (Figure 30) and the RcsAB-dependent colanic acid capsule export pathways (Figure 27). Intermediates from glycolysis (glucose-6-phosphate) and TCA pathways provide precursors for the production of trehalose and capsule exopolysaccharides. This, plus the increased transcription of ATP-dependent transporters (i.e. ProVWX) places pressure on the two main energy-producing metabolic pathways and is likely the reason the SoxRS superoxide stress response is induced (Figure 20).

The CpxR regulon may be activated in response to two distinct signals; CpxP-mediated repression of CpxA kinase activity is relieved by the detection and interaction of CpxP with misfolded or aggregated periplasmic proteins and consequent removal (or refolding) by DegP, allowing CpxR to be phosphorylated in a CpxA-dependent manner (Isaac *et al.*, 2005). Alternatively, CpxR may become phosphorylated by acetyl phosphate, an intermediate of the phosphotransacetylase-acetate kinase (Pta-AckA) pathway (Wolfe *et al.*, 2008). Thus, the CpxR regulon could be linked to both periplasmic protein quality and central metabolism.

The Sigma-E sigmulon is activated in response to misfolded proteins and the accumulation of off-pathway LPS intermediates in the periplasm (Lima *et al.*, 2013). The operon encoding the Sigma-E response regulators, *rpoE-rseABC* (Figure 18), is repressed by phosphorylated CpxR. However, transcription of *rpoE-rseABC* is decreased in response to wildtype pIV but not the leaky mutant, pIV-E292K. This suggests Sigma-E is released from RseA in response to pIV-E292K, but not wildtype pIV, and that phosphorylated CpxR and Sigma-E compete at the *rpoE* promoter, to prevent increased transcription of the Sigma-E sigmulon in response to misfolded or aggregated proteins. Even though it is tempting to speculate that overexpression of the leaky mutant, pIV-E292K, causes the stress of misfolded protein in the periplasm, however this protein permits phage assembly and export, hence it is unlikely that it has a defect in folding (Table 3; Spagnuolo *et al.* (2010)). The 4-fold increase in transcription of *cpxP* (an indicator of CpxR phosphorylation) in response to pIV-E292K, relative

to that caused by wildtype pIV could be attributable to metabolic stress caused by intrinsic mistargeting of pIV secretin to the inner membrane, exacerbated by the leaky phenotype of the pIV-E292K mutant.

4.12 Conclusion

Secretins are among the largest of all Gram-negative membrane proteins. Thus, it is not surprising that the cellular response to their production is not confined to a single stress response. Secretins are among the first proteins of type 2/3 secretion systems to be expressed and assembled when these systems are induced. Their C-terminal domain normally spans the outer membrane while N-terminal domains penetrate into the periplasm, parting the peptidoglycan layer. Thus, assembly of secretins into OM affects membrane biogenesis and integrity. Misintegration into the inner membrane will permeablise this membrane that is essential for maintenance of the proton gradient and protonmotive force.

Breaking from the findings of Lloyd *et al.* (2004) and Maxson and Darwin (2004), that the response to secretin stress is confined to the Psp regulon induction, the data presented here clearly indicates stimulation of the CpxR, Rcs and SoxRS regulons in response to secretin stress in addition to the induction of the Psp response. Additional transcriptional relays are also stimulated to fine tune the secretin stress response and prepare cells for secretin-gate disregulation/opening. This suggests that the Rcs and Cpx induction found by Walker and Miller (2009) and Carlsson *et al.* (2007) by expression of entire type 3 secretin system and its secreted effector proteins in

Y. enterocolitica could be partially or fully due to the secretin component. The transcriptomics data presented here represent a starting point for epistatic analysis to resolve which pathways and regulators act as the critical conductors of the concerted secretin stress response.

Chapter 5

5 Roles of differentially expressed pathways in the survival of secretin-producing cells

Although a number of stress response genes and regulons were induced by production of the pIV variants, these are not all necessarily required for survival of secretin-producing *E. coli*. To determine importance of differentially expressed genes in the survival of *E. coli* K-12 cells experiencing secretin stress, viability of insertional inactivation mutants of stress response regulators and effectors (obtained from the Keio collection; Baba *et al.* (2006)) under secretin stress was analysed. Mutants of the key regulators that control the stress responses (*pspF*, *rcsA*, *rcsB*, *soxR*, *soxS*) were analysed as they cover the complete corresponding stress response. In addition, the gene encoding the periplasmic chaperone, *spy*, (Quan *et al.*, 2011) was also analysed because it was highly induced in response to the leaky pIV mutant, pIV-E292K, and much less in response to wildtype pIV.

To avoid the polar effect of kanamycin resistance cassette that replaced mutated genes in the used strains, only mutants of single-gene transcription units or those of genes located at the 3'-termini of the respective operons were analysed.

Insertional inactivation of *pspF*, *soxS* and *rcsB*, the key positive regulators of the corresponding stress responses prevented growth of *E. coli* K-12 over-producing leaky mutant pIV-E292K, but not wildtype (induced by 1 mM IPTG; Figure 31A). The synthetic lethality of the inactivation mutants when pIV was only marginally expressed was also assessed. This was possible because *gIV*

is expressed at a low level in the absence of IPTG (data not shown). At this low expression level, leaky mutant pIV-E292K still reduced the viability of the *pspF* mutant by 3 orders of magnitude, but the viability of the *rcsB* and *soxS* mutants was not compromised (Figure 32A). This suggests that the component of leaky pIV stress that relieved by Psp is more detrimental to *E. coli* than the components of secretin stress relieved by the RcsB or SoxS regulons.

To determine whether the observed phenotype was solely due to the mutation in each corresponding gene and not due to the polarity on other genes, inactivation mutants were each complemented by a plasmid expressing the same gene (from the ASKA collection, Kitagawa *et al.* (2005); Figure 31B, Figure 32B). In most cases, the complementing plasmid was able to rescue the inactivation of the response regulator and restore viability. Plating efficiency of the complemented *pspF* mutants was increased by 4 orders of magnitude. However, it was still 1 order of magnitude below the *pspF*-wildtype parent strain, probably due to differences in plasmid versus chromosomal expression.

Unexpectedly, overexpression of *rcsA*, *lon* and *katE*, which were not required for survival of the cells producing pIV-E292K, specifically reduced viability of the cells expressing this mutant by 3 to 4 orders of magnitude. Overexpresion of *soxR* also caused a slight decrease in plating efficiency relative to the vector control in cells producing either of the pIV variants.

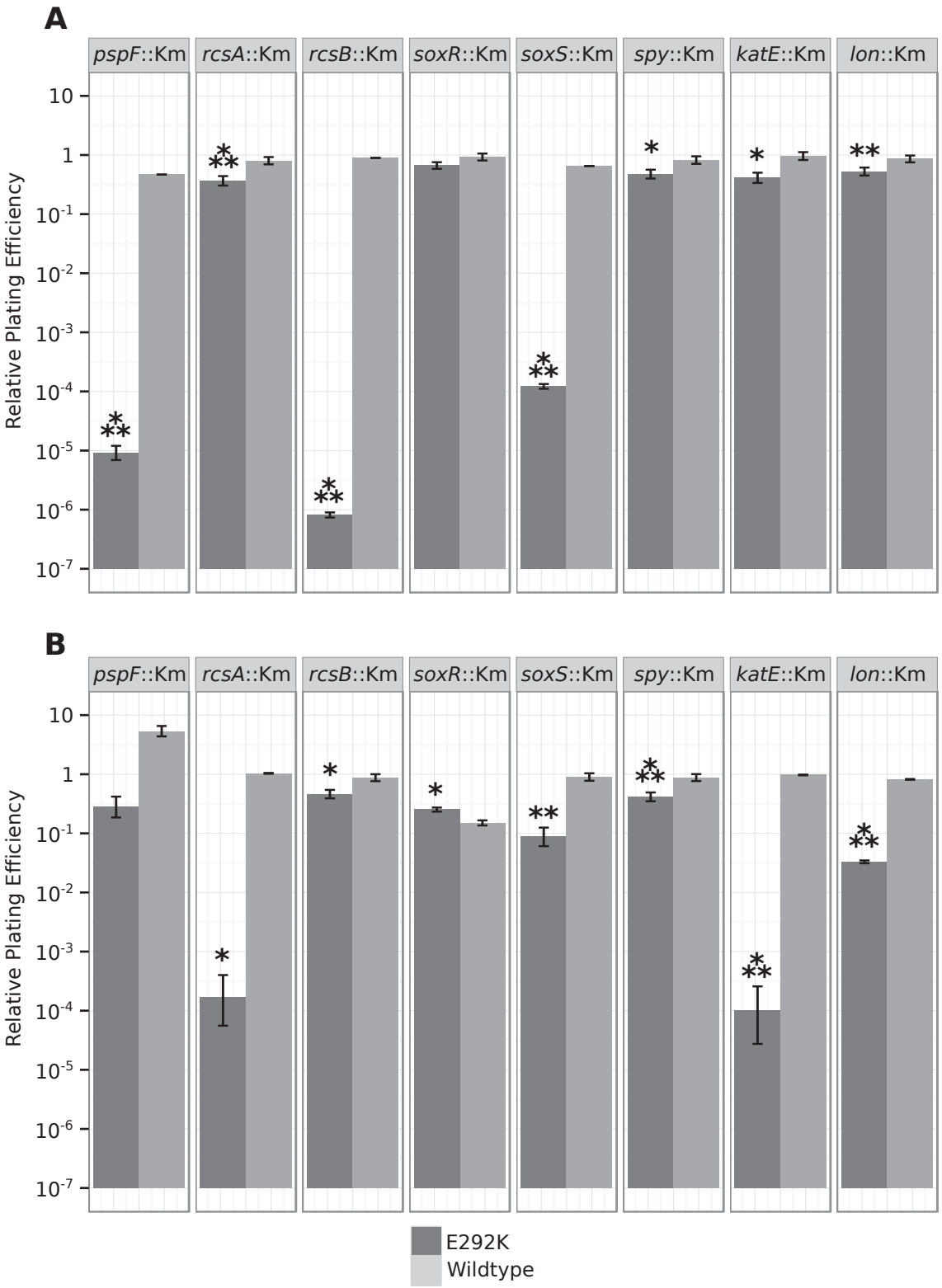

Figure 31: The roles of stress-induced or regulatory genes in survival of *E. coli* K-12 under pIV secretin stress.

(A) Plating efficiency of chromosomal deletion mutants producing the pIV variants relative to the vector control. The expression of pIV variants was induced by 1 mM IPTG. **(B)** Relative plating efficiencies of complemented chromosomal mutants co-expressing pIV variants. Both the complementing gene and *gIV* expression was induced by 1 mM IPTG. The increased plating efficiency of the complemented *pspF*::Km mutant co-expressing wildtype gIV, relative to the vector control, is likely due to a pipetting error. All chromosomal mutants contained Km cassette replacing the complete ORF (Baba *et al.*, 2006). Complementing plasmids express the ORFs from a *lac* promoter (Kitagawa *et al.*, 2005). Each expressed ORF corresponds to the gene deleted from the chromosome. In all cases, relative plating efficiency is displayed on the log-transformed y-axis, error bars are the 95% confidence intervals derived from exact Poisson tests. Welch's one-tailed t-test was used to determine whether colony counts were significantly lower than that of the vector controls with a minimum of 3 biological replicates. * indicates $p < 0.01 – 0.05$, ** indicates $p < 0.001 – 0.01$, *** indicates $p < 0.001$.

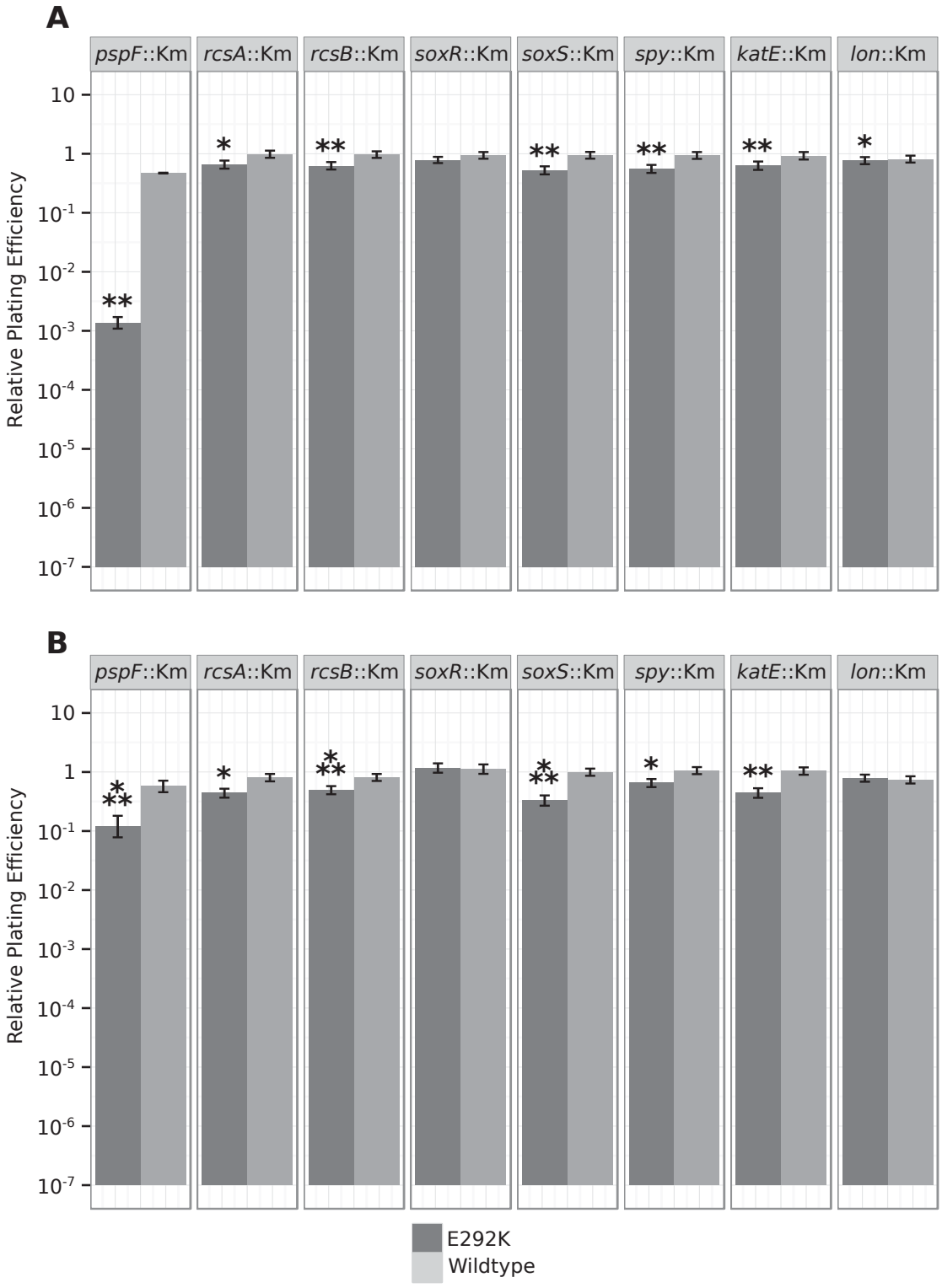

Figure 32: The roles of stress-induced or regulatory genes in survival of *E. coli* K12 under weak pIV secretin stress.

(A) Plating efficiency of chromosomal deletion mutants expressing the pIV variants at a low level (without IPTG induction), relative to the vector control. **(B)** Relative plating efficiencies of complemented chromosomal mutants co-expressing pIV variants at a low level (without IPTG induction). All chromosomal mutants contained Km cassette replacing the complete ORF (Baba *et al.*, 2006). Complementing plasmids express the ORFs from a *lac* promoter (Kitagawa *et al.*, 2005). Each expressed ORF corresponds to the gene deleted from the chromosome. In all cases, relative plating efficiency is displayed on the log-transformed y-axis, error bars are the 95% confidence intervals derived from exact Poisson tests. Welch's one-tailed t-test was used to determine whether colony counts were significantly lower than that of the vector controls with a minimum of 3 biological replicates. * indicates $p < 0.01 – 0.05$, ** indicates $p < 0.001 – 0.01$, *** indicates $p<0.001$.

5.1 Discussion

Analysis of inactivation mutants of the key regulators for phage-shock (Psp), capsular polysaccharide (Rcs) and superoxide (Sox) responses indicated the requirement for these regulons only when the integrity of the secretin gate is compromised and then, with the exception of *psp*, only when the secretin is overexpressed (Figure 31A & Figure 32A). Small (<10-fold), but significant decreases in plating efficiency were also observed in *katE*::Km, *lon*::Km, *rcsA*::Km and *spy*::Km backgrounds expressing the leaky mutant, pIV-E292K (Figure 31A). Expression of wildtype pIV did not significantly decrease plating efficiency in any genetic background analysed.

Loss of *pspF* prevents expression from the *pspA-D* and *pspG* promoters. The only Psp protein that could be produced in *pspF*::Km backgrounds is PspE. However, differential gene expression analysis indicated *pspE* is down-regulated during secretin stress (Figure 17). The growth defect, and restoration by *in trans* complementation, of *pspF*::Km cells expressing pIV-E292K (Figure 31) suggests that one, or more, of the effectors PspABCD or PspG is required for survival of leaky secretin stress. However, it is unlikely that PspA contributes to survival, since cells expressing the naturally permissive secretins YscC and InvG, tolerate loss of *pspA* (Darwin & Miller, 2001, Horstman & Darwin, 2012, Maxson & Darwin, 2006, Jovanovic *et al.*, 2009, Khanum, 2015). Similar viability analysis of the cultures producing leaky mutants of pIV in *pspA*-null backgrounds is yet to be completed; however, they are expected to behave similarly to the naturally permissive secretins that allow molecules >600 Da to cross the outer membrane.

Given the high level of *rcsA* up-regulation (Figure 25), it was expected that loss of *rcsA* would confer a growth defect in cells producing pIV-E292K. However, only inactivation of *rcsB* or *in trans* overexpression of *rcsA* caused growth defects in cells producing pIV-E292K (Figure 31). This indicates the RcsA-independent branches of the Rcs regulon are important in the survival of leaky secretin stress. Furthermore, the result suggests *yjb* or *wza* capsule production (Figure 27) during secretin stress could merely be a consequence of Rcs stimulation and may not provide any fitness benefit to cells expressing gate-permissive secretins.

Induction of superoxide stress responses is thought to be a consequence of the response to osmotic shock (Sévin & Sauer, 2014). In the absence of SoxS, cells producing pIV-E292K suffer a ~4-log reduction in plating efficiency, which is incompletely complemented by *in trans* expression of *soxS* (~10-fold reduction in plating efficiency). Given that *soxS* expression is dependent on the presence of SoxR, the lack of growth defects in *soxR*::Km cells producing pIV-E292K is highly suggestive that *de facto* basal expression of *soxS* can occur as a fail safe against loss-of-function mutations to the superoxide sensor, SoxR, that prevent it from activating the *soxS* promoter, as indicated by promoter anaylsis and clinical isolates bearing DNA-binding mutants of SoxR (Hidalgo *et al.*, 1998, Koutsolioutsou *et al.*, 2005). This data further implicates the superoxide stress response in osmotic shock.

Interestingly, *in trans* complementation of *katE* or *lon* in cells expressing pIV-E292K caused reductions in plating-efficiency (Figure 31B) However, their

absence only caused plating-efficiency of cells producing pIV-E292K to drop slightly (<10-fold; Figure 31B). Hydrogen peroxide, the substrate removed by catalase II, KatE, may provide an additional signal to activate transcription of genes required for survival of leaky secretin stress. Overexpression of *katE* could deplete this signal, preventing transcription of hydrogen-peroxide-activated genes. Lon protease, on the other hand, degrades the response regulators, SoxS, MarA, and RcsA (among others), preventing runaway activation of those responses. Overexpression of *lon* would prematurely shutdown these stress responses, preventing the amelioration of secretin stress.

5.2 Conclusion

This work represents a preliminary analysis of essentiality of the stress responses induced by secretins for preventing the lethal effect of opening a large channel (10 nm in diameter) in order to secrete virulence proteins or filamentous phage in *E. coli*. Given that only the expresson of the leaky mutant demonstrated synthetic lethality with three types of stress responses induced by pIV, it is clear that wildtype pIV secretin expressed on its own and, representing a closed or inactive state of the channel, does not impose as much stress as the “leaky” secretin mutant, pIV-E292K. This analysis showed that the extent of induction does not necessarily correlate with the requirement of this response for survival. For example, while both pIV variants equally induce the Psp regulon, only cells producing the leaky mutant pIV-E292K require the positive activator *pspF* and hence the actual expression of the *psp* stress response effectors in order to survive.

Because pIV localises to the OM and naturally mislocalises to the IM, any defect in gate integrity will cause stress at both IM and OM, effectively piercing the entire cell wall. The RcsB-dependent up-regulation of capsular polysaccharides in the leaky pIV−E292K mutant could be a means to thicken the cell wall, prevent the loss of osmolytes, and inhibit the passage of noxious compounds into the cell. Although *rcsA* expression was highly up-regulated in the cells producing the pIV-E292K mutant, the gene encoding the constitutively expressed co-regulator RcsB, but not RcsA, was essential for survival of the stress imposed by this secretin. Moreover, overexpression of *rcsA* from a *tac* promoter on a multicopy plasmid had detrimental effect on the pIV-E292K-producing cells, suggesting that at a certain concentration RcsA may inhibit expression of the secretin stress-relieving effectors of the *rcs* regulon.

Identification of response regulators critical to the amelioration of leaky secretin stress has pinpointed specific regulons as the target for more detailed analysis aimed to identify the particular effectors that mediate protection from the stress caused by the leaky pIV-E292K mutant. Once the crucial effectors of secretin stress have been identified and functionally characterised, it will be trivial to diagnose how secretins cause stress, the nature of that stress, and critically, will pinpoint novel targets for the development of potential virulence-reducing drugs through inactivation of the secretion-stress-protecting effectors.

.

Chapter 6

6 Discussion

6.1 General discussion

The transcriptional response of *Escherichia coli* K-12 to the presence of the pIV secretin, embedded within the inner- and outer-membrane was analysed by a stranded RNA-seq strategy. To investigate the additional stresses imposed by gate-permissive or substrate-secreting secretins, the transcriptome of a leaky mutant of pIV (pIV-E292K) was also sequenced using the same stranded approach. This is the first time the transcriptome of *E. coli* K-12 expressing a gate-permissive secretin has been sequenced, providing insight into the additional stress responses required by cells during secretion of substrates through a secretin. The data presented here indicates that the presence of pIV multimers in the IM and OM activates transcription of the PspF, Rcs, CpxR, and SoxS regulons (Figure 17, Figure 18, Figure 20, Figure 25), which respond to stress in all three layers of the Gram-negative cell wall (the inner- and outer-membranes and the periplasm). This breaks from previous findings that production of secretins only stimulates the PspF regulon, presumably due to their mislocalisation to the IM (Jovanovic *et al.*, 2006, Lloyd *et al.*, 2004, Seo *et al.*, 2007). Furthermore, the observation that pIV can directly stimulate activation of the CpxR and Rcs regulons (Figure 18, Figure 25) is in agreement with observations which indicate the function and assembly of the T3SS, including expression of secreted effector molecules, may be regulated by the CpxR and Rcs regulons (Carlsson *et al.*, 2007, MacRitchie *et al.*, 2012, Walker & Miller, 2009). Additional metabolic strain is imposed by leaky (or gate-permissive) secretins, such as pIV-E292K, as a consequence of increased

production of capsular polysaccharides, osmoprotectants, and ROS generation, resulting in the activation of the SoxS regulon – these observations provide further insight into the link between superoxide stress and osmotic shock (Sévin & Sauer, 2014).

Production of either wildtype pIV or the leaky mutant, pIV-E292K, induced similar transcriptional responses in the PspF regulon (Figure 17). This result was unexpected since western blots (Spagnuolo *et al.*, 2010) and ELISA experiments (Figure 9B) indicate PspA protein is produced in somewhat greater amounts in response to pIV-E292K production, compared that produced in response to wildtype pIV. A small RNA identified upstream of the *pspA* promoter (Shinhara *et al.*, 2011) may regulate the rate of translation of *pspA-D* transcripts. However, the presence and function of this small RNA is yet to be characterised in the context of secretin stress.

6.1.1 Psp response

Previous publications indicated secretins of the *Yersinia enterocolitica* T3SS require an intact Psp response to prevent growth defects during the production of pIV or T3SS secretins, YscC or YsaC in a *Yersinia enterocolitica* host (Horstman & Darwin, 2012, Seo *et al.*, 2007). However, similar experiments with wildtype pIV, have indicated they do not require an intact Psp response for survival of secretin stress in an *E. coli* K-12 host (Figure 31, Figure 32; Jovanovic *et al.* (2009)). This finding is interesting, since induction of Psp was shown to occur independently of PMF-depletion (Engl *et al.*, 2011) and by extension, independently of compromising IM integrity by secretins of different

secretin-gate permissivities. Therefore, while transcription of the PspF regulon is independent of secretin-gate integrity (Figure 17), the translation of *psp* transcripts and requirement of Psp for survival is dependent on secretin-gate integrity (Figure 9B, Figure 31A, Figure 32A; Horstman and Darwin (2012), Spagnuolo *et al.* (2010)).

According to synthetic lethality assays presented in this thesis, differences in the requirement for an intact Psp response appear to be variable, depending on permissivity of the secretin and genetic background. Subtle differences in OM composition and activities of effector proteins vary between species, therefore, production of foreign genes (e.g. pIV in *Yersinia*) could result in the imposition of additional stresses due to sub-optimal cellular machinery and conditions.

6.1.2 Cpx Response

Given the ~5-fold increase in transcription of *cpxP* in response to wildtype pIV, and nearly 20-fold increase in response to pIV-E292K (Figure 18), it is clear that the Cpx response is being stimulated by both variants of the pIV secretin studied here. However, the nature of pIV-dependent stimulation of the Cpx response is unclear. Transcription of *cpxP* requires CpxR to be phosphorylated, yet CpxP inhibits the autokinase activity of CpxA (Fleischer *et al.*, 2007, Raivio & Silhavy, 1997), unless it interacts with misfolded proteins in the periplasm and is subsequently degraded by DegP (Figure 7B; Isaac *et al.* (2005)). Alternatively, CpxR may be phosphorylated off-pathway, by acetyl phosphate from the Pta-AckA pathway (Wolfe *et al.*, 2008) allowing the response to be stimulated by changes in central metabolism and suggesting it is buffered

against stochastic stress signals. Transcription of *degP* increases 2-fold in response to pIV-E292K, but decreases 2-fold in response to wildtype pIV (Figure 18), indicating this variant stimulates the Cpx response via a periplasm signal. However, stimulation of the Cpx response by wildtype pIV (Figure 33) may result from increased turnover of glycolysis and the TCA cycle, since exposure of cells to excess carbon source can also elicit a Cpx response, independent of CpxA (Wolfe *et al.*, 2008).

Phosphorylated CpxR acts as a repressor of *rpoE-rseABC* transcription (De Wulf *et al.*, 2002). Transcription of this operon is repressed nearly 2-fold in response to wildtype pIV, relative to the vector control. Yet, no significant change in transcription of *rpoE-rseABC* was detected in response to pIV-E292K (Figure 18). Transcription of *rpoE-rseABC* should be inversely correlated with that of *cpxP*, as is the case in the wildtype pIV transcriptome. It is possible that aggregated or misfolded pIV-E292K interacting with CpxP may also interact the PDZ domain of DegS allowing the release of Sigma-E from RseA (Figure 7A). The release of Sigma-E from RseA would not increase transcription of *rpoE-rseABC* because CpxR is still activated and able to inhibit expression from this promoter. The leaky mutant, pIV-E292K, may have a folding defect sufficient to allow it to induce both Cpx and Sigma-E responses, while still forming multimers stable enough for phage production, passage of maltopentaose or vancomycin (Table 3; Spagnuolo *et al.* (2010)). However, overproduction of wildtype pIV only stimulates the Cpx response, likely through interaction with CpxP.

The Sigma-E response may also be stimulated by over-accumulation of LPS intermediates within the periplasm, by interacting with RseB. However, a full response requires that the PDZ domain of DegS also detect misfolded OMPs. Detection of either signal (OMPs or LPS) alone produces a buffered response (Lima *et al.*, 2013). It is unlikely that production of wildtype or pIV−E292K have caused defects in LPS biogenesis or assembly of the OM, since transcription of *rpoE* is not greater than the vector control. However, transcription of two CpxR-regulated genes involved in peptidoglycan stability and synthesis of undecaprenyl phosphate, *ycfS* and *bacA*, increased 8- and 3-fold in the response to pIV-E292K, respectively, relative to the vector control, (Figure 33, Figure 34, Note: *ycfS* is not shown, *ycfS* FDR = 6.1e-48). While this may be an indication of peptidoglycan disruption by the pIV N-terminal domains, production of wildtype pIV, which should similarly disrupt peptidoglycan, only stimulated a ~1.5-fold increase in each of *ycfS* and *bacA* (FDR < 1e-2) relative to the vector control. Thus, increased transcription of these genes caused by the leaky mutant pIV-E292K is likely an adaptation to the osmotic shock imposed by a leaky secretin, not interaction of the pIV N-terminal domain with the peptidoglycan layer.

6.1.3 Rcs response

Stimulation of the Rcs phosphorelay by both wildtype pIV and pIV-E292K is interesting since inactivation of the most highly transcribed regulator of the Rcs regulon, *rcsA* (Figure 25), did not cause viability to decrease more than 2.7-fold relative to the vector control in response to either pIV variant (Figure 31A, Figure 32A). Yet, ectopic overexpression of *rcsA* concomitant with pIV-E292K

production caused a nearly 4-log decrease in viability (Figure 31B). Meanwhile, inactivation of *rcsB*, the transcriptional regulator required by all branches of the Rcs regulon, caused a 6-log decrease in the viability of pIV-E292K producing cells, rescuable by complementation (Figure 31). These observations suggest the RcsA-independent branches of the Rcs regulon are responsible for mitigating stress imposed by leaky secretins.

Increased transcription of osmoprotectant import channels one of three indications cells producing wildtype pIV and the leaky mutant pIV-E292K are experiencing osmotic shock (Figure 29). Increased production of capsular polysaccharides is the second strategy to mitigate osmotic shock. It appears *E. coli* K-12 produced different capsular polysaccharides depending on the severity of osmotic stress experienced (Figure 27, Figure 28). Cells producing wildtype pIV likely experience a mild, or limited, osmotic stress, largely ameliorated through rebalance of key osmolytes (K^+ and proline/glycine-betaine uptake; Figure 29). Additionally, production of wildtype pIV resulted in a nearly 5-fold increase in expression of PGA capsular polysaccharide synthesis and export pathway, *pgaABCD*, transcriptionally regulated by sodium-sensitive NhaR (Figure 28). Rebalancing of cytoplasmic osmolytes may not be sufficient to prevent or mitigate osmotic shock in cells producing pIV-E292K. Furthermore, over-accumulation of K^+ can cause protein denaturation and DNA damage (Booth, 1985), at which point, the import and utilisation of proline or glycine-betaine as osmoprotectants becomes more important as an osmotic regulation strategy. Cells producing pIV-E292K do not express the *pga* operon as highly as those producing wildtype pIV, instead they express the RcsA-

dependent type 4 capsule synthesis and export apparatus encoded by *yjbEFGH* and *wza-wzb-wzc-wcaA-wcaB* (Figure 27). Furthermore, cells producing pIV-E292K deploy a third adaptation to osmotic shock – the production of the osmoprotectant trehalose (Figure 30).

Previous microarray experiments, identical in design, with plasmid-borne pIV being expressed from a *tac* promoter induced by 1 mM IPTG and expressed in similar copy numbers for the same period of time did not reveal differential expression of any of these genes in response to overexpression of *pspF*, Δ*pspA* or pIV production (Bury-Mone *et al.*, 2009, Lloyd *et al.*, 2004, Jovanovic *et al.*, 2006). Therefore, the data shown here for wildtype pIV are a lot more informative due to the use of RNAseq, a technology that is much more precise and covers much broader dynamic range than do microarrays. Furthermore, analysis of stress caused by isogenic leaky-gate secretin in parallel allowed comparison of the stress levels caused by a ground-state closed-gate secretin and a secretin that could serve as a model of open-gate or leaky secretins. The transcriptional and synthetic-lethality analysis performed here clearly indicates two different states of osmoprotection determined by the integrity of the secretin gate.

It is not known whether the PGA capsule directly contributes to survival of wildtype pIV expressing cells. However, lack of synthetic lethality of *rcsA*::km cultures co-expressing wildtype or pIV−E292K (Figure 31A) suggests type 4 capsule production via this pathway is not critical to survival. Alternatively, PGA may compensate for the lack of a type 4 capsule. Inactivation of each capsule

export channel and combination thereof will determine the impact of capsule production on survival during secretin expression.

6.1.4 ROS responses

While the import of osmoprotectants and production of capsule polysaccharides is a clear indication of osmotic stress, osmotic shock is thought to preceed and cause oxidative stress within the cell (Sévin & Sauer, 2014, Smirnova *et al.*, 2000). Synthesis of the osmoprotectant, trehalose, from UDP-glucose and glucose-6-phosphate is mediated by OtsA and OtsB (Giaever *et al.*, 1988), both of which show increased transcription in response to pIV-E292K, but not wildtype pIV (Figure 30). Removal of precursors from core metabolic pathways for the production of capsule polysaccharides or osmolytes would increase turnover of glycolysis and the TCA cycle, both of which result in the production of ROS in aerobic conditions. This appears to hold true, since transcription of *soxS* increases in response to wildtype pIV and pIV-E292K expression, 29- and 18-fold, respectively (Figure 20). Furthermore, inactivation of *soxS*, but curiously not *soxR*, caused the viability of cells expressing pIV-E292K to drop by 4 orders of magnitude. Cultures producing wildtype pIV were not affected and complementation with pCA24N-*soxS* (Table 2) partially restored viability pIV-E292K expressing cultures (Figure 31, Figure 32).

SoxS is able to activate transcription of *marRAB* and inhibit expression of *rob* (Martin & Rosner, 2002). Transcription of each operon was similar in response to either wildtype pIV or pIV-E292K (Figure 20, Figure 21). Many effectors of ROS removal within the SoxS/MarA/Rob regulon show increased transcription

in response to pIV-E292K compared to wildtype pIV, including *poxB*, *ydbK* and *sodA*.

Catalase II (an enzyme that removes hydrogen peroxide), encoded by the *katE* gene, is expressed when Fe^{2+} is limiting (Hoerter *et al.*, 2005), during stationary phase and hyperosmotic stress (Weber *et al.*, 2006). It shows a differential increase in transcription in response to pIV-E292K, but not wildtype pIV. Curiously, while insertional inactivation of *katE* does not cause any growth defects, ectopic expression of KatE in a *ΔkatE* host strain (Table 2) concurrent with pIV−E292K production causes a 10^4-fold decrease in viability (Figure 31B). It is possible that the peroxide substrate of KatE is a stress signal required for activating a transcriptional regulator. In that case, the observed effect of *katE* overexpression would remove peroxide without removing the superoxide (if superoxide dismutase is the bottleneck in stress mitigation), rendering the ROS stress untreated and resulting in observed reduced viability.

In summary, secretin stress affects multiple systems, from the cell surface, to core metabolism. Balancing cytoplasmic osmolytes and production of a PGA capsule largely mitigates the cellular response to the secretin stress imposed by wildtype pIV. However, the increased stress of a disregulated or permissive secretin gate, as in the pIV-E292K mutant, requires a more global response – massive production of type 4 capsule polysaccharides and production of the trehalose osmoprotectants draw sugar precursors from central metabolism, resuting in increased turnover of TCA and glycolysis. Consequently, ROS generation increases which results in activation of the superoxide stress

response. Detailed analyses of synthetic lethalities caused by inactivation of capsule export channels and key effectors of oxidative stress are required to determine their role in the adaptation to secretin stress.

6.1.5 Implications

Findings from this thesis that identified, in addition to the Psp response, three major stress responses induced by and required for survival of secretin stress, provided a lead for analyses of stress response requirements for surviving expression of the T3SS secretin, InvG, in *E. coli* (Khanum, 2015). InvG was found to be naturally leaky to maltopentaose and vancomycin, thereby qualifies as a leaky secretin. Expressed from the same vector used for pIV expression in this work, concomitant with expression of the cognate pilotin, InvH, resulted in targeting of InvG to both OM and IM (Khanum, 2015). Synthetic lethality assays demonstated toxicity of the InvG secretin in cells lacking PspF, SoxS (plating efficiency was decreased by three orders of magnitude relative to cells containing the vector) and CpxR (ten-fold), while the effect of RcsA and B mutations was milder, reducing plating efficiency by 50% or less (Khanum, 2015). The synthetic lethality caused by the production of InvG show, under the same experimental conditions, three of four regulons required for survival of the pIV-E292K leaky mutant are also required for survival of the leaky T3SS secretin, InvG. It would be interesting to examine whether stress caused by the production of the whole T3SS in enteric bacteria requires these same responses. It has been reported that T3SS results in extended generation time in *S. typhimurium* and that it is lost in a large sub-population within a bacterial culture during the time-couse of growth under laboratory conditions. At the

same time, CpxR has been reported to be a positive regulator of T3SS based on culture assays, which average findings across the whole cell population (Sturm *et al.*, 2011). It would be interesting to resolve, therefore, whether the actual effect of CpxR may be elimination of T3SS-expressing cells rather than by blocking the T3SS expression in the whole cell population (Liu *et al.*, 2012, Carlsson *et al.*, 2007).

Consistent with a lack of mucoidy, *rcs* is not required by cells for survival of InvG secretin stress. Furthermore, InvG is also less severly leaky than pIV-E292K, suggesting differences in stress caused by different secretins. This warrants further investigation into the correlation between phenotypes caused by a particular secretin and the corresponding stress responses.

6.2 Conclusion

Previous to this work, it was thought that the response to secretin stress was confined to the Psp regulon. The data presented here expands the secretin stress response to include elements of the Rcs, Cpx and Sox regulons. The extent to which these regulons are involved is dependent on the integrity of the secretin gate. Furthermore, secretin expression induces capsular polysaccharide production and import of osmolytes. The type of capsule produced and osmolytes imported is dependent on the integrity of the secretin gate.

The data presented here lends further evidence to support the hypothesis that oxidative stress is a consequence of osmotic shock. This is likely caused by the change in metabolic flux to support the import and production of osmoprotectants and capsule polysaccharides using precursors derived from core metabolism.

The infection process of pathogens that utilise the crucial virulence factors secreted through the T3SS. The expression of pili and host-cell attachment factors have been reported to be regulated by the CpxR regulon. In at least *Yersinia enterocilitica*, the activation of the RcsB regulon has also been suggested to indirectly activate the expression of the Ysa-Ysp T3SS, while the CpxR response inhibits their expression (Flores-Kim & Darwin, 2012a). Here we have shown that expression of a severely leaky pIV mutant (pIV-E292K) is able to induce the Cpx and Rcs responses and that those responses are required for survival of bacteria expressing the secretin.

Observations of the stress responses required for survival of cells expressing pIV-E292K provided leads for investigation of responses to a naturally leaky secretin, InvG of the *S. enterica* T3SS (Khanum, 2015). This analysis showed that, apart from Rcs, *E. coli* uses similar stress response pathways to survive the stress caused by the production of secretins from two very distantly related secretion systems.

We have identified several response regulators that are required for survival of cells expressing permissive or leaky secretins, one of which is linked to the redox state of the cell. Furthermore, regulatory networks governing the response to secretin expression have been identified in much greater resolution than before. The findings will have deep implications for the investigation of host-cell interactions, secretin stress and virulence factors. Further epistatic analysis of individual effectors in each regulon has the potential to reveal novel drug targets.

6.3 Future Directions

The work presented here, while it represents a starting point and a hypotheses generator in understanding the transcriptional response of *E. coli* to the production of large OM channels, requires deeper analysis to better understand the regulation of responses, changes in metabolic flux and the physical adaptations required to survive secretin stress.

Firstly, complete epistatic synthetic lethality experiments must be performed to ascertain the importance of the CpxR, SoxS and RcsB/RcsAB regulons. Moreover, such experiments will determine which of the genes within those regulons are critical to the survival of secretin producing cells. Given the lack of synthetic lethality caused by the loss of *rcsA*, it would be interesting to investigate what effect losing the specific Wza and Yjb channels would have to survival. This would be expanded to include the PGA capsule synthesis pathway and osmolyte import channels. Furthermore, it has been suggested that PspG is the crucial Psp effector required for survival of secretin stress (Jovanovic *et al.*, 2006). Complementation of a Δ*pspBC* strain by Psp effectors (PspA, PspD, PspE, PspG) would determine which of the Psp effectors is required for survival of secretin stress.

Osmoprotectant synthesis and import pathways were also among genes differentially expressed in response to either wildtype pIV, or corresponding leaky mutant, pIV-E292K. Growth rate analysis of cells producing wildtype pIV or a leaky mutant in media supplemented by osmolytes (Na^+, K^+) and osmoprotectants (Sucrose, glycine-betaine, proline, trehalose, PEG) will

determine the importance of these products in the survival of secretin stress. These experiments should be expanded to analyse growth rate in strains lacking the ability to synthesise or import osmoprotectants and osmolytes.

To extend this work further, sequencing the small regulatory RNAs transcribed in response to secretin stress would determine the critical small regulatory RNAs in regulating the secretin stress response.

The toxic effect of pIV-E292K mutant expressed at low level (in the absence of IPTG) results in less prominent synthetic lethality for all responses apart from PspF, which is still required for survival, even at very low expression of a seretin. A more detailed analysis of the “natural” level of secretin expression is required in order to compare the exact protein copy number in plasmid-expressed vs. phage-expressed pIV. Furthermore, analyses in the context of filamentous phage genome would indicate the role in survival of cells expressing pIV-E292K under its own promoter and in the context of active secretion of its substrate, filamentous phage.

To orthogonally validate the RNA-seq transcriptome, transcription of a set of genes central to IM and OM stress, biogenesis and repair will be determined using nanoCounter technology (Kulkarni, 2011). Briefly, transcripts for a set of target genes in rRNA-depleted total RNA samples (prepared as described in section 2.5) are directly counted by digital imaging. Similar analysis of pIV-$PulD_{PBS}$ fusion proteins in the presence and absence of the pilotin, PulS, will

separate the IM from OM stress responses caused by secretins inserting into the IM (or OM).

7 References

Adams, H., W. Teertstra, J. Demmers, R. Boesten & J. Tommassen, (2003) Interactions between phage-shock proteins in Escherichia coli. *J. Bacteriol.* **185**: 1174-1180.

Adams, H., W. Teertstra, M. Koster & J. Tommassen, (2002) PspE (phage-shock protein E) of *Escherichia coli* is a rhodanese. *FEBS Lett.* **518**: 173-176.

Al Mamun, A.A., M.J. Lombardo, C. Shee, A.M. Lisewski, C. Gonzalez, D. Lin, R.B. Nehring, C. Saint-Ruf, J.L. Gibson, R.L. Frisch, O. Lichtarge, P.J. Hastings & S.M. Rosenberg, (2012) Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* **338**: 1344-1348.

Alekshun, M.N. & S.B. Levy, (1999) The mar regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol.* **7**: 410-413.

Ali Azam, T., A. Iwata, A. Nishimura, S. Ueda & A. Ishihama, (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J. Bacteriol.* **181**: 6361-6370.

Altuvia, S., D. Weinstein-Fischer, A. Zhang, L. Postow & G. Storz, (1997) A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* **90**: 43-53.

Anders, S. & W. Huber, (2010) Differential expression analysis for sequence count data. *Genome Biology* **11**.

Andrews, S., (2010) FastQC: a quality control tool for high throughput sequence data. In., pp.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner & H. Mori, (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**.

Babinski, K.J., A.A. Ribeiro & C.R. Raetz, (2002) The Escherichia coli gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *J. Biol. Chem.* **277**: 25937-25946.

Baneyx, F. & G. Georgiou, (1991) Construction and characterisation of *Escherichia coli* strains deficient in multiple secreted proteases - Protease III degrades high-molecular-weight substrates in vivo. *J. Bacteriol.* **173**: 2696-2703.

Barnhart, M.M. & M.R. Chapman, (2006) Curli Biogenesis and Function. *Annual review of microbiology* **60**: 131-147.

Baslé, A., G. Rummel, P. Storici, J.P. Rosenbusch & T. Schirmer, (2006) Crystal Structure of Osmoporin OmpC from E. coli at 2.0 Å. *J. Mol. Biol.* **362**: 933-942.

Bayan, N., I. Guilvout & A.P. Pugsley, (2006) Secretins take shape. *Mol. Microbiol.* **60**: 1-4.

Beck, B.J., L.E. Connolly, A. De Las Penas & D.M. Downs, (1997) Evidence that rseC, a gene in the rpoE cluster, has a role in thiamine synthesis in Salmonella typhimurium. *J. Bacteriol.* **179**: 6504-6508.

Becker, L.A., I.S. Bang, M.L. Crouch & F.C. Fang, (2005) Compensatory role of PspA, a member of the phage shock protein operon, in rpoE mutant Salmonella enterica serovar Typhimurium. *Mol. Microbiol.* **56**: 1004-1016.

Beis, K., R.F. Collins, R.C. Ford, A.B. Kamis, C. Whitfield & J.H. Naismith, (2004) Three-dimensional structure of Wza, the protein required for translocation of group 1 capsular polysaccharide across the outer membrane of Escherichia coli. *J. Biol. Chem.* **279**: 28227-28232.

Belunis, C.J. & C.R. Raetz, (1992) Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from Escherichia coli. *J. Biol. Chem.* **267**: 9988-9997.

Benjamini, Y. & Y. Hochberg, (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289-300.

Bennion, D., E.S. Charlson, E. Coon & R. Misra, (2010) Dissection of beta-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of Escherichia coli. *Mol. Microbiol.* **77**: 1153-1171.

Benson, S.A. & A. Decloux, (1985) Isolation and characterisation of outer-membrane permeability mutants in *Escherichia coli* K-12. *J. Bacteriol.* **161**: 361-367.

Benson, S.A., J.L.L. Occi & B.A. Sampson, (1988) Mutations that alter the pore function of the OmpF porin of *Escherichia coli* K-12. *J. Mol. Biol.* **203**: 961-970.

Berg, H.C., (2003) The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* **72**: 19-54.

Bergler, H., D. Abraham, H. Aschauer & F. Turnowsky, (1994) Inhibition of lipid biosynthesis induces the expression of the *pspA* gene. *Microbiology-UK* **140**: 1937-1944.

Bertram, R. & C.F. Schuster, (2014) Post-transcriptional regulation of gene expression in bacterial pathogens by toxin-antitoxin systems. *Frontiers in Cellular and Infection Microbiology* **4**: 6.

Blanchard, J.L., W.Y. Wholey, E.M. Conlon & P.J. Pomposiello, (2007) Rapid changes in gene expression dynamics in response to superoxide reveal SoxRS-dependent and independent transcriptional networks. *PLoS ONE* **2**: e1186.

Bolger, A.M., M. Lohse & B. Usadel, (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker, H.W. Boyer, J.H. Crosa & S. Falkow, (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**: 95-113.

Booth, I.R., (1985) Regulation of cytoplasmic pH in bacteria. *MICROBIOL. REV.* **49**: 359-378.

Boysen, A., J. Moller-Jensen, B. Kallipolitis, P. Valentin-Hansen & M. Overgaard, (2010) Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J. Biol. Chem.* **285**: 10690-10702.

Braun, V. & H. Wolff, (1970) The murein-lipoprotein linkage in the cell wall of Escherichia coli. *European Journal of Biochemistry* **14**: 387-391.

Brill, J.A., C. Quinlanwalshe & S. Gottesman, (1988) Fine-structure mapping and identification of 2 regulators of capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **170**: 2599-2611.

Brissette, J.L., M. Russel, L. Weiner & P. Model, (1990) Phage shock protein, a stress protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **87**: 862-866.

Brissette, J.L., L. Weiner, T.L. Ripmaster & P. Model, (1991) Characterization and sequence of the *Escherichia coli* stress-induced *psp* operon. *J. Mol. Biol.* **220**: 35-48.

Burghout, P., R. van Boxtel, P. Van Gelder, P. Ringler, S.A. Muller, J. Tommassen & M. Koster, (2004) Structure and electrophysiological properties of the YscC secretin from the type III secretion system of *Yersinia enterocolitica*. *J. Bacteriol.* **186**: 4645-4654.

Bury-Mone, S., Y. Nomane, N. Reymond, R. Barbet, E. Jacquet, S. Imbeaud, A. Jacq & P. Bouloc, (2009) Global Analysis of Extracytoplasmic Stress Signaling in *Escherichia coli*. *PLoS Genet.* **5**: 17.

Button, J.E., T.J. Silhavy & N. Ruiz, (2007) A suppressor of cell death caused by the loss of sigma(E) downregulates extracytoplasmic stress responses and outer membrane vesicle production in *Escherichia coli*. *J. Bacteriol.* **189**: 1523-1530.

Carbonnelle, E., S. Helaine, L. Prouvensier, X. Nassif & V. Pelicic, (2005) Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol. Microbiol.* **55**: 54-64.

Carlsson, K.E., J. Liu, P.J. Edqvist & M.S. Francis, (2007) Extracytoplasmic-stress-responsive pathways modulate type III secretion in Yersinia pseudotuberculosis. *Infection & Immunity* **75**: 3913-3924.

Castanie-Cornet, M.P., K. Cam, B. Bastiat, A. Cros, P. Bordes & C. Gutierrez, (2010) Acid stress response in Escherichia coli: mechanism of regulation of gadA transcription by RcsB and GadE. *Nucleic Acids Res.* **38**: 3546-3554.

Castanie-Cornet, M.P., K. Cam & A. Jacq, (2006) RcsF is an outer membrane lipoprotein involved in the RcsCDB phosphorelay signaling pathway in Escherichia coli. *J. Bacteriol.* **188**: 4264-4270.

Chami, M., I. Guilvout, M. Gregorini, H.W. Remigy, S.A. Muller, M. Valerio, A. Engel, A.P. Pugsley & N. Bayan, (2005) Structural insights into the secretin PulD and its trypsin-resistant core. *J. Biol. Chem.* **280**: 37732-37741.

Chandran, V., R. Fronzes, S. Duquerroy, N. Cronin, J. Navaza & G. Waksman, (2009) Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**: 1011-U1066.

Christman, M.F., R.W. Morgan, F.S. Jacobson & B.N. Ames, (1985) Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in Salmonella typhimurium. *Cell* **41**: 753-762.

Christman, M.F., G. Storz & B.N. Ames, (1989) OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 3484-3488.

Chubiz, L.M., G.D. Glekas & C.V. Rao, (2012) Transcriptional cross talk within the mar-sox-rob regulon in Escherichia coli is limited to the rob and marRAB operons. *J. Bacteriol.* **194**: 4867-4875.

Clementz, T., Z. Zhou & C.R. Raetz, (1997) Function of the Escherichia coli msbB gene, a multicopy suppressor of htrB knockouts, in the acylation of

lipid A. Acylation by MsbB follows laurate incorporation by HtrB. *J. Biol. Chem.* **272**: 10353-10360.
Cohen, S.P., H. Hachler & S.B. Levy, (1993) Genetic and functional analysis of the multiple antibiotic resistance (mar) locus in Escherichia coli. *J. Bacteriol.* **175**: 1484-1492.
Collin, S., I. Guilvout, M. Chami & A.P. Pugsley, (2007) YaeT-independent multimerization and outer membrane association of secretin PulD. *Mol. Microbiol.* **64**: 1350-1357.
Collin, S., I. Guilvout, N.N. Nickerson & A.P. Pugsley, (2011) Sorting of an integral outer membrane protein via the lipoprotein-specific Lol pathway and a dedicated lipoprotein pilotin. *Mol. Microbiol.* **80**: 655-665.
Collinet, B., H. Yuzawa, T. Chen, C. Herrera & D. Missiakas, (2000) RseB binding to the periplasmic domain of RseA modulates the RseA:sigmaE interaction in the cytoplasm and the availability of sigmaE.RNA polymerase. *J. Biol. Chem.* **275**: 33898-33904.
Cowan, S.W., T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R.A. Pauptit, J.N. Jansonius & J.P. Rosenbusch, (1992) Crystal structures explain functional properties of two E. coli porins. *Nature* **358**: 727-733.
Cox, M., (2013) Map_count.pl. In., pp.
Crago, A.M. & V. Koronakis, (1998) Salmonella InvG forms a ring-like multimer that requires the InvH lipoprotein for outer membrane localization. *Mol. Microbiol.* **30**: 47-56.
Craig, L. & J. Li, (2008) Type IV pilli: paradoxes in form and function. *Curr. Opin. Struct. Biol.* **18**: 267-277.
Craig, L., R.K. Taylor, M.E. Pique, B.D. Adair, A.S. Arvai, M. Singh, S.J. Lloyd, D.S. Shin, E.D. Getzoff, M. Yeager, K.T. Forest & J.A. Tainer, (2003) Type IV pilin structure and assembly: X-ray and EM analyses of Vibrio cholerae toxin-coregulated pilus and Pseudomonas aeruginosa PAK pilin. *Mol. Cell* **11**: 1139-1150.
Craig, L., N. Volkmann, A.S. Arvai, M.E. Pique, M. Yeager, E.H. Egelman & J.A. Tainer, (2006) Type IV pilus structure by cryo-electron microscopy and crystallography: Implications for pilus assembly and functions. *Mol. Cell* **23**: 651-662.
Cronan, J.E., Jr. & R.M. Bell, (1974) Mutants of Escherichia coli defective in membrane phospholipid synthesis: mapping of the structural gene for L-glycerol 3-phosphate dehydrogenase. *J. Bacteriol.* **118**: 598-605.
Crowell, D.N., M.S. Anderson & C.R. Raetz, (1986) Molecular cloning of the genes for lipid A disaccharide synthase and UDP-N-acetylglucosamine acyltransferase in Escherichia coli. *J. Bacteriol.* **168**: 152-159.
Daefler, S., I. Guilvout, K.R. Hardie, A.P. Pugsley & M. Russel, (1997a) The C-terminal domain of the secretin PulD contains the binding site for its cognate chaperone, PulS, and confers PulS dependence on pIV(f1) function. *Mol. Microbiol.* **24**: 465-475.
Daefler, S. & M. Russel, (1998) The Salmonella typhimurium InvH protein is an outer membrane lipoprotein required for the proper localization of InvG. *Mol. Microbiol.* **28**: 1367-1380.
Daefler, S., M. Russel & P. Model, (1997b) Module swaps between related translocator proteins pIV(f1), pIV(IKe) and PulD: Identification of a specificity domain. *J. Mol. Biol.* **266**: 978-992.

Daley, D.O., M. Rapp, E. Granseth, K. Melen, D. Drew & G. von Heijne, (2005) Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308**: 1321-1323.

Danese, P.N. & T.J. Silhavy, (1998) CpxP, a stress-combative member of the Cpx regulon. *J. Bacteriol.* **180**: 831-839.

Dartigalongue, C., D. Missiakas & S. Raina, (2001) Characterization of the Escherichia coli sigma E regulon. *J. Biol. Chem.* **276**: 20866-20875.

Darwin, A.J. & V.L. Miller, (2001) The psp locus of Yersinia enterocolitica is required for virulence and for growth in vitro when the Ysc type III secretion system is produced. *Mol. Microbiol.* **39**: 429-444.

Davis, B.M., E.H. Lawson, M. Sandkvist, A. Ali, S. Sozhamannan & M.K. Waldor, (2000) Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. *Science* **288**: 333-335.

De Las Penas, A., L. Connolly & C.A. Gross, (1997) The sigma(E)-mediated response to extracytoplasmic stress in Escherichia coli is transduced by RseA and RseB, two negative regulators of sigma(E). *Mol. Microbiol.* **24**: 373-385.

De Wulf, P., A.M. McGuire, X. Liu & E.C. Lin, (2002) Genome-wide profiling of promoter recognition by the two-component response regulator CpxR-P in Escherichia coli. *J. Biol. Chem.* **277**: 26652-26661.

Diepold, A., M. Amstutz, S. Abel, I. Sorg, U. Jenal & G.R. Cornelis, (2010) Deciphering the assembly of the Yersinia type III secretion injectisome. *Embo J.* **29**: 1928-1940.

DiGiuseppe, P.A. & T.J. Silhavy, (2003) Signal detection and target gene induction by the CpxRA two-component system. *J. Bacteriol.* **185**: 2432-2440.

Doerrler, W.T., H.S. Gibbons & C.R. Raetz, (2004) MsbA-dependent translocation of lipids across the inner membrane of Escherichia coli. *J. Biol. Chem.* **279**: 45102-45109.

Doerrler, W.T. & C.R. Raetz, (2005) Loss of outer membrane proteins without inhibition of lipid export in an Escherichia coli YaeT mutant. *J. Biol. Chem.* **280**: 27679-27687.

Dong, C.J., K. Beis, J. Nesper, A.L. Brunkan-LaMontagne, B.R. Clarke, C. Whitfield & J.H. Naismith, (2006) Wza the translocon for E-coli capsular polysaccharides defines a new class of membrane protein. *Nature* **444**: 226-229.

Dorel, C., P. Lejeune & A. Rodrigue, (2006) The Cpx system of Escherichia coli, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? *Res. Microbiol.* **157**: 306-314.

Duguay, A.R. & T.J. Silhavy, (2004) Quality control in the bacterial periplasm. *Biochimica Et Biophysica Acta - Molecular Cell Research* **1694**: 121-134.

Dunstan, R.A., E. Heinz, L.C. Wijeyewickrema, R.N. Pike, A.W. Purcell, T.J. Evans, J. Praszkier, R.M. Robins-Browne, R.A. Strugnell, K.V. Korotkov & T. Lithgow, (2013) Assembly of the Type II Secretion System such as Found in <italic>Vibrio cholerae</italic> Depends on the Novel Pilotin AspS. *PLoS Pathogens* **9**: e1003117.

Dworkin, J., G. Jovanovic & P. Model, (1997) Role of upstream activation sequences and integration host factor in transcriptional activation by the constitutively active prokaryotic enhancer-binding protein PspF. *J. Mol. Biol.* **273**: 377-388.

Ebel, W. & J.E. Trempy, (1999) Escherichia coli RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions To activate its own expression. *J. Bacteriol.* **181**: 577-584.

Economou, A., (2002) Bacterial secretome: the assembly manual and operating instructions (Review). *Mol Membr Biol* **19**: 159-169.

Edgar, J.R. & R.M. Bell, (1979) Biosynthesis in Escherichia coli of sn-glycerol 3-phosphate, a precursor of phospholipid. Palmitoyl-CoA inhibition of the biosynthetic sn-glycerol-3-phosphate dehydrogenase. *J. Biol. Chem.* **254**: 1016-1021.

Engl, C., G. Jovanovic, L.J. Lloyd, H. Murray, M. Spitaler, L.M. Ying, J. Errington & M. Buck, (2009) In vivo localizations of membrane stress controllers PspA and PspG in Escherichia coli. *Mol. Microbiol.* **73**: 382-396.

Engl, C., A. Ter Beek, M. Bekker, J.T. de Mattos, G. Jovanovic & M. Buck, (2011) Dissipation of Proton Motive Force is not Sufficient to Induce the Phage Shock Protein Response in Escherichia coli. *Current Microbiology* **62**: 1374-1385.

Felder, C.B., R.C. Graul, A.Y. Lee, H.P. Merkle & W. Sadee, (1999) The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAPS PharmSci* **1**.

Ferenci, T., Z.M. Zhou, T. Betteridge, Y. Ren, Y. Liu, L. Feng, P.R. Reeves & L. Wang, (2009) Genomic Sequencing Reveals Regulatory Mutations and Recombinational Events in the Widely Used MC4100 Lineage of Escherichia coli K-12. *J. Bacteriol.* **191**: 4025-4029.

Ferrieres, L., S.N. Aslam, R.M. Cooper & D.J. Clarke, (2007) The yjbEFGH locus in Escherichia coli K-12 is an operon encoding proteins involved in exopolysaccharide production. *Microbiology* **153**: 1070-1080.

Ferrieres, L. & D.J. Clarke, (2003) The RcsC sensor kinase is required for normal biofilm formation in Escherichia coli K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol. Microbiol.* **50**: 1665-1682.

Finn, R.D., A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L.L. Sonnhammer, J. Tate & M. Punta, (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**: D222-D230.

Fleischer, R., R. Heermann, K. Jung & S. Hunke, (2007) Purification, reconstitution, and characterization of the CpxRAP envelope stress system of Escherichia coli. *J. Biol. Chem.* **282**: 8583-8593.

Flores-Kim, J. & A.J. Darwin, (2012a) Links between type III secretion and extracytoplasmic stress responses in Yersinia. *Frontiers in Cellular and Infection Microbiology* **2**.

Flores-Kim, J. & A.J. Darwin, (2012b) Phage Shock Protein C (PspC) of Yersinia enterocolitica Is a Polytopic Membrane Protein with Implications for Regulation of the Psp Stress Response. *J. Bacteriol.* **194**: 6548-6559.

Flores-Kim, J. & A.J. Darwin, (2015) Activity of a bacterial cell envelope stress response is controlled by the interaction of a protein binding domain with different partners. *J Biol Chem* **290**: 11417-11430.

Forst, S., J. Delgado & M. Inouye, (1989) Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 6052-6056.

Francez-Charlot, A., B. Laugel, A. Van Gemert, N. Dubarry, F. Wiorowski, M.P. Castanie-Cornet, C. Gutierrez & K. Cam, (2003) RcsCDB His-Asp phosphorelay system negatively regulates the flhDC operon in Escherichia coli. *Mol. Microbiol.* **49**: 823-832.

Freinkman, E., S. Okuda, N. Ruiz & D. Kahne, (2012) Regulated assembly of the transenvelope protein complex required for lipopolysaccharide export. *Biochemistry* **51**: 4800-4806.

Freundlieb, S., U. Ehmann & W. Boos, (1988) Facillitated diffusion of para-nitrophenyl-alpha-D-maltohexaoside through the outer-membrane of *Escherichia coli* - Characterization of LamB as as specific and saturable channel for maltooligosaccharides. *J. Biol. Chem.* **263**: 314-320.

Frye, S.A., R. Assalkhou, R.F. Collins, R.C. Ford, C. Petersson, J.P. Derrick & T. Tonjum, (2006) Topology of the outer-membrane secretin PilQ from Neisseria meningitidis. *Microbiology-(UK)* **152**: 3751-3764.

Fuentes, A.M., J.J. Diaz-Mejia, R. Maldonado-Rodriguez & C.F. Amabile-Cuevas, (2001) Differential activities of the SoxR protein of Escherichia coli: SoxS is not required for gene activation under iron deprivation. *FEMS Microbiology Letters* **201**: 271-275.

Fujikawa, M., K. Kobayashi & T. Kozawa, (2012) Direct oxidation of the [2Fe-2S] cluster in SoxR protein by superoxide: distinct differential sensitivity to superoxide-mediated signal transduction. *J. Biol. Chem.* **287**: 35702-35708.

Galan, J.E. & A. Collmer, (1999) Type III secretion machines: Bacterial devices for protein delivery into host cells. *Science* **284**: 1322-1328.

Galloway, S.M. & C.R. Raetz, (1990) A mutant of Escherichia coli defective in the first step of endotoxin biosynthesis. *J. Biol. Chem.* **265**: 6394-6402.

Garrett, T.A., J.L. Kadrmas & C.R. Raetz, (1997) Identification of the gene encoding the Escherichia coli lipid A 4'-kinase. Facile phosphorylation of endotoxin analogs with recombinant LpxK. *J. Biol. Chem.* **272**: 21855-21864.

Gaudu, P., N. Moon & B. Weiss, (1997) Regulation of the soxRS Oxidative Stress Regulon: Reversible oxidation of the Fe-S centers of SoxR in vivo. *J. Biol. Chem.* **272**: 5082-5086.

Geissmann, T.A. & D. Touati, (2004) Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *Embo J.* **23**: 396-405.

Gennity, J.M. & M. Inouye, (1991) The protein sequence responsible for lipoprotein membrane localization in Escherichia coli exhibits remarkable specificity. *J. Biol. Chem.* **266**: 16458-16464.

Gerlach, R.G. & M. Hensel, (2007) Protein secretion systems and adhesins: The molecular armory of Gram-negative pathogens. *Int. J. Med. Microbiol.* **297**: 401-415.

Gervais, F.G. & G.R. Drapeau, (1992) Identification, cloning, and characterization of rcsF, a new regulator gene for exopolysaccharide synthesis that suppresses the division mutation ftsZ84 in Escherichia coli K-12. *J. Bacteriol.* **174**: 8016-8022.

Giaever, H.M., O.B. Styrvold, I. Kaasen & A.R. Strom, (1988) Biochemical and genetic characterization of osmoregulatory trehalose synthesis in Escherichia coli. *J. Bacteriol.* **170**: 2841-2849.

Gonzalez-Flecha, B. & B. Demple, (1997) Transcriptional regulation of the Escherichia coli oxyR gene as a function of cell growth. *J. Bacteriol.* **179**: 6181-6186.

Gottesman, S., P. Trisler & A. Torrescabassa, (1985) Regulation of capsular polysaccharide synthesis in *Escherichia coli* K-12: Characterization of 3 regulatory genes. *J. Bacteriol.* **162**: 1111-1119.

Green, R.C. & A.J. Darwin, (2004) PspG, a new member of the Yersinia enterocolitica phage shock protein regulon. *J. Bacteriol.* **186**: 4910-4920.

Griffith, K.L., M.M. Fitzpatrick, E.F. Keen, 3rd & R.E. Wolf, Jr., (2009) Two functions of the C-terminal domain of Escherichia coli Rob: mediating "sequestration-dispersal" as a novel off-on switch for regulating Rob's activity as a transcription activator and preventing degradation of Rob by Lon protease. *J. Mol. Biol.* **388**: 415-430.

Griffith, K.L., I.M. Shah & R.E. Wolf, Jr., (2004) Proteolytic degradation of Escherichia coli transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol. Microbiol.* **51**: 1801-1816.

Gu, M. & J.A. Imlay, (2011) The SoxRS response of Escherichia coli is directly activated by redox-cycling drugs rather than by superoxide. *Mol. Microbiol.* **79**: 1136-1150.

Guilvout, I., M. Chami, E. Disconzi, N. Bayan, A.P. Pugsley & G.H.M. Huysmans, (2014) Independent Domain Assembly in a Trapped Folding Intermediate of Multimeric Outer Membrane Secretins. *Structure* **22**: 582-589.

Guilvout, I., M. Chami, A. Engel, A.P. Pugsley & N. Bayan, (2006) Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. *Embo J.* **25**: 5241-5249.

Guilvout, I., N.N. Nickerson, M. Chami & A.P. Pugsley, (2011) Multimerization-defective variants of dodecameric secretin PulD. *Res. Microbiol.* **162**: 180-190.

Gupta, S.D. & H.C. Wu, (1991) Identification and subcellular localization of apolipoprotein N-acyltransferase in Escherichia coli. *FEMS Microbiology Letters* **62**: 37-41.

Hara, T., S. Matsuyama & H. Tokuda, (2003) Mechanism underlying the inner membrane retention of Escherichia coli lipoproteins caused by Lol avoidance signals. *J. Biol. Chem.* **278**: 40408-40414.

Hayashi, S. & H.C. Wu, (1990) Lipoproteins in bacteria. *Journal of bioenergetics and biomembranes* **22**: 451-471.

Hidalgo, E., V. Leautaud & B. Demple, (1998) The redox-regulated SoxR protein acts from a single DNA site as a repressor and an allosteric activator. *Embo J.* **17**: 2629-2636.

Hoerter, J.D., A.A. Arnold, C.S. Ward, M. Sauer, S. Johnson, T. Fleming & A. Eisenstark, (2005) Reduced hydroperoxidase (HPI and HPII) activity in the Deltafur mutant contributes to increased sensitivity to UVA radiation in Escherichia coli. *Journal of Photochemistry and Photobiology. Biology* **79**: 151-157.

Horstman, N.K. & A.J. Darwin, (2012) Phage shock proteins B and C prevent lethal cytoplasmic membrane permeability in Yersinia enterocolitica. *Mol. Microbiol.* **85**: 445-460.

Huang da, W., B.T. Sherman & R.A. Lempicki, (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**: 1-13.

Huang da, W., B.T. Sherman & R.A. Lempicki, (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**: 44-57.

Hueck, C.J., (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**: 379-+.

Huysmans, G.H.M., I. Guilvout & A.P. Pugsley, (2013) Sequential Steps in the Assembly of the Multimeric Outer Membrane Secretin PulD. *J. Biol. Chem.* **288**: 30700-30707.

Isaac, D.D., J.S. Pinkner, S.J. Hultgren & T.J. Silhavy, (2005) The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 17775-17779.

Jain, S. & M.B. Goldberg, (2007) Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J. Bacteriol.* **189**: 5393-5398.

Jair, K.W., X. Yu, K. Skarstad, B. Thony, N. Fujita, A. Ishihama & R.E. Wolf, Jr., (1996) Transcriptional activation of promoters of the superoxide and multiple antibiotic resistance regulons by Rob, a binding protein of the Escherichia coli origin of chromosomal replication. *J. Bacteriol.* **178**: 2507-2513.

Johansen, J., A.A. Rasmussen, M. Overgaard & P. Valentin-Hansen, (2006) Conserved small non-coding RNAs that belong to the sigma(E) regulon: Role in down-regulation of outer membrane proteins. *J. Mol. Biol.* **364**: 1-8.

Jorgensen, M.G., D.P. Pandey, M. Jaskolska & K. Gerdes, (2009) HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J. Bacteriol.* **191**: 1191-1199.

Jovanovic, G., J. Dworkin & P. Model, (1997) Autogenous control of PspF, a constitutively active enhancer-binding protein of Escherichia coli. *J. Bacteriol.* **179**: 5232-5237.

Jovanovic, G., C. Engl & M. Buck, (2009) Physical, functional and conditional interactions between ArcAB and phage shock proteins upon secretin-induced stress in Escherichia coli. *Mol. Microbiol.* **74**: 16-28.

Jovanovic, G., L.J. Lloyd, M.P.H. Stumpf, A.J. Mayhew & M. Buck, (2006) Induction and function of the phage shock protein extracytoplasmic stress response in Escherichia coli. *J. Biol. Chem.* **281**: 21147-21161.

Jovanovic, G. & P. Model, (1997) The RIB element in the goaG-pspF intergenic region of Escherichia coli. *J. Bacteriol.* **179**: 3095-3102.

Jovanovic, G., L. Weiner & P. Model, (1996) Identification, nucleotide sequence, and characterization of PspF, the transcriptional activator of the Escherichia coli stress-induced psp operon. *J. Bacteriol.* **178**: 1936-1945.

Jubelin, G., A. Vianney, C. Beloin, J.M. Ghigo, J.C. Lazzaroni, P. Lejeune & C. Dorel, (2005) CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in Escherichia coli. *J. Bacteriol.* **187**: 2038-2049.

Kakeda, M., C. Ueguchi, H. Yamada & T. Mizuno, (1995) An Escherichia coli curved DNA-binding protein whose expression is affected by the stationary phase-specific sigma factor sigma S. *Molecular & General Genetics* **248**: 629-634.

Kanehara, K., K. Ito & Y. Akiyama, (2002) YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev.* **16**: 2147-2155.

Kazmierczak, B.I., D.L. Mielke, M. Russel & P. Model, (1994) pIV, a Filamentous Phage Protein That Mediates Phage Export across the Bacterial-Cell Envelope, Forms a Multimer. *J. Mol. Biol.* **238**: 187-198.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes & A. Drummond, (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647-1649.

Kelly, T.M., S.A. Stachula, C.R. Raetz & M.S. Anderson, (1993) The firA gene of Escherichia coli encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase. The third step of endotoxin biosynthesis. *J. Biol. Chem.* **268**: 19866-19874.

Kersey, P.J., J.E. Allen, M. Christensen, P. Davis, L.J. Falin, C. Grabmueller, D.S.T. Hughes, J. Humphrey, A. Kerhornou, J. Khobova, N. Langridge, M.D. McDowall, U. Maheswari, G. Maslen, M. Nuhn, C.K. Ong, M. Paulini, H. Pedro, I. Toneva, M.A. Tuli, B. Walts, G. Williams, D. Wilson, K. Youens-Clark, M.K. Monaco, J. Stein, X. Wei, D. Ware, D.M. Bolser, K.L. Howe, E. Kulesha, D. Lawson & D.M. Staines, (2014) Ensembl Genomes 2013: scaling up access to genome-wide data. *Nucleic Acids Res.* **42**: D546-D552.

Keseler, I.M., A. Mackie, M. Peralta-Gil, A. Santos-Zavaleta, S. Gama-Castro, C. Bonavides-Martinez, C. Fulcher, A.M. Huerta, A. Kothari, M. Krummenacker, M. Latendresse, L. Muniz-Rascado, Q. Ong, S. Paley, I. Schroder, A.G. Shearer, P. Subhraveti, M. Travers, D. Weerasinghe, V. Weiss, J. Collado-Vides, R.P. Gunsalus, I. Paulsen & P.D. Karp, (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* **41**: D605-612.

Kessels, J.M., H. Ousen & H. Van den Bosch, (1983) Facilitated utilization of endogenously synthesized lysophosphatidic acid by 1-acylglycerophosphate acyltransferase from Escherichia coli. *Biochimica et biophysica acta* **753**: 227-235.

Khanum, S., (2015) Characterization of the secretins. large outer membrane channels of Gram-negative bacteria. In: Institute of Fundamental Science. Palmerston North: Massey University, pp.

Kim, S., J.C. Malinverni, P. Sliz, T.J. Silhavy, S.C. Harrison & D. Kahne, (2007) Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**: 961-964.

Kitagawa, M., T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga & H. Mori, (2005) Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Research* **12**: 291-299.

Kleerebezem, M., W. Crielaard & J. Tommassen, (1996) Involvement of stress protein PspA (phage shock protein A) of Escherichia coli in maintenance of the protonmotive force under stress conditions. *Embo J.* **15**: 162-171.

Kleerebezem, M. & J. Tommassen, (1993) Expression of the *pspA* gene stimulates efficient protein export in *Escherichia coli*. *Mol. Microbiol.* **7**: 947-956.

Klena, J.D., R.S. Ashford, 2nd & C.A. Schnaitman, (1992) Role of Escherichia coli K-12 rfa genes and the rfp gene of Shigella dysenteriae 1 in generation of lipopolysaccharide core heterogeneity and attachment of O antigen. *J. Bacteriol.* **174**: 7297-7307.

Kobayashi, K. & S. Tagawa, (1999) Isolation of reductase for SoxR that governs an oxidative response regulon from Escherichia coli. *FEBS Lett.* **451**: 227-230.

Koebnik, R., K.P. Locher & P. Van Gelder, (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37**: 239-253.

Kol, M.A., A. van Dalen, A.I. de Kroon & B. de Kruijff, (2003) Translocation of phospholipids is facilitated by a subset of membrane-spanning proteins of the bacterial cytoplasmic membrane. *J. Biol. Chem.* **278**: 24586-24593.

Koo, J., L.L. Burrows & P.L. Howell, (2012) Decoding the roles of pilotins and accessory proteins in secretin escort services. *FEMS Microbiology Letters* **328**: 1-12.

Koo, J., S. Tammam, S.Y. Ku, L.M. Sampaleanu, L.L. Burrows & P.L. Howell, (2008) PilF Is an Outer Membrane Lipoprotein Required for Multimerization and Localization of the Pseudomonas aeruginosa Type IV Pilus Secretin. *J. Bacteriol.* **190**: 6961-6969.

Koo, M.S., J.H. Lee, S.Y. Rah, W.S. Yeo, J.W. Lee, K.L. Lee, Y.S. Koh, S.O. Kang & J.H. Roe, (2003) A reducing system of the superoxide sensor SoxR in Escherichia coli. *Embo J.* **22**: 2614-2622.

Koronakis, V., J. Li, E. Koronakis & K. Stauffer, (1997) Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol. Microbiol.* **23**: 617-626.

Korotkov, K.V., E. Pardon, J. Steyaert & W.G.J. Hol, (2009) Crystal Structure of the N-Terminal Domain of the Secretin GspD from ETEC Determined with the Assistance of a Nanobody. *Structure* **17**: 255-265.

Koutsolioutsou, A., S. Pena-Llopis & B. Demple, (2005) Constitutive soxR mutations contribute to multiple-antibiotic resistance in clinical Escherichia coli isolates. *Antimicrobial Agents and Chemotherapy* **49**: 2746-2752.

Krapp, A.R., M.V. Humbert & N. Carrillo, (2011) The soxRS response of Escherichia coli can be induced in the absence of oxidative stress and oxygen by modulation of NADPH content. *Microbiology* **157**: 957-965.

Kulkarni, M.M., (2011) Digital multiplexed gene expression analysis using the NanoString nCounter system. *Current Protocols in Molecular Biology*: 25B. 10.21-25B. 10.17.

Kumagai, H., T. Fujiwara, H. Matsubara & K. Saeki, (1997) Membrane localization, topology, and mutual stabilization of the rnfABC gene products in Rhodobacter capsulatus and implications for a new family of energy-coupling NADH oxidoreductases. *Biochemistry* **36**: 5509-5521.

Lan, C.-Y. & M.M. Igo, (1998) Differential Expression of the OmpF and OmpC Porin Proteins in Escherichia coli K-12 Depends upon the Level of Active OmpR. *J. Bacteriol.* **180**: 171-174.

Langmead, B. & S.L. Salzberg, (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357-U354.

Laubacher, M.E. & S.E. Ades, (2008) The Rcs phosphorelay is a cell envelope stress response activated by peptidoglycan stress and contributes to intrinsic antibiotic resistance. *J. Bacteriol.* **190**: 2065-2074.

Lechtzier, V., M. Hutoran, T. Levy, M. Kotler, T. Brenner & M. Steinitz, (2002) Sodium dodecyl sulphate-treated proteins as ligands in ELISA. *Journal of Immunological Methods* **270**: 19-26.

Lederberg, J., (1956) Bacterial protoplasts induced by penicillin. *Proc. Natl. Acad. Sci. U. S. A.* **42**: 574-577.

Lessl, M., D. Balzer, R. Lurz, V.L. Waters, D.G. Guiney & E. Lanka, (1992) Dissection of IncP conjugative plasmid transfer - definition of the transfer region Tra2 by mobilization of the Tra1 region in trans. *J. Bacteriol.* **174**: 2493-2500.

Levin, J.Z., M. Yassour, X. Adiconis, C. Nusbaum, D.A. Thompson, N. Friedman, A. Gnirke & A. Regev, (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nature Methods* **7**: 709-715.

Lima, S., M.S. Guo, R. Chaba, C.A. Gross & R.T. Sauer, (2013) Dual Molecular Signals Mediate the Bacterial Response to Outer-Membrane Stress. *Science* **340**: 837-841.

Lin, E.C., (1976) Glycerol dissimilation and its regulation in bacteria. *Annual review of microbiology* **30**: 535-578.

Linderoth, N.A., P. Model & M. Russel, (1996) Essential role of a sodium dodecyl sulfate-resistant protein IV multimer in assembly-export of filamentous phage. *J. Bacteriol.* **178**: 1962-1970.

Linderoth, N.A., M.N. Simon & M. Russel, (1997) The filamentous phage pIV multimer visualized by scanning transmission electron microscopy. *Science* **278**: 1635-1638.

Liochev, S.I. & I. Fridovich, (2011) Is superoxide able to induce SoxRS? *Free Radical Biology and Medicine* **50**: 1813.

Liu, J., E.J. Thanikkal, I.R. Obi & M.S. Francis, (2012) Elevated CpxR similar to P levels repress the Ysc-Yop type III secretion system of Yersinia pseudotuberculosis. *Res. Microbiol.* **163**: 518-530.

Lloyd, L.J., S.E. Jones, G. Jovanovic, P. Gyaneshwar, M.D. Rolfe, A. Thompson, J.C. Hinton & M. Buck, (2004) Identification of a new member of the phage shock protein response in Escherichia coli, the phage shock protein g (PspG). *J. Biol. Chem.* **279**: 55707-55714.

Love, M., W. Huber & S. Anders, (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

MacRitchie, D.M., N. Acosta & T.L. Raivio, (2012) DegP is involved in Cpx-mediated posttranscriptional regulation of the type III secretion apparatus in enteropathogenic Escherichia coli. *Infection & Immunity* **80**: 1766-1772.

Malinverni, J.C. & T.J. Silhavy, (2009) An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 8009-8014.

Marciano, D.K., M. Russel & S.M. Simon, (1999) An Aqueous Channel for Filamentous Phage Export. *Science* **284**: 1516-1519.

Marciano, D.K., M. Russel & S.M. Simon, (2001) Assembling filamentous phage occlude pIV channels. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 9359-9364.

Marlovits, T.C., T. Kubori, A. Sukhan, D.R. Thomas, J.E. Galan & V.M. Unger, (2004) Structural insights into the assembly of the type III secretion needle complex. *Science* **306**: 1040-1042.

Martin, R.G., E.S. Bartlett, J.L. Rosner & M.E. Wall, (2008) Activation of the Escherichia coli marA/soxS/rob regulon in response to transcriptional activator concentration. *J. Mol. Biol.* **380**: 278-284.

Martin, R.G., W.K. Gillette, S. Rhee & J.L. Rosner, (1999) Structural requirements for marbox function in transcriptional activation of mar/sox/rob regulon promoters in Escherichia coli: sequence, orientation and spatial relationship to the core promoter. *Mol. Microbiol.* **34**: 431-441.

Martin, R.G., W.K. Gillette & J.L. Rosner, (2000) Promoter discrimination by the related transcriptional activators MarA and SoxS: differential regulation by differential binding. *Mol. Microbiol.* **35**: 623-634.

Martin, R.G. & J.L. Rosner, (1995) Binding of purified multiple antibiotic-resistance repressor protein (MarR) to mar operator sequences. *Proc. Natl. Acad. Sci. U. S. A.* **92**: 5456-5460.

Martin, R.G. & J.L. Rosner, (2002) Genomics of the marA/soxS/rob regulon of Escherichia coli: identification of directly activated promoters by application of molecular genetics and informatics to microarray data. *Mol. Microbiol.* **44**: 1611-1624.

Martin, R.G. & J.L. Rosner, (2004) Transcriptional and translational regulation of the marRAB multiple antibiotic resistance operon in Escherichia coli. *Mol. Microbiol.* **53**: 183-191.

Maxson, M.E. & A.J. Darwin, (2004) Identification of inducers of the Yersinia enterocolitica phage shock protein system and comparison to the regulation of the RpoE and Cpx extracytoplasmic stress responses. *J. Bacteriol.* **186**: 4199-4208.

Maxson, M.E. & A.J. Darwin, (2006) PspB and PspC of Yersinia enterocolitica are dual function proteins: regulators and effectors of the phage-shock-protein response. *Mol. Microbiol.* **59**: 1610-1623.

Mi, H., A. Muruganujan, J.T. Casagrande & P.D. Thomas, (2013a) Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**: 1551-1566.

Mi, H., A. Muruganujan & P.D. Thomas, (2013b) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **41**: D377-386.

Miki, T., N. Okada, Y. Shimada & H. Danbara, (2004) Characterization of Salmonella pathogenicity island 1 type III secretion-dependent hemolytic activity in Salmonella enterica serovar Typhimurium. *Microb. Pathog.* **37**: 65-72.

Misra, R. & S.A. Benson, (1988a) Genetic identification of the pore domain of the OmpC porin of *Escherichia coli* K-12. *J. Bacteriol.* **170**: 3611-3617.

Misra, R. & S.A. Benson, (1988b) Isolation and characterisation of OmpC porin mutants with altered pore properties. *J. Bacteriol.* **170**: 528-533.

Missiakas, D., J.M. Betton & S. Raina, (1996) New components of protein folding in extracytoplasmic compartments of Escherichia coli SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* **21**: 871-884.

Missiakas, D., M.P. Mayer, M. Lemaire, C. Georgopoulos & S. Raina, (1997) Modulation of the Escherichia coli sigmaE (RpoE) heat-shock

transcription-factor activity by the RseA, RseB and RseC proteins. *Mol. Microbiol.* **24**: 355-371.
Miyadai, H., K. Tanaka-Masuda, S. Matsuayama & H. Tokuda, (2004) Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the Escherichia coli periplasm. *J. Biol. Chem.* **279**: 39807-39813.
Mizuno, T. & S. Mizushima, (1990) Signal transduction and gene regulation through the phosphorylation of two regulatory components: the molecular basis for the osmotic regulation of the porin genes. *Mol. Microbiol.* **4**: 1077-1082.
Moreau, P.L., (2004) Diversion of the metabolic flux from pyruvate dehydrogenase to pyruvate oxidase decreases oxidative stress during glucose metabolism in nongrowing Escherichia coli cells incubated under aerobic, phosphate starvation conditions. *J. Bacteriol.* **186**: 7364-7368.
Nakajima, H., K. Kobayashi, M. Kobayashi, H. Asako & R. Aono, (1995) Overexpression of the robA gene increases organic solvent tolerance and multiple antibiotic and heavy metal ion resistance in Escherichia coli. *Appl. Environ. Microbiol.* **61**: 2302-2307.
Narita, S. & H. Tokuda, (2009) Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. *FEBS Lett.* **583**: 2160-2164.
Nikaido, H., (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**: 593-+.
Nikaido, H., (2005) Restoring permeability barrier function to outer membrane. *Chemistry & biology* **12**: 507-509.
Nouwen, N., N. Ranson, H. Saibil, B. Wolpensinger, A. Engel, A. Ghazi & A.P. Pugsley, (1999) Secretin PulD: association with pilot PulS, structure, and ion-conducting channel formation. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 8173-8177.
Ogasawara, H., K. Yamada, A. Kori, K. Yamamoto & A. Ishihama, (2010) Regulation of the Escherichia coli csgD promoter: interplay between five transcription factors. *Microbiology* **156**: 2470-2483.
Okon, M., T.F. Moraes, P.I. Lario, A.L. Creagh, C.A. Haynes, N.C.J. Strynadka & L.P. McLntosh, (2008) Structural Characterization of the Type-III Pilot-Secretin Complex from Shigella flexneri. *Structure* **16**: 1544-1554.
Oliver, S. & T.C. Marlovits, (2011) Three-Dimensional Model of Salmonella's Needle Complex at Subnanometer Resolution. *Science* **331**: 1192-1195.
Onufryk, C., M.L. Crouch, F.C. Fang & C.A. Gross, (2005) Characterization of six lipoproteins in the sigmaE regulon. *J. Bacteriol.* **187**: 4552-4561.
Opalka, N., R. Beckmann, N. Boisset, M.N. Simon, M. Russel & S.A. Darst, (2003) Structure of the filamentous phage pIV multimer by cryo-electron microscopy. *J. Mol. Biol.* **325**: 461-470.
Otto, K. & T.J. Silhavy, (2002) Surface sensing and adhesion of Escherichia coli controlled by the Cpx-signaling pathway. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 2287-2292.
Palmer, T. & B.C. Berks, (2012) The twin-arginine translocation (Tat) protein export pathway. *Nature Reviews Microbiology* **10**: 483-496.
Peleg, A., Y. Shifrin, O. Ilan, C. Nadler-Yona, S. Nov, S. Koby, K. Baruch, S. Altuvia, M. Elgrably-Weiss, C.M. Abe, S. Knutton, M.A. Saper & I.

Rosenshine, (2005) Identification of an Escherichia coli operon required for formation of the O-antigen capsule. *J. Bacteriol.* **187**: 5259-5266.
Prigent-Combaret, C., E. Brombacher, O. Vidal, A. Ambert, P. Lejeune, P. Landini & C. Dorel, (2001) Complex regulatory network controls initial adhesion and biofilm formation in Escherichia coli via regulation of the csgD gene. *J. Bacteriol.* **183**: 7213-7223.
Quan, S., P. Koldewey, T. Tapley, N. Kirsch, K.M. Ruane, J. Pfizenmaier, R. Shi, S. Hofmann, L. Foit, G. Ren, U. Jakob, Z.H. Xu, M. Cygler & J.C.A. Bardwell, (2011) Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nat. Struct. Mol. Biol.* **18**: 262-U241.
R Development Core Team, (2013) R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing, pp.
Raffa, R.G. & T.L. Raivio, (2002) A third envelope stress signal transduction pathway in Escherichia coli. *Mol. Microbiol.* **45**: 1599-1611.
Rahav-Manor, O., O. Carmel, R. Karpel, D. Taglicht, G. Glaser, S. Schuldiner & E. Padan, (1992) NhaR, a protein homologous to a family of bacterial regulatory proteins (LysR), regulates nhaA, the sodium proton antiporter gene in Escherichia coli. *J. Biol. Chem.* **267**: 10433-10438.
Raivio, T.L., D.L. Popkin & T.J. Silhavy, (1999) The Cpx envelope stress response is controlled by amplification and feedback inhibition. *J. Bacteriol.* **181**: 5263-5272.
Raivio, T.L. & T.J. Silhavy, (1997) Transduction of envelope stress in Escherichia coli by the Cpx two-component system. *J. Bacteriol.* **179**: 7724-7733.
Randall, L.P. & M.J. Woodward, (2002) The multiple antibiotic resistance (mar) locus and its significance. *Research in Veterinary Science* **72**: 87-93.
Randall-Hazelbauer, L. & M. Schwartz, (1973) Isolation of the bacteriophage lambda receptor from Escherichia coli. *J. Bacteriol.* **116**: 1436-1446.
Ray, B.L. & C.R. Raetz, (1987) The biosynthesis of gram-negative endotoxin. A novel kinase in Escherichia coli membranes that incorporates the 4'-phosphate of lipid A. *J. Biol. Chem.* **262**: 1122-1128.
Rey, S., M. Acab, J.L. Gardy, M.R. Laird, K. deFays, C. Lambert & F.S. Brinkman, (2005) PSORTdb: a protein subcellular localization database for bacteria. *Nucleic Acids Res.* **33**: D164-168.
Rhodius, V.A., W.C. Suh, G. Nonaka, J. West & C.A. Gross, (2006) Conserved and variable functions of the sigma(E) stress response in related genomes. *PLoS Biology* **4**: 43-59.
Rizzitello, A.E., J.R. Harper & T.J. Silhavy, (2001) Genetic evidence for parallel pathways of chaperone activity in the periplasm of Escherichia coli. *J. Bacteriol.* **183**: 6794-6800.
Robert, V., E.B. Volokhina, F. Senf, M.P. Bos, P. Van Gelder & J. Tommassen, (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biology* **4**: e377.
Rock, C.O. & S. Jackowski, (2002) Forty years of bacterial fatty acid synthesis. *Biochemical and biophysical research communications* **292**: 1155-1166.
Rowley, G., M. Spector, J. Kormanec & M. Roberts, (2006) Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nature Reviews Microbiology* **4**: 383-394.

Ruiz, N., L.S. Gronenberg, D. Kahne & T.J. Silhavy, (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 5537-5542.

Ruiz, N. & T.J. Silhavy, (2005) Sensing external stress: watchdogs of the Escherichia coli cell envelope. *Curr. Opin. Microbiol.* **8**: 122-126.

Russel, M., (1994) Mutants at Conserved Positions in Gene-IV, a Gene Required for Assembly and Secretion of Filamentous Phages. *Mol. Microbiol.* **14**: 357-369.

Russel, M., (1998) Macromolecular assembly and secretion across the bacterial cell envelope: Type II protein secretion systems. *J. Mol. Biol.* **279**: 485-499.

Russel, M. & B. Kazmierczak, (1993) Analysis of the Structure and Subcellular Location of Filamentous Phage-pIV. *J. Bacteriol.* **175**: 3998-4007.

Salgado, H., M. Peralta-Gil, S. Gama-Castro, A. Santos-Zavaleta, L. Muniz-Rascado, J.S. Garcia-Sotelo, V. Weiss, H. Solano-Lira, I. Martinez-Flores, A. Medina-Rivera, G. Salgado-Osorio, S. Alquicira-Hernandez, K. Alquicira-Hernandez, A. Lopez-Fuentes, L. Porron-Sotelo, A.M. Huerta, C. Bonavides-Martinez, Y.I. Balderas-Martinez, L. Pannier, M. Olvera, A. Labastida, V. Jimenez-Jacinto, L. Vega-Alvarado, V. Del Moral-Chavez, A. Hernandez-Alvarez, E. Morett & J. Collado-Vides, (2013) RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res.* **41**: D203-213.

Salscheider, S.L., A. Jahn & K. Schnetz, (2014) Transcriptional regulation by BglJ-RcsB, a pleiotropic heteromeric activator in Escherichia coli. *Nucleic Acids Res.* **42**: 2999-3008.

Sambrook, J., E.F. Fritsch & T. Maniatis, (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sankaran, K. & H.C. Wu, (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.* **269**: 19701-19706.

Schleifer, K.H. & O. Kandler, (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological reviews* **36**: 407-477.

Schroeder, A., O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow & T. Ragg, (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* **7**: 3.

Seo, J., A. Brencic & A.J. Darwin, (2009) Analysis of Secretin-Induced Stress in Pseudomonas aeruginosa Suggests Prevention Rather than Response and Identifies a Novel Protein Involved in Secretin Function. *J. Bacteriol.* **191**: 898-908.

Seo, J., D.C. Savitzky, E. Ford & A.J. Darwin, (2007) Global analysis of tolerance to secretin-induced stress in Yersinia enterocolitica suggests that the phage-shock-protein system may be a remarkably self-contained stress response. *Mol. Microbiol.* **65**: 714-727.

Sévin, D.C. & U. Sauer, (2014) Ubiquinone accumulation improves osmotic-stress tolerance in Escherichia coli. *Nature Chemical Biology* **10**: 266-272.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski & T. Ideker, (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**: 2498-2504.

Shin, S. & C. Park, (1995) Modulation of flagellar expression in Escherichia coli by acetyl phosphate and the osmoregulator OmpR. *J. Bacteriol.* **177**: 4696-4702.

Shinhara, A., M. Matsui, K. Hiraoka, W. Nomura, R. Hirano, K. Nakahigashi, M. Tomita, H. Mori & A. Kanai, (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in Escherichia coli. *BMC Genomics* **12**: 428-428.

Sikora, A.E., S.R. Lybarger & M. Sandkvist, (2007) Compromised outer membrane integrity in Vibrio cholerae type II secretion mutants. *J. Bacteriol.* **189**: 8484-8495.

Silverman, P.M., (1985) Host-cell plasmid interactions in the expression of DNA donor activity by F+ strains of *Escherichia coli* K-12. *BioEssays* **2**: 254-259.

Six, D.A., S.M. Carty, Z. Guan & C.R. Raetz, (2008) Purification and mutagenesis of LpxL, the lauroyltransferase of Escherichia coli lipid A biosynthesis. *Biochemistry* **47**: 8623-8637.

Skarstad, K., B. Thony, D.S. Hwang & A. Kornberg, (1993) A novel binding-protein of the origin of the *Escherichia coli* chromosome. *J. Biol. Chem.* **268**: 5365-5370.

Sklar, J.G., T. Wu, D. Kahne & T.J. Silhavy, (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21**: 2473-2484.

Sledjeski, D.D. & S. Gottesman, (1996) Osmotic shock induction of capsule synthesis in Escherichia coli K-12. *J. Bacteriol.* **178**: 1204-1206.

Smirnova, G.V., N.G. Muzyka & O.N. Oktyabrsky, (2000) The role of antioxidant enzymes in response of Escherichia coli to osmotic upshift. *FEMS Microbiology Letters* **186**: 209-213.

Snyder, W.B., L.J.B. Davis, P.N. Danese, C.L. Cosma & T.J. Silhavy, (1995) Overproduction of NlpE, a new outer-membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal-transduction pathway. *J. Bacteriol.* **177**: 4216-4223.

Spagnuolo, J., (2010) Defining the Gate Domain of the Filamentous Phage Secretin pIV. In: Institute of Molecular BioSciences. Palmerston North: Massey University, pp. 119.

Spagnuolo, J., N. Opalka, W.X. Wen, D. Gagic, E. Chabaud, D. Bellini, M. Bennett, G.E. Norris, S.A. Darst, M. Russel & J. Rakonjac, (2010) Identification of the gate regions in the primary structure of the secretin pIV. *Mol. Microbiol.* **76**: 17.

Sperandeo, P., R. Cescutti, R. Villa, C. Di Benedetto, D. Candia, G. Deho & A. Polissi, (2007) Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli. *J. Bacteriol.* **189**: 244-253.

Sperandeo, P., F.K. Lau, A. Carpentieri, C. De Castro, A. Molinaro, G. Deho, T.J. Silhavy & A. Polissi, (2008) Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of Escherichia coli. *J. Bacteriol.* **190**: 4460-4469.

Sperandeo, P., R. Villa, A.M. Martorana, M. Samalikova, R. Grandori, G. Deho & A. Polissi, (2011) New insights into the Lpt machinery for lipopolysaccharide transport to the cell surface: LptA-LptC interaction and LptA stability as sensors of a properly assembled transenvelope complex. *J. Bacteriol.* **193**: 1042-1053.

Spreter, T., C.K. Yip, S. Sanowar, I. Andre, T.G. Kimbrough, M. Vuckovic, R.A. Pfuetzner, W.Y. Deng, A.C. Yu, B.B. Finlay, D. Baker, S.I. Miller & N.C.J. Strynadka, (2009) A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nat. Struct. Mol. Biol.* **16**: 468-476.

Stevenson, G., K. Andrianopoulos, M. Hobbs & P.R. Reeves, (1996) Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J. Bacteriol.* **178**: 4885-4893.

Stout, V. & S. Gottesman, (1990) RcsB and RcsC: A 2-component regulator of capsule synthesis in *Escherichia coli*. *J. Bacteriol.* **172**: 659-669.

Stout, V., A. Torres-Cabassa, M.R. Maurizi, D. Gutnick & S. Gottesman, (1991) RcsA, an unstable positive regulator of capsular polysaccharide synthesis. *J. Bacteriol.* **173**: 1738-1747.

Sturm, A., M. Heinemann, M. Arnoldini, A. Benecke, M. Ackermann, M. Benz, J. Dormann & W.D. Hardt, (2011) The Cost of Virulence: Retarded Growth of Salmonella Typhimurium Cells Expressing Type III Secretion System 1. *PLoS Pathogens* **7**.

Takeda, S., Y. Fujisawa, M. Matsubara, H. Aiba & T. Mizuno, (2001) A novel feature of the multistep phosphorelay in Escherichia coli: a revised model of the RcsC --> YojN --> RcsB signalling pathway implicated in capsular synthesis and swarming behaviour. *Mol Microbiol* **40**: 440-450.

Tam, R. & M.H. Saier Jr, (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *MICROBIOL. REV.* **57**: 320-346.

Thede, G.L., D.C. Arthur, R.A. Edwards, D.R. Buelow, J.L. Wong, T.L. Raivio & J.N.M. Glover, (2011) Structure of the Periplasmic Stress Response Protein CpxP. *J. Bacteriol.* **193**: 2149-2157.

Thomas, P.D., A. Kejariwal, N. Guo, H. Mi, M.J. Campbell, A. Muruganujan & B. Lazareva-Ulitsky, (2006) Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* **34**: W645-W650.

Tokuda, H. & S. Matsuyama, (2004) Sorting of lipoproteins to the outer membrane in E-coli. *Biochimica Et Biophysica Acta - Molecular Cell Research* **1693**: 5-13.

Tokunaga, M., J.M. Loranger & H.C. Wu, (1984a) A distinct signal peptidase for prolipoprotein in Escherichia coli. *Journal of cellular biochemistry* **24**: 113-120.

Tokunaga, M., J.M. Loranger & H.C. Wu, (1984b) Prolipoprotein modification and processing enzymes in Escherichia coli. *J. Biol. Chem.* **259**: 3825-3830.

Toledano, M.B., I. Kullik, F. Trinh, P.T. Baird, T.D. Schneider & G. Storz, (1994) Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* **78**: 897-909.

Torrescabassa, A.S. & S. Gottesman, (1987) Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. *J. Bacteriol.* **169**: 981-989.
Tosi, T., Leandro F. Estrozi, V. Job, I. Guilvout, Anthony P. Pugsley, G. Schoehn & A. Dessen, (2014) Structural Similarity of Secretins from Type II and Type III Secretion Systems. *Structure* **22**: 1348-1355.
Vianney, A., G. Jubelin, S. Renault, C. Dorel, P. Lejeune & J.C. Lazzaroni, (2005) Escherichia coli tol and rcs genes participate in the complex network affecting curli synthesis. *Microbiology* **151**: 2487-2497.
Viarre, V., E. Cascales, G. Ball, G.P.F. Michel, A. Filloux & R. Voulhoux, (2009) HxcQ Liposecretin Is Self-piloted to the Outer Membrane by Its N-terminal Lipid Anchor. *J. Biol. Chem.* **284**: 33815-33823.
Vinue, L., L.M. McMurry & S.B. Levy, (2013) The 216-bp marB gene of the marRAB operon in Escherichia coli encodes a periplasmic protein which reduces the transcription rate of marA. *FEMS Microbiology Letters* **345**: 49-55.
Voegele, R.T., G.D. Sweet & W. Boos, (1993) Glycerol kinase of Escherichia coli is activated by interaction with the glycerol facilitator. *J. Bacteriol.* **175**: 1087-1094.
Volokhina, E.B., F. Beckers, J. Tommassen & M.P. Bos, (2009) The beta-Barrel Outer Membrane Protein Assembly Complex of Neisseria meningitidis. *J. Bacteriol.* **191**: 7074-7085.
Voulhoux, R., M.P. Bos, J. Geurtsen, M. Mols & J. Tommassen, (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**: 262-265.
Walker, K.A. & V.L. Miller, (2009) Synchronous gene expression of the Yersinia enterocolitica Ysa type III secretion system and its effectors. *J. Bacteriol.* **191**: 1816-1826.
Walsh, N.P., B.M. Alba, B. Bose, C.A. Gross & R.T. Sauer, (2003) OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113**: 61-71.
Wandersman, C., M. Schwartz & T. Ferenci, (1979) *Escherichia coli* mutants impaired in maltodextrin transport. *J. Bacteriol.* **140**: 1-13.
Wang, X., Y. Kim, S.H. Hong, Q. Ma, B.L. Brown, M. Pu, A.M. Tarone, M.J. Benedik, W. Peti, R. Page & T.K. Wood, (2011) Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology* **7**: 359-366.
Wang, X., J.F. Preston & T. Romeo, (2004) The pgaABCD locus of Escherichia coli promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J. Bacteriol.* **186**: 2724-2734.
Wang, Z., M. Gerstein & M. Snyder, (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**: 57-63.
Weber, A., S.A. Kogl & K. Jung, (2006) Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli. *J. Bacteriol.* **188**: 7165-7175.
Wehland, M. & F. Bernhard, (2000) The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. *J. Biol. Chem.* **275**: 7013-7020.
Weiner, L., J.L. Brissette & P. Model, (1991) Stress-induced expression of the Escherichia coli phage shock protein operon is dependent on sigma 54

and modulated by positive and negative feedback mechanisms. *Genes Dev.* **5**: 1912-1923.
Weiss, M.S. & G.E. Schulz, (1992) Structure of porin refined at 1.8 A resolution. *J. Mol. Biol.* **227**: 493-509.
Weissenborn, D.L., N. Wittekindt & T.J. Larson, (1992) Structure and regulation of the glpFK operon encoding glycerol diffusion facilitator and glycerol kinase of Escherichia coli K-12. *J. Biol. Chem.* **267**: 6122-6131.
Whitaker, R., (2012) Analysis of the gate residues in the Type 2 secretin PulD. In: Institute of Molecular BioSciences. Palmerston North: Massey University, pp.
Wilhelm, B.T. & J.-R. Landry, (2009) RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods* **48**: 249-257.
Wolfe, A.J., N. Parikh, B.P. Lima & B. Zemaitaitis, (2008) Signal integration by the two-component signal transduction response regulator CpxR. *J. Bacteriol.* **190**: 2314-2322.
Wright, P.R., A.S. Richter, K. Papenfort, M. Mann, J. Vogel, W.R. Hess, R. Backofen & J. Georg, (2013) Comparative genomics boosts target prediction for bacterial small RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **110**: E3487-3496.
Wu, T., J. Malinverni, N. Ruiz, S. Kim, T.J. Silhavy & D. Kahne, (2005) Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121**: 235-245.
Wu, T., A.C. McCandlish, L.S. Gronenberg, S.S. Chng, T.J. Silhavy & D. Kahne, (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 11754-11759.
Yakushi, T., K. Masuda, S. Narita, S. Matsuyama & H. Tokuda, (2000) A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nature cell biology* **2**: 212-218.
Yamaguchi, K., F. Yu & M. Inouye, (1988) A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell* **53**: 423-432.
Yamaguchi, S., D.A. Reid, E. Rothenberg & A.J. Darwin, (2013) Changes in Psp protein binding partners, localization and behaviour upon activation of the Yersinia enterocolitica phage shock protein response. *Mol. Microbiol.* **87**: 656-671.
Yethon, J.A., D.E. Heinrichs, M.A. Monteiro, M.B. Perry & C. Whitfield, (1998) Involvement of waaY, waaQ, and waaP in the modification of Escherichia coli lipopolysaccharide and their role in the formation of a stable outer membrane. *J. Biol. Chem.* **273**: 26310-26316.
Yokota, N., T. Kuroda, S. Matsuyama & H. Tokuda, (1999) Characterization of the LolA-LolB system as the general lipoprotein localization mechanism of Escherichia coli. *J. Biol. Chem.* **274**: 30995-30999.
Yoshida, T., S. Cai & M. Inouye, (2002) Interaction of EnvZ, a sensory histidine kinase, with phosphorylated OmpR, the cognate response regulator. *Mol. Microbiol.* **46**: 1283-1294.
Young, K., L.L. Silver, D. Bramhill, P. Cameron, S.S. Eveland, C.R. Raetz, S.A. Hyland & M.S. Anderson, (1995) The envA permeability/cell division gene of Escherichia coli encodes the second enzyme of lipid A

biosynthesis. UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase. *J. Biol. Chem.* **270**: 30384-30391.

Yu, N.Y., M.R. Laird, C. Spencer & F.S. Brinkman, (2011) PSORTdb--an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res.* **39**: D241-244.

Zhang, A., S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis & G. Storz, (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *Embo J.* **17**: 6061-6068.

Zhang, Y.M. & C.O. Rock, (2008) Membrane lipid homeostasis in bacteria. *Nature Reviews Microbiology* **6**: 222-233.

Zheng, M., F. Aslund & G. Storz, (1998) Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**: 1718-1721.

Zielke, R.A., R.S. Simmons, B.R. Park, M. Nonogaki, S. Emerson & A.E. Sikora, (2014) The Type II Secretion Pathway in Vibrio cholerae Is Characterized by Growth Phase-Dependent Expression of Exoprotein Genes and Is Positively Regulated by σ(E). *Infection & Immunity* **82**: 2788-2801.

Appendix

8 Appendix 1: Differential Expression of the CpxR regulon in response to wildtype pIV

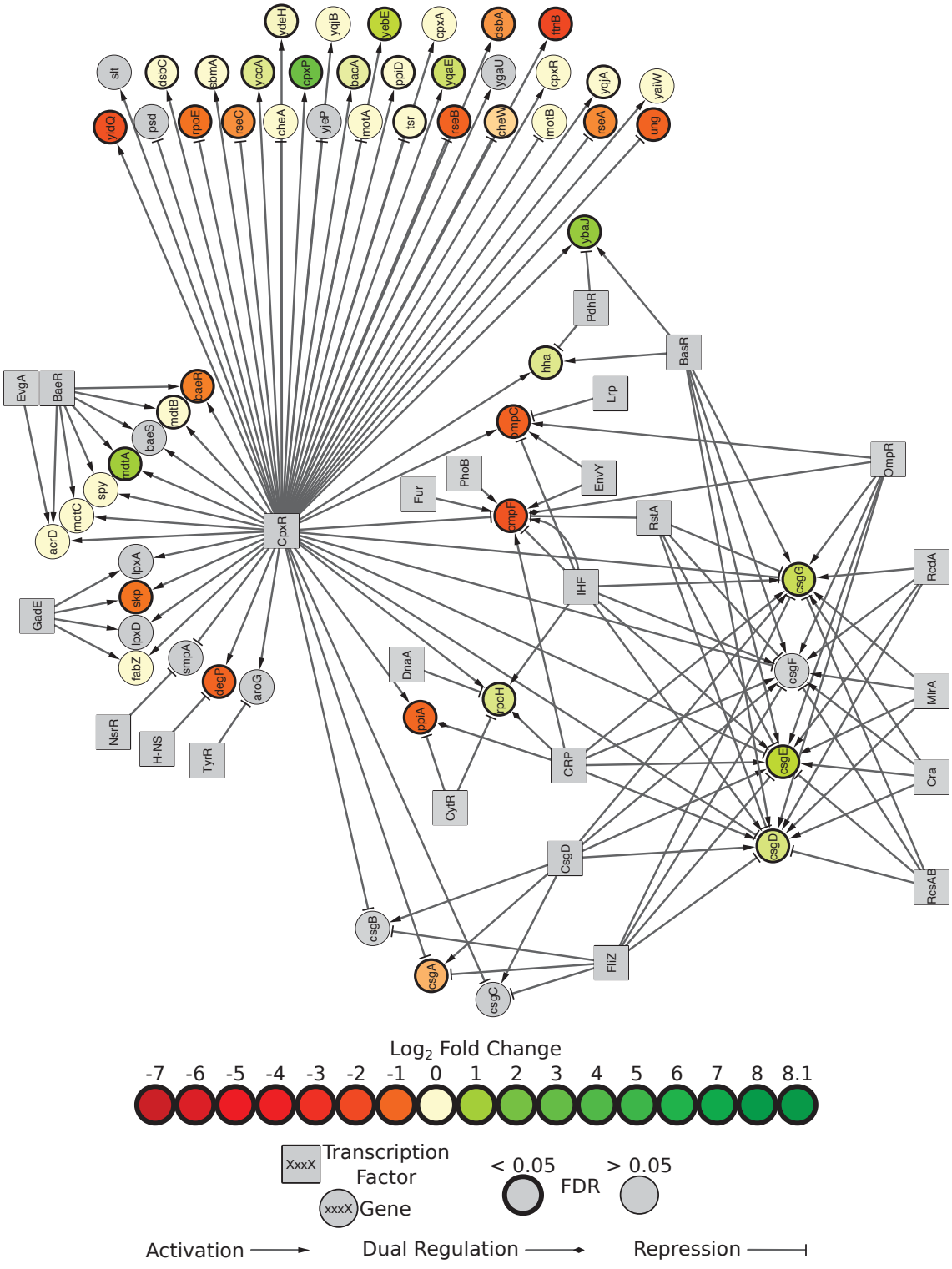

Figure 33: Differential expression of the CpxR regulatory network in response to wildtype pIV.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 fold change relative to the vector control reported by DESeq2 is represented by the colour of the node. Colour scale is limited to the minimum and maximum fold-changes of the genes mapped to the entire RegulonDB network. False discovery rate is represented by the node border thickness. Grey-filled nodes are genes that were removed from the primary analysis.

9 Appendix 2: Differential expression of the CpxR regulon in response to pIV-E292K

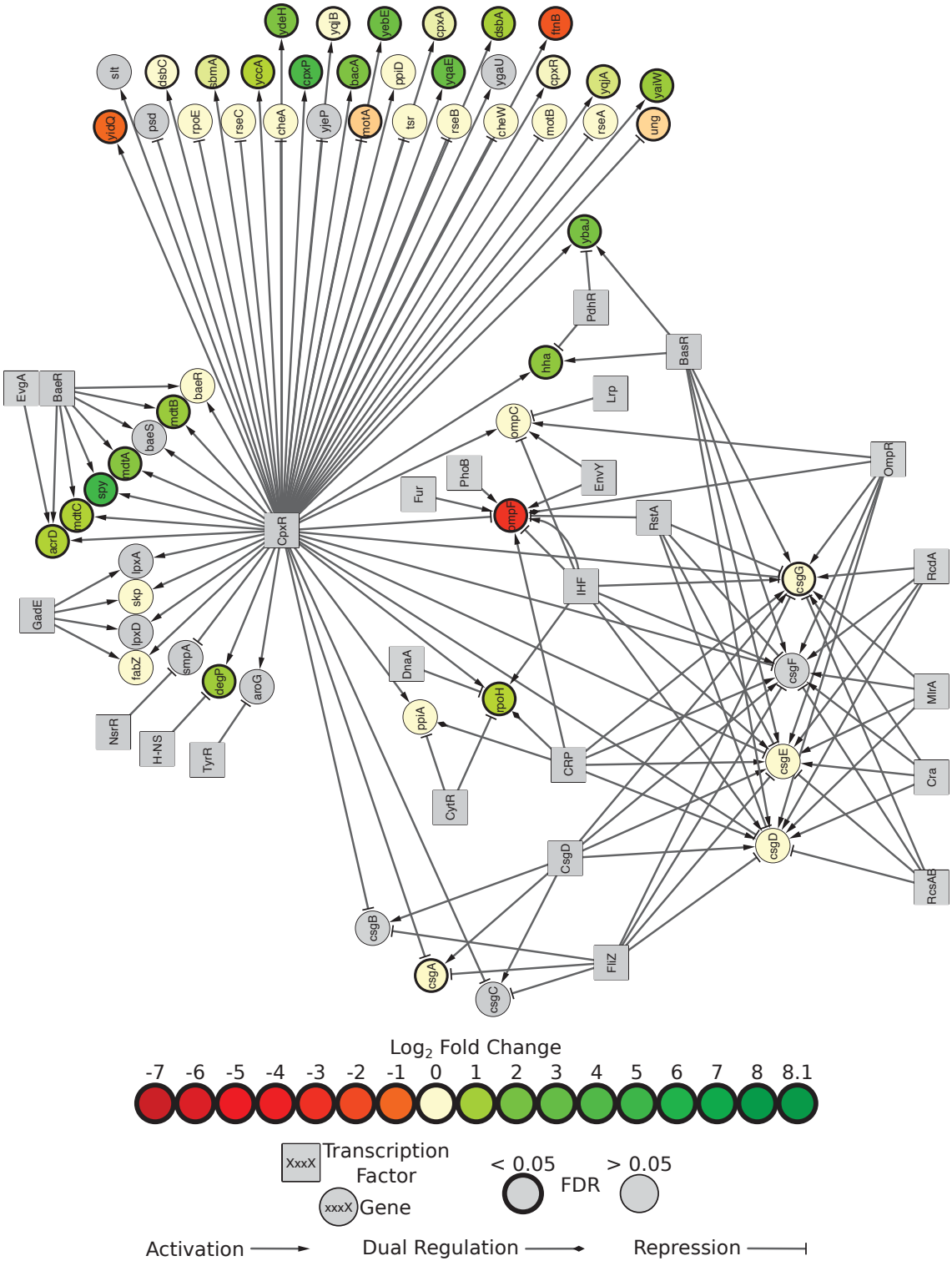

Figure 34: Differential expression of the CpxR regulatory network in response to leaky mutant pIV-E292K.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 fold change relative to the vector control reported by DESeq2 is represented by the colour of the node. Colour scale is limited to the minimum and maximum fold-changes of the genes mapped to the entire RegulonDB network. False discovery rate is represented by the node border thickness. Grey-filled nodes are genes that were removed from the primary analysis.

10 Appendix 3: Differential expression of the Sigma-E regulatory network in response to wildtype pIV

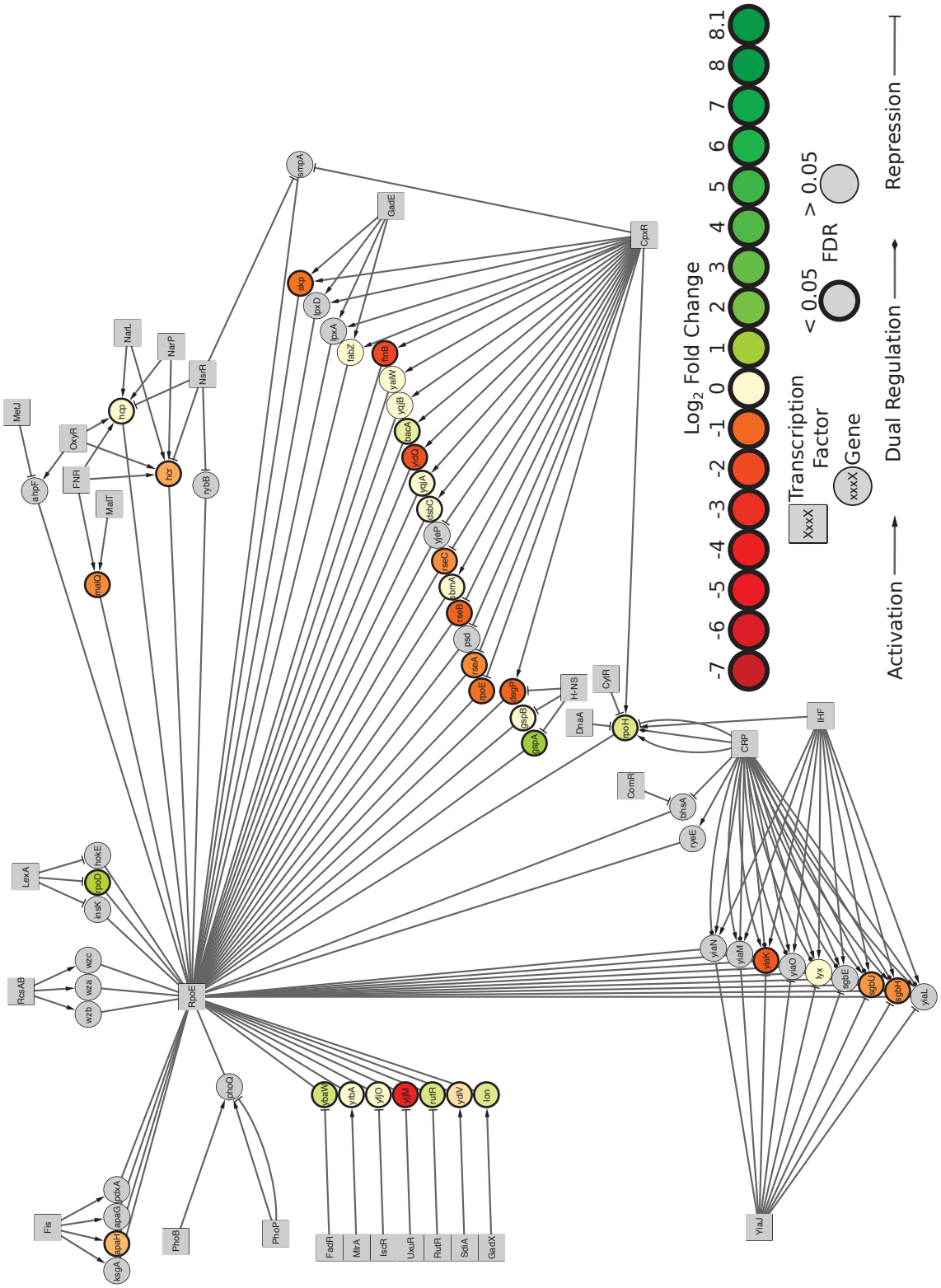

Figure 35: Differential expression of genes regulated by the alternative sigma factor, Sigma-E (RpoE) in response to wildtype pIV.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 fold change relative to the vector control reported by DESeq2 is represented by the colour of the node. Colour scale is limited to the minimum and maximum fold-changes of the genes mapped to the entire RegulonDB network. False discovery rate is represented by the node border thickness. Grey-filled nodes are genes that were removed from the primary analysis.

11 Appendix 4: Differential expression of the Sigma-E regulatory network in response to pIV-E292K

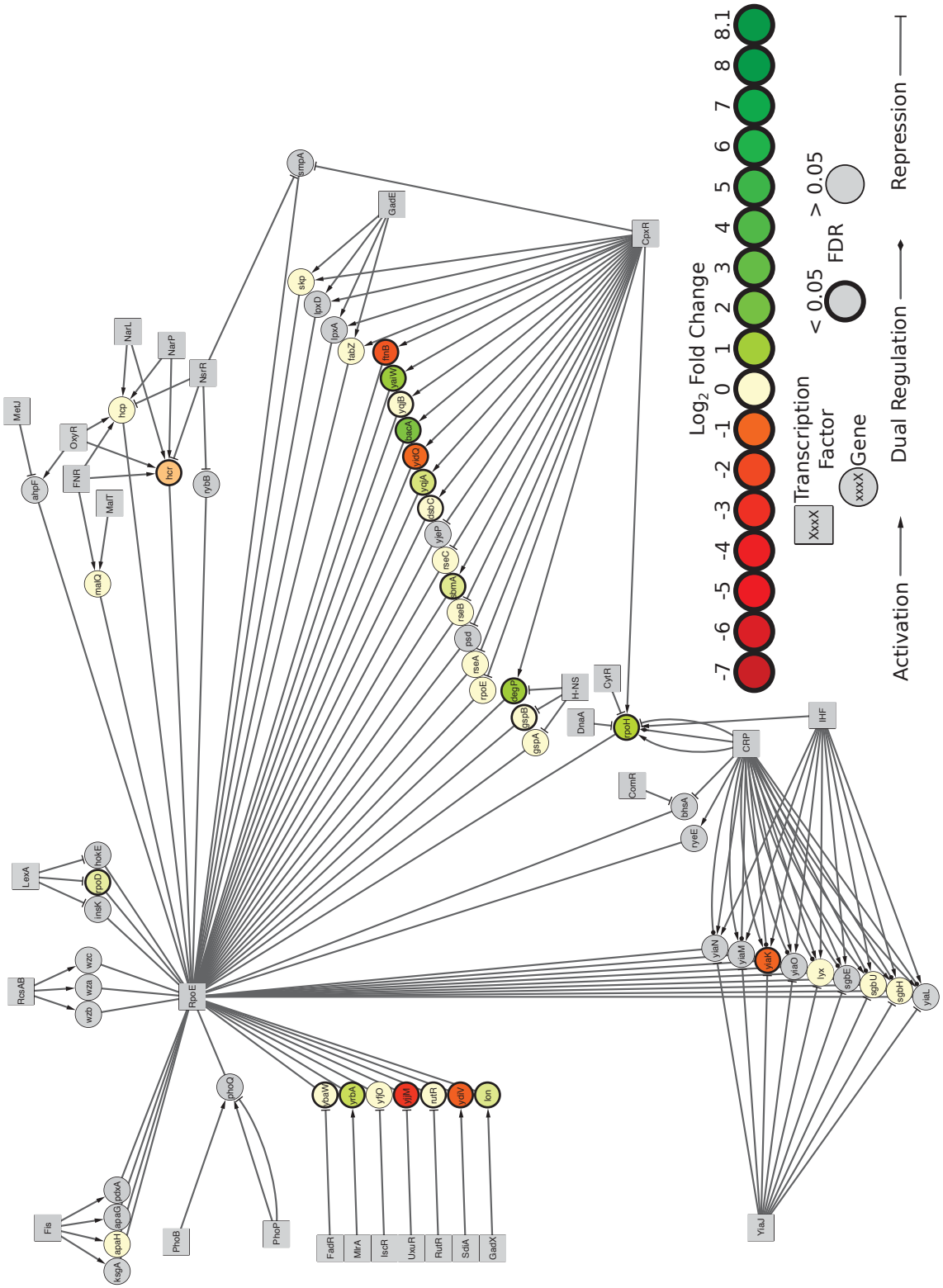

Figure 36: Differential expression of genes regulated by the alternative sigma factor, Sigma-E (RpoE) in response to leaky mutant pIV-E292K.

Transcription factor-to-gene networks for *E.* coli MG1655 were obtained from RegulonDB (Salgado *et al.*, 2013). Gene identifiers from *E. coli* BW2952 were mapped to equivalent MG1655 genes and the network analysis software, Cytoscape v3.2.1 (Shannon *et al.*, 2003), was used to create network views. Log_2 fold change relative to the vector control reported by DESeq2 is represented by the colour of the node. Colour scale is limited to the minimum and maximum fold-changes of the genes mapped to the entire RegulonDB network.. False discovery rate is represented by the node border thickness. Grey-filled nodes are genes that were removed from the primary analysis.