PagP-MEDIATED SIGNAL TRANSDUCTION

PagP-MEDIATED SIGNAL TRANSDUCTION: LINK TO THE *dmsABC* OPERON IN *Escherichia coli.*

BY LISET MALDONADO ALVAREZ. B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirement for the Degree Doctor of Philosophy

McMaster University

DESCRIPTIVE NOTE

DOCTOR OF PHILOSOPHY (2018) McMaster University
(Chemical Biology) Hamilton, Ontario

TITLE: PagP-mediated Signal Transduction: Link to the *dmsABC* operon in *Escherichia coli.*

SUPERVISOR: Dr. Russell Bishop
NUMBER OF PAGES: 172

ABSTRACT

In *Escherichia coli*, the integral outer membrane (OM) enzyme PagP covalently modifies lipid A by incorporating a phospholipid-derived palmitate chain to fortify the OM permeability barrier. We perturbed the bacterial OM in order to activate PagP and examined if it exerts transcriptional regulation through either of its extracellular or periplasmic active sites. Data from RNA-seq revealed the differential expression of 50 genes upon comparing the *E. coli imp4213* (*lptD4213*) strain NR760Δ*pagP*λInCh*pagP* (shortened to NR760*λp* in this work), in which PagP is constitutively activated, and the mutant NR760*λpY87F* carrying the periplasmic residue mutation Y87F. 40 genes were upregulated, and encoded proteins related to anaerobic processes, whereas 10 genes were downregulated and encoded proteins related to aerobic processes. RNA-seq was followed by a study of differential gene expression using the NanoString nCounter system. Results confirmed a 2.7-fold upregulation of *dmsA* when we compared the strains NR760*λp* to NR760*λpY87F*. We also found a 2.5-fold repression of *dmsA* transcription when we compared the *lptD*$^{+}$ parental strain NR754*λp* to NR754*λpS77A* carrying the mutation S77A in the extracellular active site of PagP. We then investigated *dmsA* transcription using a *lacZ* reporter gene in plasmid pRS551-*lacZ*. High basal β-galactosidase activity became attenuated in *pagP* null mutants. OM perturbation using pentamidine showed that *dmsA* transcription was repressed. Complementation of the chromosomal Δ*pagP* deletion with a single copy pBADGr plasmid, expressing PagP under the control of an arabinose-inducible promoter,

restored *dmsA* β-galactosidase activity. We observed partial complementation with the downstream *cspE* gene, which identified a polar effect of the Δ*pagP* allele. Through deletion of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB*, and *fadD* in *E. coli* MC4100, we showed that β-galactosidase activity of p*dmsA*-*lacZ* was affected by all of these regulators. Our results indicate that these regulators are involved in PagP-mediated regulation of *dmsA* transcription under aerobic conditions.

Acknowledgements

Foremost, I would like to thank God for showing me the path to my goals without giving up. Thank you for showing me that no obstacles are too large to overcome and for helping me to stay focused no matter what happens.

I would like to thank my supervisor Dr. Russell Bishop for giving me the opportunity to work in his lab and to work on this project, which I absolutely enjoyed. I learned so much. I am greatly indebted to you. None of this would have been possible without you. I would also like to express my gratitude to my committee members, Dr. Lori Burrows and Dr. Michael Surette for their support and their insightful discussions. I would like to thank Dr. Burrows for her encouragement and for being there for me everytime I needed it, for helping me to get my head out of the weeds during confusing times. Thank you Sara Kilmury, Hanjeon Harvey and Ryan Lamers for all your help. You are amazing!

I would also like to express my gratitude to Dr. Brian Golding not only for adopting me in his lab for a while to analyze the results from the RNA-seq, but also for allowing me to use his own computer everyday that I was there. Thank you to Dr. Golding's graduate student Wilson Sung for helping me day after day.

I would like to thank Michelle Mendonca and Jake Szamosi for helping me with the transcriptome analysis.

I would also like to express my gratitude to Tammy Feher for her support throughout all these years, to my lab mates Charneal Dixon and Sanchia Miller for their help and support during this long journey. Thank you to Laxman Pandey for

his help when I started in the lab and for his contributions to this project and thank you Xiaodan Ni for being there for me.

Last, but not least, I would like to extend my heartfelt gratitude to my family; this research could not have been completed without their continued support, encouragement and love throughout my graduate degree. Thank you also to all my long-time friends. Thank you.

PREFACE

This thesis contains results from four and a half years of research. The main objective of this work was to study the role of outer membrane enzyme PagP in signal transduction. All the experiments were designed by my supervisor Dr. Russell Bishop and myself. The transcriptome analysis in Chapter 2 was carried out in collaboration with Dr. Brian Golding's laboratory, from the Department of Biology at McMaster University and in collaboration with Michelle Mendonca and Jake Szamosi, members of Dr. Surette's laboratory. Chapter 3 was in collaboration with members of the Farncombe Metagenomic Laboratory and members of Dr. Surette's and Dr. Burrows' laboratories. Chapter 4 was in collaboration with Dr. Tracy Raivio and Yun Peng from the University of Alberta.

TABLE OF CONTENTS

Chapter 1: General Introduction

List of Figures

List of tables

List of Abbreviations

ACoA	Acyl Coenzyme A
ACP	Acyl Carrier Protein
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
Amp	Ampicillin
BMP	bis(monoacylglycero)phosphate
CAMPs	Cationic Antimicrobial Peptides
cDNA	Complementary DNA
CRP	cAMP receptor protein
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FHL	Formate Hydrogenlyase Complex
FhIA	Formate DNA binding transcriptional regulator
FNR	Fumarate Nitrate Reductase regulator
Gen	Gentamicin
IHF	Integration host factor
IM	Inner membrane
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
L-Ara4N	L-4-aminoarabinose

LPS	Lipopolysaccharide
MIC	Minimum inhibitory concentration
Mo-bisPGD	Molybdobis (pyranoterin guanine dinucleotide)
ModE	Molybdate-responsive transcription factor
OD	Optical density
OM	Outer membrane
OMPs	Outer membrane proteins
ONPG	2-Nitrophenyl β-D-galactopyranoside
ORI	Origin of replication
PCR	Polymerase Chain Reaction
PG	Phosphatidylglycerol
PPG	Palmitoyl-PG
PtdEtn	Phosphatydylethanolamine
RBS	Ribosomal binding site
RIN	RNA integrity number
RNA	Ribonucleic acid
Tat	Twin arginine translocator
TCA	Tricarboxylic acid cycle
Tet	Tetracycline
TF	Transcription factor
TLC	Thin Layer Chromatography
TLR4	Toll-like receptor 4

Chapter 1. General Introduction

1.1 *Escherichia coli*

Escherichia coli is a Gram-negative rod-shaped bacterium that populates the lower intestine of mammals. In the human host, *E. coli* lives as a commensal species metabolizing important micronutrients and contributing to the digestion of different foods. Although most *E. coli* strains are commensal species, the enterotoxicgenic *E. coli* can infect approximately 200 000 people annually and is responsible for the death of 100 000 people a year (Crofts et al., 2018). The laboratory strain *E. coli* K12 is the most widely studied, and it has been used as a model organism to study gene regulation and cell physiology (Van Elsas et al., 2011).

One interesting aspect of *E. coli* is that in order to colonize the human gut it has to overcome a wide range of environmental threats before it reaches the human intestine. This can include surviving on soil or abiotic surfaces, followed by passages through the mouth, throat, stomach, small intestine and large intestine. *E. coli* has developed a number of cellular systems capable of detecting changes in the environment and adapting through gene expression. Depending on the magnitude of the threat or stimulus adaptation involves changes in gene expression at transcriptional and/or post transcriptional levels (Van Elsas et al., 2011).

1.2 Gram-negative bacterial cell envelope

The unpredictable and hostile environment that bacterial cells are often found in have led to the evolution of a sophisticated cell envelope that confers protection and survival (Grabowicz and Silhavy, 2017; Lima et al., 2013; Ruiz and Silhavy, 2005; Walsh, 2003). Gram-negative bacteria possess a unique cell envelope composed of an inner membrane (IM) and an outer membrane (OM) separated by an aqueous periplasmic space (Bos et al., 2007). The OM is the key distinguishing feature of Gram-negative bacteria and is an asymmetrical lipid bilayer composed of phospholipids in the inner leaflet and lipopolysaccharides (LPS) in the outer leaflet (Bos et al., 2007; Nikaido, 1985) (Figure 1.1).

1.3 LPS biosynthesis and its role in signal transduction

Because of its signal transduction role in the innate immune system, particular interest is placed in studying the LPS as it contains a membrane-anchored hydrophobic lipid A, which determines the permeability barrier of the Gram-negative cell envelope (Raetz et al., 2007). Attached to lipid A are also core oligosaccharides and a distal polysaccharide (O-antigen), which makes up the LPS molecule. Lipid A is a glucosamine-based saccharolipid with a characteristic β-1',6-linked disaccharide backbone that can be modified with L-4-aminoarabinose (L-Ara4N), phosphoethanolamine (pEtN), and palmitate groups by enzymes present in *E. coli* K12 (Raetz et al., 2007).

Lipid A biosynthesis is initiated in the cytosol and the inner leaflet of the IM and is acylated with *R*-3-hydroxymyristate at the 2, 3, 2', and 3'-positions, and

phosphorylated at the 1 and 4'-positions (Figure 1.2). Two units of 3-deoxy-D-*manno*-2-octulosonic acid (Kdo) are attached to the 6'-position, after which acyloxyacyl linkages are formed with laurate and myristate chains at their respective 2' and 3'-positions (Raetz and Whitfield, 2002). The resulting lipid A is hexa-acylated and continues to undergo glycosylation and phosphorylation to complete the core oligosaccharide prior to its translocation to the periplasmic leaflet of the IM (Raetz and Whitfield, 2002). The periplasmic leaflet of the IM is where the O-antigen is attached to the distal sugar of the core oligosaccharide (Raetz and Whitfield, 2002). Additionally, modifications that occur at the periplasmic leaflet of the IM include the incorporation of L-Ara4N and pEtN. The LPS is transported to the OM, where lipid A can be further modified. Modification of lipid A by the OM enzyme PagP occurs in response to perturbations of OM lipid asymmetry, and results in the incorporation of a palmitate chain in acyloxyacyl-linkage at position-2 (Figure 1.2) (Raetz et al., 2007).

Lipid A or endotoxin is the only region of LPS that is recognized by the mammalian innate immune system. LPS recognition occurs through the toll-like receptor 4 (TLR4), which is found in complex with myeloid differentiating factor-2 (MD2) and are together found in dendritic cells and macrophages. TLR4-MD2 recognizes bacterial lipid A in order to activate the inflammatory response (Miller et al, 2005). Hepta-acylated lipid A, which has been palmitoylated by PagP, attenuates the inflammatory response through TLR4-MD2 and provides resistance to host cationic antimicrobial peptides (CAMPs) (Guo et al., 1997; Guo et al.,1998).

LPS and lysophospholipids are also involved in bacterial cell signalling (May and Silhavy, 2018). When the permeability barrier of the OM has been compromised, phospholipids aberrantly migrate into the outer leaflet and occupy space that normally should be occupied by LPS. The hydrolysis of these misplaced phospholipids by the OM phospholipase PldA produces fatty acids that travel across the cell envelope to signal that the OM needs to be repaired and more LPS needs to be synthesized. Acyl-CoAs were recently identified as signaling molecules processed by FadD to inhibit the degradation of LpxC, a critical enzyme of LPS biosynthesis (May and Silhavy, 2018). Defects in LPS modification and lipid A underacylation can also be involved in activation of envelope stress response regulators such as σ^E, Cpx, and Rcs (Lima et al., 2013; Tam and Missiakas, 2005; Grabowicz and Silhavy, 2017; Konovalov et al., 2014). Several genes coding for proteins involved in LPS biosynthesis are transcribed by the *rpoE* regulon, which supports the hypothesis that intermediate molecules from LPS biosynthesis may activate σ^E (Tam and Missiakas, 2005). Lipid A palmitoylation essentially serves as a front-line of defense for Gram-negative bacteria as PagP activity is triggered in response to OM lipid redistribution at the bacterial cell surface (Jia et al., 2004).

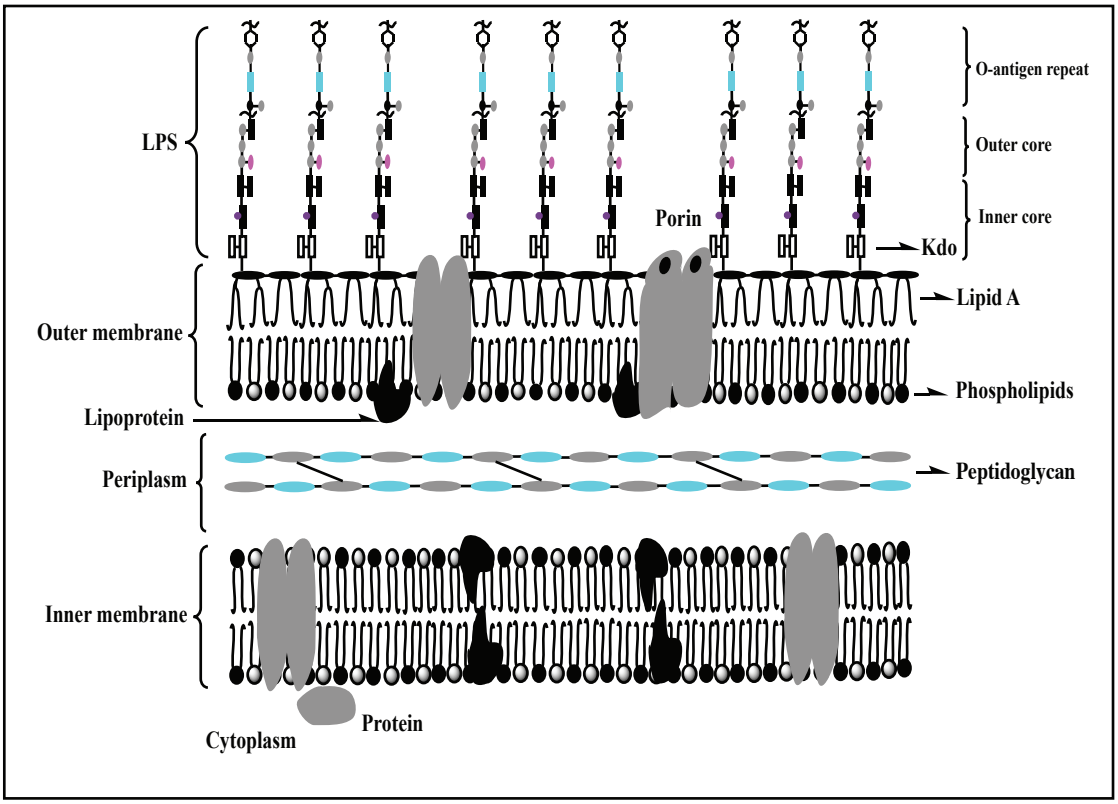

Figure 1.1. Gram-negative bacteria cell envelope. The Gram-negative cell envelope consists of two membranes. The IM is a phospholipid bilayer. Many proteins involved in energy production, lipid biosynthesis, and transport are located in the IM (Silhavy et al., 2010). The OM is an asymmetric bilayer with phospholipids localized on the inner leaflet and lipopolysaccharides (LPS) in the outer leaflet. Lipid A, core oligosaccharides, and O-antigen make up LPS. Lipoproteins and beta barrels can also be found in the OM. The IM and the OM are separated by the periplasm and the peptidoglycan layer which protects the cells against osmotic variations.

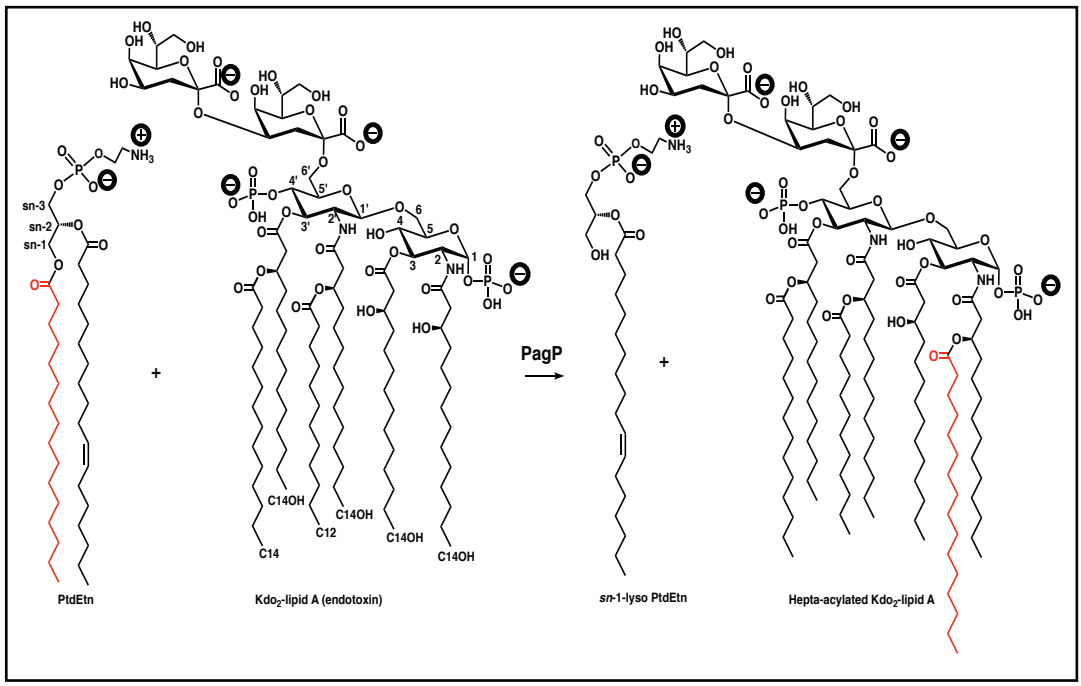

Figure 1.2. Lipid A palmitoylation by PagP. PagP catalyzes the transfer of a palmitate chain (red) from the *sn*-1 position of phosphatidylethanolamine (PtdEtn) to the hydroxyl group of the N-linked R-3-hydroxymyristate chain at the proximal glucosamine unit of lipid A, converting a hexa-acylated lipid A to hepta-acylated and producing lyso-PtdEtn.

1.4 Envelope stress response regulators σ^E, Cpx, and Rcs

Due to the compartmentalization in the cell envelope, Gram-negative bacteria have stress response systems capable of sending signals across the cell envelope to the cytoplasm (Grabowicz and Silhavy, 2017; Lima et al., 2013; Ruiz and Silhavy, 2005; Walsh, 2003). There are several well-known envelope stress response systems that are interconnected in *E. coli* (Grabowicz and Silhavy, 2017), such as the sigma factor σ^E, and two-component phosphorelay systems Cpx and Rcs. The sigma factor σ^E is activated by misfolded OM proteins, accumulation of underacylated lipid A, or defective LPS. Sigma factor σ^E responds to proteolytic steps that degrade an anti-sigma factor, releasing σ^E into the cytoplasm (Grabowicz and Silhavy, 2017) where the σ^E regulon is transcribed (Klein and Raina, 2015). Sigma factor σ^E regulates its own expression and is the first gene in an operon, which also encodes for *rseA*, *rseB* and *rseC* genes. The *rseA* gene encodes a hydrophobic polypeptide, RseA, the anti-sigma factor, which possesses a cytoplasmic N-terminal domain that binds σ^E in the absence of extracytoplasmic stress, and a C-terminal domain, which binds to RseB, a periplasmic protein that senses the concentration of misfolded OM proteins (Ades et al., 1999). Additionally, RseB binds accumulated LPS in the periplasm (Lima et al., 2011; Grabowicz and Silhavy, 2017). Members of the RpoE regulon include genes whose products are folding factors in the periplasm, proteases, OM proteins such as the beta-barrel assembly proteins BamA-E, and genes that encode proteins involved in LPS synthesis and translocation, such as *lptA*, *lptB* and *lptD,* as well as for proteins

involved in phospholipid and lipid A synthesis (Klein et al., 2011). Other members of σ^E regulon include regulatory small RNAs (sRNAs), such as MicA and RybB. RpoE along with sRNAs have the ability of generate glycoforms with LPS truncations in the outer core, which interfere in the incorporation of the O-antigen (Klein and Raina, 2015).

The Cpx two-component system responds to numerous signals in the periplasm that can be identified by the CpxA sensory domain. Cpx is activated by misfolded IM proteins and alterations in protein translocation (Grabowicz and Silhavy., 2017). The σ^E, Cpx, and Rcs systems are interconnected by regulatory sRNAs such as RprA, MicA and MicF (Grabowicz and Silhavy., 2017). The regulation mediated by sRNAs often involves feedforward and feedback circuits that control the transcription of genes in complex biological processes, such as the regulation of RpoS by RprA (Klein and Raina., 2015). The σ^E and the Cpx systems overlap because Cpx downregulates σ^E and σ^E-dependent genes (Grabowicz and Silhavy., 2017).

The Rcs phosphorelay system is more complex than prototypical two-component systems; this phosphorelay system controls the transcription of genes related to capsular polysaccharides (Majdalani and Gottesman, 2005). The Rcs system consist of RcsC, the sensor kinase, RcsB, the response regulator, and RcsD, needed in the phosphorelay as it contains an HPt domain (Huang et al., 2006). The Rcs system also possesses an OM component, RcsF, which forms complexes with β barrel proteins. RcsF is a lipoprotein that plays a role in signaling by monitoring

levels of LPS and interactions among LPS molecules (Konovalova et al., 2014). The Rcs is involved in the expression of genes related to cell division, of extracellular polysaccharides, the formation of biofilms, and it also controls the synthesis of RprA RNA, which is a strong post transcriptional regulator of RpoS (Huang et al., 2006). These three systems σ^E, Cpx, and Rcs, are interconnected by the regulatory sRNAs RprA, MicA and MicF (Grabowicz and Silhavy., 2017).

1.5 The two-component system PhoPQ

The PhoPQ two-component system consists of an IM-bound sensor kinase PhoQ and cytoplasmic transcriptional regulator PhoP. The system responds to levels of Mg^{+2} and Ca^{+2} (Groisman, 2001), and it plays a role in Gram-negative virulence by resisting CAMPs (Miller et al., 2005). The PhoPQ system regulates expression of genes involved in modification of cell surface structures including lipid A. In response to environmental stress the autophosphorylation of PhoQ leads to phosphorylation of cytosolic regulator PhoP, triggering the transcription of numerous genes (Groisman, 2001). The PhoPQ system interacts with other two-component systems to give a more general response within the bacterial cell. A number of PhoP-dependent genes in *Salmonella* are regulated by *pmrAB,* which is induced under mildly acidic conditions or in high concentrations of Fe^{+3}. However, *pmrAB*-dependent genes in *E. coli* are independent of PhoPQ. *E. coli* PhoPQ is controlled by EvgAS (Eguchi et al., 2004).

1.6 The two-component system PmrAB.

The PmrAB two-component system in *E. coli* consists of a membrane sensor kinase PmrA, which transfers a phosphoryl group to the response regulator PmrB. This system is sometimes referred as BasSR (Ogasawara et al., 2012). PmrAB also plays an important role in the regulation of stress response regulators such as *elbA*, an RpoS control factor, and also is known to regulate the toxin-antitoxin pair YafQ/DinJ (Ogasawara et al., 2012). In addition, PmrAB controls the *pmrHFIJKLM* (*arnBCADTEF*) operon involved in lipid A modification (Hagiwara et al., 2004; Yan et al., 2007).

1.7 Transcriptional regulation in *E. coli* K12

Bacteria possess an efficient system to express proteins whenever they are needed in response to different types of stress. This allows bacteria to survive in limited nutrient sources, metabolize the most energy-efficient carbon source available, survive in aerobic and anaerobic conditions, manage moderate DNA damage, and also restore its OM. They do so by regulating gene expression to restore optimal conditions in the bacterial cell with the assistance of transcription factors (TFs) and the integration of these TFs to more complex networks of operons (Gottesman, 1984). Operons are scattered in the bacterial genome and they consist of several genes, that can participate in the same or different functions and in different metabolic pathways. Regulons consist of more than one operon and they depend on global regulators (Gottesman, 1984). There are transcriptional regulators such as CRP, FNR, IHF, FIS, ArcA, NarL and Lrp, which can regulate

approximately 51% of the genes in *E. coli.* However, there are cases in which the regulation is done by several TFs, which occurs in approximately 49% of the genes (Martinez-Antonio and Collado-Vides., 2003).

Bacteria monitor intracellular and extracellular conditions to restore the optimal physicochemical conditions by differential gene expression. A bacterial cell will respond to a stimulus or signal by inducing or repressing the expression of different genes. Sometimes the messenger can be a small molecule that can be rapidly synthesized and degraded or it could be an intermediate metabolite from a biosynthetic pathway (Martinez-Antonio and Collado-Vides., 2003). A reduced growth rate, high cell density, low temperature, high osmolarity, carbon, nitrogen or amino acids starvation, can trigger the expression of sigma factors such as σ^S, a sigma subunit of RNA polymerase with the consequent expression of numerous genes, which help the bacterial cells adapt to the new conditions (Hengge-Aronis R., 2002). The global regulators and the TFs upregulate or downregulate the expression of certain genes in response to the concentration of certain metabolites, different levels of divalent cations, absence of oxygen, etc. The global regulators respond to a larger number of signals, whereas the local TFs are more specific (Martinez-Antonio and Collado-Vides., 2003).

The genes in prokaryotic genomes are mostly organized in operons with polycistronic mRNA, which allows the co-transcription of several genes. The transcriptional control of an operon depends on a promoter binding site of RNA polymerase and may include one or more operators (binding sites for transcriptional

factors), which modulate the activity or binding of the RNA polymerase. (Keseler et al., 2011). There are two main mechanisms of gene regulation in bacteria: transcriptional regulation and post-transcriptional regulation. Transcriptional regulation works by allowing or preventing the transcription of certain genes by physical interaction between the TF, and the RNA polymerase at the promoter of the gene in question. Numerous TFs are members of two-component systems including a response regulator that depends on a cognate sensor kinase and examples include the aforementioned PhoPQ, PmrAB, EvgAS, CpxAB, and Rcs. There are approximately 37 two-component systems in *E. coli*. They generally possess a periplasmic domain, which senses the stress in the envelope, and an inner membrane domain, which responds to the signal and transmits it to the cytoplasm (Ruiz and Silhavy., 2005).

1.8 Post-transcriptional regulation and sRNAs

Post-transcriptional regulation involves a regulator, which could enhance or inhibit the translation of an mRNA into protein. This can be achieved by a regulatory RNA structure embeded within the gene that is being regulated, or by small non-coding RNAs, which can interact with the target mRNA by base pairing (Picard et al., 2009). This regulation also depends on the mRNA and protein stability, mRNA degradation rate, riboswitches, sRNA, toxin-antitoxin systems, and ribosome activity (Picard et al., 2009).

1.9 The stationary phase global stress regulator RpoS

Gram-negative bacteria show an exponential growth in the presence of sufficient nutrients. However, this exponential growth decreases when the nutrients are depleted, and this stage is called stationary phase where the cells develop resistance to starvation in order to survive. The starvation response triggers the activation of the sigma factor, σ^s, also called RpoS (Yamada et al., 1999). Stationary phase is comparable to what occurs under natural conditions where bacterial cells have to coexist with numerous microorganisms. In these cases, numerous genes are expressed to help the cells to adapt (Yamada et al., 1999). The master regulator of gene expression in stationary phase, RpoS controls numerous genes that respond to starvation and stress responses. Regulation of RpoS-dependent genes is frequently mediated by other regulatory proteins such as FNR, OxyR, and Fur (Lacour and Landini., 2004). RpoS, not only regulates genes in the stationary phase, but also negatively regulates genes in the exponential phase such as FNR (Patten et al., 2004). In spite of the low levels of RpoS in exponential phase, it plays an important role in regulating genes whose products are enzymes from the tricarboxylic acid cycle, genes that encode proteins involved in iron and molybdenum uptake, and phosphotransferases (Dong et al., 2008; Rahman et al., 2006).

1.10 The Fumarate Nitrate Regulatory Protein FNR

The DNA-binding protein FNR, along with two-component system ArcAB, are major regulators of the anaerobic switch, which responds to the absence of

oxygen in the bacterial cell (Gunsalus and Park., 1994). An O_2-sensitive [4Fe-4S] cluster within FNR directly senses oxygen availability in the cell and regulates FNR activation (Kiley et al., 2003). FNR controls 297 genes in *E. coli* (Kang et al., 2005). FNR also represses the expression of genes that encode enzymes involved in aerobic respiration, such as NADH dehydrogenase and other oxidases, and it positively regulates enzymes involved in anaerobic oxidation of carbon sources such as glycerol and formate coupled to anaerobic reduction of terminal electron acceptors such nitrate, fumarate and dimethyl sulfoxide (Kang et al., 2005).

1.11 FNR and RpoS feedback regulation

There is a complex relationship between FNR and the master regulator RpoS, which can affect FNR and/or RpoS-dependent genes. Under anaerobic conditions, FNR downregulates the master regulator RpoS, which decreases the expression of RpoS-dependent genes (Kang et al., 2005). Under aerobic conditions, RpoS represses expression of FNR, which negatively regulates some FNR-dependent genes (Patten et al., 2004).

1.12 OM enzyme PagP

PagP is among the few integral OM enzymes and the first OM enzyme to be involved in lipid A modification (Raetz and Whitfield, 2002). PagP transfers a palmitate chain from the *sn*-1 position of a phospholipid to the hydroxyl group of the *N*-linked *R*-3-hydroxymyristate chain at the proximal glucosamine unit of lipid A, thereby converting hexa-acylated lipid A to a hepta-acylated form (Figure 1.2) (Guo et al., 1997; Raetz and Whitfield, 2002; Raetz, 2007; Bishop, 2005). The

structure of PagP comprises an eight-stranded β-barrel and an amphipathic α-helix with the active site located at the cell surface (Ahn et al., 2004) (Figure 1.3). Serine 77 plays an important role in the catalytic activity. The β-barrel can be divided into a lower and an upper half. The lower half is accommodated by the inner leaflet of the OM, and consists of polar residues that line its hydrophilic core (Ahn et al., 2004)). The upper hydrophobic half is where the lipid binding pocket, otherwise known as the hydrocarbon ruler, can be found. The depth of the hydrocarbon ruler determines the length of the selected acyl chain (Ahn et al., 2004).

PagP can be found in a dormant or an activated state. It remains in its dormant state unless the OM experiences perturbations. Under exogenous envelope stress conditions, which results in the translocation of phospholipids from the inner to the outer leaflet, lipid A palmitoylation occurs (Jia et al., 2004). The migration of phospholipids into the outer leaflet of the OM serve as points of access for lipophilic antibiotics and detergents (Figure 1.4). The migration of phospholipids into the outer leaflet allows access to the lipid binding pocket of PagP (Ahn et al., 2004; Hwang et al., 2004; Cuesta-Seijo, 2010). The *pagP* gene, also known as *crcA* in *E. coli* K12, is regulated by PhoP. During biofilm formation *pagP* is upregulated by SlyA and downregulated by HN-S (Chalabaev et al., 2014). These two regulatory proteins, SlyA and HN-S, are also regulated by the Rcs phosphorelay system, and it has been demonstrated that overexpression of RcsB increases the expression of SlyA (Navasa et al., 2013), which in turn can increase expression of *pagP*.

The gene *pagP* and its adjacent downstream gene *cspE* overlap for 22

nucleotides. The gene *cspE*, whose product is a soluble cytoplasmic cold shock protein, belongs to the CspA family of nine homologues (from CspA to CspI). CspE is constituvely expressed and functions as an antiterminator through its RNA binding capabilities (Johnston et al., 2006), and it has also been reported as a reporter to activate RpoS translation (Hu et al., 2012).

In *E. coli* O157:H7, it has been shown that PagP can be activated by the deletion of *msbB* encoding the myristoyltransferase MsbB (Smith et al., 2008). The *msbB*-deficient mutants demonstrated an insufficient supply of the myristate chain on the distal glucosamine unit of lipid A, leading to a defect in OM lipid asymmetry (Smith et al., 2008). This perturbation to the OM leads to activation of PagP demonstrated by lipid A palmitoylation and initiation of signal transduction, which effectively exerts negative control on cytoplasmic enzymes of R3 core biosynthesis resulting in core truncation (Smith et al., 2008). These findings were the first to suggest that PagP might function as an apical sensory transducer that is sensitive to alterations in the OM permeability barrier (Smith et al., 2008).

Considering that *E. coli* O157:H7 is a pathogenic strain, we shifted our focus to the laboratory strain *E. coli* K12. However, a lipid A myristoylation mutant of *E. coli* K12 does not generate the same effect as observed in *E. coli* O157:H7 (Smith et al., 2008). The creation of *E. coli lptD4213* strains, therefore, provided an alternative means of OM permeabilization for the activation of PagP in a non-pathogenic model organism.

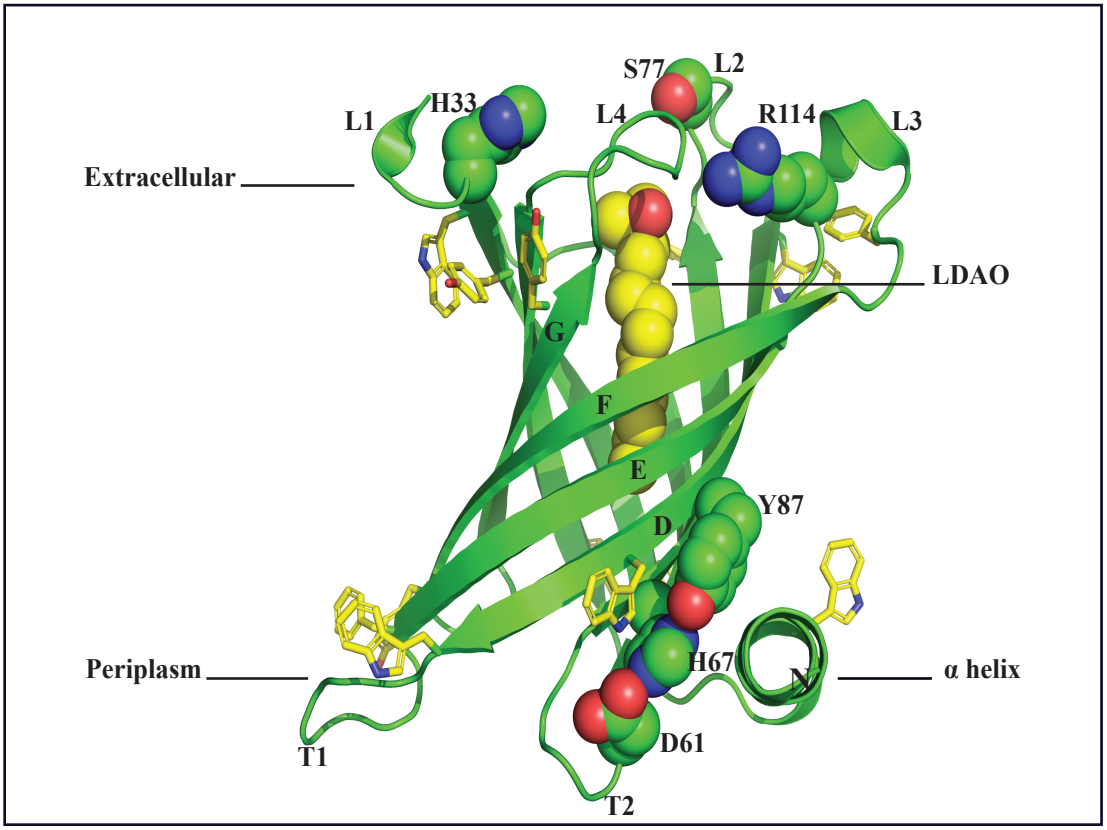

Figure 1.3. PagP Structure. The crystal structure of PagP (pdb:1THQ) shows an eight-stranded β-barrel preceded by an amphipathic N-terminal (N) α-helix. The PagP barrel axis is tilted by ~25º with respect to the membrane normal. The lipid binding pocket revealed by a crystallographic LDAO detergent molecule is shown in yellow. Aromatic belt residues (yellow) are localized at the membrane boundaries. The amino acid residues involved in PagP catalytic activity in the extracellular active site (H33, S77 and R114) are represented by spheres. The amino acid residues (Y87, H67 and D61) involved in PagP putative catalytic active site are also shown as spheres.

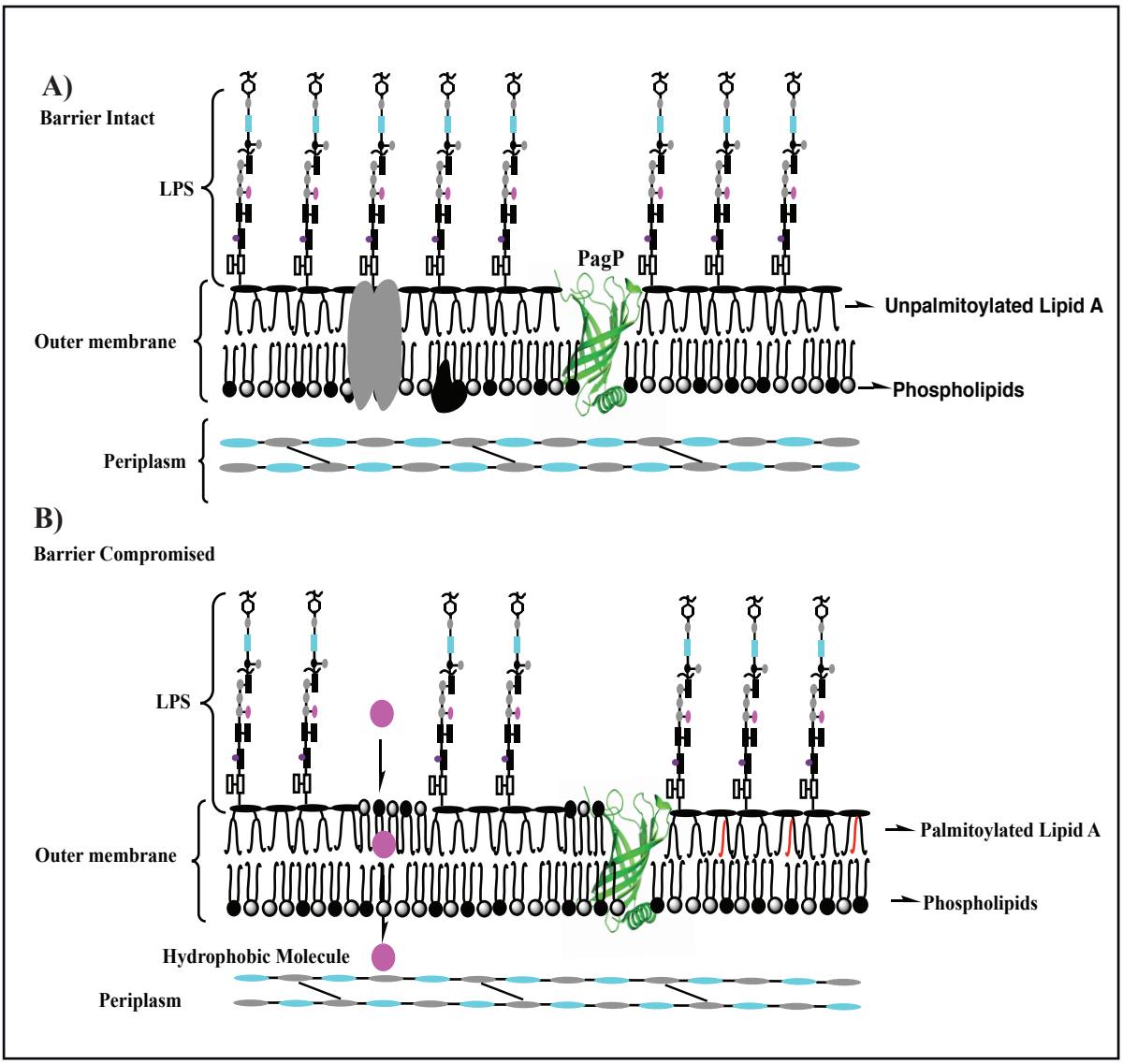

Figure 1.4. Perturbations in the OM and PagP. A) PagP remains dormant unless there is a breach in OM lipid asymmetry, which can result when Mg^{2+} is stripped off, causing LPS molecules to repel each other. B) The migration of phospholipids from the inner leaflet to the outer leaflet serves as a point of access for lipophilic antibiotics and detergents. This compromise in the OM permeability barrier activates PagP, which palmitoylates lipid A as a way to restore lipid asymmetry.

1.13 LptD

The *lptD* gene (formerly known as *imp* or *ostA)* encodes a large OM protein that is one of 7 essential proteins, termed LptA-G (for LPS transport), which transports LPS from the outer leaflet of the IM to the cell surface (Bishop, 2014). LptD forms a complex with LptE in the OM that functions in the final stages of LPS assembly, and consists of two main domains, an N-terminal β-jellyroll and a larger C-terminal β-barrel (Bishop, 2014). LptD expression is σ^E-dependent (Dartigalongue et al., 2001; Braun and Silhavy, 2002). The *lptD4213* mutation involves a 23-amino-acid deletion within the L4 loop, which reduces the rate of LPS transport and renders cells vulnerable to lipophilic antibiotics and detergents that a continuous Mg^{2+}-bridged sheet of LPS would be able to repel (Bishop, 2014). Similar to *msbB*-deficient *E. coli* O157:H7, *E. coli* K12 possessing the *lptD4213* allele has a severe OM defect that triggers lipid A palmitoylation by PagP (Wu et al., 2006; Malinverni and Silhavy, 2009).

On the periplasmic side of PagP there is a cluster of amino acid residues, D61, H67 and Y87, believed to be involved in PagP cell signalling (Figure 1.5). According to its similarities with the chymotrypsin active site, this putative catalytic triad in PagP was initially suspected to be involved in proteolytic cleavage of a lipoprotein that could be released into the periplasm, thus initiating PagP signal transduction. Chymotrypsin possesses a catalytic serine residue, which acts as nucleophile (Stroud et al., 1972). The putative catalytic triad in PagP possesses a tyrosine residue instead of serine, but the hydroxyl group of tyrosine was expected

to function as nucleophile, so a mutant of Y87F was prepared to test the role of tyrosine 87 in cell signalling.

The amino acid residue serine 77, located in the L2 loop of PagP, plays an important role in lipid A palmitoylation. Mutation S77A did not show any palmitoylation of lipid A or of phosphatidylglycerol (PG) (Dalebroux et al., 2014). Therefore, S77A is a catalytically inactive mutant serving as a control to test the signalling properties of PagP. Phenotypically, S77A PagP is indistinguishable from the Δ*pagP* allele, but the latter has the potential to exert a polar effect on the downstream *cspE* gene (Hu et al., 2012). However, the Keio collection of gene deletions in *E. coli* led to the construction of a Δ*pagP* allele that does not extend to the end of *pagP* and therefore does not include the 22 nucleotides overlapping with *cspE* (Baba et al., 2006).

1.14 Research application

E. coli K12 has been selected as a model organism to study the structure and function of the proteins and metabolic pathways in Gram-negative bacteria. The cell envelope characteristic of Gram-negative bacteria is the first barrier to a hostile environment. One defense mechanism by which bacteria become resistant to antibiotics involves membrane lipid modification. It has been demonstrated that the outer membrane enzyme PagP participates in membrane lipid modification, palmitoylating lipid A (Bishop, 2005) and phosphatidylglycerol (Dalebroux et al., 2014). Smith et al., 2008, demonstrated that PagP is involved in signal transduction

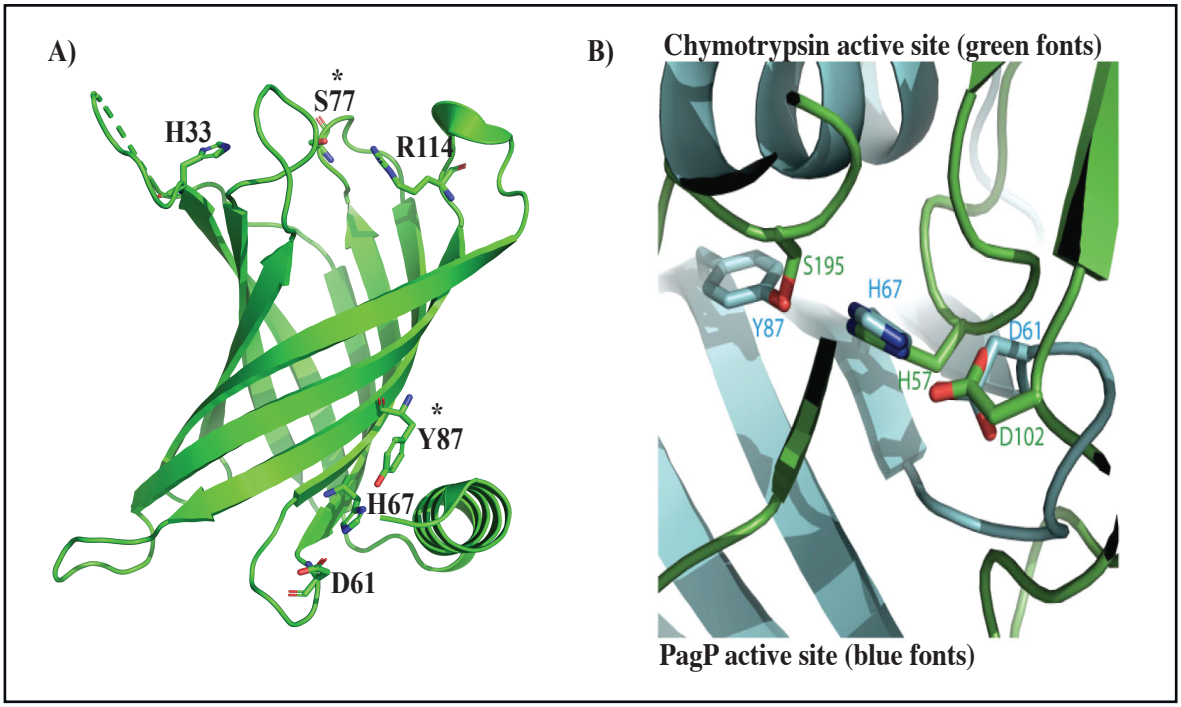

Figure 1.5. PagP's putative catalytic triad and Chymotrypsin active site. A) The putative catalytic triad in the periplasmic site of PagP consist of amino acids D61, H67 and Y87. The latest expected to function as a nucleophile, similar to S195 in the Chymotrypsin active site. B) Superposition of PagP's putative periplasmic active site and Chymotrypsin's active site.

and, additionally, is suspected to have influence on other enzymes that play important roles in the bacterial cell. It is known that outer membrane proteins are key molecules that interface between the cell and the environment, have diverse roles in transport and metabolism, and function as receptors or signal transducers.

1.15 Objectives

Previous results suggest that OM enzyme PagP might be involved in cell signaling. PagP possesses one extracellular active site and one putative active site on its periplasmic side. One objective of my thesis is to test the hypothesis that when PagP activity is triggered in the outer membrane by perturbation of lipid asymmetry, it functions as a sensory transducer to communicate outer membrane disruption to the gene expression machinery in the cytoplasm. Another objective of this thesis is to determine what active site is involved in signal transduction mechanism. Additionally, I aimed to determine the mechanism by which PagP exerts transcriptional regulation on genes encoding proteins from different compartments of the bacterial cell and to elucidate the mediators involved in this transcriptional regulation; the findings of these investigations are included in chapters 2, 3 and 4. While this work was in progress, unpublished discoveries in our lab have demonstrated that PagP can expand a family of novel glycerophosphoglycerol phospholipids in a process that depends on the PagP Y87F mutation. I speculate in chapter 5 how these novel lipids might function as second messengers in a PagP sensory transduction pathway.

Chapter 2. Investigating the role of PagP in OM signal transduction using RNA-seq gene expression analysis

2.1 Preface

The research in this chapter was carried out from 2014 to 2015. Experiments were designed by my supervisor Dr. Russell Bishop and myself. Some of the strains employed in this study were first described in the M.Sc. thesis of Laxman Pandey. I carried out the RNA isolation and the RNA-seq. The transcriptome analysis was performed with the collaboration of Dr. Brian Golding's laboratory, in the Department of Biology at McMaster University. Michelle Mendonca and Jake Szamosi, members of Dr. Surette's laboratory helped with the transcriptome analysis.

2.2 Abstract

The OM enzyme PagP possesses an extracellular catalytic active site mainly involved in lipid A palmitoylation. It also possesses a cluster of amino acid residues D61, H67 and Y87, which are believed to be involved in PagP cell signalling. To test whether or not the Y87 residue is involved in signal transduction, we carried out an RNA-sequencing study of differential gene expression by comparing *E. coli* strains with the *lptD4213* mutation, in which PagP is constituvely activated, with wild-type and Y87F PagP strains, NR760*λp* and NR760*λpY87F*, respectively. Results showed that 50 genes were differentially expressed; 40 were upregulated and encode proteins whose functions are related to processes that occur under anaerobic conditions, and 10 genes were downregulated and encode proteins whose

functions are related to processes that occur under aerobic conditions. 26 genes have FNR as a common transcription factor regulating their expression. Four genes, *yhjY*, *ssnA*, *yfcE* and *yfcF*, are regulated by global response regulator RpoS, and the rest of the genes are regulated by CpxAR or by interconnected effects of σ^E, Cpx, and Rcs. Since the *lptD4213* mutation is known to cause insufficient delivery of LPS to the cell surface, which perturbs OM lipid asymmetry and activates PagP, we suggest that a lipid product derived from the periplasmic active site controlled by residue Y87 might be implicated as a second messenger in sensory transduction from the OM.

2.3 Introduction

Previous findings demonstrated that OM enzyme PagP might have a role in signal transduction in *E. coli* O157:H7. Results showed that PagP can be activated by the deletion of *msbB* encoding the myristoyltransferase MsbB (Smith et al., 2008). The *msbB*-deficient mutants demonstrated an insufficient supply of the myristate chain on the distal glucosamine unit of lipid A, leading to a defect in OM lipid asymmetry (Smith et al., 2008). This perturbation to the OM led to activation of PagP as demonstrated by lipid A palmitoylation and initiation of signal transduction, which effectively exerted negative control on cytoplasmic enzymes of R3 core biosynthesis resulting in core truncation (Smith et al., 2008). These findings were the first to suggest that PagP might function as an apical sensory transducer that is sensitive to alterations in the OM permeability barrier (Smith et al., 2008).

Here, the model organism was shifted to *E. coli* K12 MC4100. However, a lipid A myristoylation mutant of *E. coli* K12 did not generate the same effect as observed in *E. coli* O157:H7 (Smith et al., 2008). The creation of *E. coli lptD4213* strains provided an alternative means of OM permeabilization for the activation of PagP in a non-pathogenic model organism. The *lptD4213* mutation involves a 23-amino-acid deletion within the L4 loop, which causes insufficient LPS delivered to the cell surface, thus compromising the OM permeability barrier.

On the periplasmic side of PagP there is a cluster of amino acid residues, D61, H67 and Y87, believed to be involved in PagP cell signalling (Figure 1.5). According to its similarities with the chymotrypsin active site it was initially thought that this putative catalytic triad was involved in the proteolytic cleavage of a lipoprotein that could be released into the periplasm, thus initiating PagP signal transduction. Chymotrypsin possesses a catalytic serine residue, which acts as nucleophile (Stroud et al., 1972). The putative catalytic triad of PagP possesses a tyrosine residue instead of serine. The hydroxyl group of tyrosine was expected to function as nucleophile, so a mutant Y87F was prepared to test the role of this tyrosine in cell signalling. Another mutant was prepared with the mutation S77A, which affects the extracellular active site. We first carried out RNA-seq to study differential gene expression by comparing the *lptD4213* strains NR760*λp* and NR760*λpY87F.*

To study differential gene expression at a global level in *E. coli* K12 we used RNA-seq. Illumina, a high-throughput RNA-seq analysis is a sequence-based

method that allows whole transcriptome analysis of differentially expressed genes under various conditions. This unbiased technology allowed us to detect even weakly expressed genes in order to understand gene regulation at a global level (Marguerat and Bahler, 2010). Illumina requires RNA isolation and conversion to a library of cDNA, in which each nucleotide is sequenced to obtain short sequences read. The reads are aligned to a reference genome or transcriptome, analyzed and graphed (Wang et al., 2010). Read counts are an approximation of the transcript abundance at a particular growth stage or physiological condition (Anders and Huber, 2010).

2.4 Materials and Methods

To convert *E. coli* NR754 (MC4100, *ara*$^{+}$) and *E. coli* NR760 (NR754*lptD4213*) to NR754Δ*p* and NR760Δ*p*, respectively, the *pagP* deletion from the Keio collection was transduced (Baba et al., 2006). The *pagP* gene was then subcloned from pACPagP, pACS77A, and pACY87F into pBad18, thus generating plasmids suitable for use in the lambda InCh system to insert genes into the chromosomal *attB* locus (Boyd et al., 2000). The resulting strains were provided by Laxman Pandey (Pandey, 2013). The bacterial strains, and their abbreviated names, are listed in Table 2.1.

Table 2.1 *E. coli* strains used in this project.

Abbreviated name	Strain name	Description and genotype	Source
	NR754	MC4100 *ara*$^+$	(Ruiz et al., 2008)
	NR760	MC4100 *ara*$^+$ *lptD4213*	(Ruiz et al., 2005)
	NR754 *ΔpagP*	MC4100 *ara*$^+$*ΔpagP*	(Pandey, 2013)
	NR760 *ΔpagP*	MC4100 *ara*$^+$ *lptD4213 ΔpagP*	(Pandey, 2013)
NR754*λp*	NR754 *ΔpagP λInChpagP*	MC4100 *ara*$^+$ *ΔpagP λInChpagP*	(Pandey, 2013)
NR754*λpY87F*	NR754 *ΔpagP λInChY87F*	MC4100 *ara*$^+$ *ΔpagP λInChY87F*	(Pandey, 2013)
NR754*λpS77A*	NR754 *ΔpagP λInChS77A*	MC4100 *ara*$^+$ *ΔpagP λInChS77A*	(Pandey, 2013)
NR754*λpBAD 18*	NR754 *ΔpagP λInChBAD18*	MC4100 *ara*$^+$ *ΔpagP λInChBAD18*	(Pandey, 2013)
NR760*λp*	NR760 *ΔpagP λInChpagP*	MC4100 *ara*$^+$ *lptD4213 ΔpagP λInChpagP*	(Pandey, 2013)
NR760*λpY87F*	NR760 *ΔpagP λInChY87F*	MC4100 *ara*$^+$ *lptD4213 ΔpagP λInChY87F*	(Pandey, 2013)
NR760*λpS77A*	NR760 *ΔpagP λInChS77A*	MC4100 *ara*$^+$ *lptD4213 ΔpagP λInChS77A*	(Pandey, 2013)
NR760*λpBAD 18*	NR760 *ΔpagP λInChBAD18*	MC4100 *ara*$^+$ *lptD4213 ΔpagP λInChBAD18*	(Pandey, 2013)
DH5α*λpir*		80*dlacZ* M15 (*lacZYA-argF*)*U196 recA*1 *hsdR17 deoR thi*-1 *supE*44 *gyrA*96 *relA*1/*pir*	(Miller and Mekalanos, 1988)

2.5 RNA extraction protocol

E. coli strains were streaked onto Luria Bertani (LB) agar plates and single colonies were inoculated into 5 ml of LB media and grown overnight. Overnight cultures were subcultured (1:100) and grown at 37 °C to mid-log phase with an optical density (OD 600nm) between 0.6 - 0.8. 1 ml of subculture was centrifuged at 12, 000 xg for 3 minutes, and the supernatant was discarded. The pellet was resuspended in 700 µl of RNase free water with 50 µl of 100 mg/ml lysozyme. The mixture was incubated at 37 °C for 45 minutes and centrifuged at 12,000 xg at 4 °C for 15 minutes. The pellet was resuspended in 1 ml of trizol and vortexed at high speed for 10-20 seconds, then the mixture was incubated at room temperature for 5 minutes. 200 µl of chloroform was then added to the mixture and vortexed briefly. After 3 minutes of incubation, the mixture was centrifuged at 12, 000 xg for 20 minutes at 4 °C and 400 ul of the aqueous phase was extracted without touching the other phase.

After extracting the aqueous phase, one volume of 70% ethanol was added to one volume of aqueous phase and mixed. 800 µl of the mixture was loaded onto the column for RNA purification from the RNAeasy Mini Kit (Qiagen, Venlo, Netherlands) following the manufacturer's instructions. The samples were stored at -80 °C. The RNA concentrations were determined with a NanoDrop and the RNA integrity (RIN) was determined using an Agilent BioAnalyzer 2100. The mRNA enrichment and the cDNA were carried out using Ribo-zero (Epicentre, Madison, WI, USA). and Omnicript RT kit (Qiagen, Venlo, Netherlands), respectively. RIN,

dsDNA, and RNA-seq were carried out in the Farncombe Genomics Facility (McMaster University, Hamilton, Canada). The RNA-seq was carried out using Illumina Hiseq1000). The RNA-seq data was aligned, assembled and analyzed with the software DESeq2, following the established protocol established (Love et al., 2014). The genes were classified into pathways based on Ecocyc data base (Keseler et al., 2011). Approximately 16 million reads were obtained per sample per biological replicate which were mapped back to the genome of *E. coli* k12.

2.6 Results.

Previous studies (Smith et al., 2007) suggested that PagP has a signal transduction role in response to membrane perturbation. This activity is postulated to be mediated by an active site independent of the extracellular active site involved in lipid A palmitoylation. In order to investigate this, I sought to determine whether or not there was differential gene expression in response to membrane perturbation depending on the particular active site.

Results from the RIN using the Agilent Bioanalyzer showed the high quality of the RNA in all the samples examined (Figure 2.1). The RIN was measured from 1 to 10, with 1 representing completely degraded RNA and 10 the highest quality. The amount of RNA is separated according to their molecular weight and detected using fluorescence. The results are visualized as electropherograms in which the amount of fluorescence correlates to a certain amount and size of RNA (Schroeder et al., 2006).

The analysis of the RNA-seq showed that 50 genes were differentially expressed when we compared the strains NR760*λp* and NR760*λpY87F*. 10 were downregulated and 40 genes were up regulated. The genes considered differentially expressed were those with a Log2FoldChange < -2 and > +2 and an adjusted p-value <0.05 (Figure 2.2). The MA-plot is a visual representation of the genomic data with M representing the log ratio of level counts for each gene between two samples and A represents the mean average (Love et al., 2014). The majority of the upregulated genes encode proteins whose function are related to processes that occur under anaerobic conditions and the downregulated genes encoded proteins whose functions are related processes that occur under aerobic conditions. Numerous genes belonged to operons or families (Tables 2.2 and 2.3).

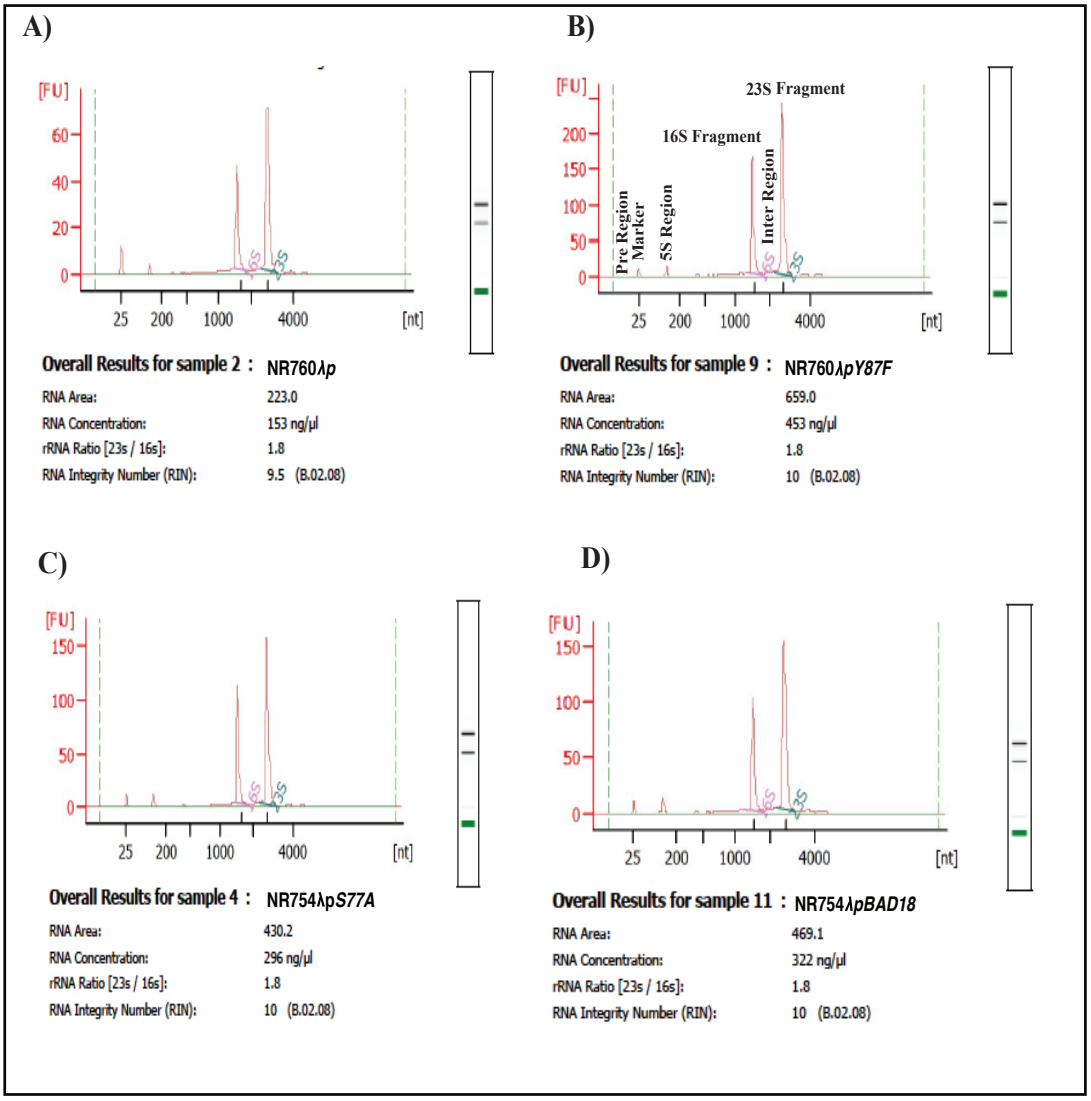

Figure 2.1. Overview of RIN electropherograms using the Agilent 2100 for the RNA-seq samples. This figure is representative of one of the three replicates. The RIN was measured from 1 to 10, with 1 representing completely degraded RNA and 10 representing RNA of the highest quality. The RNA extracted from strain NR760*λpY87F* (Chart B) was selected as a representation of the different regions that indicate the highest RNA quality. The chart plots fluorescence in the Y axis and nucleotides in the X axis and it shows the pre region, marker, 5S region, 16S fragment, inter region and 23S fragment. The Agilent software automatically calculates the 23s/16s ratio and the RNA area.

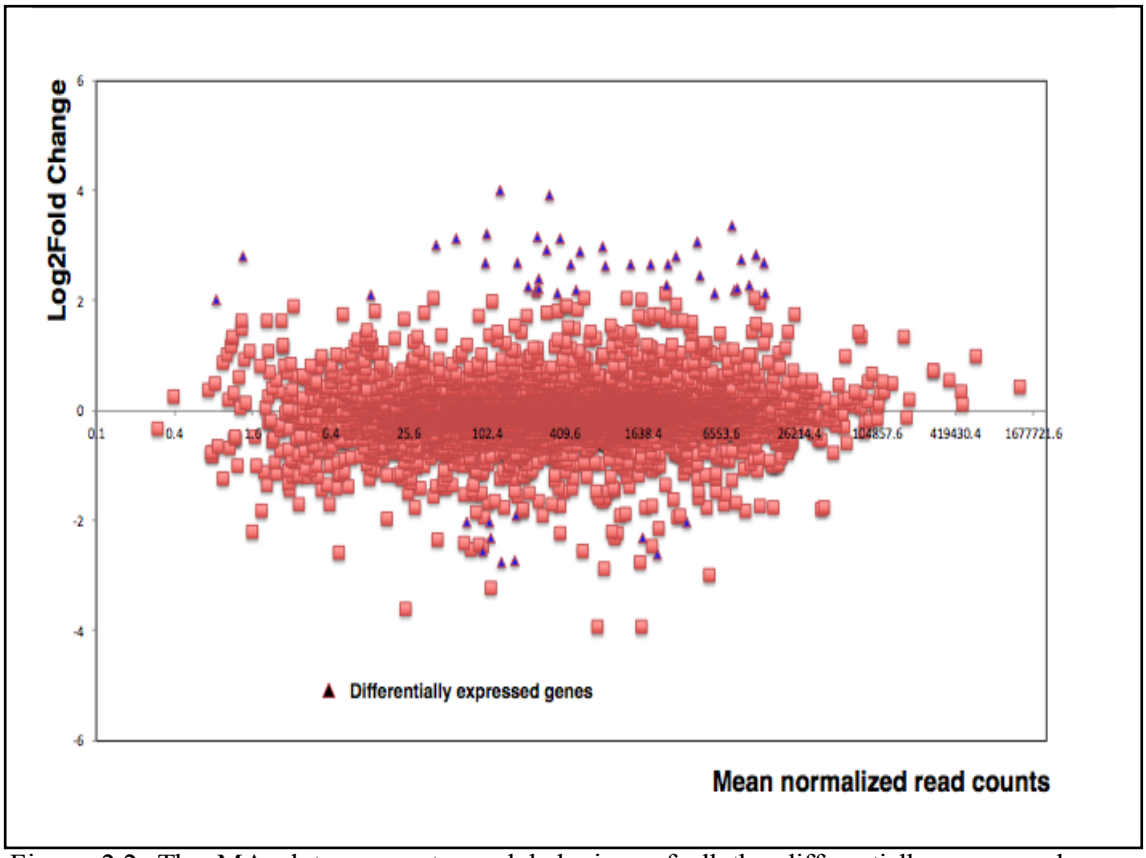

Figure 2.2. The MA-plot represents a global view of all the differentially expressed genes between two samples, strains NR760*λp* and NR760*λpY87F*. The chart plots the Log2Fold Change over the mean read count expression. The majority of the values scatter around cero. Fifty genes were differentially expressed with a Log2FoldChange <-2 and >+2 an adjusted p value <0.05. Of those genes 40 were upregulated and 10 were downregulated.

Table 2.2 Downregulated genes. Comparison of NR760*λp* (WT) and NR760*λpY87F* (Y87F). Based on RNA-seq data. Genes selected have a Log2FoldChange <-2 and >+2 and a padjust ≤0.05. Genes marked with an asterisk represent those that have FNR as common regulator.

Gene name	Accession	Gene Product/Function	Log2FoldChange
*sdhB**	b0724	Succinate dehydrogenase Fe-S subunit.	-2.04
cysE	b3607	Serine acetyltransferase.	-2.16
lldD	b3605	L-lactate dehydrogenase.	-2.64
ansP	b1453	L-asparagine permease.	-2.35
efeB/ycdB	b1019	Dimeric protein that contains a heme cofactor.	-2.75
efeO/ycdO	b1018	Monomeric membrane associated protein.	-2.75
ybdA	b0591	Putative inner membrane sensory kinase.	-2.02
yghK	b2975	Glycolate permease	-2.05
*fepE**	b0587	Polysaccharide co-polymerase family.	-2.36
paaI	b1396	Catalysis of aerobic phenyl acetic acid degradation.	-2.60

Table 2.3. Upregulated genes. Comparison of NR760*λp* (WT) and NR760*λpY87F* (Y87F). Based on RNA-seq data. Genes selected have a Log2FoldChange <-2 and >+2 and a padjust ≤0.05. The genes were grouped by operons and/or families. Genes marked with an asterisk represent those with FNR as common regulator.

Gene name	Accession	Gene Product/Function	Log2FoldChange
*tdcA**	b3118	TdcA transcriptional activator of the *tdcABC* operon in *Escherichia coli*.	2.69
*tdcD**	b3115	Propionate kinase acetate kinase anaerobic.	2.23
*tdcE**	b3114	Formate acetyltransferase	2.00
*tdcC**	b3116	Propianate kinase/acetate kinase anaerobic	2.74
*tdcR**	b3119	Threonine dehydratase operon activator	2.19
tag	b3549	3 methyl-adenine DNA glycosylase	3.25
ogt	b1335	O-6-alkylguanine-DNA/cysteine-protein methyl transferase.	2.38
yiaC	b3550	Putative acetyl transferase	3.08
*dcuC**	b0621	Anaerobic C4 dicarboxylate transport	2.13
*dcuB**	b4123	Anaerobic dicarboxylate transport protein.	2.65
*dcuR**	b4124	Response regulator in two-component regulatory system.	2.63
yhjY	b3548	Involved in bacterial adhesion and biofilm formation.	2.44
*ycaC**	b0897	Putative cysteine hydrolase	2.80
*garP**	b3127	Putative transport protein	2.13
deoC	b4381	Deoxiribose-5-phosphate aldolase	2.10
*dmsA**	b0894	Anaerobic dimethyl sulfoxide reductase subunit.	2.29

Table 2.3. Continued

*nrfE**	b4074	Formate dependent nitrite reductase, involved in attachment of heme to cytochrome C552.	2.64
*nrfG**	b4076	Formate dependent nitrite reductase complex	2.67
*nrfF**	b4075	*nrfEGF* is believed to code for proteins involved in c cytochrome assembly.	3.11
*ynfH**	b1590	Putative dimethyl sulfoxide reductase anchor subunit.	2.91
*ynfG**	b1589	Putative dimethyl sulfoxide reductase Fe-S subunit.	3.16
*ynfF**	b1588	Putative dimethyl sulfoxide reductase major subunit	2.96
*ynfI**	b1591	Putative dimethyl sulfoxide reductase component.	2.68
*ydhV**	b1673	Putative oxidoreductase.	2.89
*ydhY**	b1674	Putative oxidoreductase Fe-S subunit	3.92
ydhZ	b1675	Fumarase D	4.00
yfcF	b2301	Glutathione transferase activity.	2.63
yfcE	b2300	Putative metallo dependent phosphatase.	2.21
*nikE**	b3480	Nickel ATP binding protein nickel transport system.	2.18
*nikR**	b3481	Nickel responsive transcription regulator.	2.16
mrp	b2113	Putative ATP binding protein	2.21
ygeY	b2872	Putative deacetylase.	2.13
ssnA	b2879	Putative Aminohydrolase. Involved in cell viability in the stationary phase.	2.38

Tabble 2.3. Continued

gltP	b4077	Glutamate-aspartate symport protein.	2.18
*hypB**	b2727	Guanine nucleotide binding protein function as nickel donor.	3.20
*hypC**	b2728	Hydrogenase expression/formation protein	3.13
hycA	b2725	Regulatory protein for hycE (part of the FHL complex.	2.99
hycB	b2724	Hydrogenase Fe-S subunit (part of the FHL complex).	2.75
hycC	b2723	Dehydrogenase membrane subunit (part of the FHL complex).	2.25
umuD	b1183	Forms part of the complex with *umuC*.	2.09

2.7 Discussion

The analysis of the differentially expressed genes revealed that there were numerous genes that have unrelated function and belong to different pathways. Even though the majority encode proteins related to anaerobic respiration, some of them are related to transport, nutrient limitation, others to extracellular stress. Some of these genes can be grouped in operons or families such as *tdcACDER*, *dcuBCR*, *nrfEGF*, *ynhHGFI, ydhVYZ, nikER*, *hypBC* and *hycAB* while others appeared over-expressed on their own. However, in order to find a common feature among all these genes, which could justify their simultaneous expression, we studied their regulators.

2.7.1 Downregulated genes

Among the downregulated genes caused by the mutation Y87F are *sdhB, cysE, lldD*, *ansP, ycdB, ycdO, ybdA, yghK, fepE* and *paaI* (Table 2.2). Gene *sdhB*, encodes the SdhB subunit of succinate dehydrogenase. This aerobic respiratory enzyme is a member of the Krebs Cycle; it is embedded in the inner membrane and plays an important role in cell metabolism because it contributes to the generation of the proton-motive force (Salmon et al., 2005). The *sdh* operon is positively regulated by CRP and Fur under aerobic conditions and negatively regulated by ArcA under anaerobic conditions. Along with the σ^E-dependent small regulatory RNA RyhB, which inhibits SdhB mRNA translation, FNR also inhibits *sdhB* transcription initiation (Cecchini et al., 2002).

The *cysE* gene encodes serine acetyltransferase CysE, which carries out the conversion of L-serine to O-acetyl-L-serine, the first step of cysteine biosynthesis (Keseler et al., 2011). Small σ^E-dependent RyhB sRNA base pairs *in vitro* with the *cysE* mRNA, inhibiting its translation (Salvail et al., 2010). Based on microarray data, extraordinary upregulation of *cysE* was found in a *rpoS* mutant in exponential phase in *E. coli* (Rahman et al., 2006).

The *lldD* gene encodes L-lactate dehydrogenase, which functions in aerobic respiration and also in anaerobic nitrate respiration (Nishimura, 1983). The enzyme is located in the inner membrane and interconverts pyruvate and lactate in *E. coli.* LldD is downregulated by transcriptional regulator ArcA (Keseler et al., 2011).

Another downregulated gene is *ansP,* which encodes AnsP located in the inner membrane of *E. coli*. AnsP belongs to the APC (Amino Acid-Polyamine Organocation) superfamily of transporters and it is believed to function as an L-asparagine proton symporter (Keseler et al., 2011). The *ansP* gene possesses two experimentally demonstrated consensus sequences for RpoS binding, and *ansP* expression depends on RpoS in stationary phase (Lacour and Landini, 2004).

The gene *ycdB,* also called *efeB,* encodes a dimeric protein that contains a heme cofactor. The function of this protein in the periplasm is believed to be related to detoxification reactions under certain conditions (Sturm et al., 2006). The gene *ycdO,* also called *efeO,* belongs to the same operon as *efeB,* and encodes a monomeric membrane-associated protein that might be involved in bacterial envelope stress responses since both are downregulated by CpxAR (Cao et al., 2007).

The gene *ybdA*, also called *entS*, encodes a protein that reduces the lethal effects of stress. It is regulated by Fur-Fe^{2+}, a DNA-binding protein that represses and activates bacterial iron uptake (Martinez-Antonio and Collado-Vides, 2003).

Gene *yghK,* also called *glcA*, belongs to the LctP family and it encodes a glycolate/lactate:H^+ transporter located in the inner membrane, which is regulated by Fis DNA-binding protein (Keseler et al., 2011).

The gene product of *fepE* is a polysacaccharide co-polymerase family protein FepE, which is located in the inner membrane (Ozenberger et al., 1987).

Transcription of this gene is upregulated by FNR and H-NS and downregulated by Fe^{+2} and Fur (Keseler et al., 2011).

The downregulated genes also included *paaI*, which encodes the phenylacetyl-CoA thioesterase enzyme located in the cytosol. It is believed to be involved in the catalysis of the aerobic phenylacetate degradation pathway (Ferrandez et al., 1998). The transcription of *paaI* is positively regulated by cyclic-AMP, CRP, IhfB, IhFA, SlyA and negatively regulated by phenylacetyl-CoA (Keseler et al., 2011).

2.7.2 Upregulated genes

Listed in Table 2.3 are the upregulated genes organized by operons and in a decreasing order of read abundance. Among the upregulated genes are *tdcA,* which encodes the first transcriptional activator component of the *tdcABC* operon, while *tdcR* encodes a protein that allows the proper expression of all the *tdc* genes (Hagewood et al., 1994). The product of *tdcD* is a propionate kinase/acetate kinase in anaerobic respiration. In *E. coli*, under anaerobic conditions, the *tdcABC* operon is involved in transport and metabolism of threonine and serine. This operon is not only regulated by FNR and ArcA under anaerobic conditions, but also by TdcA and TdcR (Sawers, 2001).

The *ogt* gene encodes for a methyl transferase located in the cytoplasm, which is responsible for the direct repair of DNA alkyl lesions in *E. coli*. It repairs O_6 methylguanine and O_4 methylthymine lesions by transferring the methyl groups

to its own cysteine residue (Ferrezuelo et al., 1998). This protein is regulated by NarL and Fis transcriptional regulators (Keseler et al., 2011).

The gene *dcuC*, which is transcribed divergently from *pagP*, encodes an inner membrane protein involved in anaerobic C4 dicarboxylate transport whose primary function is succinate transport during glucose fermentation. The *dcuB* gene encodes an inner membrane protein, which belongs to Dcu family and is involved in anaerobic dicarboxylate transport. Both proteins, along with DcuR, are regulated by the DNA binding transcriptional regulator FNR (Zientz, 1999). The *dcuR* gene encodes a cytosolic DcuR response regulator protein, which belongs to the two-component signal transduction system DcuRS.

Gene *yhjY* encodes an outer membrane protein involved in bacterial biofilm formation and adhesion (Casadio, 2003). The gene *ycaC*, encodes a putative cytosolic hydrolase monomer, which is upregulated by BaeR and repressed by FNR (Nishino, 2005). The gene *garP*, also called *yhaU,* encodes a galactarate/glutarate/glycerate transporter, localized in the inner membrane, and believed to be a member of the major facilitator superfamily. This protein is positively regulated by CdaR and FNR and negatively by Fe^{+2}-Fur and H-NS (Pao, 1998).

The gene *deoC* encodes a deoxyribosephosphate aldolase, which catalyzes the reversible aldol reaction where the acetaldehyde donor and glyceraldehyde-3-phosphate acceptor produce deoxyribose 5-phosphate. DeoC is located in the cytosol and is part of the superfamily of pyrimidine deoxyribonucleoside

degradation enzymes. It is regulated positively by Fis and CRP and negatively by DeoR, ModE and CytR (Keseler et al., 2011).

The gene *dmsA* encodes the DmsA subunit of dimethyl sulfoxide reductase from the *dmsABC* operon, which encodes an iron-sulfur-molybdoenzyme complex (Sambasivarao et al., 1991). This complex allows *E. coli* to grow under anaerobic conditions using a variety of terminal electron acceptors (Weiner et al., 1991). DmsA is the catalytic subunit and contains a molybdo-*bis*(pyranoterin guanine dinucleotide)(Mo-*bis*PGD) cofactor and a [4Fe-4S] cluster (Weiner et al., 1991). In the early 1990s the DmsA subunit was reported to be located in the cytoplasm. Although most proteins that possess a twin arginine leader are periplasmic proteins, histological, immunological and biochemical evidences supported that DMSO was the exception (Sambasivarao et al., 1991; Weiner et al., 1992; Weiner et al., 1993). However, later research using reporter proteins demonstrated that the complex DmsAB is translocated to the periplasm. The twin arginine peptide leader allows the complex to be translocated in its folded state through the inner membrane to the periplasm (Berk, 1996; Stanley et al., 2002; Tang et al., 2011; Tang et al., 2013). DmsA expression is positively regulated by FNR under anaerobic conditions and negatively regulated by IhfA and IhfB, Fis, ModE and NarL (Bearson et al., 2002; Lubitz and Weiner, 2003).

Also upregulated are the genes *nrfE*, *nrfG* and *nrfF*, which belong to the *nrfEGF* operon and encode proteins involved in the addition of a heme group to apocytrochrome c552. Gene *nrfE* encodes the inner membrane protein NrfE, which

is a putative cytochrome c-type biogenesis protein. The gene *nrfG* encodes a putative formate-dependent nitrite reductase complex subunit NrfG, which is anchored in the inner membrane and exposed to the periplasm. This operon is positively regulated by FNR and enhanced by NarL; however, it is negatively regulated by Fis and IHF (Browning et al., 2005).

The genes belonging to the *ynfIHGF* operon were upregulated and they encode subunits of an oxidoreductase complex; this operon is a parologue of the *dmsABC* operon in *E. coli* where *ynfF* is the paralogue of *dmsA*. Gene *ynfF* encodes a putative selenite reductase maturation protein located in the inner membrane and is exposed to the periplasmic space. The *ynfG* gene is a paralogue of *dmsB* and *ynfH* is a parologue of *dmsC*. Gene *ynfH* codes for a menaquinol dehydrogenase, which is believed to be the anchor subunit located in the inner membrane (Lubitz and Weiner., 2003). The *ynfIHGF* operon is regulated by FNR and NarL (Lubitz and Weiner., 2003; Constantinido et al., 2006).

The *ydhY* gene is the first open reading frame of the *ydhYVWXUT* operon, and it encodes a putative ferredoxin-like protein possessing four iron-sulfur clusters. Expression of this operon depends on DNA binding transcription factor FNR (Kang et al., 2005). It is believed that YdhY is involved in electron transfer along with YdhV (Partridge et al., 2008). Gene *ydhZ,* or *fumD,* codes for a predicted Fumarase D located in the cytosol (Sevin et al., 2017).

Genes *yfcF* and *yfcE* encode a putative glutathione-*S*-transferase and a putative metal-dependent phosphatase, respectively. They are located in the cytosol and are regulated by the global stress regulator RpoS (Ito et al., 2008).

NikE, which is also called *hydD* or *hydC*, belongs to NikABCDE superfamily of transporters in which *nikE* encodes one of the ATP binding components of the NikABC transporter. NikR encodes a protein located in the cytosol, which is positively regulated by FNR and negatively regulates itself. It is also downregulated by NarL (Eitinger, 2000).

The gene *mrp,* encoding a putative ATP binding protein, is located in the cytosol. Mrp is a member of the Mrp/NBP35 subfamily of the Mrp/MinD family (Leipe et al., 2002). Gene *ygeY* encodes a putative peptidase located in the inner membrane and cytosol, which is believed to be involved in purine catabolism (Xi et al., 2000).

The gene *ssnA* encodes a soluble protein SsnA involved in cell viability at the beginning of stationary phase. It has been demonstrated that there is a relationship between SsnA and RpoS regarding bacterial cell death at early stationary phase. Expression of *ssnA* is negatively regulated by RpoS (Yamada et al., 1998). Repression of SsnA leads to a significant decrease of cell death while expression of SsnA leads to inhibition of cell growth (Nitta et al., 2000). RpoS exerts a negative regulation on SsnA and expression of σ^E decreases growth inhibition caused by SsnA (Nitta et al., 2000).

The gene *gltP* encodes the glutamate/aspartate H^+ symporter GltP. It is known as a proton-dependent transporter for glutamate and aspartate in *E. coli.* This protein is located in the inner membrane (Jacobs et al., 1995).

The gene *hycA* is the first gene of the *hyc* operon, which encodes a protein that acts as transcriptional regulator of FhlA located in the cytosol (Hasona et al., 1998). The gene *hycB* encodes a subunit of the formate hydrogenlyase, which is responsible for anaerobic oxidation of formic acid to carbon dioxide and molecular hydrogen (Hasona et al., 1998). The gene *hycC* encodes a membrane-bound dehydrogenase subunit of the formate hydrogenlyase (FHL) complex located in the inner membrane. The expression of all these genes is positively regulated by formate DNA-binding transcriptional regulator (FhlA), integration host factor (IHF), ModE DNA-binding transcriptional dual regulator, and negatively by the transcriptional repressor NsrR (Bohm et al., 1990).

The gene *hypB* encodes a hydrolase involved in nickel incorporation into subunits of hydrogenase apoproteins. This protein plays an important role in the process of maturation of all hydrogenases (Jacobi et al., 1992). The gene *hypC* encodes a hydrogenase expression/formation protein. The gene *hypA* encodes a hydrogenase-3 nickel incorporation protein involved in hydrogenase isozyme maturation and participates in biosynthesis of hydrogenase-3 under anaerobic conditions. The transcription of these genes is upregulated by FhlA, FNR, IhfB and IhfA, and downregulated by NsrR (Keseler et al., 2011).

The gene *umuD* encodes the UmuD subunit of error prone DNA polymerase Pol V. UmuD is located in the cytosol and it is regulated by LexA (Keseler et al., 2011). The *tag* gene encodes a protein 3-methyladenine DNA glycosylase that is believed to remove 3-methyladenines from DNA. It catalyzes the hydrolysis of *N*-glycosylic bonds between the backbone deoxyribose-phosphate and the methylated base. Lastly, *yiaC* encodes a predicted uncharacterized acyltransferase with an acyl-CoA *N*-acyltransferase domain (Keseler et al., 2011).

2.7.3 Common features among all differentially expressed genes

Through RNA-sequencing we showed that 50 genes were differentially expressed when we compared *E. coli* NR760*λp* and NR760*λpY87F*. Ten genes were downregulated, and forty genes were upregulated. 26 genes have FNR as a common transcription factor regulating their expression. The genes regulated by FNR are marked with asterisks in Tables 2.2 and 2.3. Four genes, *yhjY*, *ssnA*, *yfcE* and *yfcF*, are directly regulated by global response regulator RpoS and the rest of the genes are regulated by CpxAR or indirectly regulated by interconnected effects of σ^E, Cpx, and Rcs.

LptD expression depends on σ^E (Dartigalongue et al., 2001) and the *lptD4213* mutation was used here to trigger PagP activity in the OM. *E. coli* NR760λ*p* and NR760*λpY87F* both carry the *lptD4213* mutation known to cause insufficient delivery of LPS to the cell surface (Bishop, 2014), which can cause accumulation of LPS or its metabolites activating σ^E (Tam and Missiakas, 2005) and also activating the Rcs system (Grabowicz and Silhavy, 2017). Numerous

changes in the periplasm can activate CpxA/CpxR system, (Grabowicz and Silhavy, 2017). However, these effects will be cancelled out by the presence of the *lptD4213* allele in both NR760*λp* and NR760*λpY87F*. It has been previously established in the Bishop lab that Y87F PagP is fully functional in the palmitoylation of lipid A (Smith et al., 2008). Only the mutation Y87F on the periplasmic side of PagP could have caused the differential expression of all these genes.

Chapter 3. Role of OM PagP in signal transduction using NanoString nCounter gene expression

3.1 Preface

The research in this chapter was carried out from 2015 to 2016. The experiments were designed by Dr. Russell Bishop and myself. All experiments were conducted and analyzed by myself. Members of Farncombe Metagenomic Laboratory and members of Dr. Surette's and Dr. Burrows' laboratory helped with the transcriptome analysis.

3.2 Abstract

The OM enzyme PagP possesses an extracellular active site mainly involved in the palmitoylation of lipid A and phosphatidylglycerol. The amino acid residue Ser77 located in extracellular loop L2 of PagP plays an essential role in the transpalmitoylation reaction. PagP also possesses a periplasmic cluster of amino acid residues D61, H67 and Y87, which are implicated in cell signalling. To test if Y87 is involved in PagP signaling, we carried out RNA-sequencing to study differential gene expression by comparing *E. coli* NR760*λp* to the mutant NR760*λpY87F*, both of which carry the *lptD4213* allele that constituvely activates PagP. To validate the RNA-seq results, we used NanoString nCounter to study differential gene expression. Here we included not only the *lptD4213* strains NR760*λp*, NR760*λpY87F,* and NR760*λpS77A,* but also the *lptD*$^{+}$ parental strains NR754*λp,* NR754*λpY87F,* and NR754*λpS77A.* Of the 50 genes found to be differentially expressed during RNA sequencing we selected 45 to be validated

using NanoString nCounter. Comparison of NR754*λp* and NR754*λpY87F* showed that only *dmsA*, *dcuC*, *paaI*, and *sdhB* were overexpressed in the wild-type NR754*λp* with a ≥ 2-fold change and a $p \leq 0.05$. Similarly, the expression of *dmsA*, *sdhB, tdcA* and *dcuC* was different (i.e., a fold change ≥ 2 and a $p \leq 0.05$) when NR760*λp* was compared to NR760*λpY87F*. We also found that only *dmsA* was differentially expressed between NR754*λp* and the catalytically inactive strain NR754*λpS77A*. Thus, the inactivation of PagP caused a 2-fold repression in *dmsA* expression. Our results suggest that PagP activity in the OM controls transcription in the cytoplasm and provides a target gene *dmsA* to monitor PagP-mediated sensory transduction across the *E. coli* cell envelope.

3.3 Introduction

Our previous results obtained from RNA-seq showed that 50 genes were differentially expressed upon comparing the leaky-membrane strains, NR760*λp* and NR760*λpY87F*. Of those genes, 10 were downregulated and encoded proteins with unrelated functions albeit involved with processes occurring under aerobic respiration. The remaining 40 genes were upregulated, with the majority belonging to operons encoding proteins that function under anaerobic respiration. Genes were considered differentially expressed with a Log2FoldChange <-2 and >+2 and a padjust value of ≤0.05. To validate these results and determine whether or not PagP exerts transcriptional regulation on different genes we used NanoString nCounter. Of those 50 genes found differentially expressed in the RNA-seq we selected 45 to

perform the nCounter. These genes were selected according to the read abundance and the relevance of their function.

The OM enzyme PagP possesses an extracellular catalytic active site involved in the palmitoylation of lipid A and PG (Bishop, 2005; Raetz and Whitfield, 2002; Raetz et al., 2007; Dalebroux et al., 2014). The amino acid residue Ser77 located in the extracellular loop L2 of PagP plays an essential role in the transpalmitoylation reaction. PagP also possesses a periplasmic cluster of amino acid residues D61, H67 and Y87, believed to be involved in cell signalling. In order to determine whether or not the extracellular active site or the periplasmic active site is involved in PagP mediated transcriptional regulation we used NanoString nCounter to analyze differential gene expression. In this part of the project we used not only the *lptD*4213 background strains NR760*λp,* NR760*λpY87F* and NR760*λpS77A,* in which PagP is constituvely activated, but also the *lptD*$^{+}$ parental strain NR754*λp* and its mutants NR754*λpS77A* and NR754*λpY87F*.

The NanoString nCounter gene expression assay (NanoString Technologies, Seattle, WA) is a highly reproducible method that can detect the expression of up to 800 genes simultaneously using a colour code set. It does not require the use of an enzymatic reaction, which makes it advantageous compared to other methods since it eliminates biases involved in reverse transcription and cDNA amplification (Geiss et al., 2008). NanoString is highly sensitive since it allows the detection of very low amounts of nucleic acid and has a detection limit between 1fM to 5fM (Geiss et al., 2008). nCounter can be used as a bridge (follow-

up) between RNA sequencing or microarrays and real-time PCR. The mRNA quantification provided by nCounter is more accurate than PCR-based methods (Reiss et al., 2011). The absolute transcript abundance is measured, and each sample is normalized against house keeping genes; therefore, control samples are not required (Geiss et al., 2008).

3.4 Codeset development for nCounter gene expression assay.

NanoString is based on the direct imaging of mRNA molecules. The target mRNAs are detected using reporter and capturer probes. Each probe has a specific sequence that recognizes an ~100 bp region within the target mRNA. A colour code is attached to the 5'-end of the reporter probe, which allows the identification of the mRNA for each gene of interest (Kulkarni, 2011). The 3'-end of the reporter probe contains a biotin label, which facilitates the attachment of the target molecular complex to a streptavidin-coated slide for identification (Geiss et al., 2008). A diagram of the nCounter workflow is depicted in (Figure 3.1).

3.5 Materials and Methods

Bacterial strains used for this project are shown in Table 2.1, while the genes, the target sequences and universal capture sequences and reporter sequences for nCounter analysis are listed in Table 3.1, 3.2 and 3.3 respectively.

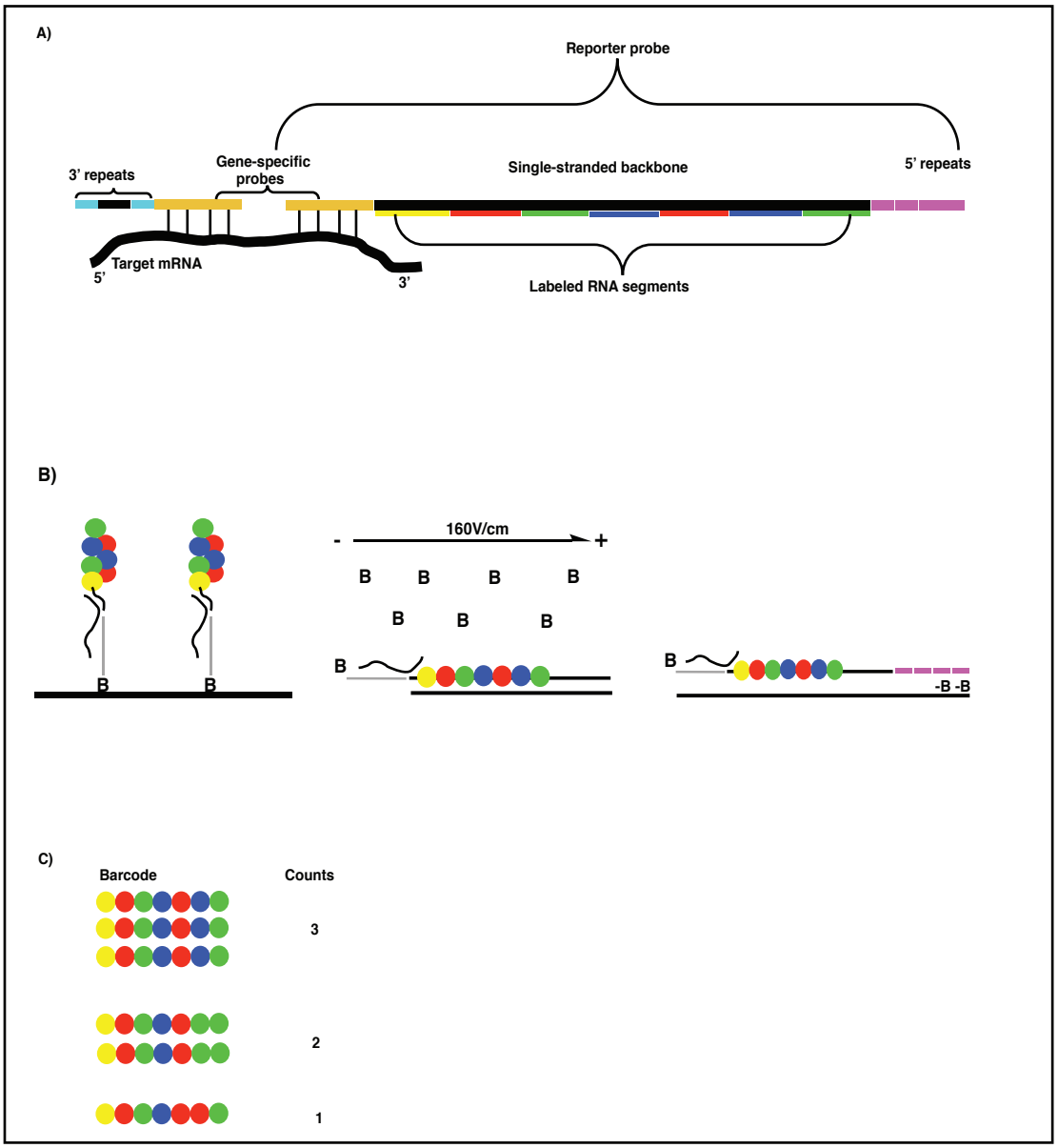

Figure 3.1. Representation of NanoString nCounter gene expression system. A) Hybridized complex. The capture probe and the reporter probe hybridize to a complementary target mRNA in solution, via the gene specific sequences. The complex (tripartite molecules is then purified by affinity. B) Represents the binding, electrophoresis and immobilization. The purified complex is attached to a streptaviding-coated slide via biotnylated capture probes. To elongate and align the molecules voltage is applied. Voltage is turned off and the immobilized reporters are ready for imaging and counting (images not to scale).

Table 3.1 List of the target sequence for each gene used for NanoString nCounter analysis.

Gene name	Accession	Position	Target sequence
sdhB	b0724.1	296-395	TGATCCGCGATTTGGTGGTAGACATGGGACAATTC TATGCGCAATATGAGAAAATTAAGCCTTACCTGTT GAATAATGGACAAAATCCGCCAGCTCGCGA
cysE	b3607.1	374-473	GTCGCGCACTGGCAATCTTTCTGCAAAACCAGGTT TCTGTGACGTTCCAGGTCGATATTCACCCGGCAGC AAAAATTGGTCGCGGTATCATGCTTGACCA
ycdB	b1019.1	629-728	GGGAAGGGTTTATTTCCGATCACGCGGCGCGTAGT AAAGGCAAAGAGACGCCGATTAATTTGCTGGGTTT CAAAGACGGCACTGCCAATCCCGATAGCCA
tdcA	b3118.1	562-661	ACGCTGGAGTCGTTGAAGAACGAACAGTGGGTGTT GCCACAAACTAATATGGGGTACTACAGCGAACTGC TTACTACGTTACAAAGAAATGGCATCAGTA
dmsA	b0894.1	1192-1291	TTACCGTTTGTCCGTATGCCGACCTTGGAAAACCC GATCCAGACCAGCATTTCGATGTTTATGTGGACCG ATGCCATTGAACGTGGCCCGGAAATGACGG
dcuC	b0621.1	856-955	TTCCTCCGCAGCTTTAATACCCAGAAAGTTTTCTCT GGTCTGGAAGTGGCTTATCGCGGGATGGCAGATGC GTTTGCTAACGTGGTGATGCTGCTGGTTG
dcuR	b4124.1	221-320	CGCGTTGCAAAAGTGATGTGATTGTCATCTCCTCC GCAGCCGATGCGGCAACCATTAAAGATTCGCTGCA TTACGGTGTCGTGGATTACCTGATCAAACC
rsmC	b4371.1	550-649	GGTAGCCAGTTGCTGCTCTCGACGTTAACTCCGCA CACGAAAGGTAAAGTGCTGGATGTCGGCTGTGGC GCGGGCGTGCTTTCAGTTGCCTTTGCGCGCC
rpoD	b3067.1	334-433	ACCCGCGAAGGCGAAATTGACATCGCTAAGCGTAT TGAAGACGGGATCAACCAGGTTCAATGCTCCGTTG CTGAATATCCGGAAGCGATCACCTATCTGC
ihfB	b0912.1	41-140	AATCGCACATTCCCGCCAAGACGGTTGAAGATGCA GTAAAAGAGATGCTGGAGCATATGGCCTCGACTCT TGCGCAGGGCGAGCGTATTGAAATCCGCGG
ycdO	b1018.1	548-647	TTGAACCGATTGCTGAACTGTTCTCCGATCTGGAT GGCAGCATTGACGCCCGTGAAGATGATTACGAGC AAAAAGCCGCCGACCCAAAATTCACTGGTTT
ybdA	b0591.1	313-412	CTGCCGGAGCCGTCATTGCTGGCAATCTATTTACTT GGTTTATGGGATGGTTTTTTCGCATCGCTTGGCGTT ACGGCGCTATTGGCGGCGACACCAGCAC
yghK	b2975.1	728-827	ACTATATTGGTCCGGAACTGCCGGATATTACTTCG GCGCTGGTGAGTATCGTCTCACTCGCTTTATTCCTT AAAGTCTGGCGGCCGAAAAATACCGAAAC
fepE	b0587.1	543-642	TGAAGAGGCGCAGACCGTTTTGAGCGGGTATATCG ATTATATCTCTACGTTGGTGGTGAAAGAGTCGCTA GAAAACGTCCGTAATAAACTGGAGATCAAA

Table 3.1. Continued

paaI	b1396.1	305-404	ATCAGGGCAAGCAAACCGGTGTTTACGACATCGA AATTGTTAACCAACAACAAAAAACGGTTGCGCTGT TTCGCGGTAAATCTCACCGCATCGGCGGCAC
tdcD	b3115.1	779-878	TGGGTGACCTGGAACGCGTAGTGAATAAAGAGTC GGGATTATTAGGTATTTCCGGTCTTTCTTCGGATTT ACGTGTTCTGGAAAAAGCCTGGCATGAAGG
tdcE	b3114.1	1221-1320	CGATCTGATGCGTACTGACTTCAACAGCGACGATT ACGCGATTGCCTGCTGCGTCAGCCCAATGGTGATT GGTAAGCAAATGCAGTTCTTTGGTGCACGC
tdcC	b3116.1	357-456	TACCAATACCTTTATGACGTTCTGGGAAAACCAGC TCGGCTTTGCACCGCTGAATCGCGGCTTTGTGGCG CTGTTCCTGTTGCTGCTGATGGCTTTCGTC
tdcR	b3119.1	1-100	ATGACGGGGATAACCATATTTTATGGCGATAACAT CATTCGTTATGTGGTTAATACAAAAAAGGGGCTGA GACCATATTTCAAGCAATTACCGGATAATT
ogt	b1335.1	157-256	CGCATTTCTGCCACCAATCCAGGCGGTTTAAGCGA CAAGCTTCGTGAATATTTTGCCGGTAATCTTAGCA TTATTGATACGCTTCCCACTGCTACGGGGG
dcuB	b4123.1	793-892	AAACCGCTGTCGATGGTACTGGTTATTCAGATGTT TATGCTGCTGACCGGGGCGCTGATTATTATCCTGA CCAAAACCAATCCCGCGTCTATCTCAAAAA
yhjY	b3548.1	276-375	TATGAGTATGGGCTGGAATTTTCCTTTATATGAAC AGGTTACAACCGGCCCGGTCGCGGCATTACATTAC GATGGCACAACCACCTCGATGTATAACGAG
ycaC	b0897.1	250-349	CGCCCGGGAAATATTAACGCCTGGGATAACGAAG ATTTTGTAAAAGCTGTCAAAGCGACAGGTAAAAA ACAGTTAATTATTGCCGGTGTGGTAACCGAAG
garP	b3127.1	588-687	CAAGTTGATTCATAACCCGACAGATCACCCACGTA TGTCTGCGGAAGAGCTGAAGTTTATCTCTGAAAAT GGCGCGGTGGTCGATATGGACCACAAAAAG
nrfE	b4074.1	847-946	TCGGTTCATGCGTTCGCGCTGGATAACGTCCGCGC CGTGCCGTTGTTCAGCCTGTTTGCACTGATTAGCCT TGCGTCTCTGGCTCTGTATGGCTGGCGAG
nrfG	b4076.1	339-438	GGCGACGGTGCTTTATTACCAGGCCAGCCAGCATA TGACCGCCCAGACTCGCGCAATGATCGACAAAGCC CTCGCGCTGGACAGTAATGAAATCACCGCC
nrfF	b4075.1	115-214	GCCAGCCAGTTACGTTGTCCGCAGTGCCAGAATCA AAACTTACTGGAATCCAACGCGCCGGTGGCTGTCA GTATGCGCCATCAGGTTTACAGCATGGTGG
ynfG	b1589.1	333-432	CGGCGCGCCACAGTACAATGCTGAAAAAGGGCAC ATGACGAAGTGCGATGGTTGTTATTCGCGCGTCGC CGAGGGGAAACAACCCATATGTGTCGAATCC
ynfF	b1588.1	1682-1781	GTGAAGTCGCCAAACGCTTAGGACCAGACGTTTAT CAAACCTTTACTGAAGGTCGCAGTCAGCATGAATG GATCAAATATCTCCATGCGAAAACGAAGGA

Table 3.1. Continued

ynfI	b1591.1	288-387	TCGCGAATCTGTGCTGTTTGGCGATTCAACATTGG CACTTCGTCAGTGGATGCGCGAGAAAGGCATTCAG TTTGAAATGAAGCAAAACGAACCGGAAGAT
ydhV	b1673.1	1217-1316	CAGAAATTCGCTGGGATCAACTGGAAGCGGGTGA CGTTAACTTCATTAAAGATTTTTACTACCGTCTGGC GCATCGTGTGGGTGAGCTGAGTCACCTGGC
ydhY	b1674.1	234-333	TACCAACTTCAACGATGGCTCAGTAGGGACATTCT TCTCCCGTATCAAAATCCATCGCAATTATTTCTTTG GCGACAACGGGGTTGGCTCTGGCGGCGGC
ydhZ	b1675.1	36-135	AGAAATGTGCAGAGTTGTCGGTAAAGTTGTGCTGG AAATGCGCGATCTGGGGCAGGAACCAAAGCATAT TGTTATCGCAGGCGTATTACGTACAGCATTA
yfcF	b2301.1	361-460	CCGACGGATGTTGTCTTTGCGGGGGCGAAAAAAGC GCCACTAACGGCCGAGGGAAAAGCCAGTGCAGAG AAACTGTTCGCGATGGCAGAACATTTGTTAG
yfcE	b2300.1	205-304	GCTGTGCGCGGCAACTGCGACAGCGAAGTGGATC AAATGCTGCTGCATTTCCCGATAACCGCGCCGTGG CAACAGGTATTACTGGAAAAACAACGTCTGT
nikE	b3480.1	227-326	TGGCGAAACTCAATCGCGCCCAGCGTAAAGCGTTC CGCCGCGATATTCAGATGGTATTTCAGGACTCCAT CAGCGCCGTGAATCCGCGCAAAACCGTGCG
nikR	b3481.1	292-391	ATCGCCGTGTTGAAAGGTGACATGGGTGACGTGCA GCATTTTGCCGATGACGTTATCGCCCAGCGCGGCG TGCGGCACGGGCATTTGCAGTGCTTGCCGA
mrp	b2113.1	607-706	ATGGTGTGGCGTGGACCGATGGCGAGCAAGGCGT TAATGCAGATGTTGCAGGAAACCTTGTGGCCGGAT CTCGACTATCTGGTGCTTGATATGCCGCCAG
ssnA	b2879.1	850-949	GTCGCTCATGGGCTGTACTTGTCGAAAGATGACAT CACCCTACTCAATCAGCGCGATGCGTTCCTGGTGC ATAACGCCCGTTCAAACATGAACAACCATG
gltP	b4077.1	415-514	ATTATGGGCACGATTTTGTCGCTGGTGCCGACGAA CATTGTGGCGTCGATGGCGAAAGGCGAAATGCTGC CGATCATCTTTTTCTCGGTGCTGTTTGGTC
hycA	b2725.1	219-318	CAGCCACACCATCGAGTATTACGTCGAAACAAAA GATGGCGAAGACAAACAGCGGATTGCGCAGGCGC AACTGAGCATTGACGGCATGATTGATGGCAAG
hycB	b2724.1	196-295	AACGCCATCACCCGCGTCGATGGGGCCGTGCAGTT GAATGAAAGCCTGTGCGTAAGCTGCAAGCTGTGCG GCATCGCCTGCCCGTTTGGCGCAATTGAAT
hypB	b2727.1	397-496	CCGTGCGCAGTTATTGAAGGCGACCAGCAAACCGT GAACGATGCCGCACGCATTCGCGCTACCGGCACAC CAGCGATTCAGGTGAACACCGGTAAAGGCT
hycC	b2723.1	949-1048	CTGATTGCTCTTGGCCTGGTCGGTGGTCTGTACCAT CTGCTTAACCATAGCCTGTTCAAAAGCGTACTGTT CCTCGGGGCGGGGAGCGTCTGGTTCCGTA

Table 3.1. Continued

umuD	b1183.1	108-207	GCGCATCGATCTGAATCAACTGTTGATCCAGCATC CCAGCGCGACTTACTTCGTCAAAGCAAGTGGTGAT TCTATGATTGATGGTGGAATTAGTGACGGT
tag	b3549.1	248-347	CCGGGATTATCCGCCATCGAGGGAAAATTCAGGCA ATTATTGGTAATGCGCGGGCGTACCTGCAAATGGA ACAGAACGGCGAACCGTTTGTCGACTTTGT
yiaC	b3550.1	173-272	AGCTTCTCGGTTTTGTCAGCATTATGGAAGGCCGA TTTCTGGCAGCGATGTTTGTCGCACCGAAGGCCGT CAGGCGCGGTATTGGTAAGGCGCTGATGCA
deoC	b4381.1	278-377	CAATCGCCTACGGTGCTGATGAAGTTGACGTTGTG TTCCCGTACCGCGCGCTGATGGCGGGTAACGAGCA GGTTGGTTTTGACCTGGTGAAAGCCTGTAA

Table 3.2. Sequences of the universal capture sequence.

Gene name	**Universal Capture Sequence (3'- 5')**
sdhB	CGAAAGCCATGACCTCCGATCACTCTCGCGAGCTGGCGGATTTTGT CCATTATTCAACAGGTAAGGCTTAATTTT
cysE	CGAAAGCCATGACCTCCGATCACTCTGGTCAAGCATGATACCGCG ACCAATTTTTGCTGCCGGGTGAATATCGAC
ycdB	CGAAAGCCATGACCTCCGATCACTCTGGCTATCGGGATTGGCAGTG CCGTCTTTGAAACCCAGCAAATTAATCGG
tdcA	CGAAAGCCATGACCTCCGATCACTCTACTGATGCCATTTCTTTGTAA CGTAGTAAGCAGTTCGCTGTAGTACCCC
dmsA	CGAAAGCCATGACCTCCGATCACTCCCGTCATTTCCGGGCCACGTTC AATGGCATCGGTCCACATAAACATCGAA
dcuC	CGAAAGCCATGACCTCCGATCACTCCAACCAGCAGCATCACCACGTT AGCAAACGCATCTGCCATCCCGCGATAA
dcuR	CGAAAGCCATGACCTCCGATCACTCGGTTTGATCAGGTAATCCACGA CACCGTAATGCAGCGAATCTTTAATGGT
rsmC	CGAAAGCCATGACCTCCGATCACTCCGCGCAAAGGCAACTGAAAGC ACGCCCGCGCCACAGCCGACATCCAGC
rpoD	CGAAAGCCATGACCTCCGATCACTCGCAGATAGGTGATCGCTTCCGG ATATTCAGCAACGGAGCATTGAACCTGG
ihfB	CGAAAGCCATGACCTCCGATCACTCCGGATTTCAATACGCTCGCCCT GCGCAAGAGTCGAGGCCATATGCTC
ycdO	CGAAAGCCATGACCTCCGATCACTCAAACCAGTGAATTTTGGGTCGG CGGCTTTTTGCTCGTAATCATCTTCACG
ybdA	CGAAAGCCATGACCTCCGATCACTCGTGCTGGTGTCGCCGCCAATAG CGCCGTAACGCCAAGCGATGCGAAAAAA
yghK	CGAAAGCCATGACCTCCGATCACTCGTTTCGGTATTTTTCGGCCGCCA GACTTTAAGGAATAAAGCGAGTGAGAC
fepE	CGAAAGCCATGACCTCCGATCACTCTTTGATCTCCAGTTTATTACGGA

Table 3.2. Continued

	CGTTTTCTAGCGACTCTTTCACCACCA
paaI	CGAAAGCCATGACCTCCGATCACTCGTGCCGCCGATGCGGTGAGATT TACCGCGAAACAGCGCAACCGTTTTTTG
tdcD	CGAAAGCCATGACCTCCGATCACTCCCTTCATGCCAGGCTTTTTCCAG AACACGTAAATCCGAAGAAAGACCGGA
tdcE	CGAAAGCCATGACCTCCGATCACTCGCGTGCACCAAAGAACTGCATTT GCTTACCAATCACCATTGGGCTGACGC
tdcC	CGAAAGCCATGACCTCCGATCACTCGACGAAAGCCATCAGCAGCAAC AGGAACAGCGCCACAAAGCCGCGATTCA
tdcR	CGAAAGCCATGACCTCCGATCACTCAATTATCCGGTAATTGCTTGAAA TATGGTCTCAGCCCCTTTTTTGTATTA
ogt	CGAAAGCCATGACCTCCGATCACTCCCCCCGTAGCAGTGGGAAGCGTA TCAATAATGCTAAGATTACCGGCAAAA
dcuB	CGAAAGCCATGACCTCCGATCACTCTTTTTGAGATAGACGCGGGATTG GTTTTGGTCAGGATAATAATCAGCGCC
yhjY	CGAAAGCCATGACCTCCGATCACTCCTCGTTATACATCGAGGTGGTTG TGCCATCGTAATGTAATGCCGCGACCG
ycaC	CGAAAGCCATGACCTCCGATCACTCCTTCGGTTACCACACCGGCAATA ATTAACTGTTTTTTACCTGTCGCTTTG
garP	CGAAAGCCATGACCTCCGATCACTCCTTTTTGTGGTCCATATCGACCA CCGCGCCATTTTCAGAGATAAACTTCA
nrfE	CGAAAGCCATGACCTCCGATCACTCCTCGCCAGCCATACAGAGCCAG AGACGCAAGGCTAATCAGTGCAAACAGG
nrfG	CGAAAGCCATGACCTCCGATCACTCTGATTTCATTACTGTCCAGCGCG AGGGCTTTGTCGATCATTGCGC
nrfF	CGAAAGCCATGACCTCCGATCACTCCACCATGCTGTAAACCTGATGGC GCATACTGACAGCCACCGGCGCGTTG
ynfG	CGAAAGCCATGACCTCCGATCACTCGGATTCGACACATATGGGTTGTT TCCCCTCGGCGACGCGCGAATAACAAC
ynfF	CGAAAGCCATGACCTCCGATCACTCTCCTTCGTTTTCGCATGGAGATA TTTGATCCATTCATGCTGACTGCGACC
ynfI	CGAAAGCCATGACCTCCGATCACTCATCTTCCGGTTCGTTTTGCTTCAT TTCAAACTGAATGCCTTTCTCGCGCA
ydhV	CGAAAGCCATGACCTCCGATCACTCGCCAGGTGACTCAGCTCACCCAC ACGATGCGCCAGACGGTAGTAAAAATC
ydhY	CGAAAGCCATGACCTCCGATCACTCGCCGCCGCCAGAGCCAACCCCGT TGTCGCCAAAGAAATAATTGCGATGGA
ydhZ	CGAAAGCCATGACCTCCGATCACTCTAATGCTGTACGTAATACGCCTG CGATAACAATATGCTTTGGTTCCTGCC
yfcF	CGAAAGCCATGACCTCCGATCACTCCTAACAAATGTTCTGCCATCGCG AACAGTTTCTCTGCACTGGCTTTTCCC
yfcE	CGAAAGCCATGACCTCCGATCACTCACAGACGTTGTTTTTCCAGTAAT ACCTGTTGCCACGGCGCGGTTATCGGG

Table 3.2. Continued

nikE	CGAAAGCCATGACCTCCGATCACTCCGCACGGTTTTGCGCGGATTCAC GGCGCTGATGGAGTCCTGAAATACCAT
nikR	CGAAAGCCATGACCTCCGATCACTCCAAGCACTGCAAATGCCCGTGCC GCACGCCGCGCTGGGCGATAACG
mrp	CGAAAGCCATGACCTCCGATCACTCCGGCATATCAAGCACCAGATAGT CGAGATCCGGCCACAAGGTTTCC
ssnA	CGAAAGCCATGACCTCCGATCACTCCATGGTTGTTCATGTTTGAACGG GCGTTATGCACCAGGAACGCATCGCGC
gltP	CGAAAGCCATGACCTCCGATCACTCGACCAAACAGCACCGAGAAAAA GATGATCGGCAGCATTTCGCCTTTCGCC
hycA	CGAAAGCCATGACCTCCGATCACTCCATCAATCATGCCGTCAATGCTC AGTTGCGCCTGCGCAATCCGCT
hycB	CGAAAGCCATGACCTCCGATCACTCATTCAATTGCGCCAAACGGGCAG GCGATGCCGCACAGCTTGCAGCTTACG
hypB	CGAAAGCCATGACCTCCGATCACTCCTTTACCGGTGTTCACCTGAATC GCTGGTGTGCCGGTAGCGCGAATG
hycC	CGAAAGCCATGACCTCCGATCACTCTACGGAACCAGACGCTCCCCGC CCCGAGGAACAGTACGCTTTTGAACAGG
umuD	CGAAAGCCATGACCTCCGATCACTCACCGTCACTAATTCCACCATCAA TCATAGAATCACCACTTGCTTTGACGA
tag	CGAAAGCCATGACCTCCGATCACTCACAAAGTCGACAAACGGTTCGCC GTTCTGTTCCATTTGCAGGTACGCCCG
yiaC	CGAAAGCCATGACCTCCGATCACTCCATCAGCGCCTTACCAATACCG CGCCTGACGGCCTTCGGTGCGACAAA
deoC	CGAAAGCCATGACCTCCGATCACTCTTACAGGCTTTCACCAGGTCAAA ACCAACCTGCTCGTTACCCGCCATCAG

Table 3.3. Reporter tag sequences.

Gene name	**Reporter Tag sequences (5'-3')**
sdhB	CTCATATTGCGCATAGAATTGTCCCATGTCTACCACCAAATCGCGGA TCACCTCAAGACCTAAGCGACA GCGTGACCTTGTTTCA
cysE	CTGGAACGTCACAGAAACCTGGTTTTGCAGAAAGATTGCCAGTGCGC GACCATCCTCTTCTTTTCTTGGTGTTGAGAAGATGCTC
ycdB	CGTCTCTTTGCCTTTACTACGCGCCGCGTGATCGGAAATAAACCCTTC CCCACAATTCTGCGGGTTAGCAGGAAGGTTAGGGAAC
tdcA	ATATTAGTTTGTGGCAACACCCACTGTTCGTTCTTCAACGACTCCAGC GTCTGTTGAGATTATTGAGCTTCATCATGACCAGAAG
dmsA	ATGCTGGTCTGGATCGGGTTTTCCAAGGTCGGCATACGGACAAACGG TAACAAAGACGCCTATCTTCCAGTTTGATCGGGAAACT
dcuC	GCCACTTCCAGACCAGAGAAAACTTTCTGGGTATTAAAGCTGCGGAG GAACGAACCTAACTCCTCGCTACATTCCTATTGTTTTC
dcuR	TGCCGCATCGGCTGCGGAGGAGATGACAATCACATCACTTTTGCAAC GCGCCAATTTGGTTTTACTCCCCTCGATTATGCGGAGT

Table 3.3. Continued

rsmC	ACTTTACCTTTCGTGTGCGGAGTTAACGTCGAGAGCAGCAACTGGCTA CCCTTTCGGGTTATATCTATCATTTACTTGACACCCT
rpoD	TTGATCCCGTCTTCAATACGCTTAGCGATGTCAATTTCGCCTTCGCGGG TCAACAGCCACTTTTTTTCCAAATTTTGCAAGAGCC
ihfB	CAGCATCTCTTTTACTGCATCTTCAACCGTCTTGGCGGGAATGTGCGAT TCACCGTGTGGACGGCAACTCAGAGATAACGCATAT
ycdO	GGCGTCAATGCTGCCATCCAGATCGGAGAACAGTTCAGCAATCGGTTC AACCTGGAGTTTATGTATTGCCAACGAGTTTGTCTTT
ybdA	CCATCCCATAAACCAAGTAAATAGATTGCCAGCAATGACGGCTCCGGC AGCAGATAAGGTTGTTATTGTGGAGGATGTTACTACA
yghK	GATACTCACCAGCGCCGAAGTAATATCCGGCAGTTCCGGACCAATATA GTCTTCCTTCCTGTGTTCCAGCTACAAACTTAGAAAC
fepE	ACGTAGAGATATAATCGATATACCCGCTCAAAACGGTCTGCGCCTCTT CACATAAAATTGGTTTTGCCTTTCAGCAATTCAACTT
paaI	TTGTTGGTTAACAATTTCGATGTCGTAAACACCGGTTTGCTTGCCCTGA TCTGGTCAAGACTTGCATGAGGACCCGCAAATTCCT
tdcD	AATACCTAATAATCCCGACTCTTTATTCACTACGCGTTCCAGGTCACCC ACTTTCGTTGGGACGCTTGAAGCGCAAGTAGAAAAC
tdcE	AGCAGGCAATCGCGTAATCGTCGCTGTTGAAGTCAGTACGCATCAGAT CGCCAGCAGACCTGCAATATCAAAGTTATAAGCGCGT
tdcC	GCGGTGCAAAGCCGAGCTGGTTTTCCCAGAACGTCATAAAGGTATTGG TACCTGCCAATGCACTCGATCTTGTCATTTTTTTGCG
tdcR	ACCACATAACGAATGATGTTATCGCCATAAAATATGGTTATCCCCGTC ATCAAACTGGAGAGAGAAGTGAAGACGATTTAACCCA
ogt	TATTCACGAAGCTTGTCGCTTAAACCGCCTGGATTGGTGGCAGAAATG CGCGATTGCTGCATTCCGCTCAACGCTTGAGGAAGTA
dcuB	CCGGTCAGCAGCATAAACATCTGAATAACCAGTACCATCGACAGCGGT TTCTGAGGCTGTTAAAGCTGTAGCAACTCTTCCACGA
yhjY	GGCCGGTTGTAACCTGTTCATATAAAGGAAAATTCCAGCCCATACTCA TACTAGGACGCAAATCACTTGAAGAAGTGAAAGCGAG
ycaC	ACAGCTTTTACAAAATCTTCGTTATCCCAGGCGTTAATATTTCCCGGGC GCCACGCGATGACGTTCGTCAAGAGTCGCATAATCT
garP	GCTCTTCCGCAGACATACGTGGGTGATCTGTCGGGTTATGAATCAACTT GCATTTGGAATGATGTGTACTGGGAATAAGACGACG
nrfE	CTGAACAACGGCACGGCGCGGACGTTATCCAGCGCGAACGCATGAAC CGACACAAGAATCCCTGCTAGCTGAAGGAGGGTCAAAC
nrfG	GAGTCTGGGCGGTCATATGCTGGCTGGCCTGGTAATAAAGCACCGCTT GACGTAGATTGCTATCAGGTTACGATGACTGC
nrfF	GATTCCAGTAAGTTTTGATTCTGGCACTGCGGACAACGTAACTGGCTGG CCTTACAGATCGTGTGCTCATGACTTCCACAGACGT
ynfG	CATCGCACTTCGTCATGTGCCCTTTTTCAGCATTGTACTGTGCTTGGAG GAGTTGATAGTGGTAAAACAACATTAGC
ynfF	TTCAGTAAAGGTTTGATAAACGTCTGGTCCTAAGCGTTTGGCGACTTCA CCCTACGTATATATCCAAGTGGTTATGTCCGACGGC
ynfI	TCCACTGACGAAGTGCCAATGTTGAATCGCCAAACAGCACAGATTCGC GACAGCAAGAAGGAGTATGGAACTTATAGCAAGAGAG
ydhV	TTTAATGAAGTTAACGTCACCCGCTTCCAGTTGATCCCAGCGAATTTCT

Table 3.3. Continued

	GCACCCCTCCAAACGCATTCTTATTGGCAAATGGAA
ydhY	TTTTGATACGGGAGAAGAATGTCCCTACTGAGCCATCGTTGAAGTTGG TACCCGAAGCAATACTGTCGTCACTCTGTATGTCCGT
ydhZ	CCAGATCGCGCATTTCCAGCACAACTTTACCGACAACTCTGCACATTTC TCCGGGAATCGGCATTTCGCATTCTTAGGATCTAAA
yfcF	TCGGCCGTTAGTGGCGCTTTTTTCGCCCCCGCAAAGACAACATCCCCGA TCTTCATAACGGACAAACTGAACGGGCCATT
yfcE	AAATGCAGCAGCATTTGATCCACTTCGCTGTCGCAGTTGCCGCGCCGC TATGCAGACGAGCTGGCAGAGGAGAGAAATCA
nikE	CTGAATATCGCGGCGGAACGCTTTACGCTGGGCGCGATTGAGTTTCGC CACATTCGCAACCATGTGAAGTAATGTGAGCGTACTT
nikR	TCATCGGCAAAATGCTGCACGTCACCCATGTCACCTTTCAACACGGCG ATCACCAGTTAGCGTGGCGTATACCATGTTGTTAACA
mrp	TGCAACATCTGCATTAACGCCTTGCTCGCCATCGGTCCACGCCACCCTG AATCAATAGAACAATATCAGTTATGGCGGTG
ssnA	TGATTGAGTAGGGTGATGTCATCTTTCGACAAGTACAGCCCATGAGCG ACCGGTTGTTAATATGACAGGCCGCTAAAGACGTTCT
gltP	ATCGACGCCACAATGTTCGTCGGCACCAGCGACAAAATCGTGCCCATA ATCCGTCTCAGATGAGTGGGTTAATCAATCAAGTATG
hycA	GTTTGTCTTCGCCATCTTTTGTTTCGACGTAATACTCGATGGTGTGGCT GCTGACACATTAGTAACGTCGGCAAGCACTTAGTCG
hycB	CACAGGCTTTCATTCAACTGCACGGCCCCATCGACGCGGGTGATGCGT GAACCAGATTATGTATGGACGCGCAATAGATA
hypB	CGTGCGGCATCGTTCACGGTTTGCTGGTCGCCTTCAATAACTGCGCAT ACGAAATTTGAGCAAGCAATTGAAGGCTTAGA
hycC	CTATGGTTAAGCAGATGGTACAGACCACCGACCAGGCCAAGAGCAAT CAGCTATCAGCTAATAGGGTCGGCTCAACAGTGTATCC
umuD	AGTAAGTCGCGCTGGGATGCTGGATCAACAGTTGATTCAGATCGATGC GCCTATCAATTCGTGACCCCGATCATCCAGTCCAGAA
tag	CGCATTACCAATAATTGCCTGAATTTTCCCTCGATGGCGGATAATCCCG GCTTGAGCTCTAGGCCCAAAACGACCTTAATGGTCA
yiaC	CATCGCTGCCAGAAATCGGCCTTCCATAATGCTGACAAAACCGAGAAG CTCTAGCCCAGATCCTACGAGATGAGCTACGTAACTA
deoC	CGCGCGGTACGGGAACACAACGTCAACTTCATCAGCACCGTAGGCGAT TGCAAATGCACTCTATATGGAGGGAGAGTAGCTGGAT

3.5.1 RNA extraction

The gene expression was analyzed from three biological replicates using NanoString nCounter. Total RNA was isolated using Trizol reagent (Invitrogen), following the manufacture's protocol. The experimental and control strains (Table 1) were streaked onto LB agar and a single colony was inoculated into 5 ml LB

media overnight at 37 ºC. Three biological replicates were grown under aerobic conditions. Overnight cultures were subcultured (1:100) and grown at 37 ºC to mid-log phase with an optical density (O.D.- 600 nm) between 0.6 - 0.8. Following growth to mid-log phase, 1 ml was centrifuged at 12,000 xg for 3 minutes and the supernatant was discarded. The pellet was resuspended in 700 µl of RNase-free water with 50 µl of 100 mg/ml lysozyme. The mixture was incubated at 37 °C for 45 minutes and centrifuged at 12,000 xg and 4 °C for 15 minutes. The resulting pellet was resuspended in 1 ml trizol and vortexed on high speed for 10-20 seconds. The mixture was then incubated at room temperature for 5 minutes. Following incubation, 200 µl of chloroform was added to the mixture and vortexed briefly. After 3 minutes of incubation, the mixture was centrifuged at 12, 000 xg for 20 minutes at 4 °C and 400 ul of the aqueous phase was extracted. RNA was extracted from the aqueous phase by mixing one volume of 70 % ethanol to one volume of aqueous phase. The 800 µl mixture was loaded onto a column from an RNAeasy Mini Kit (50) for RNA purification following the manufacturer's instructions. The RNA concentrations were determined using a NanoDrop and samples were stored at -80 °C. RNA integrity and NanoString assays were carried out in the Farncombe Genomic Facility, McMaster University.

3.5.2 Quality control of extracted RNA

The concentration and the quality of the RNA samples were determined by measuring the absorbance at 260 and 280 nm using a Nanodrop system (Thermo

Scientific) (Figure 3.2). Hybridization required 100 to 150 ng of total RNA, and RNA integrity (RIN) was verified using a Bioanalyzer (Model 2100 Agilent Technologies). To normalize transcript levels, three housekeeping genes with different ranges of expression were included (i.e., low, medium and high expression). Among the housekeeping genes were *ihfB* (Integration host factor; low expression), *rpoD* (RNA polymerase sigma factor; medium expression), and *rsmC* (RNA small ribosomal subunit methyltransferase; high expression).

3.5.3 Data analysis

For NanoString data processing the RCC files were imported into nSolver 2.6 for quality control analyses. The positive control geometric mean normalization factor (without the negative control background subtraction) was determined for each sample. A normalized global background threshold was calculated by multiplying the positive geometric mean normalization factor of each sample by the global background threshold and calculating the mean +2 standard deviations. The normalized data were exported to Microsoft Excel and filtered to include only genes where all samples had counts above the normalized global background threshold. The data were then exported in .CSV format for gene expression analyses. Statistical significance was determined using the Student's t-test in Statplus for Mac.

3.6 Results

The Bioanalyzer generated values based on an electropherogram, where values ranged from 1 to 10 with 1 representing completely degraded RNA and 10 representing RNA of the highest quality. The peaks of 18S and 28S ribosomal subunits show the intactness of the RNA samples. The RIN of our RNA samples ranged from 8 to 10 (Figure 3.2) with exception of sample NR754*λpBAD18*, which showed a value of 7.90 with a little degradation. From the results of the RNA sequencing, 45 differentially expressed genes were selected to perform a NanoString gene expression analysis. Comparison of NR754*λp* and NR754*λpY87F* showed that only *dmsA*, *dcuC*, *paaI*, and *sdhB* had a differential expression with a fold change ≥ 2 and a $p \leq 0.05$ between samples. Similarly, the expression of *dmsA*, *sdhB, tdcA* and *dcuC* was different (i.e., a fold change ≥ 2 and a $p \leq 0.05$) when *E. coli* NR760*λp* was compared to NR760*λpY87F* (Figure 3.3). We also found that only the *dmsA* gene was 2.5-fold downregulated when we compared the wild-type strain NR754*λp* and the catalytically inactive NR754*λpS77A* (Figure 3.4).

Normalized read count values for the samples and the house keeping genes are shown in Table 3.4, Table 3.5 and Table 3.6. Internal positive controls (POS A, B, C and D) used for sample normalization were within the normal parameters and their values are also included in the aforementioned tables.

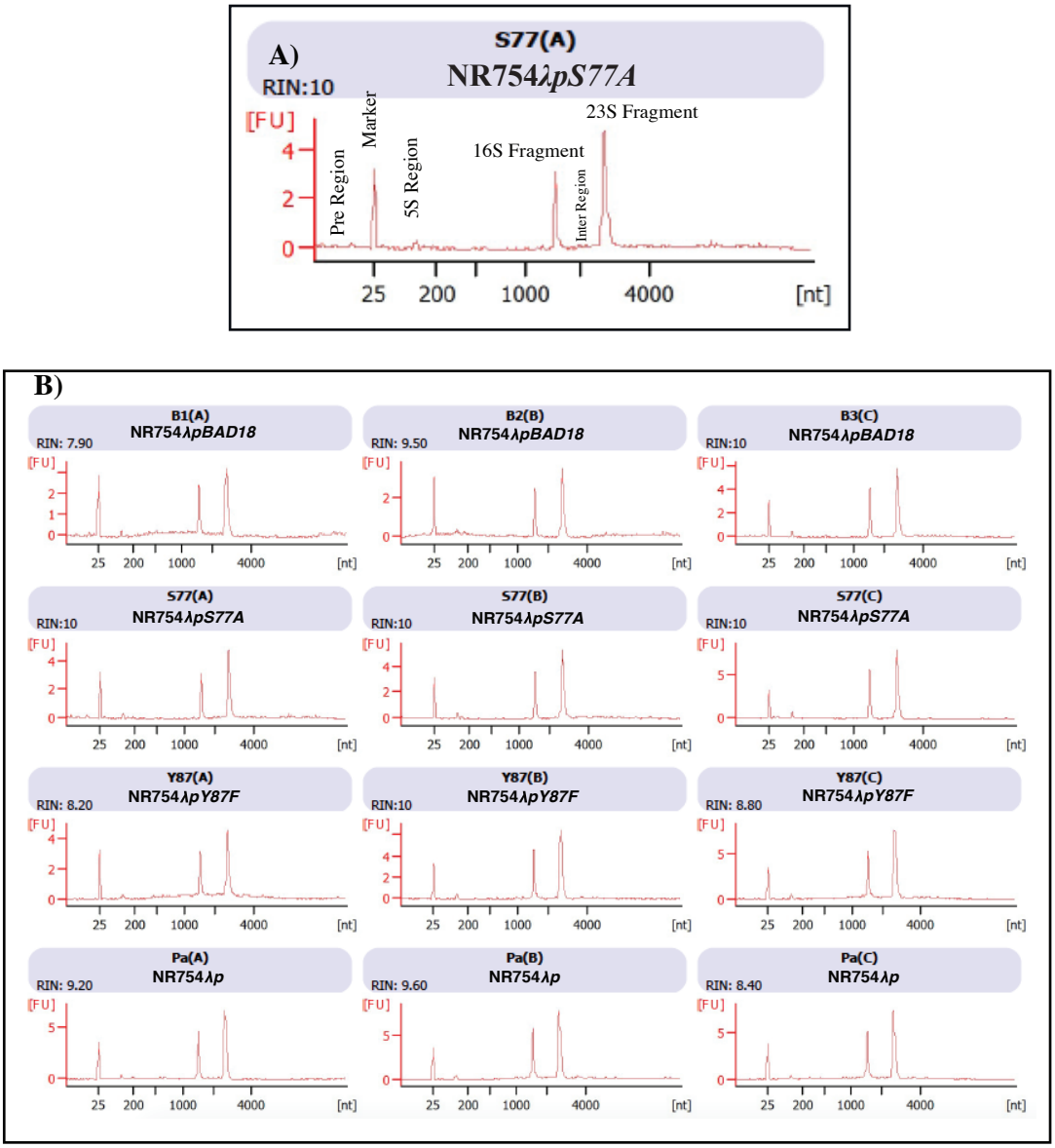

Figure 3.2. Overview of RIN electropherograms using the Agilent 2100 Expert Software. The RIN was measured from 1 to 10, with 1 representing completely degraded RNA and 10 representing RNA of the highest quality. A) The RNA extracted from strain NR754*λpS77A* showed the highest quality and it was selected as a representation of the different regions that indicate the RNA quality. The charts plot fluorescence in the Y axis and nucleotides in the X axis and they show the Pre-region, marker, 5S region, fast region, 16S fragment, inter region, 23S fragment for all the samples. B) Shows the electropherograms for all 12 strains examined. NR754*λpBAD18* showed the lowest value of RIN in this case a reduction of the 16S and 23S peaks is observed along with an increase of the baseline between the two peaks and the marker.

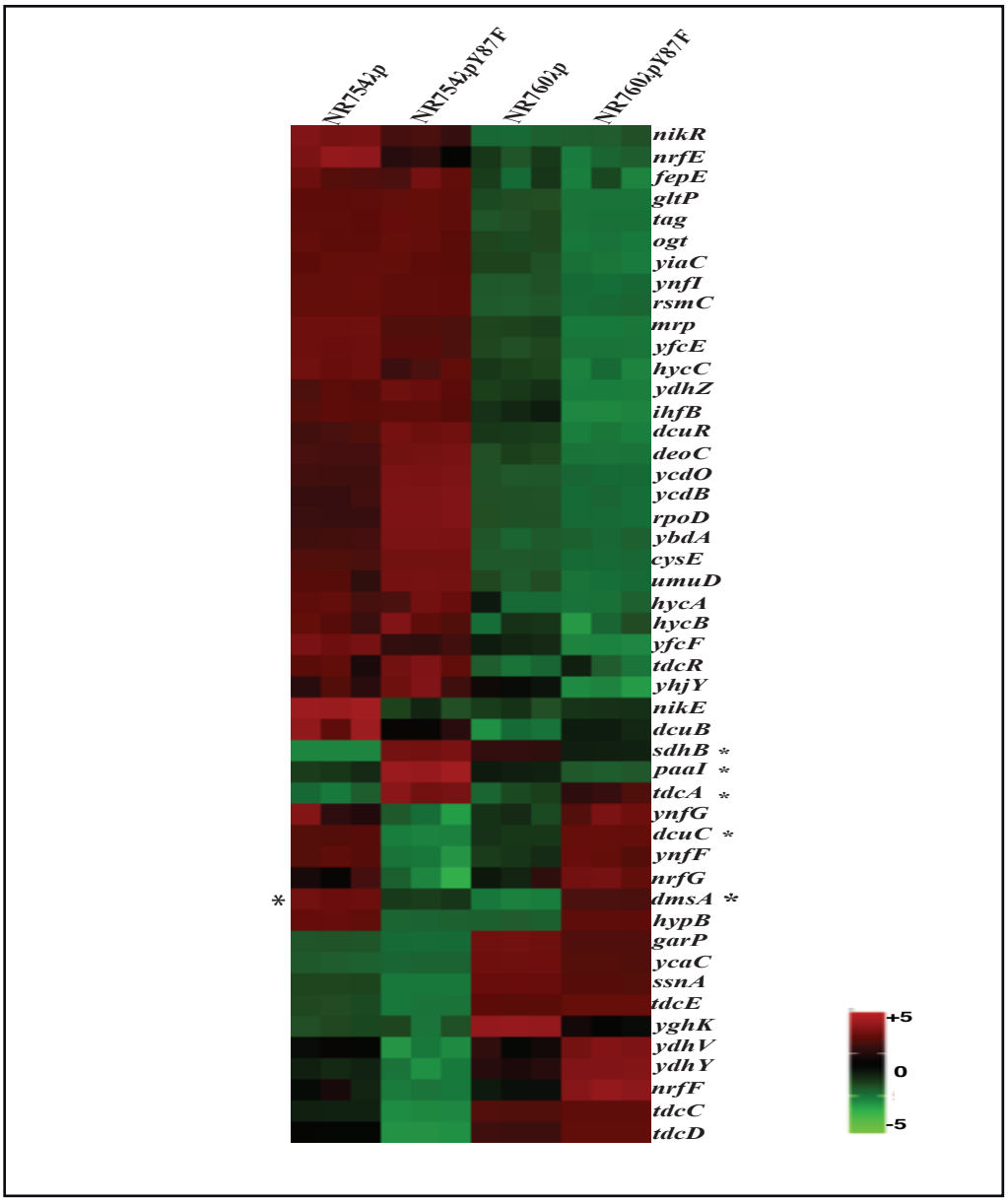

Figure 3.3. Agglomerative Cluster Heat Map. Each column of the grid represents a sample, while each row represents a gene. Red and green color represent upregulated and downregulated genes respectively. The black color represents median values (not absolute values). Based on the similarity of the gene expression pattern the genes are shown in cluster. On the left side of the map a comparison of the differential gene expression between the intact membrane strains NR754*λp* and its mutant NR754*λpY87F* is shown. On the right side the differential gene expression between the *lptD4213* background strains NR760*λp* and NR760*λpY87F* is shown. Genes *sdhB*, *paaI*, *tdcA*, *dcuC* and *dmsA*, marked with asterisks showed a differentially expression of ≤-2 and ≥+2 Log2FoldChange and a p≤0.05.

Table 3.4. Normalized read counts for strains NR754*λp,* NR754*λpY87F,* housekeeping genes, negative and positive controls.

Gene name	NR754*λp*	NR754*λp*	NR754*λp*	NR754*λp Y87F*	NR754*λp Y87F*	NR754*λp Y87F*
tag	11602.3	11564.6	11428.6	11901.17	11844	11542.2
cysE	18374.37	18148.32	17791.22	22342.99	22278.69	22533.37
dcuB	683.54	615.59	703.57	522.2	522.06	553.39
dcuC*	34278.67	34552.06	34833.82	13953.13	13549.94	13650.95
dcuR	8210.99	8381.95	8487.86	9558.85	9298.28	9421.68
deoC	72168.41	71563.91	71641.7	80862.92	81111.1	82417.89
dmsA*	21504.54	21596.62	21819.07	10639	10484.79	10859.3
fepE	140.51	122.61	120.65	118.09	146.34	131.87
garP	11847.41	11927.32	11790.56	9717.03	9614.68	9691.3
gltP	31549.77	31669.53	32134.39	32567.05	33077.25	31999.79
hycA	264.12	266.92	242.76	247.02	283.44	273.16
hycB	140.51	134.1	123.56	153.84	138.43	131.87
hycC	379.28	363.99	370.68	297.94	317.72	350.87
hypB	11151.86	11203.27	10756.67	10577.25	10648.26	10690.93
mrp	63186.18	63281.58	63819.6	57239.35	57126.47	56445.29
nikE	322.23	318.01	334.34	101.84	118.65	95.37
nikR	4195.3	3928.52	3980.09	2862.35	2962.32	2618.57
nrfE	3446.25	3682.03	3628.31	2835.26	2900.36	2626.81
nrfF	812.44	868.46	770.43	629.46	639.4	631.09
nrfG	557.82	545.34	601.81	459.36	427.14	392.08
ogt	2177.41	2102.19	2084.53	2202.56	2172.63	2046.35
paaI*	824.06	841.64	901.26	2498.32	2433.66	2623.28
sdhB*	42612.21	42467.85	42615.21	114587.04	115216.66	117952.28
ssnA	384.56	360.16	344.52	328.27	340.13	339.1
tdcA	4660.15	4477.69	4769.43	7174.28	6854.07	6925.56
tdcC	77670.57	76985.42	76648.06	26587.75	26829.62	27155.91
tdcD	7434.48	7361.5	7329.3	2098.55	2201.63	2221.78
tdcE	7211.56	6951.54	7069.1	3900.25	4030.18	4077.39
tdcR	75.01	76.63	63.96	80.17	83.06	76.53
umuD	1843.56	1831.44	1574.3	2077.96	2108.03	2074.61
ybdA	1538.24	1591.33	1644.08	3991.25	3977.44	3960.82
ycaC	1099.8	1090.69	1109.14	510.28	518.11	512.18
ycdB	1826.66	1840.38	1995.86	3815.74	3779.69	3958.47
ycdO	2873.64	2842.94	2821.53	5576.27	5506.72	5748.14
ydhV	236.65	255.43	244.21	137.59	156.88	143.64
ydhY	295.82	284.8	292.18	198.26	170.07	191.92
ydhZ	8263.82	8707.62	8544.56	9549.1	9401.11	8632.81
yfcE	9283.32	9166.12	9301.91	8225.18	8225.15	7922.83
yfcF	2435.2	2365.28	2417.42	2035.71	2013.11	2102.86
yghK	255.67	261.82	260.2	263.27	220.16	253.14
yhjY	894.84	968.08	895.45	1019.48	1052.04	937.22
yiaC	3466.32	3590.07	3581.79	3587.14	3501.52	3436.88

Table 3.4. Continued

ynfF	1575.22	1629.65	1585.93	941.48	926.8	844.21
ynfG	1634.38	1463.62	1439.11	1236.16	1201.01	1111.48
ynfI	11204.01	11291.3	11406.79	11100.53	11041.13	10929.95
ihfB	80134.29	82152.77	81325.91	82937.63	82641.69	80413.93
rpoD	24492.46	24157.32	24261.42	36146.61	36204.36	36906.08
rsmC	11085.68	11106.11	11049.19	10563.16	10660.13	10398.93
NEG_A	13	11	7	9	8	8
NEG_B	4	11	6	5	9	9
NEG_C	12	10	6	5	10	12
NEG_D	5	3	4	6	11	3
NEG_E	15	14	6	11	15	13
NEG_F	16	12	10	9	15	7
POS_A	122009	111050	101283	109084	98747	111565
POS_B	18351	16654	15393	16459	15059	16569
POS_C	7977	7316	6791	7260	6367	7067
POS_D	952	903	865	841	761	911

Table 3.5. Normalized read counts for strains NR760*λp,* NR760*λpY87F*, housekeeping genes, negative and positive controls.

Gene name	**NR760*λp***	**NR760*λp***	**NR760*λp***	**NR760*λp Y87F***	**NR760*λp Y87F***	**NR760*λp Y87F***
tag	6061.02	6230.38	6430.51	5335.8	5394.23	5395
cysE	7318.69	7326.44	7417.68	6554.25	6574.29	6662.48
dcuB	385.19	419.62	412.87	489.96	490.12	480.37
dcuC*	19776.93	19394.71	19445.55	47241.02	47283.96	46988.61
dcuR	5803.77	5748.73	5671.58	4596.57	4741.22	4596.42
deoC	48637.92	50596.03	49783.35	43479.38	43363.79	43639.61
dmsA*	6308.22	6887.55	6088.47	17393.47	17339.58	17319.38
fepE	62.61	49.42	64.95	43.69	59.36	42.84
garP	56532.36	56947.55	56165.61	42985.12	42631.14	42798.2
gltP	11810.35	11583.72	11522.22	8867.22	8891.68	8895.29
hycA	179.67	136.64	136.39	131.8	132.48	141.51
hycB	73.5	89.16	88.14	62.32	75.29	82.61
hycC	175.58	168.62	166.07	120.34	137.55	118.56
hypB	10732.35	10970.28	10754.01	10558.85	10551.79	10389.02
mrp	34682.45	34960.48	35456.41	28184.7	28257.63	28656.3
nikE	104.81	109.51	94.63	108.88	110.04	110.91
nikR	963.67	958.45	1015	1043.67	1054.09	1145.85
nrfE	2313.89	2164.98	2302.77	1960.54	2082.11	2125.72
nrfF	785.36	802.42	804.39	1109.57	1144.58	1130.56
nrfG	519.94	508.78	581.72	650.41	652.29	630.3
ogt	868.39	844.09	862.84	634.65	653.01	620.35
paaI	975.92	957.48	947.27	709.15	690.66	709.08
sdhB*	90240.25	89709.13	88495.38	6171.4	5779.05	5483.51
ssnA	3065.22	3098.23	3052.43	2254.23	2248.62	2209.86
tdcA*	4686.3	4940.5	5023.05	26076.47	26197.1	26410.82

Table 3.5. Continued

tdcC	122523.40	121627.59	121286.26	153417.39	152994.56	152840.75
tdcD	23118.46	22823.41	22724.36	26396.79	26698.94	26458.68
tdcE	46009.61	44733.91	44557.11	51941.16	52161.35	52176.15
tdcR	47.64	44.58	46.39	55.87	47.78	44.37
umuD	982.72	905.14	962.12	805.13	819.52	845.24
ybdA	145.64	108.54	132.67	116.76	105.7	119.33
ycaC	10803.13	10760.95	10990.6	8242.59	8200.3	7924.6
ycdB	449.17	440.94	436.06	324.49	346.78	316.68
ycdO	560.78	525.26	528.84	437.67	405.42	419.94
ydhV	295.36	242.28	264.42	375.35	398.9	393.17
ydhY	396.08	377.95	395.24	593.11	596.54	605.82
ydhZ	4682.22	4831.96	5028.62	3499.18	3537.27	3522.47
yfcE	4024.8	3860.92	4033.1	3153.92	3153.57	3133.12
yfcF	1671.44	1643.6	1617.14	1251.4	1261.14	1240.7
yghK	552.61	558.2	550.18	350.28	334.47	319.74
yhjY	858.86	820.83	811.82	642.53	653.01	621.88
yiaC	1936.86	1932.4	1832.39	1592.36	1556.52	1514.55
ynfF	1099.78	1118.35	1146.75	1675.45	1650.63	1581.1
ynfG	1298.5	1315.08	1261.79	1528.61	1614.43	1588.74
ynfI	4235.78	4208.83	4363.4	3766.36	3717.54	3860.57
ihfB	55806.89	57943.79	59433.29	42525.25	42447.25	43086.57
rpoD	11529.97	11198.01	11299.55	9486.11	9615.64	9392.49
rsmC	2369.69	2330.7	2470.7	2014.98	2103.11	2166.26
NEG_A	5	11	9	19	21	15
NEG_B	12	8	9	15	16	8
NEG_C	3	9	10	12	14	13
NEG_D	4	5	7	3	10	12
NEG_E	5	10	11	10	16	12
NEG_F	19	19	19	29	28	25
POS_A	89983	97913	99442	109649	105129	86255
POS_B	13634	14353	14628	16569	15814	12936
POS_C	6016	6507	6676	7228	6919	5559
POS_D	716	810	803	927	867	691

Table 3.6. Normalized read counts for strains NR754*λp,* NR754*λpS77A*, housekeeping genes, negative and positive controls.

Gene name	NR754*λp*	NR754*λp*	NR754*λp*	NR754*λp S77A*	NR754*λp S77A*	NR754*λp S77A*
tag	8190.04	8128.3	8062.08	11600.28	11187.71	11185.68
cysE	12970.43	12755.73	12550.48	9451.29	9032.14	9026.52
dcuB	482.51	432.67	496.32	1155.17	1149.89	1193.22
dcuC	24197.25	24285.25	24572.85	50333.56	50480.04	51098.7
dcuR	5796.12	5891.33	5987.6	9868.23	9846.8	10341.22
deoC	50943.53	50299.4	50538.25	336726.05	36902.87	36056.81
dmsA*	21504.54	21596.62	21819.07	8626.89	8515.26	8822.48
fepE	99.19	86.18	85.11	108.52	101.2	115.6
garP	8363.06	8383.23	8317.42	66543.1	66621.5	67372.7
gltP	22270.92	22259.24	22668.59	18626.99	18537.39	19152.22
hycA	186.44	187.61	171.25	202.76	184.69	325.24
hycB	99.19	94.25	87.16	131.37	106.26	103.84
hycC	267.73	255.83	261.49	274.16	260.59	307.61
hypB	14931.06	14902.93	14642.39	23643.21	24135.04	24150.41
mrp	44603	44478.08	45020.31	47037.96	45956.38	46161.25
nikE	227.46	223.52	235.85	98.53	97.41	107.76
nikR	2961.45	2761.2	2807.68	1850.56	1750.77	1833.91
nrfE	2432.7	2587.95	2559.52	4243.73	4147.95	4294.8
nrfF	573.5	610.41	543.49	1657.79	1630.59	1816.28
nrfG	393.77	383.3	424.54	1158.03	1186.57	1185.38
ogt	1537.03	1477.55	1470.49	1148.03	1271.33	1252
paaI	581.7	591.56	635.78	523.05	523.68	538.64
sdhB	30079.88	29848.94	30062.08	30076.79	43030.31	44073.52
ssnA	6271.46	6253.14	6243.03	7001	7003.07	6955.54
tdcA	3289.59	3147.19	3364.5	5571.67	5486.33	5621.25
tdcC	54827.5	54109.96	54069.9	235615.3	234182.84	238020.42
tdcD	5247.98	5174.1	5170.32	39111.67	38597.84	38435.71
tdcE	5090.62	4885.96	4986.76	103962.72	101705.16	103711.95
tdcR	52.95	53.86	45.12	91.39	78.43	113.64
umuD	1301.37	1287.24	1110.56	1567.84	1506.62	1728.11
ybdA	1085.84	1118.48	1159.78	129.94	144.21	133.23
ycaC	776.35	766.6	782.42	21625.58	21689.78	21762.01
ycdB	1289.44	1293.53	1407.94	432.65	423.78	470.23
ycdO	2028.49	1998.19	1990.4	446.93	523.71	625.02
ydhV	167.05	179.53	172.28	698.24	761.53	838.58
ydhY	208.82	200.18	206.12	1198.01	1046.16	1232.4
ydhZ	5833.41	6120.24	6027.59	9508.4	9416.7	9349.81
yfcE	6553.08	6442.5	6561.85	6692.58	6731.09	6953.58
yfcF	1719	1662.46	1705.32	2687.31	2709.64	2707.76
yghK	1800.48	1840.02	1803.56	1390.78	1325.73	1418.54
yhjY	631.67	680.43	631.68	2090.45	2007.56	2127.81
yiaC	2446.87	2523.32	2526.71	2944.33	2977.82	3129.01

Table 3.6. Continued

ynfF	1111.94	1145.41	1118.77	1792.02	1838.05	1722.23
ynfG	1153.71	1028.72	1015.2	2079.03	1901.3	1945.59
ynfI	7908.89	7936.2	8046.7	6409.85	6660.25	6469.63
ihfB	56566.63	57741.88	57369.8	105687.62	104570.4	104133.2
rpoD	17289.18	16979.21	17114.75	18215.75	18028.86	18147.09
rsmC	7825.36	7806.04	7794.44	3975.28	4059.4	4049.89
NEG_A	13	11	7	14	11	13
NEG_B	4	11	6	11	16	6
NEG_C	12	10	6	14	14	11
NEG_D	5	3	4	8	7	3
NEG_E	15	14	6	7	11	5
NEG_F	16	12	10	23	21	16
POS_A	122009	111050	101283	108270	105479	102494
POS_B	18351	16654	15393	16209	15703	15360
POS_C	7977	7316	6791	7171	7087	6731
POS_D	952	903	865	889	951	842

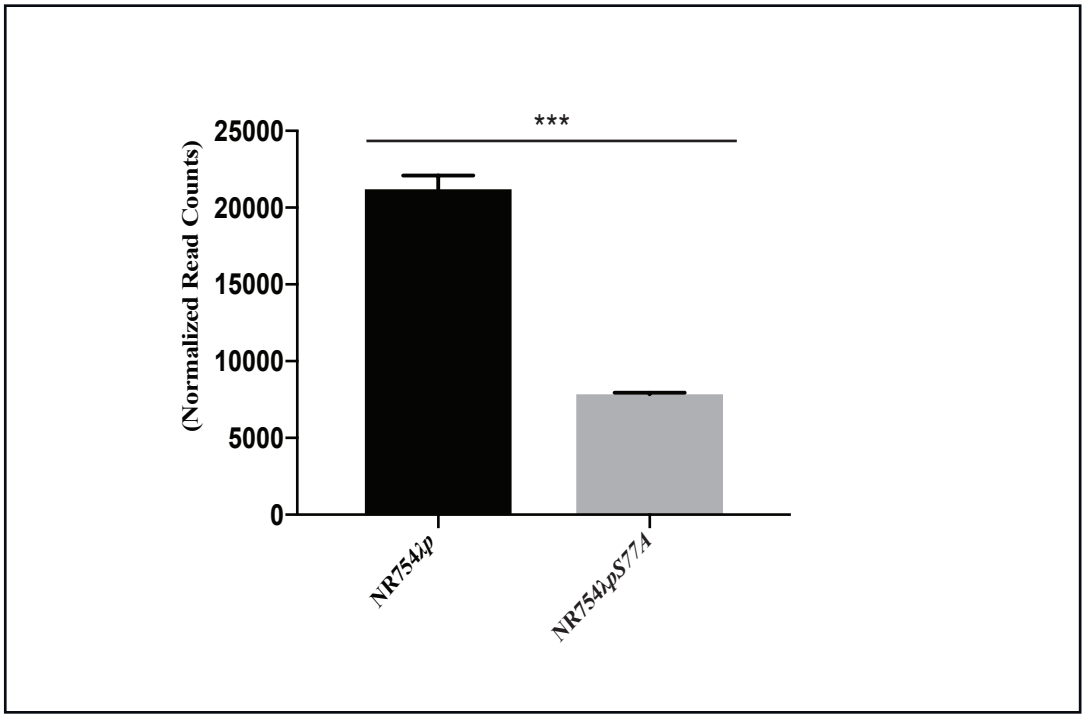

Figure 3.4. Differential gene expression between intact membrane strain NR754*λp* and catalytically inactive NR754*λpS77A*. Results show the mean value of the normalized transcript read abundance of three biological replicas. Gene *dmsA* is upregulated in NR754*λp*. Mutation in the extracellular active site of PagP represses *dmsA* expression by 2.5-fold.

3.7 Discussion

Based on the RNA-seq data we found that 50 genes were differentially expressed between NR760*λp* and the PagP Y87F mutant NR760*λpY87F*. However, NanoString analyses revealed that only *dmsA*, *dcuC*, *tdcA*, *sdhB* and *paaI* were differentially expressed. RNA-seq and NanoString nCounter are reliable platforms for studying gene expression and are generally highly correlative (Passalacqua et al., 2012). Despite the high quality of the extracted RNA, which can be observed in Figure 3.2, the accuracy of the methods, and the identical protocol to extract RNA for both RNA-seq and NanoString nCounter showed that only 5 genes had statistically different expression. The lack of a perfect correlation between the two approaches might indicate that the samples were not collected at the exact same optical density or that the cultures were not perfectly synchronized. In addition, the analysis pipeline used to analyze the RNA-seq data could have affected the results due to the fact that it is complicated to use DEseq2 with a small targeted dataset like ours.

To test gene expression using NanoString nCounter, *E. coli* NR760*λp,* NR760*λpY87F,* NR760*λpS77A* and NR760*λpBAD18* were examined. These strains carry the *lptD4213* mutation that perturbs OM lipid asymmetry and maintains PagP in a constitutively active state. We also used the *lptD*$^{+}$ parental strains NR754*λp*, NR754*λpY87F*, NR754*λpS77A* and NR754*λpBAD18* was used as a negative control.

On the left side of the heatmap in Figure 3.3 we compared *E. coli* NR754*λp* and NR754*λpY87F*. We observed that the *dmsA* and *dcuC* genes were upregulated in NR754*λp*, whereas *sdhB* was downregulated and *paaI* had no change in its expression. The PagP periplasmic Y87F mutation in NR754*λpY87F* triggered the upregulation of *paaI* and *sdhB*. There was no change in the expression of *dmsA* and transcription of *dcuC* was repressed. On the right side of the heatmap we compared the strains with the *lptD4213* allele in which PagP is constituvely activated. We observed that *dmsA*, *dcuC* and *tdcA* are upregulated, whereas *sdhB* is downregulated.

In order to explain the relationship among the differentially expressed genes encoding proteins with unrelated functions we looked for similarities in their modes of regulation. The gene *dmsA* encodes the DmsA subunit of dimethyl sulfoxide reductase, from the *dmsABC* operon, which is an iron-sulfur-molybdoenzyme complex (Sambasivarao et al., 1991). DmsA is the catalytic subunit, reported to be located in the cytoplasm (Sambasivarao et al., 1991). However, further studies reported that DmsA is translocated to the periplasmic side of the inner membrane since it possesses a Tat (Twin arginine translocation) signal peptide; this allows the folded protein to be translocated by the Tat translocon to the periplasm (Berks, B., 1996; Tang et al., 2011; Tang et al., 2013; Stanley et al., 2002). DmsA contains a molybdo-*bis*(pyranoterin guanine dinucleotide) (Mo-*bis*PGD) cofactor and a [4Fe-4S] cluster. DmsA expression is positively regulated by FNR under anaerobic

conditions, and is negatively regulated by IhfA, IhfB, Fis, ModE and NarL (Bearson et al., 2002).

The *dcuC* gene, transcribed divergently from *pagP*, encodes an inner membrane protein involved in anaerobic C4 dicarboxylate transport whose primary function is in succinate transport during glucose fermentation. It is controlled by the DNA-binding transcriptional regulators FNR and ArcA (Zientz, 1996).

The *sdhB* gene encodes the SdhB subunit of succinate dehydrogenase and is positively regulated by CRP and Fur under aerobic conditions. SdhB is activated by phosphorylated ArcA and it is also regulated by the σ^E-dependent small regulatory RNA RyhB, which inhibits SdhB translation (Keseler et al., 2011). SdhB is also regulated by transcriptional regulator FNR (Cecchini et al., 2001) in exponential phase (Rahman et al., 2006).

The *paaI* gene encodes the phenylacetyl-CoA thioesterase PaaI. This protein is involved in the catalysis of the aerobic phenylacetate degradation pathway (Keseler et al., 2011; Ferrandez et al., 1998). It is located in the cytoplasm and its expression is positively regulated by SlyA, IhfB, IhfA and CRP.

The *tdcA* gene encodes the first gene of the *tdcABC* operon. It is the transcriptional activator of the operon (Hagewood et al, 1994). Thus, TdcA regulates its own transcription along with TdcR and the transcriptional regulator FNR. All aforementioned genes encoded for different proteins and perform seemingly unrelated functions. However, they have as a common feature in that they are positively or negatively regulated by transcription factors FNR or CRP.

The remaining genes analyzed did not show any statistically significant difference in expression.

Our results suggest that mutation of PagP in either of its two active sites affected the expression of a similar set of genes but in an opposite manner. The increased expression of *dmsA* found in *E. coli* NR754*λp* indicates that PagP controls DmsA expression under aerobic conditions. Indeed, the PagP S77A mutation eliminates palmitoylation of lipid A and PG, and mutation of this residue repressed *dmsA* expression by 2-fold suggesting that palmitoylated lipid A, PPG, and lysophospholipids might be involved in PagP-mediated regulation of *dmsA*. Additionally, the combination of the *lptD4213* mutation with the periplasmic mutation Y87F in NR760*λpY87F* also affected the expression of the *dmsA* gene. Since PagP is a multifunctional enzyme we speculate that it might control a metabolic pathway where the active sites located on opposite sides of the membrane catalyze discreet enzymatic steps leading to the production of a lipid second messenger capable of diffusing across the cell envelope. Mutational inactivation of either active site would affect the production of this same second messenger. We presume that known envelope stress response regulators like σ^E, Cpx, and Rcs could function as downstream effectors of such a sensory transduction pathway. However, further research must be performed in order to confirm these suspicions.

The *lptD4213* mutation activates σ^E (Lima et al, 2013), which forms an interconnected network with the Cpx and Rcs phosphorelay systems (Grabowicz

and Silhavy, 2017). The expression of LptD is also regulated by σ^E (Dartigalongue et al., 2001). The *lptD4213* mutation affects the translocation and assembly of LPS, which causes the stress response to activate σ^E (Lima et al., 2013). σ^E regulates the transcription of genes responsible for LPS and OMP biogenesis and transport to the OM; thus, defects in either pathway will activate σ^E (Lima et al., 2013). The σ^E and Cpx regulators are interconnected, but they also antagonize each another under specific circumstances (Grabowicz and Silhavy, 2017). σ^E responds to signals of OM dysfunction as well as to accumulation of LPS. Any LPS derivative such as phosphorylated *N*-acetyl glucosamine and *N*-linked acyl chains can bind to RseB inhibiting the binding between RseB and RseA, allowing the RseA cleavage by DegS, which results in the release of σ^E to the cytoplasm with its consequent activation (Lima et al., 2017).

To maintain adequate cell energy, Cpx downregulates the expression of the *rpoE* operon, which represses σ^E expression. Cpx also releases small regulatory RNAs to inhibit the synthesis of outer membrane proteins (Grabowicz and Silhavy, 2017). Among the regulatory RNAs induced by Cpx and Rcs is RprA, which regulates RpoS, the master stationary phase regulator (Majdalani et al., 2001). RpoS and FNR are related in a network, which allows for feedback inhibition and feedforward activation of RpoS and FNR and vice versa with consequent activation or repression of FNR and RpoS dependent genes (Patten et al., 2004). This could explain why the same genes were affected in our samples, and why some genes were upregulated while others were downregulated. Thus, it is likely that *dmsA*,

dcuC, *tdcA*, *sdhB* and *paaI* are the genes affected most during exponential growth under aerobic conditions. These genes belong to operons and we were expecting the entire operon to be differentially expressed; however, only one member of each operon showed transcript read values above the cut off, which suggest that only these genes were affected the most.

The genes *dmsA*, *dcuC* and *tdcA* encode proteins that function during anaerobic respiration and are regulated by FNR along with other TFs (Keseler et al., 2011). With the exception of *paaI,* all these genes are regulated by a common transcription factor, FNR, known to have a close feedback inhibition and feedforward activation with RpoS (Patten et al., 2004). The particular case of *paaI*, whose product belongs to the acyl-CoA thioesterase subfamily, is the only gene that is not directly regulated by FNR, however, RpoS is known to regulate genes which encode enzymes related to Acyl-CoA metabolism and the TCA cycle during exponential phase growth (Rahman et al., 2006).

PagP-mediated signal transduction is likely to occur via its connection with the envelope stress regulators σ^E, Cpx, and Rcs because LPS triggers the activation of σ^E (Lima et al., 2011; Grabowicz and Silhavy, 2017; Tam and Missiakas, 2005). Our results demonstrate that *E. coli* NR754*λpS77A* repressed *dmsA* expression by 2-fold. We suspect that PagP activation is sensed by σ^E and is followed by a cascade of events, which include activation of Cpx, Rcs, and RprA with the consequent regulation of RpoS and FNR-RpoS feedback. Conversely, if the periplasmic PagP

Y87F mutation could be sensed by Cpx, or one of the aforementioned envelope stress response regulators, the same genes could be differentially expressed, but in an opposite manner.

Chapter 4. Reporter genes. Role of PagP on DmsA transcriptional regulation

4.1 Preface

The research in this chapter was carried out from 2016 to 2018. The experiments were designed by Dr. Russell Bishop and myself. All experiments were conducted and analyzed by myself. I supervised an undergraduate, Daniella Bui, who helped me prepare the pRS551-*dmsA-lacZ* reporter gene. Dr. Tracy Raivio and Yun Peng from the University of Alberta provided us with the regulatory deletion mutants to study the potential regulators involved in PagP mediated transcriptional regulation.

4.2 Abstract

Gram-negative bacteria have developed defense mechanisms to send signals across the cell envelope to the cytoplasm. In *E. coli*, the outer membrane enzyme PagP covalently palmitoylates lipid A and phosphatidylglycerol. PagP possesses an extracellular catalytic active site residue Ser77 mainly involved in the transpalmitoylation reaction. PagP also possesses a periplasmic active site residue Tyr87 believed to be involved in cell signalling. Differential gene expression was measured by RNA-sequencing and confirmed using NanoString nCounter. Differential expression of the gene *dmsA* was confirmed by both methods, so *dmsA* was selected to study PagP transcriptional regulation using a *lacZ* reporter gene. Transcriptional *dmsA-lacZ*-reporter genes were constructed using the low copy number plasmid pRS551-*lacZ*. Strain NR754Δ*pagP* displayed low β-galactosidase activity from the *dmsA-lacZ* reporter, but complementation using the single copy

plasmid pBADGr, expressing PagP under the control of an arabinose-inducible promoter, restored the significant β-galactosidase activity observed in the *pagP*$^{+}$ parental strain under aerobic conditions. β-galactosidase activity from the *dmsA-lacZ* fusion provided evidence that membrane perturbation with pentamidine represses *dmsA* transcription in *E. coli* MC4100. Since the *pagP* gene has been previously reported to be coupled phenotypically to the downstream *cspE* gene, we tested for any polar effects associated with the Keio Δ*pagP* deletion. While partial complementation could be traced to *csp*E, full complementation required only *pagP*. The construction of *pagP-lacZ* and *cspE-lacZ* reporter genes allowed us to monitor the expression of these genes in our various mutant constructs. Through chromosomal deletions of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB*, and *fadD* in *E. coli* MC4100 we showed that the products of these genes are all involved in PagP-mediated regulation of *dmsA*. The *dmsA-lacZ* reporter provides us with a useful probe to screen for intermediates in the signal transduction pathway that connects the activation of PagP in the OM to the transcription of *dmsA* in the cytoplasm.

4.3 Introduction

Gram-negative bacteria have developed defense mechanisms to send signals across to cell envelope to the cytoplasm. In *E. coli*, the OM enzyme PagP incorporates a phospholipid-derived palmitate chain onto lipid A and phosphatidylglycerol in order to fortify the OM permeability barrier. Previous research suggests that PagP has a role in signal transduction (Smith et al., 2008). It is known that PagP possesses an extracellular catalytic active site mainly involved

in the transpalmitoylation reaction. The amino acid residue Ser77 located in extracellular loop L2 of PagP plays an important role in this catalytic activity (Bishop, 2005). PagP also possesses a cluster of periplasmic amino acid residues D61, H67 and Y87, which are believed to be involved in PagP cell signalling. The objective of this study is to test whether or not PagP is involved in transcriptional regulation and also to determine how the two active sites are involved in this mechanism. To test if Tyr87 is involved in PagP signaling we carried out an RNA-sequencing study of differential gene expression by comparing *E. coli* NR760*λp* to the mutant NR760*λpY87F*, where PagP in both cases is constituvely activated by the *lptD4213* allele. Results showed that 50 genes were differentially expressed; 40 were upregulated and encoded proteins whose functions are related to anaerobic processes, and 10 were downregulated and encoded proteins whose function are related to aerobic processes. To validate the results from RNA-seq we used NanoString nCounter to study gene expression. Of those 50 genes found differentially expressed from the RNA-seq we selected 45 to be analyzed using NanoString nCounter. Comparison of NR760*λp* and NR760*λpY87F* showed that only *dmsA*, *dcuC*, *tdcA*, and *sdhB* were overexpressed in NR760*λpY87F* with a $\geq$ 2-fold change and a p value of ≤ 0.05. We also found a 2-fold repression of *dmsA* expression between NR754*λp* and the catalytically inactive strain NR754*λpS77A*. Since by all measures the differential expression of *dmsA* was identified, we selected the *dmsA* gene to study PagP transcriptional regulation using *lacZ* transcriptional fusion assay.

The differential expression due to the PagP Y87F mutation was observed in *lptD4213* strains, which constitutively activates PagP in the OM. In order to activate PagP in the *lptD*$^{+}$ parental strain, we perturbed the OM with pentamidine. Pentamidine was recently shown to permeabilize the OM through its association with LPS by disrupting lateral interactions between lipid A molecules without affecting bacterial growth (Stokes et al., 2017).

Transcription in this work was monitored by β-galactosidase activity quantified using Miller Units (MU). β-galactosidase is encoded by the *lacZ* gene of the *lac* operon and cleaves lactose into glucose and galactose. Our gene of interest for this project is *dmsA* from the *dmsABC* operon, which encodes dimethyl sulfoxide reductase. DmsA is the catalytic subunit and contains a molybdo-*bis*(pyranoterin guanine dinucleotide)(Mo-bisPGD) cofactor and a [4Fe-4S] cluster (Weiner et al., 1992). The DmsA subunit was reported to be located in the cytoplasm based on histological, immunological and biochemical results (Sambasivarao & Weiner, 1991; Sambasivarao et al., 1991; Weiner et al., 1992; Weiner et al., 1993; Weiner et al., 1998). However, DmsA carries a twin-arginine leader peptide and reporter protein experiments later demonstrated that the DmsAB complex is translocated to the periplasm. The twin-arginine leader peptide allows the folded complex DmsAB to be translocated through the inner membrane to the periplasm (Berks, 1996; Stanley et al., 2002; Tang et al., 2011; Tang et al., 2013). DmsA expression is regulated by a molybdate-responsive transcriptional regulator and by FNR, a DNA-binding transcriptional regulator that mediates the transition

from aerobic to anaerobic growth (Bearson et al., 2002). To determine the transcriptional influence of OM enzyme PagP on DmsA, a low copy plasmid pRS551 was used as the *lacZ*-reporter gene (Simons et al., 1987).

Our results indicate that PagP exerts regulation on *dmsA* transcription. Given the location of PagP in the OM we suspect that this mechanism of regulation identifies a novel OM sensory transduction pathway. It is known that cold shock protein CspE is the product of *cspE* gene, located downstream of *pagP,* and whose promoter overlaps with *pagP* unidirectionally for 22 nucleotides (Hu et al., 2012). The repression of expression of CspE by DinJ, directly binding to its palindromic sequence located in the last portion of *pagP*, represses the stationary phase general stress regulon RpoS (Hu et al., 2012). Overexpression of CspE along with CrcA(PagP) and CrcB, protects the cells from camphor, which influences chromosomal condensation (Hu et al., 1996; Sand et al, 2003). CspE also has the ability to form dimers and, together with its single-stranded nucleic acid-binding abilities, might play a role in DNA condensation (Jonhston et al, 2006). CrcB, encoded by the *crcB* gene, located downstream of *cspE*, is an integral membrane protein with dual topology and it has been reported to function as a fluoride transporter, thus reducing the concentration of this anion inside the cell (Baker et al., 2012; Stockbridge et al., 2015). The soluble cold shock protein CspE is constituvely expressed and functions as an antiterminator through its RNA-binding capabilities (Johnston et al., 2006), and it has also been reported to be a RpoS regulator (Hu et al., 2012).

It remains a challenge to understand the response of the signal transduction pathways to each stress condition in bacterial cells, even more so when multiple pathways are integrated. Our previous results demonstrated that PagP regulates the transcription of *dmsA* under aerobic conditions. However, there are many questions as to how this is happening and what are the possible second messengers from the cell surface in this regulation. Recently, sensory transduction was implicated from the OM phospholipase PldA, which liberates fatty acids believed to function as second messengers by virtue of their transport to the cytoplasm for lipid recycling (May and Silhavy, 2018). We address similar ideas and also tested the two-component systems CpxAR, PmrAB, PhoPQ, and Rcs, along with the global stress regulator RpoS, and co-enzyme A thioesterification enzyme FadD, to test whether they are involved as regulators of PagP-mediated transcriptional control of the *dmsA* promoter.

4.4 Materials and Methods

Plasmids and genetic constructs used in this project are shown in Table 4.1.

Table 4.1. Plasmids used in this project.

Plasmids	Description	Source
pNPTS-*efp-lacZ*	*lacZ* cloned in the *Psp*OMI and *Spe*I sites of pNPTS138-R6KT, 7.5 kb; Kan^R	(Fried et al., 2012)
pRS551-*lacZ*	ColE1 *ori*, *bla-kan-T1*$_4$-*Bam*HI-*Sma*I-*Eco*RI-*lacZ*$^+$, 12.5 kb	(Simons et al., 1987)
pRS551-*dmsA-lacZ*	ColE1 *ori*, *bla-kan-T1*$_4$-*Bam*HI-*Sma*I-*Eco*RI-*lacZ*$^+$, 12.5 kb	This study
pRS551-*cspE-pagP*-1-*lacZ*	ColE1 *ori*, *bla-kan-T1*$_4$-*Bam*HI-*Sma*I-*Eco*RI-*lacZ*$^+$, 12.5 kb	This study
pRS551-*cspE-pagP*-2-*lacZ*	ColE1 *ori*, *bla-kan-T1*$_4$-*Bam*HI-*Sma*I-*Eco*RI-*lacZ*$^+$, 12.5 kb	This study
pRS551-*cspE-pagP*-3-*lacZ*	ColE1 *ori*, *bla-kan-T1*$_4$-*Bam*HI-*Sma*I-*Eco*RI-*lacZ*$^+$, 12.5 kb	This study
pMLBAD	*pBBR*1 *ori*, *araC*-PBAD *Tpr mob*	(Guzman et al.,1995)
pBADGr	BR1 *ori*, *araC*-PBAD *Gm mob*	(Guzman et al.,1995)
pBADGr-*pagP*	BR1 *ori*, *araC*-PBAD *Gm mob*	This study
pBADGr-*cspE*	BR1 *ori*, *araC*-PBAD *Gm mob*	This study

4.4.1 DNA manipulation

Transformations, digestions, ligations were performed as described in Sambrook et al., 1989. Plasmid and genomic DNA were purified using QIAprep and Geneaid Kit respectively.

4.4.2 Construction of pRS551-*dmsA-lacZ*

To verify whether or not PagP exerts regulation on *dmsA,* reporter genes were prepared using the pRS551-*lacZ* with a background β-galactosidase activity of 50 MU. To meet this objective, it was necessary to create the following genetic constructs shown in Table 4.2. Abbreviated names for the constructs are also shown in Table 4.2. The positive control DH5α/*efp*-pNPTS138 was kindly provide by Fried et al., 2012.

Table 4.2. Genetic constructs used to test PagP regulation on *dmsA* transcription.

Genetic construct	Abbreviated name
MC4100/pRS551-*dmsA-lacZ*	MC4100/p*dmsA-lacZ*
NR754/pRS551-*dmsA-lacZ*	NR754/p*dmsA-lacZ*
NR754Δ*pagP*/pRS551-*dmsA-lacZ*	NR754Δ*pagP*/p*dmsA-lacZ*
NR754Δ*pagP*/pRS551-*lacZ*	
NR754*λpagP*/p*dmsA-lacZ*	

The promoter region selected to amplify *dmsA* from *E. coli* MC4100 is shown in Figure 4.1. Two sets of primers were designed to amplify the *dmsA* promoter. The first set was used to amplify the *dmsA* promoter region, which included the FNR binding site, along with consensus sequences for the molybdate-responsive transcription factor ModE, and the integration host factor IHF. The forward primer was designed to include an *Eco*RI site (5'-CGG GAA TTC TTT GTC TCC CTT TGA TAC-3') while the reverse primer included a *Bam*HI site 5'-CGG GGA TCC TCA TCG GAT ATT TCA-3' (Figure 4.2) to yield a 567 bp product (Figure 4.2A). The second set of primers was designed to verify that *dmsA-lacZ* junction. Forward primer 5'-CGG GAA TTC TAA TTA CTC CTC ACT TAC ACG-3' and reverse 5'-CCT GAT CTT CCA AAC TGC CGT C-3' yielded a 1164 bp PCR product (Figure 4.2 B). The amplified *dmsA* promoter (567 bp) was inserted in front of the promoterless *lacZ* between *Eco*RI and *Bam*HII sites in the plasmid pRS551-*lacZ* (Figure 4.3).

4.4.3 Construction of pRS551-*pagP-cspE-lacZ* fusions.

To test whether or not PagP exerts transcriptional regulation on *cspE,* three sets of primers were designed to create three *lacZ* reporters with different fragments of *pagP* and *cspE* promoters. Fragment *pagP-cspE*-1 (-186 to +307) was amplified with the forward primer 5'- CGG GAA TCC GAC TAT TCC CAT CG-3', which includes the *pagP* promoter region and covers from position -186 bp upstream of *pagP*, the consensus sequence of ArcA, and the PhoP, -35 and -10 boxes. The

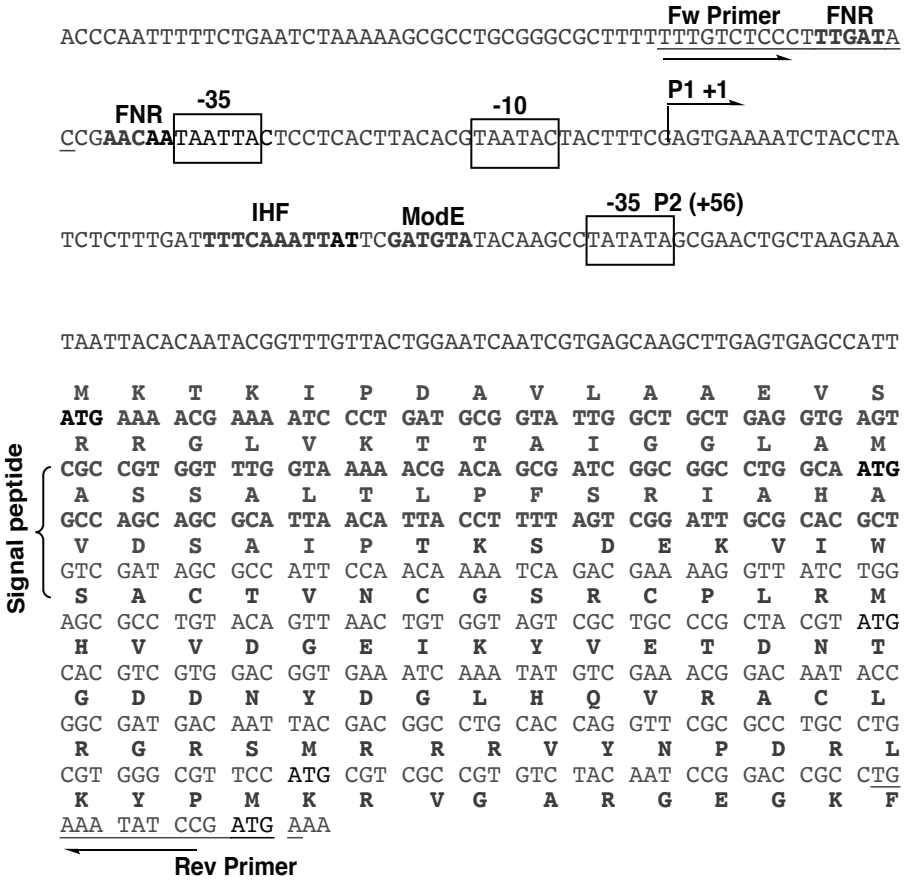

Figure 4.1. Fragment of the MC4100 *dmsA* promoter used for the construction of the pRS551-*dmsA-lacZ*. Forward and reverse primers used to amplify *dmsA* promoter are underlined. The consensus sequence of the recognition sites for RNA polymerase, -35 and -10, are shown in boxes. Consensus sequences of FNR, ModE and IHF are shown on bold. Promoters 1 and 2 are labelled P1 and P2 as reported by McNicholas et al., 1998.

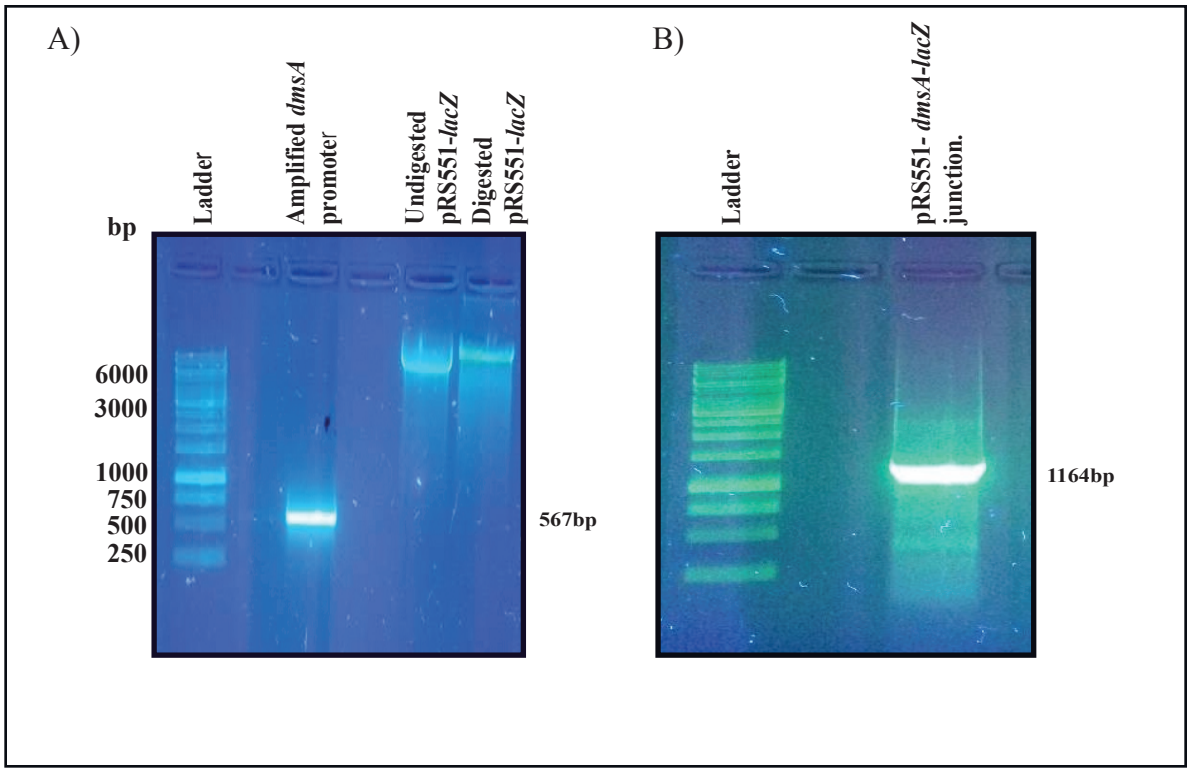

Figure 4.2. PCR amplification of the *dmsA* promoter and *dmsA-lacZ* fusion in the pRS551-*lacZ* plasmid. A) The *dmsA* fragment of the promoter amplified yielded a 567 bp PCR product. On the right side of the panel the undigested and digested plasmids are shown. B) The amplified *dmsA-lacZ* fusion, yielded a product of 1164 bp. The PCR products were separated on 1% agarose gel at 100 V for 30 minutes in 1X TAE buffer and stained with RedSafe Nucleic Acid Staining solution. Visualization occurred using the BLook LED Transilluminator from GeneDireX.

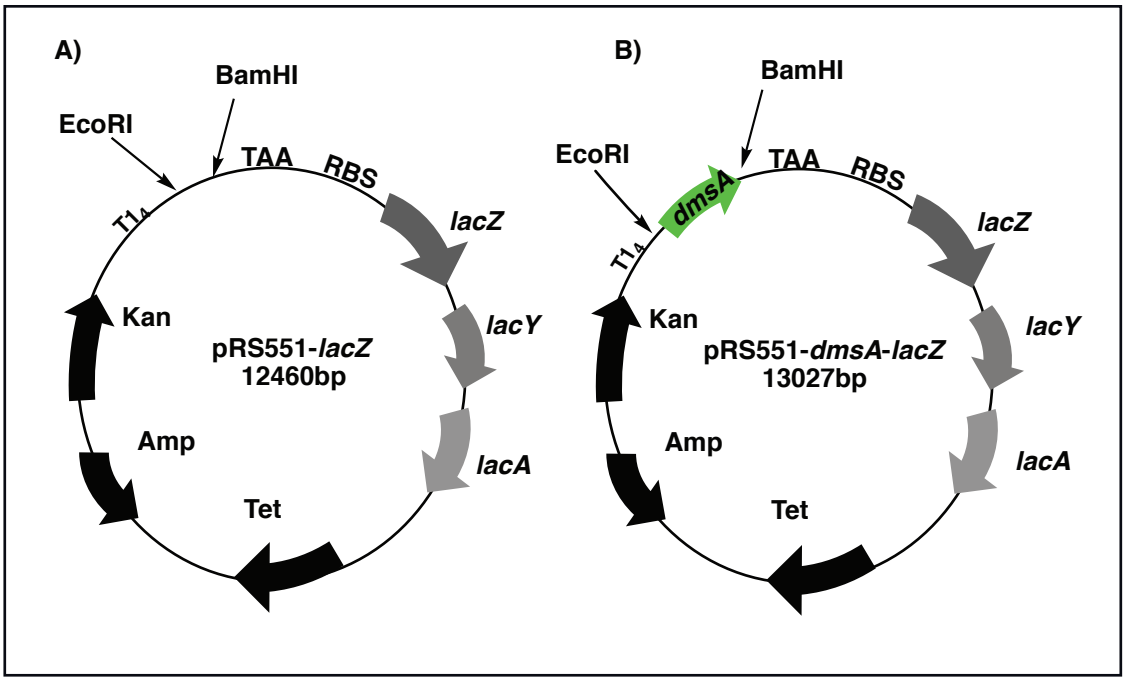

Figure 4.3. Plasmid pRS551-*lacZ* with and without insert. On the left empty pRS551. On the right pRS551-*lacZ* plus the 567 bp insert. Construction of p*dmsA-lacZ* was performed by cloning *dmsA* promoter fragment between the restriction sites *Eco*RI and *Bam*HI.

reverse primer 5'- CGG GGA TCC ACG AGT CCT TAA ATG -3' reaches position +307 bp downstream of *pagP* (Figure 4.4).

To amplify fragment *pagP-cspE* -2 (+390 to +46) forward primer 5'- CGG GAA TTC AAG TAC CTG GCG AC -3' was used, it starts 390 bp upstream *cspE* translational start site, near the *Sph*I site, and includes the binding sites for DinJ, CRP, and the -35 and -10 boxes. The reverse primer is 5'- CGG GGA TCC CTT TGG ACT CAT TAA AC -3'. Fragment *pagP-cspE* -3 (-186 to +46) covers the region from position -186 bp upstream *pagP* to +46 bp downstream *cspE* and yields a 929 bp PCR product (Figure 4.4).

The forward primer used was 5'- CGG GAA TCC GAC TAT TCC CAT CG -3' and the reverse primer was 5'- CGG GGA TCC CTT TGG ACT CAT TAA AC -3'. A diagram of the three fragments of *pagP* and *cspE* promoters is shown in Figure 4.5. The amplified promoter fragments *pagP-cspE-lacZ*-1, *pagP-cspE-lacZ*-2 and *pagP-cspE-lacZ*-3, were cloned between *Eco*RI and *Bam*HI restriction sites into pRS551-*lacZ*, generating a fusion plasmid of 12953, 12896 and 13389 bp, respectively (Figure 4.6).

The different genetic constructs were transformed into the *pagP* deleted strain NR754Δ*pagP*. The list of constructs and their abbreviated names are shown in Table 7. Only the construct that includes *cspE* promoter p*pagP-cspE* -2-*lacZ* was transformed into NR754 and NR754*λp*.

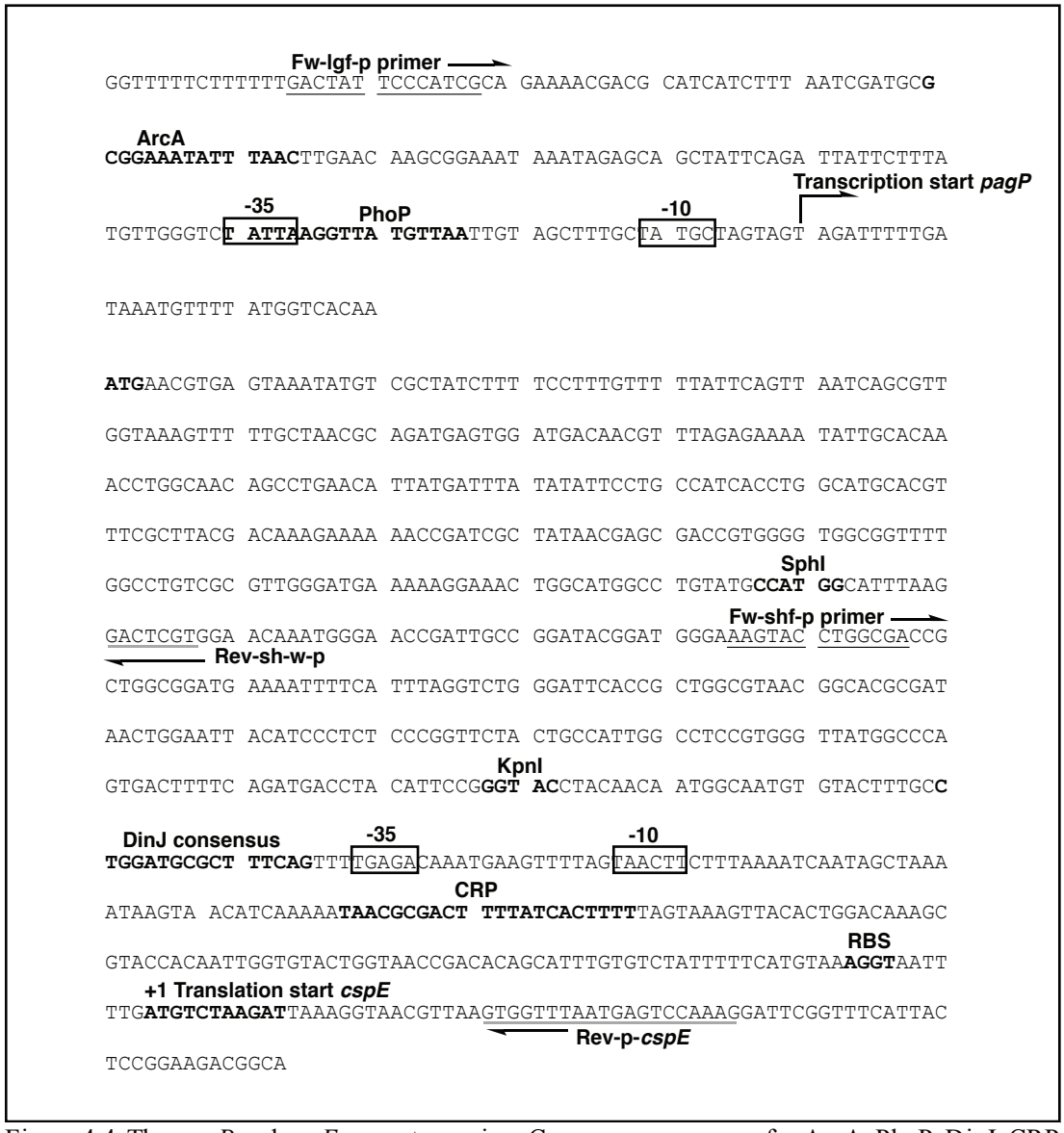

Figure 4.4. The *pagP* and *cspE* promoter region. Consensus sequences for ArcA, PhoP, DinJ, CRP binding and the translational start site of *cspE* are shown in bold. DinJ binding site to regulate *cspE* is located within *pagP*, overlaping by 22 nucleotides (Hu et al., 2011). Last codon of *pagP* gene overlaps with the -35 box of *cspE*. Primers used to amplify the different regions of promoters are underlined. -35 and -10 consensus sequences for *pagP* and *cspE* promoters are shown in boxes. The -35 and -10 consensus sequences of *pagP* were reported by Eguchi et al., 2004.

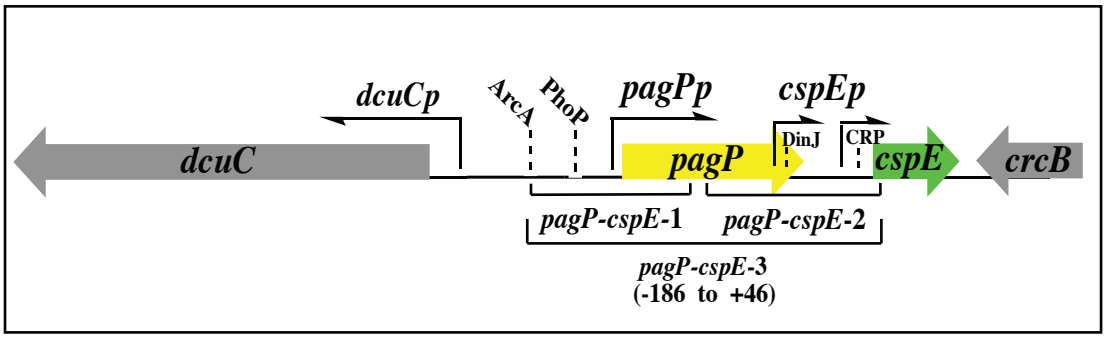

Figure 4.5. Illustration of three different fragments of promoters amplified for the construction of pRS551-*lacZ* reporter genes. Fragment *pagP-cspE*-1 (-186 bp to +307 bp) includes the region upstream of the consensus sequence of ArcA to position +307 bp downstream *pagP*. Fragment *pagP-cspE*-2 (+390 bp to +46 bp), starts 390 bp upstream *cspE* translational start site. It includes the consensus sequence for DinJ binding site and it reaches downstream *cspE* to position +46 bp. Fragment *pagP-cspE*-3 (-186 bp to +46 bp) covers the region from upstream *pagP* to downstream *cspE*.

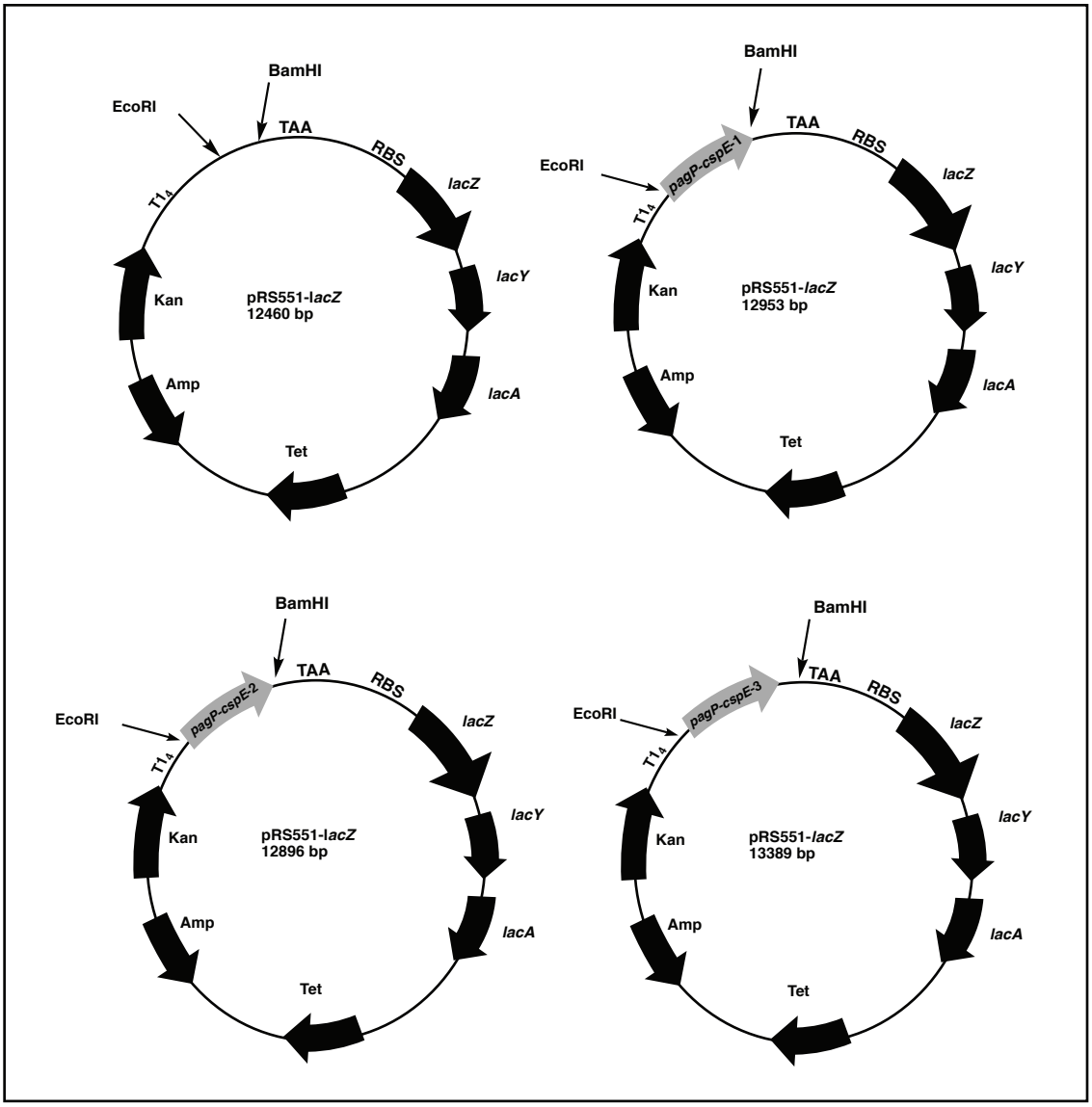

Figure 4.6. p*pagP-cspE-lacZ* reporter. A) Represents the pRS551-*lacZ* transcriptional, low copy vector without the insert, 12460 bp. B, C and D represent the vectors generated by the insertion of fragments *pagP-cspE-lacZ*-1, *pagP-cspE-lacZ*-2 and *pagP-cspE-lacZ*-3, cloned between EcoRI and BamHI restriction sites into pRS551-*lacZ*, generating a fusion plasmid of 12953, 12896 and 13389 bp respectively.

Table 4.3. Genetic constructs used to test PagP regulation on *cspE* transcription.

Genetic construct	Abbreviated name
MC4100/pRS551-*pagP-cspE*-2-*lacZ*	MC4100/p*pagP-cspE*-2-*lacZ*
NR754*λp*/pRS551-*pagP-cspE*-2-*lacZ*	NR754*λp*/p*pagP-cspE*-2-*lacZ*
NR754Δ*pagP*/pRS551-*pagP-cspE*-1-*lacZ*	NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ*
NR754Δ*pagP/*pRS551-*pagP-cspE*-2-*lacZ*	NR754Δ*pagP/*p*pagP-cspE*-2-*lacZ*
NR754Δ*pagP*/pRS551-*pagP-cspE*-3-*lacZ*	NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ*
NR754Δ*pagP/*pRS551-*lacZ*	NR754Δ*pagP/*pRS551-*lacZ*
NR754*pagP*/pRS551-*lacZ*	NR754*pagP*/pRS551-*lacZ*

4.4.4. Construction of pBADGr-*pagP* and pBADGr-*cspE*

To complement the phenotype two constructs were prepared as followed: The amplification of *pagP* gene was carried out from wild-type MC4100, by PCR. The *pagP* gene was amplified using 5'-CGG GAA TCC GAC TAT TCC CAT CG-3' as forward primer and 5'-CGG AAG CTT CTT TGG ACT CAT TAA ACC AC-3' as reverse primer. This 966 bp insert was cloned between *Eco*RI and *Hin*dIII (Figure 4.7).

To create a construct with *pagP* in the reverse orientation the forward primer 5'- CGG AAG CTT GAC TAT TCC CAT CG -3' was used and 5'-CGG GAA TTC CTT TGG ACT CAT TAA ACC-3' as reverse primer.

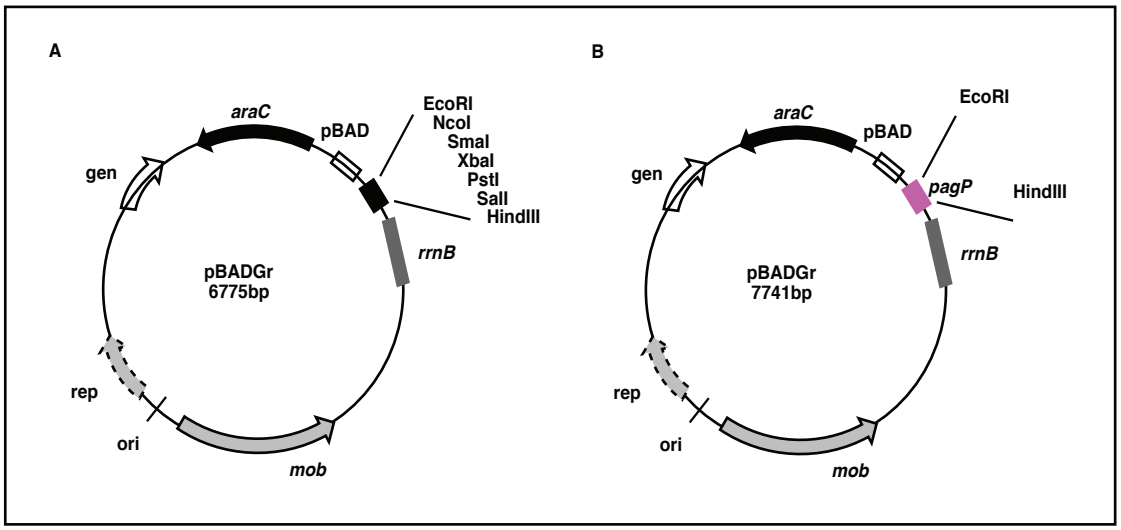

Figure 4.7. Phenotype complementation using single copy plasmid pBADGr. A) Represents the empty plasmid pBADGr. B) Represents *pagP* gene cloned in front of the arabinose promoter, between *Eco*RI and *Hin*dIII in the pBADGr plasmid.

Forward primer (5'-CGG CCA TGG CAT TTA AGG ACT C-3') was designed to include a *Nco*I restriction site and the reverse primer 5'-CGG AAG CTT GCT GAC GTA TCT TAC AG -3' included a *Hin*dIII site, to yield a 663 bp PCR product, containing the *cspE* gene, which was inserted into the inducible arabinose promoter plasmid pBADGr (Figure 4.8). To insert *cspE* in the reverse orientation in pBADGr, we used the primers 5'- CGG AAG CTT CAT TTA AGG ACT C -3' and 5'- CGG CCA TGG GCT GAC GTA TCT TAC AG -3'. These genetic constructs were transformed into NR754Δ*pagP*/p*dmsA-lacZ* (Table 8).

Table 4.4. Genetic constructs used for phenotype complementation.

Genetic construct	Abbreviated name
NR754Δ*pagP*/pRS551-*lacZ*	
NR754Δ*pagP*/pRS551-*dmsA-lacZ*	NR754Δ*pagP*/p*dmsA-lacZ*
NR754λ*pagP*/pRS551-*dmsA-lacZ*	NR754λ*pagP*/p*dmsA-lacZ*
NR754Δ*pagP*/pRS551-*dmsA-lacZ*/pBADGr-*cspE*	NR754Δ*pagP*/p*dmsA-lacZ*/pBADGr-*cspE*
NR754Δ*pagP*/pRS551-*dmsA-lacZ*/pBADGr-*cspE*R	NR754Δ*pagP*/p*dmsA-lacZ*/pBADGr-*cspE*R
NR754Δ*pagP*/pRS551-*dmsA-lacZ*/pBADGr-*pagP*	NR754Δ*pagP*/p*dmsA-lacZ*/pBADGr-*pagP*
NR754Δ*pagP*/pRS551-*dmsA-lacZ*/pBADGr-*pagP*R	NR754Δ*pagP*/p*dmsA-lacZ*/pBADGr-*pagP*R

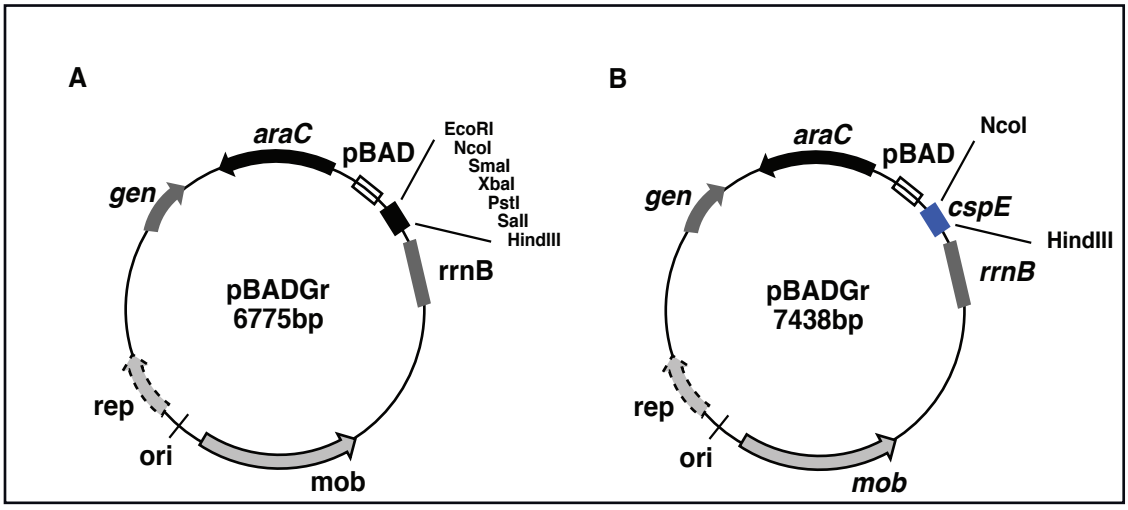

Figure 4.8. Phenotype complementation using single copy plasmid pBADGr. A) Represents the empty plasmid pBADGr. B) Represents *cspE* gene cloned in front of the arabinose promoter, between *Nco*I and *Hin*dIII sites.

4.4.5 PCR

Amplification of promoter fragments was performed using the following conditions: 1minute denaturation at 95 ºC and 35 cycles of amplification (30 seconds denaturation at 95 ºC, 30 seconds annealing at 52.5 ºC, 5 minutes of elongation at 72 ºC), followed by holding at 4 ºC. The PCR amplified products were separated on a 1 % agarose gel at 100 V for 30 minutes in 1X TAE buffer and stained with RedSafe Nucleic Acid Staining solution. Visualization occurred using the BLook LED Transilluminator from GeneDireX. Plasmid DNA was isolated using QIAprep Spin Miniprep kit from Qiagen. The amplified promoter fragments were cloned between the restriction sites *Eco*RI and *Bam*HI. The final plasmids were validated by sequencing.

4.4.6 β-galactosidase activity

A single colony of each transformation was inoculated in LB supplemented with kanamycin and incubated at 37 ºC with shaking at 200 rpm. The cells were grown to OD_{600} 0.2, at which point 80 µg/mL pentamidine was added to the appropriate samples. After 2 hours of incubation at 37 ºC with shaking at 200 rpm, 1 mL of each sample was extracted and immediately incubated on ice to stop cell growth. OD_{600} readings were taken and the samples were normalized to OD_{600} 0.6.

β-galactosidase activity was determined in accordance with Miller, 1972. The normalized samples were first centrifuged at 13,000 xg for 1 minute. The supernatant was discarded, and the pellet was resuspended in 400 uL of permeabilization solution (pH 7.0), followed by the immediate addition of 5.4 µl

/ml of β-mercaptoethanol. All samples were vortexed at the highest speed for 10 seconds and then incubated for 20 minutes on ice. Next, 1 mL of 2-nitrophenyl β-*D*-galactopyranoside (ONPG) (pH 7.0), followed by 2.7 uL of β-mercaptoethanol was added. The reaction was incubated for 7 minutes before adding 700 uL of stop solution (Na_2CO_3, pH 11.0). The samples were centrifuged for 15-20 minutes at 13,000 xg to remove particulate cell debris to perform OD readings at 420 nm.

4.4.7 Analysis of Lipid A by Thin Layer Chromatography (TLC)

Analysis of lipid A by mild acid hydrolysis was adapted from (Jia et al., 2004). An overnight culture grown at 37 °C was diluted 100-fold into 5 ml of LB broth containing 2.5 uCi/ml ^{32}P orthophosphate. Cells were grown at 37 °C until they reached OD_{600} of 0.3. The ^{32}P-labeled cells were harvested by centrifugation and washed once with 5 ml of PBS. The pellet was resuspended in 0.8 ml of PBS and converted into a single-phase Bligh/Dyer mixture by adding 2 ml of methanol and 1 ml of chloroform. After one hour of incubation at room temperature, the pellet was collected by centrifugation in a clinical centrifuge and then was washed once with 5 ml of a fresh single-phase Bligh/Dyer mixture, consisting of chloroform/methanol/water (1:2:0.8, v/v) and centrifuged for 10 minutes. The pellet was then dispersed in 1.8 ml of 12.5 mM sodium acetate, pH 4.5, containing 1% SDS, with sonic irradiation in a bath apparatus. The mixture was incubated at 100 °C for 30 min to cleave the ketosidic linkage between Kdo and the distal glucosamine sugar of lipid A. After cooling, the boiled mixture was converted to a two-phase Bligh/Dyer mixture by adding 2 ml of chloroform and 2 ml of methanol.

Partitioning was made by centrifugation, and the lower phase material was collected and washed once with 4 ml of the upper phase derived from a fresh neutral two-phase Bligh/Dyer mixture, consisting of chloroform/methanol/water (2:2:1.8, v/v). The lower phase lipid A sample was collected and dried under a stream of nitrogen gas. The lipid A sample was dissolved in 100 ul of chloroform/methanol (4:1, v/v), and a 3000 cpm portion of the sample was applied to the origin of a Silica Gel 60 TLC plate. TLC was conducted in a developing tank in the solvent of chloroform, pyridine, 88 % formic acid, water (50:50:16:5, v/v). The plate was dried and visualized with a PhosphorImager (Amersham Biosciences).

4.4.8 Regulators involved in PagP-mediated regulation of *dmsA* transcription

To investigate the regulators involved in PagP-mediated control of *dmsA* transcription, *E. coli* MC4100 carrying the Keio knockout genes Δ*rpoS*, Δ*phoQ*, Δ*rcsB*, Δ*cpxA*, Δ*cpxR*, Δ*pmrA*, Δ*pmrB,* and Δ*fadD* were transformed with the pRS551-*dmsA*-*lacZ* reporter to determine β-galactosidase activity. The knockout strains MC4100 Δ*rpoS*, Δ*phoQ*, Δ*rcsB*, Δ*cpxA*, Δ*cpxR*, Δ*pmrA*, Δ*pmrB* and Δ*fadD* were provided by Dr. Tracy Raivio. These mutants were prepared following published protocols (Sambrook and Russell, 2001; Silhavy et al., 1984).

4.4.9 Lysate preparation

Overnight cultures were prepared for each regulatory mutant containing deletions from the Keio library. The overnight culture was diluted to 1:50 with fresh 5 mL LB supplemented with 5 mM $CaCl_2$ and 0.2% glucose. Cells were grown for 30 min at 37 ˚C with shaking. 100 µL of P1 phage lysate was added to

the culture. The cells were grown for another 2-3 hr at 37 ℃ with shaking. The cultures were checked every 30 min until they lysed. 100 uL of chloroform was added and the samples were vortexed thoroughly. The samples were centrifuged to eliminate the debris at 1,500 xg for 10 min. The supernatant was transferred to a fresh tube. 100 uL of chloroform was added and the samples were stored at 4 ℃.

4.4.10 Transduction

The samples of overnight culture recipients were centrifuged at 1500 xg for 10 min and 5 mL was resuspended in 2.5 mL of 10 mM MgSO4 and 5 mM CaCl2. 100 µL of the recipient cells were added to a tube containing P1 lysate. The samples were incubated for 30 min at 30 ℃ without shaking. 1 mL of LB with 10 mM Na-citrate was added. The samples were incubated for 30 min at 37 ℃ without shaking. The samples were centrifuged 1,500 xg for 10 min. The supernatant was discarded, and the samples were resuspended in 100 µL 1 M Na-citrate and plated on Kan (30 mg/mL) and incubated at 37 ℃ overnight. Colonies were picked and re-streaked on plates Kan (50 mg/mL) and incubated at 37 ˚C overnight. The genotype of the re-streaked colonies was verified by PCR.

4.5 Results

In this work, we focused on the *dmsA* gene because its differential expression was verified both through RNA-seq (Chapter 2) and NanoString (Chapter 3). We previously demonstrated that PagP exerts regulation on *dmsA* transcription. Here, we tested this regulation one more time using the p*dmsA-lacZ* reporter.

4.5.1 Role of PagP on *dmsA* transcription: Phenotype complementation

To investigate the transcriptional regulation of *dmsA* by PagP, we prepared reporter genes using the plasmid pRS551-*lacZ* (Simons et al., 1987). Through the use of the pRS551-*dmsA*-*lacZ* transcriptional fusion we were able to corroborate the results from RNA-seq and NanoString nCounter. Genetic constructs used to analyze differential *dmsA* expression are listed in Table 6. As observed in Figure 4.9 *dmsA* is upregulated in the wild-type strain NR754, showing a β-galactosidase value of 831 MU. The NR754Δ*pagP* strain containing the *pagP* deletion showed significantly lower β-galactosidase activity (60 MU) compared to NR754. The results support the hypothesis that PagP is involved in *dmsA* transcriptional regulation under aerobic conditions (Figure 4.9).

To complement the phenotype, the *E. coli pagP* gene was cloned into the inducible single copy plasmid, pBADGr, in which the gene expression is under the control of the arabinose promoter. The construct was then transformed into NR754Δ*pagP,* which had a β-galactosidase activity of 60 MU. The list of these genetic constructs and their abbreviated names are shown in Table 8. β-galactosidase activity was restored in the NR754Δ*pagP*/p*dmsA*-*lacZ*/*p*BADGr-*pagP* strain, yielding 894 MU, similar to 830 MU found in NR754/p*dmsA*-*lacZ* (Figure 4.9). The strain NR754Δ*pagP*/p*dmsA*-*lacZ*/pBADGr-*pagP*R, carrying the pBADGr plasmid with the *pagP* gene in the reverse orientation, did not restore the activity, showing low β-galactosidase values similar to pRS551-*lacZ* (Figure 4.9).

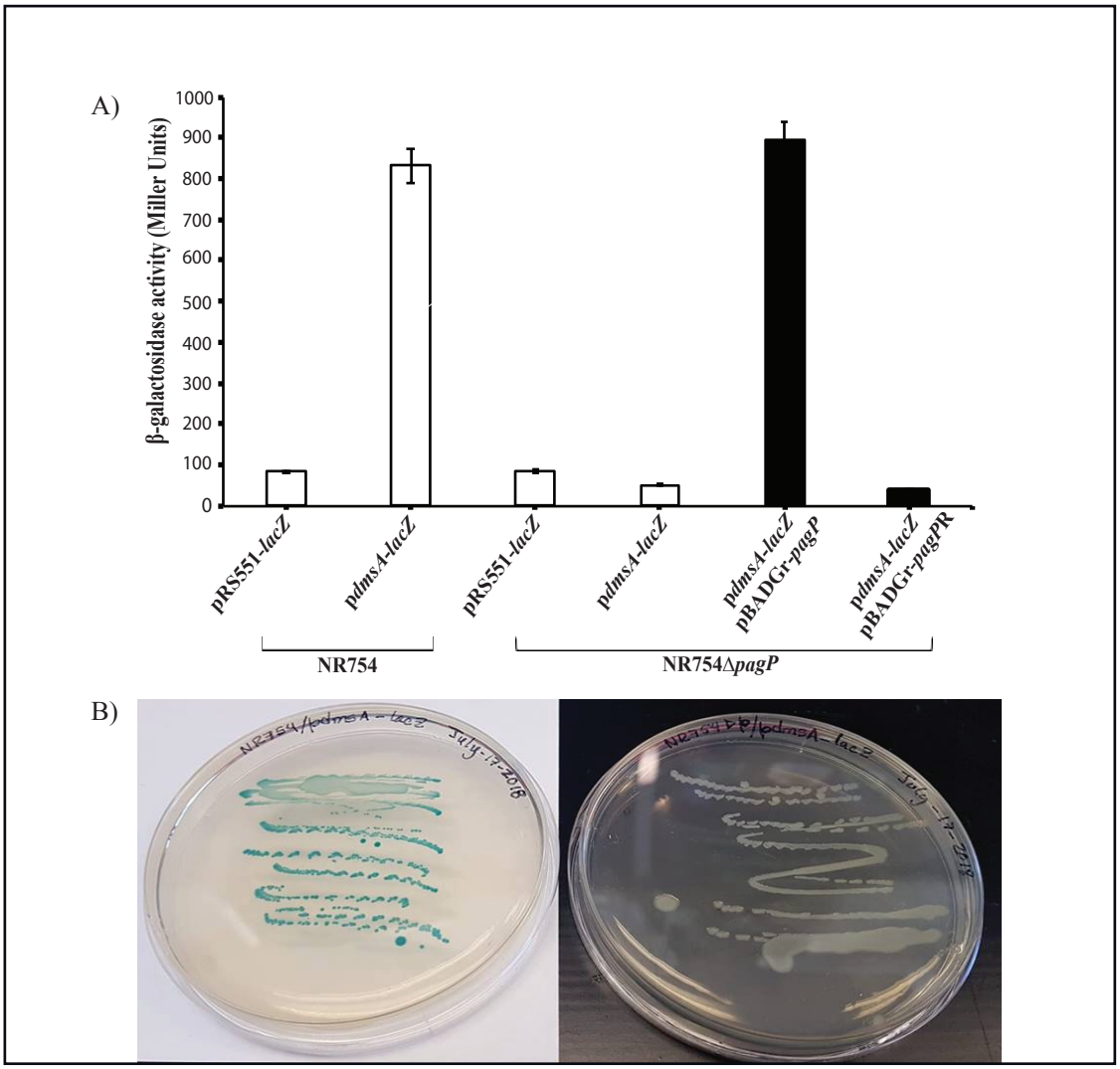

Figure 4.9. Comparison of the β-galactosidase activity between wild-type NR754 and deleted *pagP* strain NR754Δ*pagP*. A) The *dmsA* gene is upregulated in wild-type strain NR754 showing a β-galactosidase value of 831 MU. The strain NR754Δ*pagP* containing the *pagP* deletion showed a significantly lower value of expression (60 MU). The β-galactosidase activity was restored in the strain NR754Δ*pagP*/p*dmsA-lacZ*/pBADGr-*pagP*, by using a pBADGr plasmid carrying the *pagP* gene. The activity was restored without the addition of arabinose, yielding 894 MU. B) The strains NR754/p*dmsA-lacZ* and NR754Δ*pagP*/p*dmsA-lacZ* were grown on X-Gal plates. The blue colonies observed in the plate on the left represent the hydrolysis of the X-Gal by the β-galactosidase expressed in the wild-type NR754/p*dmsA-lacZ*. The white colonies observed in the plate on the right represent no hydrolysis caused by the absence of expression of the β-galactosidase in the *pagP* deleted strain NR754Δ*pagP*/p*dmsA-lacZ*.

4.5.2 Disruption of the bacterial OM with pentamidine

Previously, based on RNA-seq and the nCounter analysis, we had used strains in which the OM was disrupted by the *lptD4213* mutation to keep PagP constituvely activated. Here, we decided to use the *lptD*$^{+}$ strain NR754 and disrupt the membrane using 80 μg/mL pentamidine, which is below the minimum inhibitory concentration (MIC) 200 μg/mL reported for *E. coli* cells (Stockes et al., 2017). Pentamidine contains two aromatic rings linked by an alkyl chain. It is known to have a high affinity for LPS molecules, where the amidine groups interact favorably with the phosphate groups of Kdo_2-lipid A (David et al., 1994) (Figure 4.10). To investigate whether or not β-galactosidase activity from the pRS551-*dmsA-lacZ* transcriptional fusion was affected by pentamidine we measured it for each of the genetic constructs for 2 hours after the addition of 80 μg/mL pentamidine. *E. coli* MC4100/p*dmsA-lacZ* and NR754/p*dmsA-lacZ* (also shown in Table 6) had similar β-galactosidase activities of 743 and 831 MU, respectively (Figure 4.11). The response to 80 μg/mL pentamidine treatment was similar for MC4100/p*dmsA-lacZ* and NR754/p*dmsA-lacZ,* with both strains having a 2.8-fold repression of transcription ($p<0.05$) (Figure 4.11). The positive control, DH5α/pNPTS-*efp-lacZ,* carrying the pNPTS-*lacZ* plasmid driven by the elongation factor P *efp* promoter, showed an activity of 3000 MU and did not exhibit sensitivity to pentamidine treatment where it showed an activity of 2990 MU.

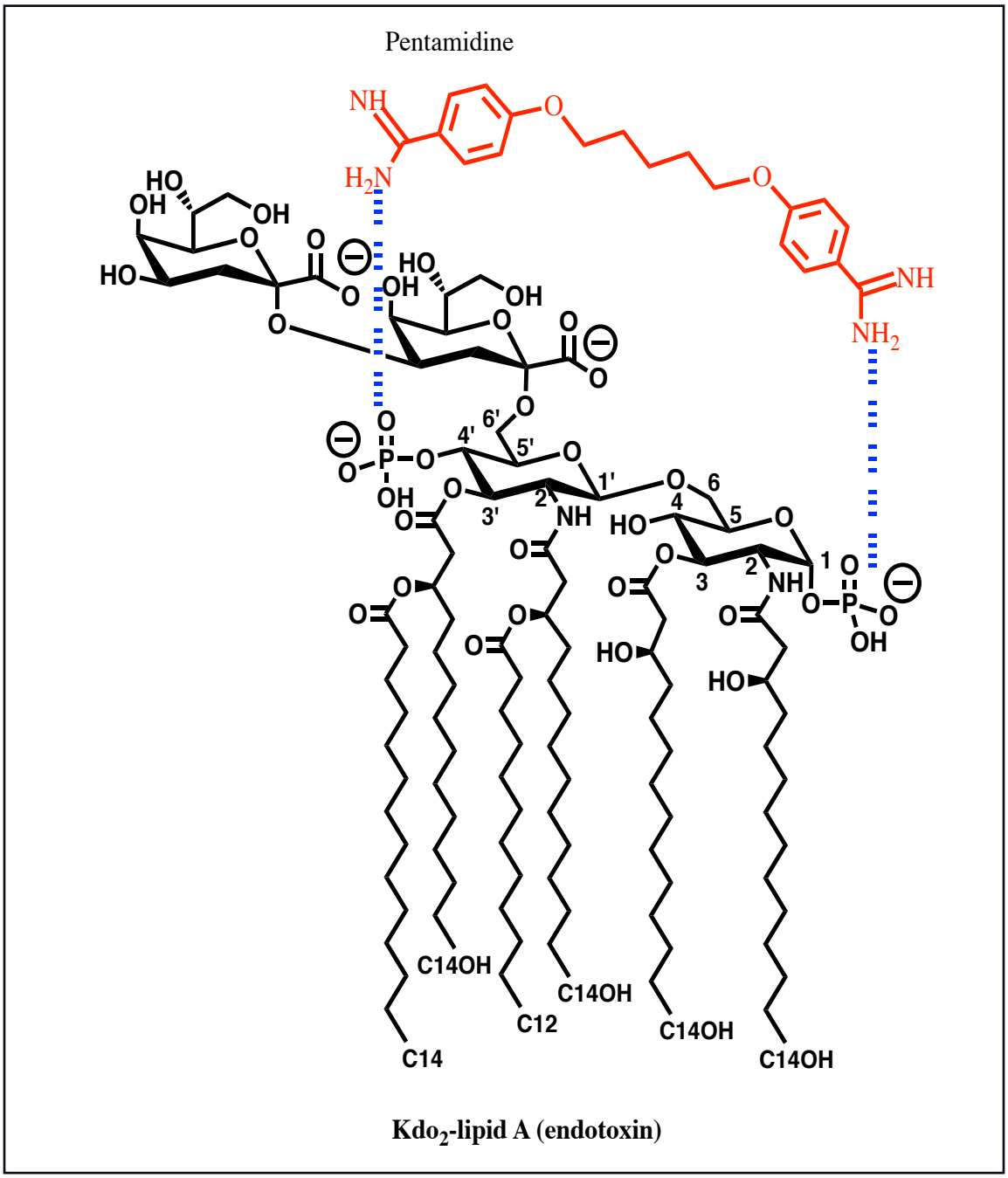

Figure 4.10. Pentamidine interaction with lipid A. Pentamidine is known to have a high affinity for LPS molecules. It contains two aromatic rings linked by an alkyl chain. The amidine groups of pentamidine form electrostatic interactions with the phosphate groups of Kdo2-lipid A (David et al., 1994), disrupting the interaction between LPS molecules.

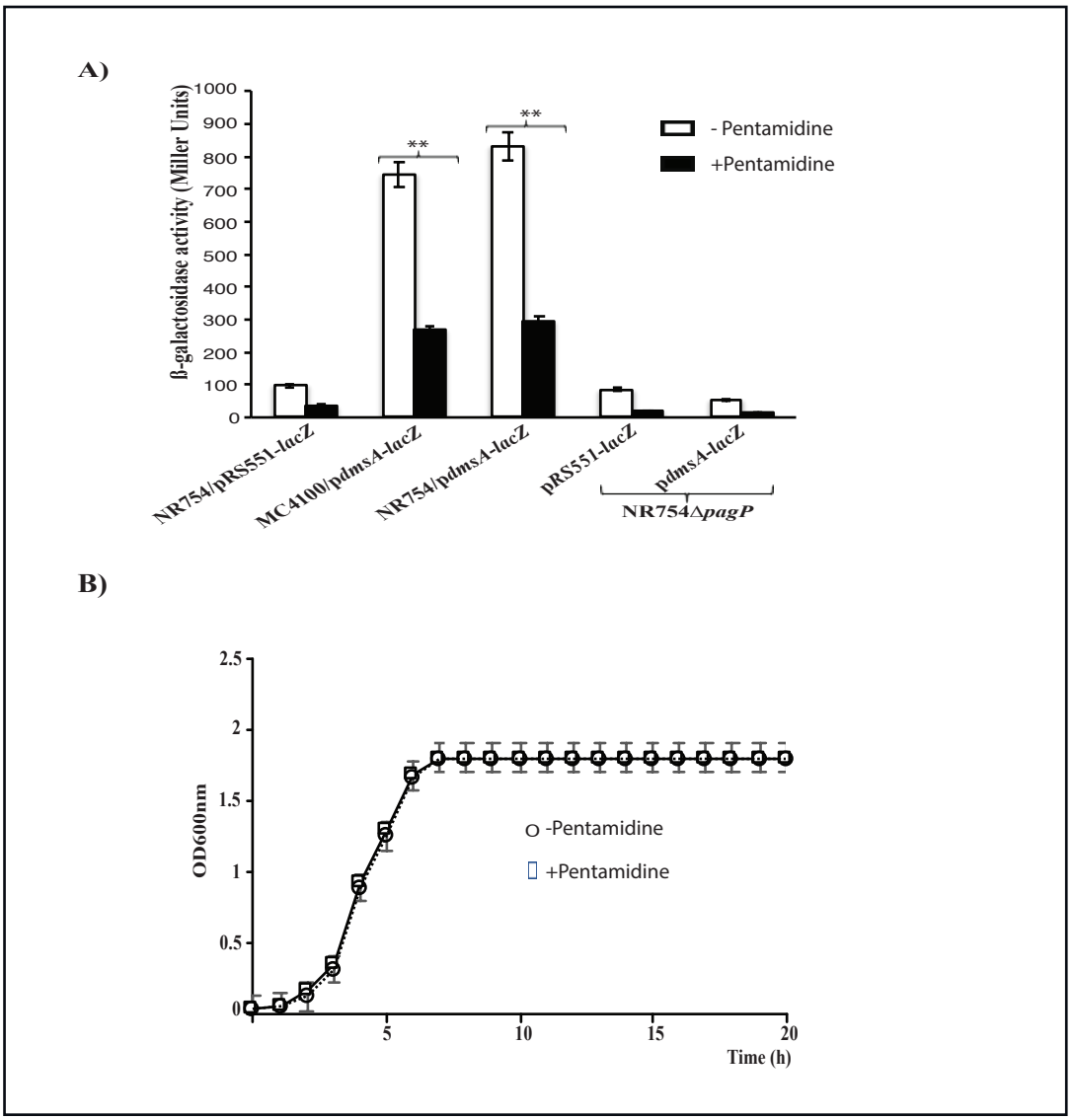

Figure 4.11. Pentamidine treatment represses expression of *dmsA* without affecting the bacterial cell growth. A) The white rectangles represent *dmsA* over expression in strains MC4100 and NR754. The black rectangles represent the repression of *dmsA* expression by pentamidine treatment. 80 μg/mL pentamidine was added to the appropriate samples after cells were grown to OD_{600}=0.2. After an additional 2 hours of growth, the bacterial cell count was normalized to OD_{600} =0.6 and the β-galactosidase activity was determined. The reactions were stopped after 7 minutes using Na_2CO_3. OD_{420nm} readings were performed following the removal of cell debris via centrifugation. An unpaired student's t-test was carried out. Values of β-galactosidase activity that are significantly different are indicated with asterisks ** $p<0.05$. B) The graph shows the growth curve for strain NR754/p*dmsA-lacZ* with and without pentamidine treatment. Pentamidine 80 μg/mL did not affect the cell growth.

4.5.3 PagP regulation on *cspE* transcription: Phenotype complementation

We decided to study the *cspE* gene for two reasons. First, the gene *cspE* is oriented in the same direction downstream of *pagP* and they overlap by 22 base-pairs (Hu et al., 2012). We had initially suspected the possibility of a polar effect, but the Keio *pagP* deletion that we used to construct NR754Δ*pagP* preserved the 22 base-pairs upstream of *cspE* that overlap with *pagP*. Second, as previously demonstrated by Hu et al., 2012, cold shock protein CspE is an enhancer of translation of RpoS, which regulates FNR and several FNR-dependent genes (Patten et al., 2004). To investigate whether or not CspE is involved in PagP-mediated regulation of *dmsA*, three different reporter genes were constructed as shown in Table 7; these include p*pagP-cspE*-1-*lacZ,* carrying the isolated *pagP* promoter, p*pagP-cspE*-2-*lacZ,* carrying the isolated *cspE* promoter, and p*pagP-cspE*-3-*lacZ,* carrying the *pagP* and *cspE* promoters together (Figure 4.12). β-galactosidase activities showed that strain NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ,* carrying only the *pagP* promoter, had a value of 1,100 MU. NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ,* carrying only the *cspE* promoter, showed values similar to those observed for NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ* (Figure 4.12); however, β-galactosidase expression in the NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ* strain was 3-fold lower than that of NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ*, suggesting that the transcription of *pagP* in tandem with *cspE* affected transcription of the latter.

To further demonstrate that PagP supressed expression of CspE, the p*pagP-cspE*-2-*lacZ* reporter was transformed into the NR754 and control MC4100 strains.

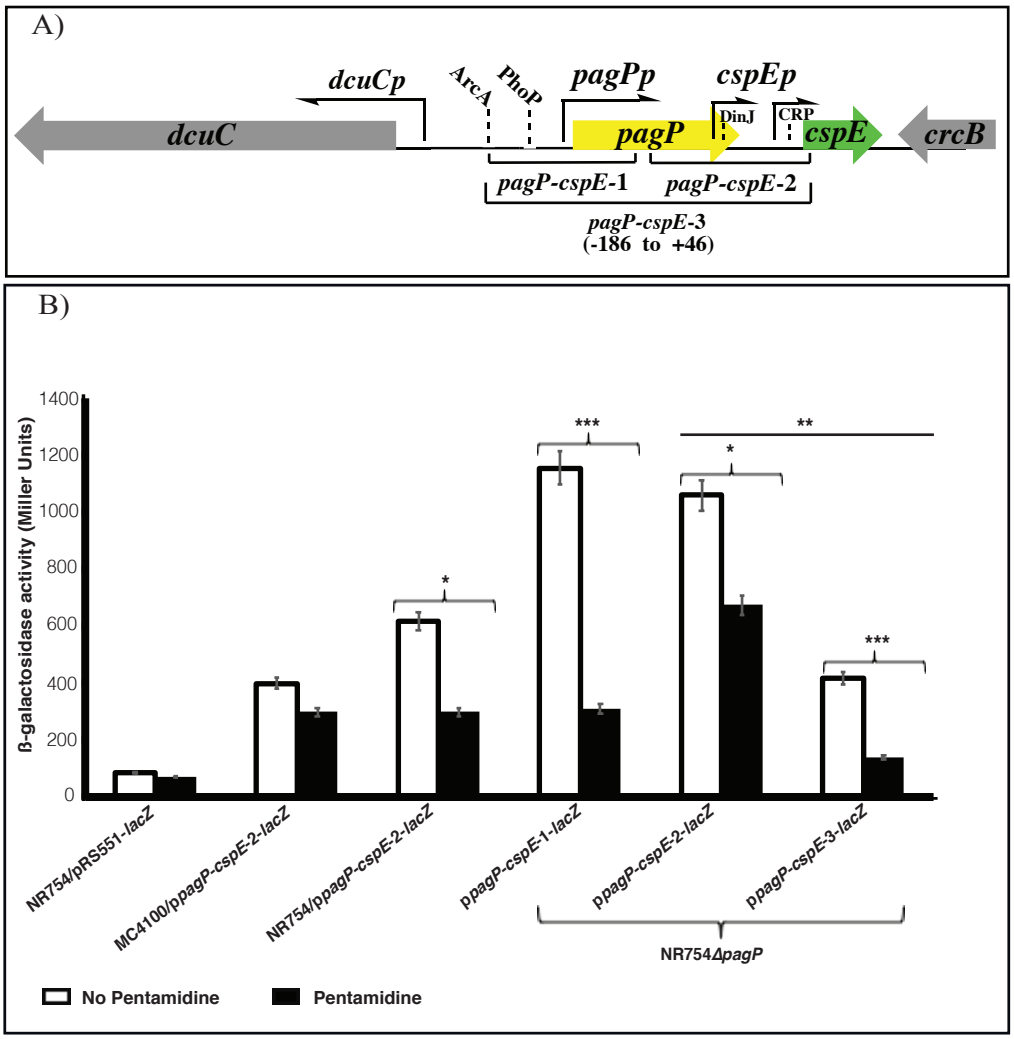

Figure 4.12. β-galactosidase activity for the three genetic constructs NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ*, NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ* and NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ*. A) Diagram representing the fragments of the promoter region used for the construction of the three *lacZ* reporters. B) On the top of this panel are the fragments of the diagram showing the section of the each promoter used for the construction of the reporter genes. The white rectangles represent the β-galactosidase activity of the untreated samples. The black rectangles represent the samples treated with pentamidine. Strain NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ* and NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ* showed similar values of β-galactosidase activity 1100 and 1059 MU. Strain NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ* bearing the reporter gene that carries *pagP* and *cspE* promoter together showed significative lower β-galactosidase values. Tandem transcription of *pagP* and *cspE* downregulates the transcription of the latest. NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ*, carrying the reporter with *pagP* promoter fragment showed to be the most sensitive to pentamidine treatment. Treatment with 80 ug/mL pentamidine was performed to the appropriate samples after cells were grown to OD_{600} 0.2. β-galactosidase activity was determined. An unpaired student's t-test was performed between the samples with and without pentamidine treatment and also between untreated strains NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ* and NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ*. Values of β-galactosidase activity that are significantly different are indicated with asterisks as follow for a p value *$p<0.05$, **$p<0.01$ and *** $p<0.001$.

NR754/p*pagP-cspE*-2-*lacZ* had a β-galactosidase activity of 630 MU, 1.8-fold lower than that of the same construct in the Δ*pagP* strain NR754Δ*pagP*/p*pagP-cspE*-2-*lacZ.* The repression of *cspE* transcription in the control strain MC4100/p*pagP-cspE*-2-*lacZ* (2.7-fold) was similar to that found for NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ,* in which there was a tandem transcription of *pagP* and *cspE* (Figure 4.12).

All three constructs, NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ,* NR754Δ*pagP*/ p*pagP-cspE*-2-*lacZ,* and NR754Δ*pagP*/p*pagP-cspE*-3-*lacZ* showed sensitivity to 80 μg/ml pentamidine treatment. However, the NR754Δ*pagP*/p*pagP-cspE*-1-*lacZ* construct carrying the *pagP* promoter showed a 3.8-fold difference between the sample with and without pentamidine treatment. Thus, these data indicate that pentamidine treatment supressed the transcription of *pagP* (Figure 4.12).

The *E. coli cspE* gene was cloned into the arabinose-inducible plasmid pBADGr. The genetic construct was then co-transformed into the NR754Δ*pagP*/p*dmsA-lacZ* with a β-galactosidase activity of 60 MU. Results showed a partial complementation with a β-galactosidase activity 354 MU. The activity was not restored with the *cspE* gene in reverse orientation.

4.5.4 Defect in lipid A acylation involved in cell signalling

PagP palmitoylates lipid A at the *R*-3-hydroxymirystate chain of the proximal glucosamine unit (Bishop, 2005). The Ser77 residue located in extracellular loop L2 of PagP plays an important role in lipid A palmitoylation. The S77A mutation did not show any palmitoylation when analyzed by TLC using the

strains NR754*λp* and NR760*λp* (Pandey, 2013). PagP also possesses a periplasmic cluster of amino acids D61, H67, and Y87 believed to be involved in signaling. Mutation Y87F was used to test for signaling from the periplasmic side of PagP in *E. coli* NR754*λpY87F* in addition to the catalytically inactive mutant S77A PagP in *E. coli* NR754*λpS77A*. To test whether the lipids modified or synthesized by PagP are involved in the regulation of DmsA, the p*dmsA-lacZ* reporter was transformed into *E. coli* NR754*λpY87F* and NR754*λpS77A*.

Results showed that β-galactosidase activities for NR754*λp* and NR754*λpY87F* were similar, at 1,100 MU; however, the catalytically-inactive strain NR754*λpS77A* showed a phenotype with significantly lower β-galactosidase activity of 300 MU $p<0.05$ (Figure 4.13). The NR754*λp* and NR754*λpY87F* strains had a similar response to pentamidine treatment, with a 4-fold reduction of *dmsA* transcription, while the catalytically inactive NR754*λpS77A* did not show any sensitivity to pentamidine (Figure 4.13).

4.5.5 ^{32}P Palmitoylated Lipid A extraction.

To test the effect of pentamidine on PagP function, strains MC4100, NR754*λp* and NR754*λpY87*F were grown at 37 °C, and 2.5 uCi/ml ^{32}Pi was added to the culture to label the lipid A. Pentamidine was added at OD_{600nm}-0.3 and the cells were grown for an additional 90 min. Mild acid hydrolysis was used to cleave the labile ketosidic bond and release the first Kdo sugar from lipid A. In parallel with the addition of 80 μg/mL pentamidine, 25 mM EDTA treatment was also carried out for 5 minutes.

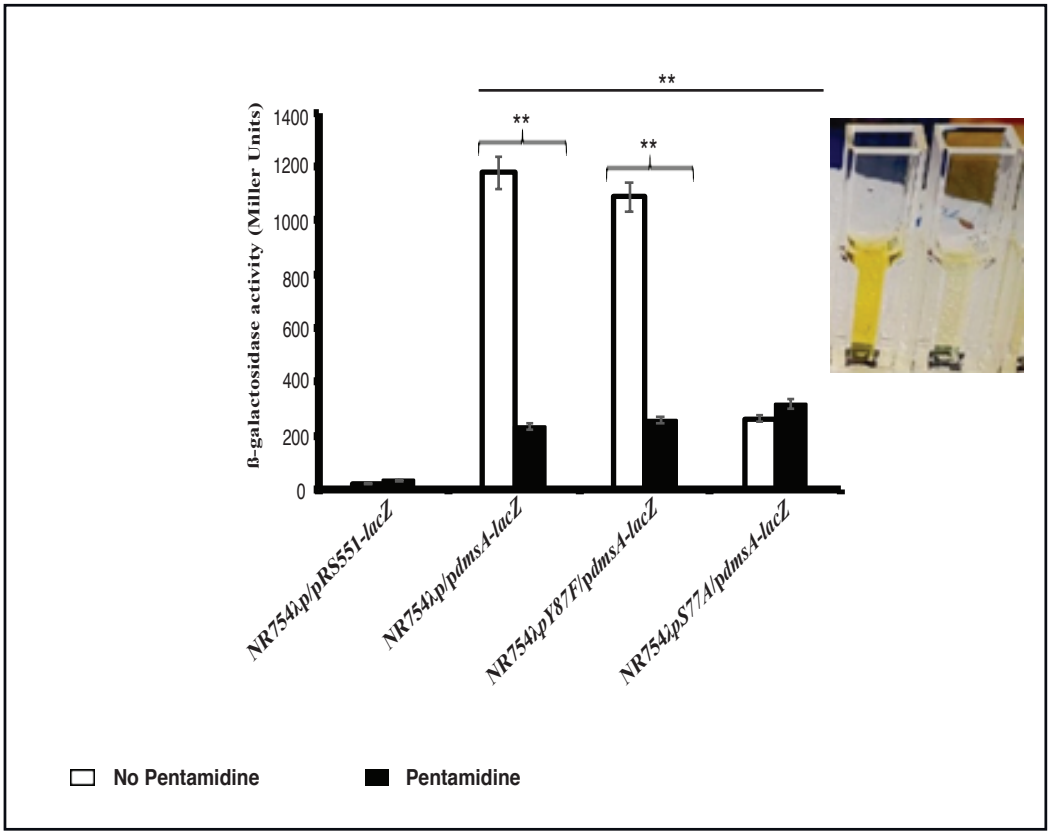

Figure 4.13. NR754*λS77A*/p*dmsA-lacZ* phenotype. Strain NR754*λpY87F*/p*dmsA-lacZ* carrying the chromosomal mutation Y87F showed similar β-galactosidase activity to that found for wild-type NR754*λp*/p*dmsA-lacZ*. The response to pentamidine treatment was statistically significant (**p<0.01) (4-fold reduction of activity). Catalytically-inactive NR754*λpS77A*/p*dmsA-lacZ*, showed 4-fold lower β-galactosidase activity and no response to pentamidine treatment.

Our results indicate that the two different OM perturbants had different effects on lipid A palmitoylation. We used the chelating agent EDTA, which strips Mg^{2+} and Ca^{2+} ions from the membrane and promotes phospholipid migration into the OM external leaflet, thereby activating PagP. We also used pentamidine, which interacts with lipid A and permeabilizes the membrane (Stockes et al., 2017). *E. coli* MC4100 showed that approximately 50% of the hexa acyl lipid A species became palmitoylated when the sample was treated with 25 mM EDTA for 5 minutes. Approximately 75% of the lipid A became palmitoylated in NR754*λp*, which indicates that *pagP* located at the *attB* site exhibits slightly higher expression. Similar levels of palmitoylation were observed in NR754*λpY87F*. These results are slightly higher than those reported by Pandey, 2013, in which the strain with *lptD4213* background NR760*λp* palmitoylated 45% lipid A even before the addition of EDTA and this value increased to 55% after the addition of EDTA.

When NR754*λp* and NR754*λpY87F* samples were treated with 80 µg/mL pentamidine, low levels of approximately 25% palmitoylation occurred; however, 5% palmitoylation is reported to be the basal level in *E. coli* (Jia et al., 2004) (Figure 4.14). Lipid A palmitoylation in those samples was low compared to that of EDTA-treated samples reported by Smith et al., 2008. Similar to MC4100 and NR754, NR754Δ*p* was also treated with pentamidine. As expected, no lipid A palmitoylation was observed in the *pagP* deletion strain (Figure 4.14A).

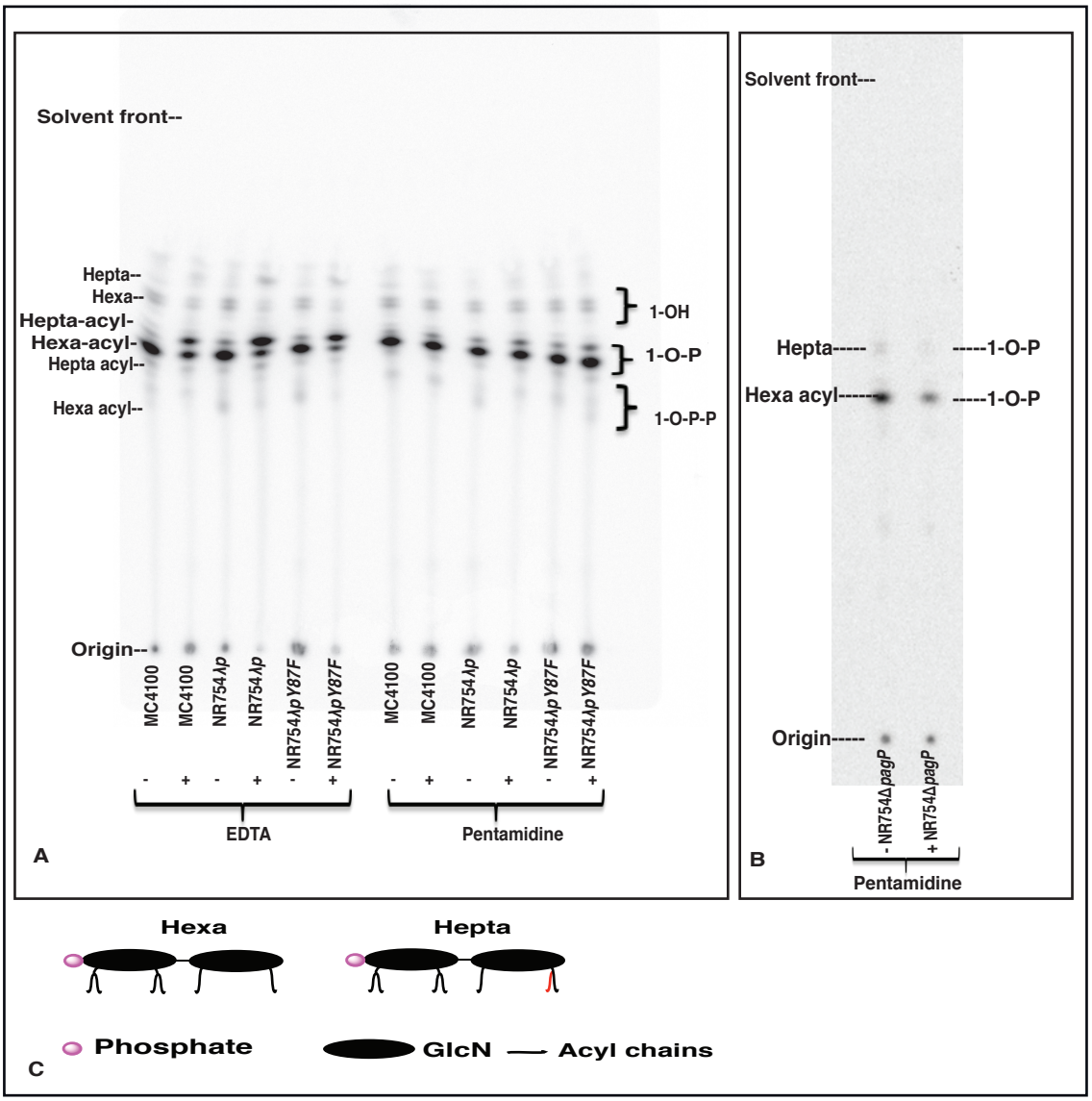

Figure 4.14. TLC analysis of ^{32}P labeled lipid A after mild acid hydrolysis isolation. A) Strains MC4100, NR754*λp* and NR754*λpY87F* were treated with 80 µg/mL pentamidine and EDTA (25 mM). The cultures were grown for 2 hours after the treatment with pentamidine, while EDTA-treated samples were only grown for 5 minutes. The isolated lipid A species were separated by TLC using G-60 silica plates. The solvent system used was 50:50:16:5, chloroform: pyridine: 88 % formic acid: H_2O. The lipid A species were visualized with a PhosphorImager (Typhoon). B) No lipid A palmitoylation was observed in the PagP deletion strain, NR754Δ*pagP*, when treated with pentamidine. C) Illustration of no palmitoylated and palmitoylated lipid A.

4.5.6 Deletion of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB*, and *fadD* downregulates *dmsA* transcription under aerobic conditions.

Using the Keio deletions of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB*, and *fadD* in *E. coli* MC4100, we observed decreased transcription of *dmsA*. Deletion of *phoQ* affected the transcription of *dmsA* very little. Results show that values of β-galactosidase activity for strain MC4100Δ*phoQ*/p*dmsA-lacZ* are similar to wild-type MC4100/p*dmsA-lacZ*, whereas a significant difference was found when we compared the wild-type MC4100 to the MC4100Δ*pmrA*, MC4100Δ*pmrB,* MC4100Δ*fadD,* MC4100Δ*rcsB,* MC4100Δ*cpxA*, MC4100Δ*cpxR,* and MC4100Δ*rpoS* (Figure 4.15).

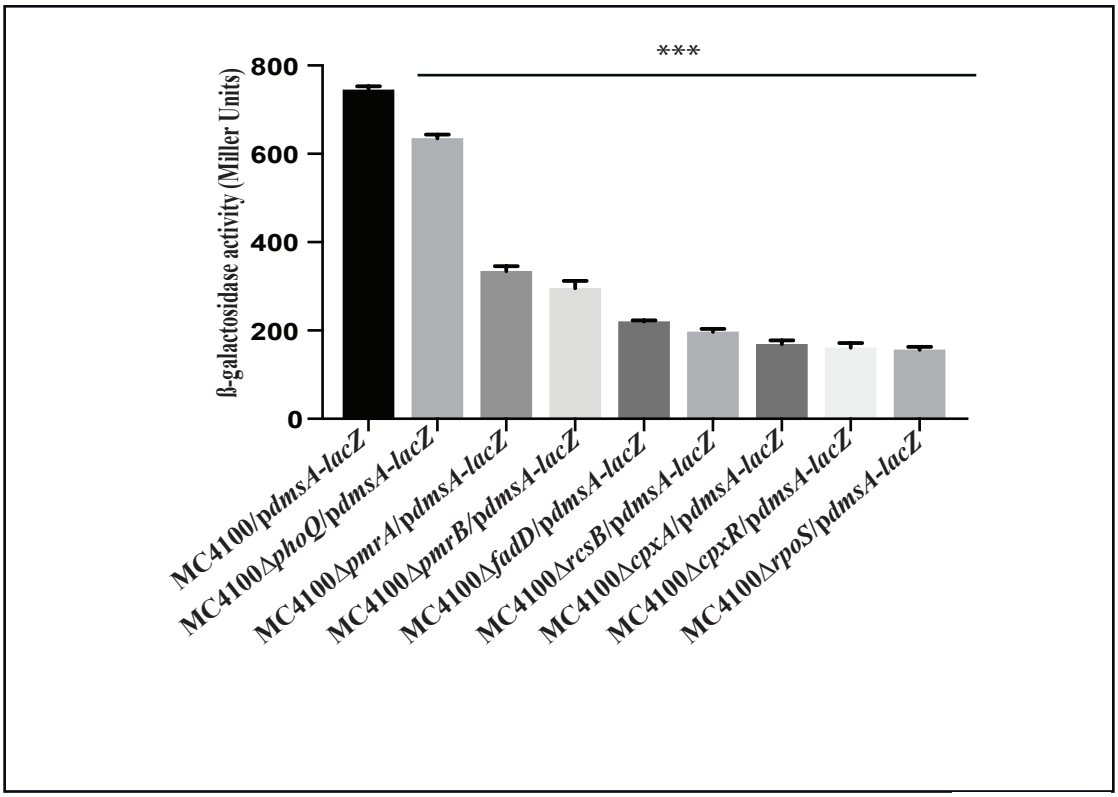

Figure 4.15. Second messengers involved in PagP regulation of *dmsA* transcription. Lack of PhoQ affected very little *dmsA* transcription. Transcription of *dmsA* gene was downregulated in strains MC4100Δ*pmrA*/p*dmsA-lacZ*, MC4100Δ*pmrB*/p*dmsA-lacZ*, MC4100Δ*fadD*/p*dmsA-lacZ*, MC4100Δ*rcsB*/p*dmsA-lacZ*, MC4100Δ*cpxA*/p*dmsA-lacZ*, MC4100Δ*cpxR*/p*dmsA-lacZ* and MC4100Δ*rpoS*/p*dmsA-lacZ*. Values of β-galactosidase activity that were significantly different are shown with asterisks. ***$p<0.001$ by Student's t-test.

4.6 Discussion

4.6.1 Role of PagP on *dmsA* transcription

The product of *dmsA* is the DmsA catalytic subunit of dimethyl sulfoxide reductase. It contains a molybdo-*bis*(pyranoterin guanine dinucleotide) (Mo-*bis*PGD) cofactor and a [4Fe-4S] cluster (Sambasivarao et al., 1991). DmsA allows bacterial growth under anaerobic conditions in the presence of dimethyl sulfoxide as oxidant (Bogachev et al., 1996). In *E. coli*, the expression of DmsA is positively regulated by FNR under anaerobic conditions and it is negatively regulated by IhfA and IhfB, Fis, ModE and NarL (Bearson et al., 2002). To further verify the transcriptional regulation of *dmsA* by PagP using *lacZ* reporter fusions, primers were designed to flank a region of the *dmsA* promoter that included the consensus sequences for FNR, ModE and IHF, which regulate DmsA transcription (Bearson et al., 2002).

Results showed that *dmsA* was overexpressed in the wild-type strain NR754/p*dmsA-lacZ* where PagP is dormant (Figure 4.9). This finding suggests that PagP controls the basal levels of *dmsA* transcription under aerobic conditions. Very low expression of *dmsA* was observed in the *pagP* deletion strain NR754Δ*pagP* where the β-galactosidase activity was 60 MU. Our values for β-galactosidase activity (830 MU) are higher than previously reported during growth under aerobic conditions, but lower than reported under anaerobic conditions, by Bearson et al., 2002. However, these investigators inserted their fusions into the chromosome (Bearson et al., 2002).

We used pBADGr-*pagP* under control of the arabinose promoter for phenotype complementation. Expression of PagP in NR754Δ*pagP* restored β-galatosidase activity, thus demonstrating that PagP is needed for transcription of *dmsA* under aerobic conditions. Numerous questions have arisen from these results: Why is PagP exerting regulation on *dmsA*? What is the biological significance of this regulation? What are the second messengers? PagP is an OM enzyme involved in lipid metabolism and *dmsA* encodes the catalytic subunit of dimethylsulfoxide reductase. DmsA is terminal reductase, which allows bacterial growth under anaerobic conditions (Bogachev et al., 1996), and its expression from the chromosome under anaerobic conditions is 22,200 MU (Bearson et al., 2002). We found in the presence of PagP that *dmsA* exhibits a basal expression of 831 MU under aerobic conditions. However, in the PagP deleted strain NR754Δ*pagP* the transcription of *dmsA* was reduced to 60 MU. We do not yet fully comprehend the physiological significance of this regulation. However, PagP-dependent transcription of the *dmsA* promoter provides us with a very useful reporter to study PagP-mediated signal transduction from the OM to the cytoplasm.

The detection of a change in β-galactosidase activity was achieved after 2 hours of treatment with 80 μg/mL pentamidine. These results demonstrated that transient membrane perturbation with pentamidine repressed *dmsA* transcription by 2.8-fold in *E. coli* NR754/p*dmsA-lacZ* (Figure 4.11). The decision to use 80 μg/mL pentamidine as a transient OM perturbant was based on results of a dose-response study on the activation of the PhoPQ system using pentamidine. In that study, the

transcript levels of the PhoPQ-dependent gene *mgtA* were quantified relative to the housekeeping gene *rsmC* using quantitative reverse transcription PCR (Stokes et al., 2017).

4.6.2 PagP regulation on *cspE* transcription

The strain NR754Δ*pagP/*p*pagP-cspE-1-lacZ* carrying the *pagP-cspE*-1 fragment of the *pagP* promoter (-186 bp to +307 bp), which includes the region upstream of the ArcA consensus sequence and encompasses the PhoP binding site, showed a β-galactosidase activity of 1,200 MU; this value is higher than that reported by Shprung *et al*., 2012, who similarly constructed a pRS551*pagP-lacZ* fusion using the *Salmonella pagP* promoter region. Membrane perturbation with pentamidine greatly repressed *pagP* transcription by 3.8-fold (Figure 4.12). The reduction in transcription of *pagP* might be explained by the fact that defects in LPS causes activation of σ^E and its translational repressors MicA and RybB (Klein and Raina, 2015). Since PagP is regulated by PhoP, which is repressed by MicA (Grabowics and Silhavy, 2017), post-transcriptional repression of PhoP by MicA could have occurred, consequently affecting the expression of genes from the PhoPQ operon including *pagP* (Coornaert et al., 2010). Post-transcriptional repression of PhoP by MicA occurs by base pairing with PhoP mRNA (Coornaert et al., 2010).

In order to determine whether or not PagP has regulation on *cspE* transcription the pRS551-2-*cspE* construct was transformed into the *pagP* deleted strain NR754Δ*pagP*. Transcription of *cspE* in strain, NR754Δ*pagP/*p*pagP-cspE*-2-

lacZ was 1,100 MU. The gene *cspE*, which encodes the cytosolic cold shock protein CspE, showed the least sensitivity to pentamidine treatment (Figure 4.12). NR754Δ*pagP/*p*pagP-cspE*-3-*lacZ,* carrying the plasmid with both promoters in tandem, from -186 bp upstream of *pagP* to +46 bp downstream of *cspE,* showed β-galactosidase activities 3-fold lower than NR754Δ*pagP/*p*pagP-cspE*-2-*lacZ,* indicating that the upstream *pagP* promoter supresses transcription of *cspE.*

Our results showed that the *pagP* deletion could influence *cspE* transcription. This finding suggests a plausible mechanism for the apparent PagP control of *dmsA* could be derived from a secondary effect that depends on CspE. However, the Keio Δ*pagP* allele only deletes the majority of the open reading frame without affecting the surrounding promoters. Nevertheless, CspE is an enhancer of RpoS translation (Hu et al., 2012). RpoS regulates more than 1000 genes in *E. coli* O157 H7 (Dong et al., 2008) and it is known to indirectly regulate some FNR-regulated genes (Dong et al., 2009), such as *dmsA* (Patten et al., 2004). A significant downregulation of the *dmsABC* operon and the *ynfEFGH* operon by RpoS was reported by Patten et al., 2004, with p-values of *$p<0.05$ and ***$p<0.001$, respectively. Activation of *cspE* transcription by the *pagP* deletion might cause the enhancement of RpoS translation with consequent repression of *dmsA*. A close genetic linkage between *pagP* and *cspE* has been observed previously during studies of a multicopy phenotype involving resistance to camphor vapours that affect chromosome condensation (Hu et al., 1996; Sand et al., 2003).

We found that phenotype complementation using pBADGr-*cspE* restored the β-galactosidase activity only partially to 354 MU, indicating that *cspE* might be involved in PagP-mediated regulation of *dmsA* transcription. However, results indicate that PagP plays a main role in this regulation since complementation with the *pagP* gene completely restored the β-galactosidase activity, yielding 894 MU.

4.6.3 Defects of lipid A acylation involved in cell signalling.

Based on the β-galactosidase activity the level of *dmsA* transcription was very similar in *lptD*$^{+}$ *E. coli* NR754*λp* and NR754*λpY87F*. However, *dmsA* was poorly transcribed in the catalytically inactive NR754*λpS77A* strain (Figure 4.13). Following membrane perturbation by pentamidine, the levels of *dmsA* transcription were reduced to 230 MU in NR754*λp* and NR754*λpY87F* strains. NR754*λpS77A* did not show any sensitivity to pentamidine, suggesting that lipid A palmitoylation is involved in PagP mediated regulation (Figure 4.13). Mutation S77A and pentamidine treatment both similarly downregulate *dmsA* transcription. The replacement of Ser77 does not allow PagP to palmitoylate lipid A, while the addition of pentamidine might not allow PagP to have the sufficient access to lipid A due to pentamidine interactions with lipid A phosphate groups. Both types of stress might, conceivably, be sensed by the σ^{E} envelope stress response.

4.6.4 ^{32}P palmitoylated lipid A is affected by the addition of pentamidine

Mild acid hydrolysis was used to break the labile ketosidic bond and release lipid A from LPS. From the origin to the first species represents the hexa acyl lipid A, which possesses a 4+2 acyl chain distribution in the distal and proximal

glucosamine units of lipid A, respectively (Figure 4.14). The second species represents the hepta acylated lipid A where incorporation of C16 (palmitate chain) to the hydroxyl group of the *R*-3-hydroxymyristate chain at the proximal glucosamine unit of lipid A has occurred. Our results indicate that the use of two different OM perturbants resulted in different outcomes (Figure 4.14). The chelating agent EDTA activates PagP through OM perturbation of lipid asymmetry (Jia et al., 2004). Lipid A palmitoylation by PagP was activated in the presence of EDTA as shown previously (Smith et al., 2008), where approximately 90 % of the substrate was palmitoylated (Figure 4.14). Treatment with pentamidine, on the other hand, did not activate PagP. This drug has a very high affinity for lipid A and can interact favourably with the phosphate groups (Stokes et al., 2017). Such interactions might cause lipid A to be less accessible to PagP. Lipid A palmitoylation was affected by pentamidine treatment and *pagP* transcription was affected by 3.8-fold.

4.6.5 Second messengers involved in PagP regulation of *dmsA* transcription

PagP regulation of *dmsA* transcription could be mediated by different transcriptional regulators involved in the envelope stress response and by sensory transduction from PagP through its influence on lipid metabolism. Considering that β-galactosidase activity in wild-type *E. coli* MC4100/p*dmsA-lacZ* was 741 MU, we would expect that deletion of regulators triggered by PagP might affect *dmsA* transcription. Deletion of *phoQ* did not affect *dmsA* transcription (Figure 4.16), which means that PhoQ does not seem to be required for this regulation. However,

deletion of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB,* and *fadD* all showed a downregulation of *dmsA* expression (Figure 4.16).

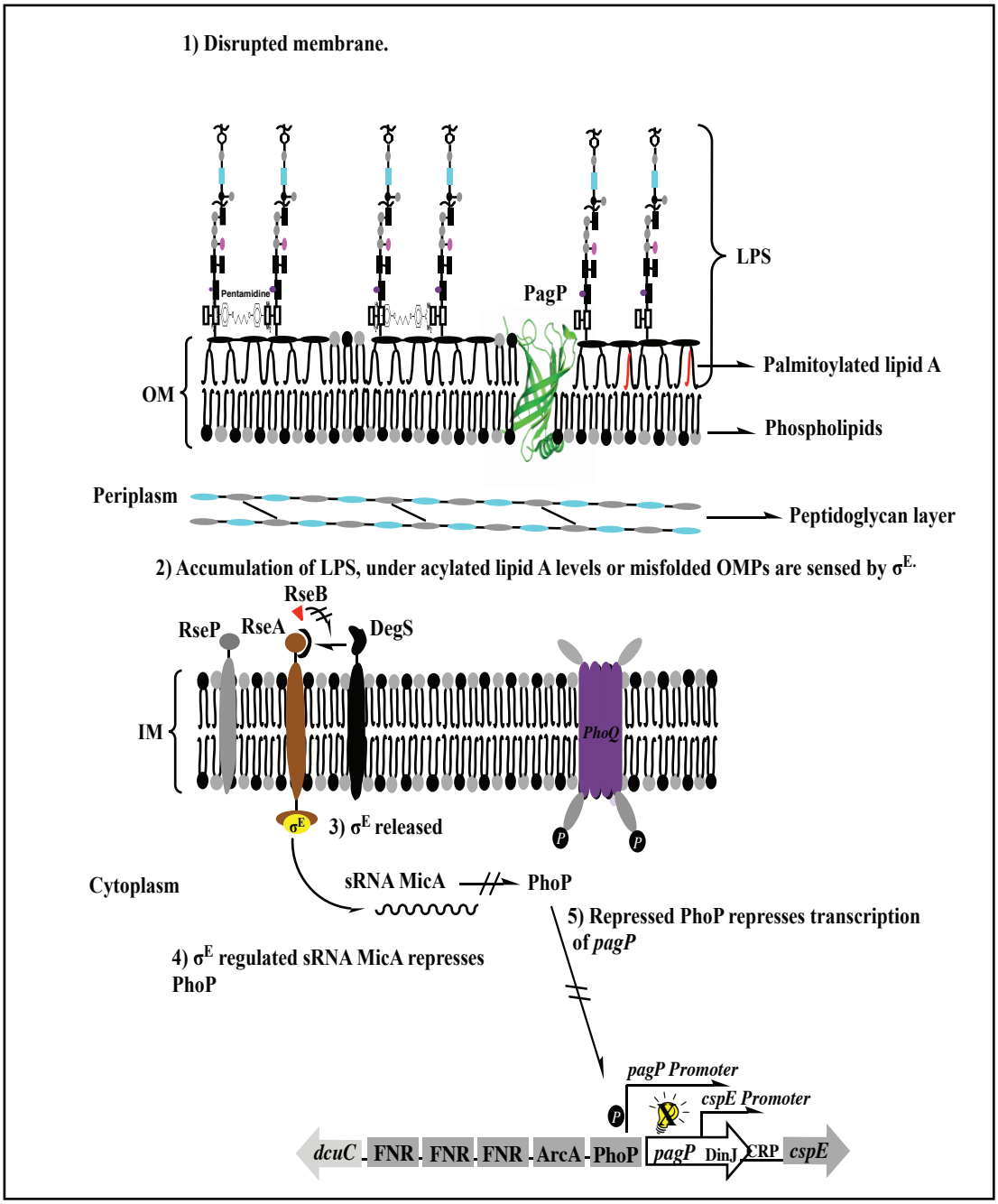

Figure 4.16. Model of pentamidine treatment downregulating *pagP* transcription. Based on our results pentamidine downregulates transcription of *pagP*. The amidine groups of pentamidine interact electrostatically with the phosphate groups of LPS to disrupt the lateral interactions between lipid A molecules (David et al., 1994). Accumulation of LPS, its metabolites and unfolded proteins are sensed by RseB and DegS, which allows σ^E to be released into the cytoplasm. σ^E-dependant sRNA MicA represses translation of PhoP (Grabowicz and Silhavy, 2017), which downregulates transcription of *pagP*.

Since we found that the β-galactosidase activity was affected in *E. coli* NR754*λpS77A* we suspect that lipid A, palmitoyl phosphatidylglycerol (PPG) and lyso-phospholipids could be part of the second messenger pathway in PagP-mediated regulation. Our model starts with PagP activation by perturbation of OM lipid asymmetry leading to modified lipids, which are transported to the cytoplasm where they influence a combination of RpoS and FNR to exert transcriptional regulation of the *dmsA* promoter.

Deletion of *rpoS*, the global stress regulator, caused a 4.7-fold repression of *dmsA* expression. The reciprocal regulation between RpoS and FNR allows RpoS to negatively regulate FNR-dependent genes (Patten et al., 2004) among them *dmsA*. Deletion of the *cpxA*, which encodes the sensor kinase CpxA, and deletion of *cpxR,* which encodes the response regulator CpxR, downregulated *dmsA* transcription by 4.4 and 4.6-fold, respectively. In the absence of the response regulator of the Rcs phosphorelay system *dmsA* transcription was downregulated by 3.8-fold. Rcs is known to intensively regulate the release of small RNA RprA, known to regulate transcription of *rpoS* (Klein and Raina, 2015). A model proposed by (Konovalova et al., 2014) shows that RcsF, the sensor response component of the Rcs system, monitors levels and defects of LPS at the bacterial cell surface. The σ^E, Cpx and Rcs are interconnected by sRNA such as RprA, MicA and MicF (Grabowicz and Silhavy et al., 2017). Lack of the membrane sensor kinase PmrB and the response regulator PmrA, members of the two-component system PmrAB, caused a decrease in *dmsA* transcription of 2.52 and 2.22-fold, respectively. The

PmrAB pair control transcription of genes in charge of modifying lipid A and genes that control levels of RpoS (Hagiwara et al., 2004). Lastly, deletion of FadD downregulated *dmsA* expression by 3.4-fold. This Acyl-CoA synthetase, which activates and transports fatty acids prior to their incorporation into phospholipids (Keseler et al., 2011), is similarly involved in PagP mediated regulation and might be connected to the OM phospholipase PldA (May and Silhavy, 2018).

Our findings revealed multiple complexities in the series of events by which PagP regulates *dmsA* transcription; however, we demonstrated that PagP is necessary for *dmsA* transcription in the wild-type strain (NR754) under aerobic conditions. Collectively, our results suggest that there are several ways by which PagP mediated regulation on *dmsA* transcription occurs. One is the activation of the σ^E, Cpx, Rcs envelope stress response regulators by defective lipid A palmitoylation or accumulation of LPS intermediate metabolites. In this case, underacylated lipid A, intermediate metabolites from LPS biosynthetic pathway, or misfolded OMPs could be activating the σ^E, Cpx, Rcs and dependent sRNA, RprA, which regulates the stationary stress regulator, RpoS (Majdalani et al., 2001). On the other hand, PmrAB, which controls the *arnBCADTEF* operon involved in lipid A modification (Hagiwara et al., 2004), plays an important role on the regulation of stress response regulators such as *elbA*, a RpoS control factor (Ogasawara et al., 2012). The Acyl-CoA synthase FadD appeared to respond to the same stress according to levels of LPS needed at the cell surface. Taking everything together our results suggest that there are several pathways involved in PagP mediated

regulation on *dmsA*; however, they all require PagP extracellular active site, lipid A palmitoylation, PG palmitoylation or their respectives products lyso-PE and lyso PG as a starting point and RpoS-FNR as an end point. Two models of what could be happenening are depicted in Figure 4.17 and Figure 4.18.

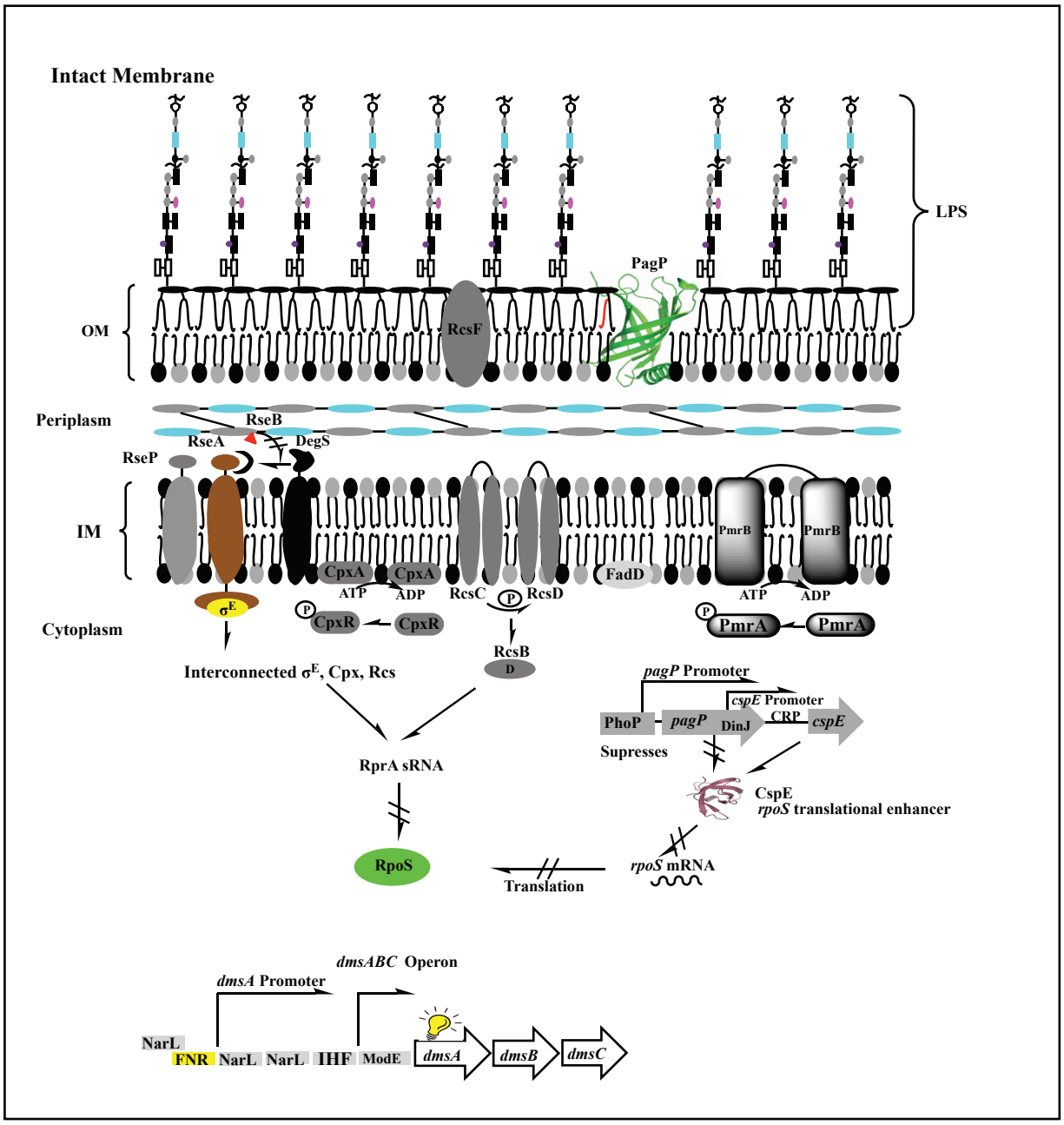

Figure 4.17. Model of PagP-mediated regulation on *dmsA* transcription in an intact membrane. Based on our results several systems appear to be involved in PagP mediate regulation of *dmsA*. The first step of this mechanism is PagP and very low levels of lipid A palmitoylation. The last step is FNR/RpoS. If PagP is present and there is a low level of lipid A palmitoylation, normal levels of LPS will not trigger the activation of σ^E, Cpx or Rcs. There will not be induction of RprA and there will not be activation of RpoS, therefore there will not be repression of FNR and *dmsA* will be transcribed. Similarly, should occur with the PmrA/PmrB system and FadD. Another pathway could be through CspE. If PagP is present, it represses the expression of CspE, expression of RpoS expression is repress, therefore FNR can up-regulate transcription of *dmsA*.

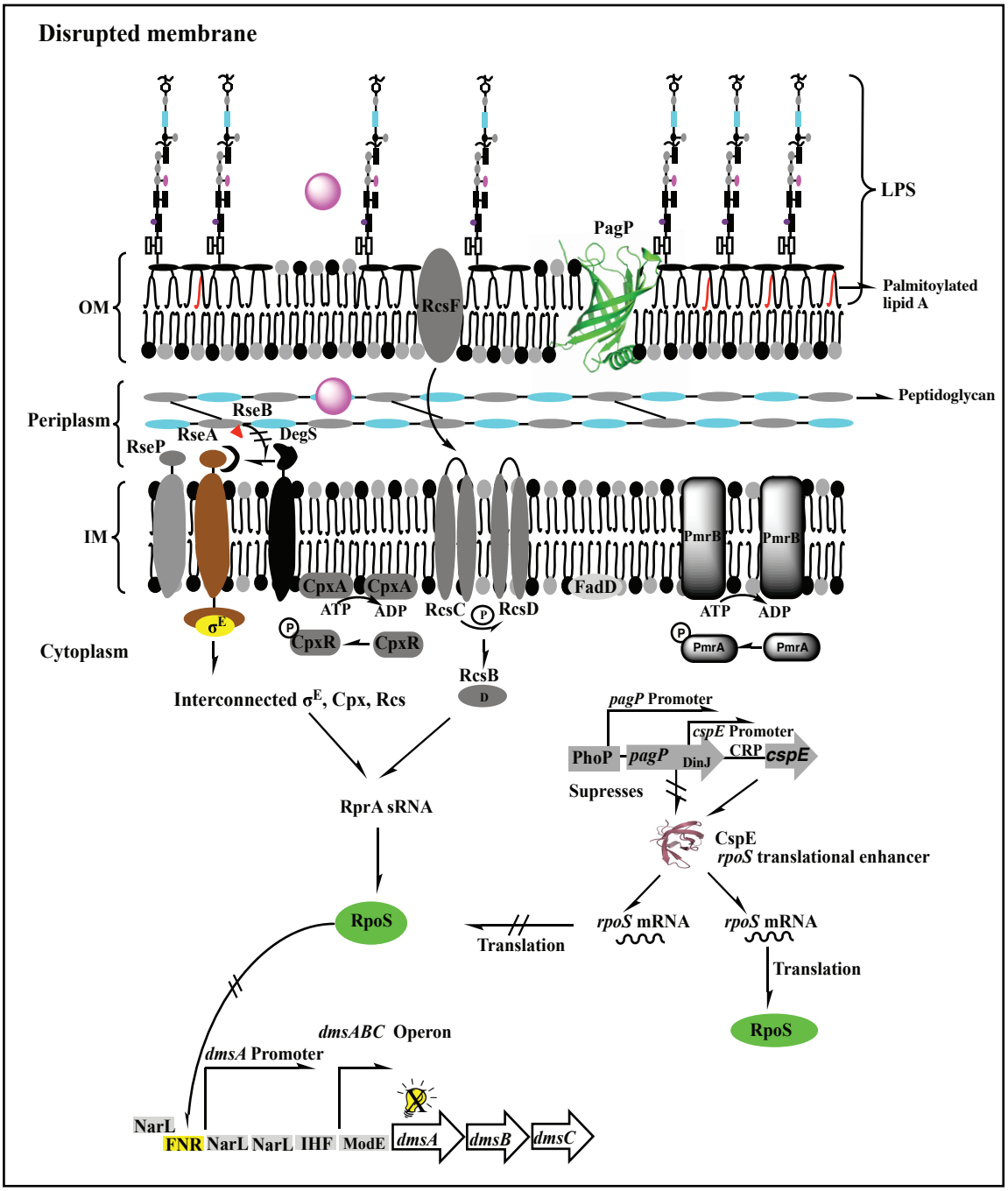

Figure 4.18. Model of PagP-mediated regulation in a disrupted membrane. Our model starts with PagP activation by perturbation of OM lipid asymmetry and ends with the combination of RpoS and FNR to regulate *dmsA* transcription. In a disrupted membrane the phospholipids migrate to the outer leaflet occupying the LPS space. Phospholipids make themselves more accessible to PagP for lipid A and PG palmitoylation. Levels of lipid A palmitoylation, defective LPS or its metabolites accumulate, activate the expression of σ^E, Cpx and Rcs, which induce the RprA sRNA, to trigger the activation of RpoS, with the consequent repression of FNR and downregulation of *dmsA*. Similarly, should occur with Acyl-CoA synthetase FadD and the two-component system PmrAB, which respond to levels of phospholipids and LPS.

Chapter 5. General Conclusions and Future Directions

5.1 Speculation on a novel sensory transduction process in *E. coli.*

The mechanisms by which OM proteins in Gram-negative bacteria communicate with other enzymes in different compartments of the bacterial cell are not fully understood. This thesis focussed on understanding the role of OM enzyme PagP in cell signaling. Using three methods to study gene regulation (RNA-sequencing, NanoString nCounter, and *lacZ* reporter genes), we demonstrated that PagP exerts transcriptional regulation on *dmsA*; this cytoplasmic response from PagP is presumably mediated by a sensory transduction mechanism across the *E. coli* cell envelope. The nature of any second messenger capable of communicating the activation of PagP by perturbed OM lipid asymmetry remains to be ascertained. However, recent observations of PagP acticity in the expansion of a novel class of glycerophosphoglycerol phospholipids suggests that *bis*(monoacylglycero)phosphate (BMP) and/or lyso-BMP are candidate second messengers since they are produced in bacterial cells when PagP is expressed, but not in Y87F PagP (unpublished observations). These novel lipids are derived from a series of metabolic steps catalyzed by PagP, which starts with the palmitoylation of phosphatidylglycerol (PG) (Dalebroux et al., 2014). Palmitoyl-PG (PPG) likely emerges from PagP in the extracellular leaflet of the OM because it depends on the S77 residue associated with the cell surface active site. However, we believe that PPG migrates to the inner leaflet of the OM where it then encounters Y87, which likely controls a novel lipase activity that involves hydrolysis of PPG to generate

BMP, followed by hydrolysis of BMP to generate lyso-BMP (Figure 5.1). Although the exact role these lipids play in *E. coli* remains to be fully ascertained, the possibility that BMP and lyso-BMP can diffuse across the cell envelope suggests they might be connected to the transcriptional control mechanisms in order to communicate the presence of a compromised OM permeability barrier. Recent advances in understanding the mechanism of lipid transport between the IM and OM provide a molecular basis for future studies of lipid-mediated sensory transduction across the *E. coli* cell envelope (Ekiert et al., 2017; May and Silhavy., 2018; Sherman et al., 2018; Zhang et al., 2018; Ho et al., 2018;).

5.2 General conclusions

Of the 50 differentially expressed genes initially identified by RNA -seq, 40 were upregulated and encode proteins whose functions are related to anaerobic processes, whereas 10 genes were downregulated and encode proteins whose function are related to aerobic processes. Similarly, comparison of strains with the *lptD4213* background NR760*λp* and NR760*λpY87F* using Nanostring nCounter showed that only *dmsA*, *dcuC*, *tdcA*, and *sdhB* were differentially expressed. Only *dmsA* was differentially expressed when the catalytically inactive strain, NR754*λpS77A,* and wild-type NR754*λp* were compared. Comparison of strain NR754*λp* to its mutant NR754*λpY87F* showed differential expression of *dmsA*, *dcuC*, *sdhB* and *paaI*.

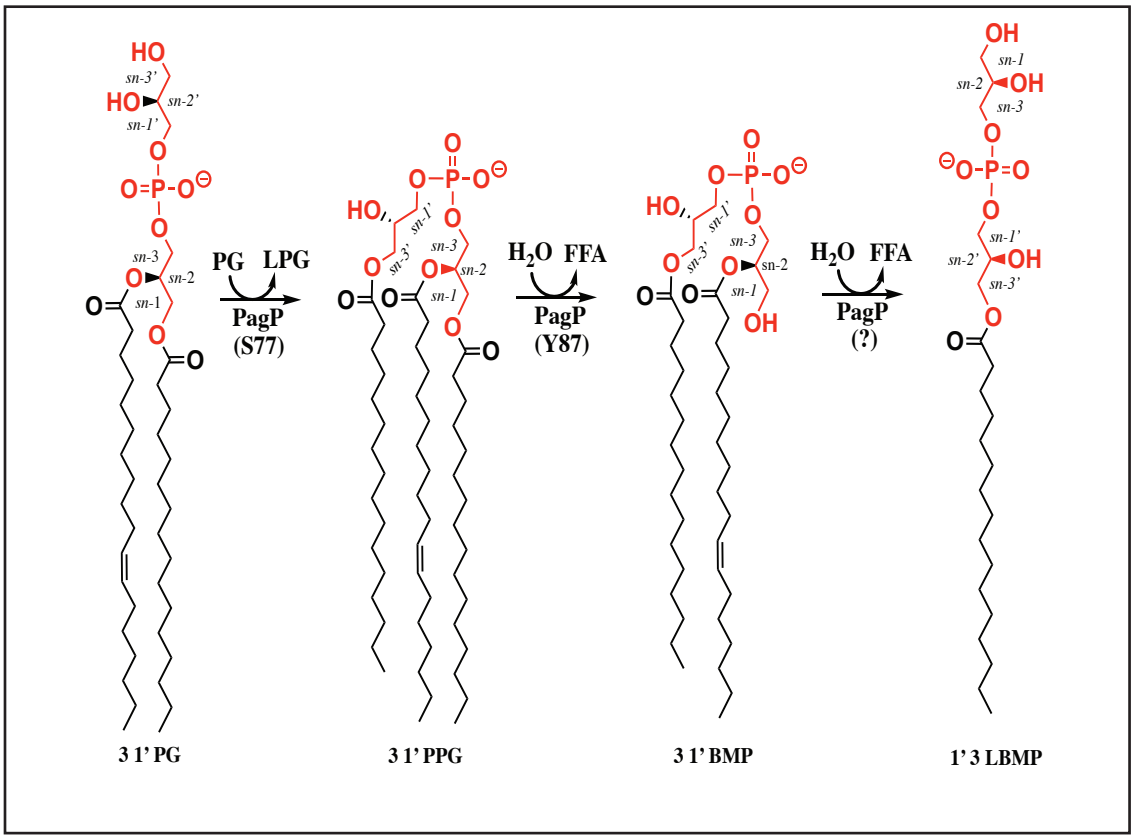

Figure 5.1. Phospholipids modified by PagP. PG is likely to be palmitoylated at the cell surface because its palmitoylation depends on S77, amino acid residue from the extracellular active site. The PPG is believed to migrate to the inner leaflet of the OM, in which is hydrolysed by PagP periplasmic site Y87, generating BMP, which is further hydrolysed to LBMP.

Using transcriptional fusion genes our results confirmed that OM enzyme PagP mediates regulation of *dmsA* transcription. Through deletion of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *pmrA*, *pmrB*, *fadD* in *E. coli* MC4100, we showed that the products of these genes are involved in PagP-mediated regulation of *dmsA*. Results indicate that PagP-mediated regulation of *dmsA* might occur through its extracellular active site via lipopolysaccharide cross talk in an interconnected network between σ^E, Cpx, and Rcs envelope stress regulators, including participation of two-component system PmrAB, Acyl-CoA synthetase FadD, and CspE. Further processing of products emerging from the cell surface active site by a novel periplasmic lipase active site in PagP could explain our observations if these novel lipids behave as second messengers.

5.2 Future directions

We demonstrated that deletion of a number of two-component systems and enzymes affect the transcription of *dmsA*; however, the transcription was not completely abolished. To determine if other enzymes or two-component systems are involved in PagP regulation of *dmsA*, the pRS551-*dmsA*-*lacZ* reporter plasmid can be transformed into strains carrying the deletion of any gene of interest. We speculate that PagP mediates regulation of *dmsA* transcription by triggering FNR-RpoS. To further analyze the remaining differentially expressed genes, *lacZ* reporters should be prepared for *tdcA*, *dcuC*, *sdhB* and *paaI* genes following the same protocol as used for *dmsA* in this thesis.

It has been established that PagP plays an important role in lipid modification and also that phospholipids and LPS serve as signal transducers. Our results demonstrate that *dmsA* is expressed in wild-type *E. coli* NR754 under aerobic conditions. We also demonstrated that in the NR754Δ*pagP* strain expression of *dmsA* is highly repressed; therefore, PagP is necessary for *dmsA* expression under aerobic conditions. Given that DMSO reductase is an enzyme of anaerobic respiration, the physiological significance of its expression under aerobic conditions remains to be explained. However, *E. coli* living in its host environment most likely transitions between periods of aerobic and anaerobic growth and adaptation to these changes might conceivably hinge on the status of the OM permeability barrier. We also demonstrated that transcription of *dmsA* is highly affected when wild-type *E. coli* is compared to the catalytically inactive strain NR754λ*pS77A*. We proposed a model to explain how the different pathways by which PagP regulates *dmsA* expression will work when the membrane is disrupted. To trigger PagP in wild-type cells we used the lipid A-interactive OM permeabilizing agent pentamidine (Stokes et al., 2017). However, while pentamidine clearly reduced transcription from the *dmsA* promoter, we did not observe lipid A palmitoylation mediated by PagP in the OM in response to pentamidine.

We propose that PagP exerts transcriptional regulation on *dmsA* via LPS or its metabolites using envelope stress response regulators. Through deletion of *rpoS*, *rcsB*, *cpxA*, *cpxR*, *basS*, *basR*, *fadD* in *E. coli* MC4100, we showed that these genes

are involved in PagP-mediated regulation of DmsA via an interconnected network between σ^E, Cpx, and Rcs envelope stress regulators, with participation of the two-component system PmrAB, Acyl-CoA synthetase FadD, and CspE.

PagP possesses a putative periplasmic active site believed to be involved in cell signalling; it consists of a cluster of amino acid residues D61, H67 and Y87. In our research studying gene expression using NanoString nCounter we detected that genes *paaI*, *sdhB* and *dcuC* were differentially expressed when we compared the strain NR754*λp* to NR754*λpY87F*. Although the over expression of these genes was not confirmed through reporter gene methods, there exists a possibility in the wild-type strain NR754 that OM perturbation triggers the periplasmic active site and modifies phospholipids that function as second messengers. Recent unpublished results from the Bishop laboratory have revealed a novel class of glycerophosphoglycerol phospholipids in *E. coli*. PagP not only transfers a palmitate chain from phosphatidylethanolamine to the proximal glucosamine unit of lipid A, but also transfers it to the PG polar head-group through its extracellular active site. The resulting PPG must flip to the periplasmic leaflet of the OM where it can be hydrolyzed to BMP; this reaction is completely abolished when the Y87F mutant is expressed *in vivo*. These results suggest that the PagP periplasmic active site functions as a novel lipase, which not only converts PPG to BMP, but further hydrolyzes BMP to lyso-BMP, both of which might be candidate second messengers.

Our results have established the role of PagP as a sensory transducer, but further studies will be necessary to establish BMP and lyso-BMP as second messengers. The use of purified PagP to generate BMP and lyso-BMP *in vitro* could be exploited to test if the *dmsA-lacZ* fusion plasmids expressed in wild-type and Δ*pagP E. coli* strains can detect any colorimetric changes on XGal plates using standard disc diffusion assays. As such, these studies provide a useful tool for further exploring PagP mediated sensory transduction in response to perturbations in the OM permeability barrier.

6 References

Ahn VE, Lo EI, Engel CK, Chen L, Hwang PM, Kay LE, Bishop RE, Prive GG (2004) A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J* 23: 2931-41

Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biology* 11: R106

Ades SE, Connolly LE, Alba BM, & Gross, CA (1999) The *Escherichia coli* sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. Genes Dev, 13(18), 2449–2461

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli K-12* in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2.

Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB, Breaker RR (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science (New York, NY)* 335: 233-235

Bearson SM, Albrecht JA, Gunsalus RP (2002) Oxygen and nitrate-dependent regulation of *dmsABC* operon expression in *Escherichia coli*: sites for Fnr and NarL protein interactions. *BMC Microbiology* 2: 13

Berks BC (1996) A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* 22: 393-404

Bishop RE (2005) The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Mol Microbiol* 57: 900-12

Bishop RE (2014) Emerging roles for anionic non-bilayer phospholipids in fortifying the outer membrane permeability barrier. *J Bacteriol* 196: 3209-13

Bohm R, Sauter M, Bock A (1990) Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. *Mol Microbiol* 4: 231-43

Bogachev AV, Murtazina RA, Skulachev VP (1996) H+/e- stoichiometry for NADH dehydrogenase I and dimethyl sulfoxide reductase in anaerobically grown *Escherichia coli* cells. *J Bacteriol* 178: 6233-7

Boyd D, Weiss DS, Chen JC, Beckwith J (2000) Towards single-copy gene expression systems making gene cloning physiologically relevant: lambda InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. *J Bacteriol* 182: 842-7

Browning DF, Grainger DC, Beatty CM, Wolfe AJ, Cole JA, Busby SJ (2005) Integration of three signals at the *Escherichia coli nrf* promoter: a role for Fis protein in catabolite repression. *Mol Microbiol* 57: 496-510

Bos MP, Robert V, Tommassen J (2007) Biogenesis of the gram-negative bacterial outer membrane. *Annu Rev Microbiol* 61: 191-214

Braun M, Silhavy TJ (2002) Imp/OstA is required for cell envelope biogenesis in *Escherichia coli*. *Mol Microbiol* 45: 1289-302

Cao J, Woodhall MR, Alvarez J, Cartron ML, Andrews SC (2007) EfeUOB (YcdNOB) is a tripartite, acid-induced and CpxAR-regulated, low-pH Fe^{2+} transporter that is cryptic in *Escherichia coli K-12* but functional in *E. coli* O157:H7. *Mol Microbiol* 65: 857-75

Casadio R, Fariselli P, Finocchiaro G, Martelli PL (2003) Fishing new proteins in the twilight zone of genomes: the test case of outer membrane proteins in *Escherichia coli K12*, *Escherichia coli* O157:H7, and other Gram-negative bacteria. *Protein Sci: a publication of the Protein Society* 12: 1158-68

Cecchini G, Schroder I, Gunsalus RP, Maklashina E (2002) Succinate dehydrogenase and fumarate reductase from *Escherichia coli. Biochim Biophys Acta* 1553: 140-157

Chalabaev S, Chauhan A, Novikov A, Lyer P, Szczesny M, Beloin C, Caroff M, Ghigo JM (2014) Biofilms formed by gram-negative bacteria undergo increased lipid a palmitoylation, enhancing in vivo survival. *MBio* 5

Constantinidou C, Hobman JL, Griffiths L, Patel MD, Penn CW, Cole JA, Overton TW (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli K12* adapts from aerobic to anaerobic growth. *J Biol Chem* 281: 4802-15

Coornaert A, Lu A, Mandin P, Springer M, Gottesman S, Guillier M (2010) MicA sRNA links the PhoP regulon to cell envelope stress. *Mol Microbiol* 76: 467-79

Cuesta-Seijo JA, Neale C, Khan MA, Moktar J, Tran CD, Bishop RE, Pomes R, Prive GG (2010) PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler. *Structure* 18: 1210-9

Crofts, AA, Giovanettia SG, Rubina, EJ, Polyb, FM, Gutiérrez, RL, Talaatc, KR, Porterb, CK, Riddleb, MS, DeNearingc, B, Brubakerc, J, Jr. MM, Alcalab, AN, Chakrabortyc, S, Proutyb, MG, Savarinob, SJ, Daviesa, BD, Trentf, MS (2018) Enterotoxigenic *E. coli* virulence gene regulation in human infections. *PNAS* 1-9

Dalebroux ZD, Miller SI (2014) Salmonellae PhoPQ regulation of the outer membrane to resist innate immunity. *Curr Opin Microbiol* 17: 106-13

Dalebroux ZD, Matamouros S, Whittington D, Bishop RE, Miller SI (2014) PhoPQ regulates acidic glycerophospholipid content of the *Salmonella Typhimurium* outer membrane. *Proc Natl Acad Sci U S A* 111: 1963-8

Dartigalongue C, Missiakas D, Raina S (2001) Characterization of the *Escherichia coli* sigma E regulon. *J Biol Chem* 276: 20866-75

David SA, Bechtel B, Annaiah C, Mathan VI, Balaram P (1994) Interaction of cationic amphiphilic drugs with lipid A: implications for development of endotoxin antagonists. *Biochim Biophys Acta* 1212: 167-75

Dong T, Kirchhof MG, Schellhorn HE (2008) RpoS regulation of gene expression during exponential growth of *Escherichia coli K12*. *Mol Genet Genomics* 279: 267-77

Dong T, Schellhorn HE (2009) Global effect of RpoS on gene expression in pathogenic *Escherichia coli* O157:H7 strain EDL933. *BMC Genomics* 10: 349

Eguchi Y, Okada T, Minagawa S, Oshima T, Mori H, Yamamoto K, Ishihama A, Utsumi R (2004) Signal Transduction Cascade between EvgA/EvgS and PhoP/PhoQ Two-Component Systems of *Escherichia coli*. *Journal of Bacteriology* 186: 3006-3014

Eitinger T, Mandrand-Berthelot MA (2000) Nickel transport systems in microorganisms. *Arch Microbiol* 173: 1-9

Ferrandez A, Minambres B, Garcia B, Olivera ER, Luengo JM, Garcia JL, Diaz E (1998) Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. *J Biol Chem* 273: 25974-86

Ferrezuelo F, Prieto-Alamo MJ, Jurado J, Pueyo C (1998) Influence of DNA repair by (A)BC excinuclease and Ogt alkyltransferase on the distribution of mutations induced by n-propyl-N-nitrosourea in *Escherichia coli*. *Environ Mol Mutagen* 31: 82-91

Fried L, Lassak J, Jung K (2012) A comprehensive toolbox for the rapid construction of *lacZ* fusion reporters. *J Microbiol methods* 91: 537-43

Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T, James JJ, Maysuria M, Mitton JD, Oliveri P, Osborn JL, Peng T, Ratcliffe AL, Webster PJ, Davidson EH, Hood L et

al. (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* 26: 317-25

Gottesman S (1984) Bacterial regulation: global regulatory networks. *Annu Rev Genet* 18: 415-41

Grabowicz M, Silhavy TJ (2017) Envelope Stress Responses: An Interconnected Safety Net. *Trends Biochem Sci* 42: 232-242

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

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

Guo L, Lim KB, Gunn JS, Bainbridge B, Darveau RP, Hackett M, Miller SI (1997) Regulation of lipid A modifications by Salmonella typhimurium virulence genes phoP-phoQ. *Science (New York, NY)* 276: 250-3

Guo L, Lim KB, Poduje CM, Daniel M, Gunn JS, Hackett M, Miller SI (1998) Lipid A Acylation and Bacterial Resistance against Vertebrate Antimicrobial Peptides. *Cell* 95: 189-198

Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177: 4121-30

Hagiwara D, Yamashino T, Mizuno T (2004) A Genome-wide view of the *Escherichia coli* BasS-BasR two-component system implicated in iron-responses. *Biosci Biotechnol Biochem* 68: 1758-67

Hagewood BT, Ganduri YL, Datta P (1994) Functional analysis of the *tdcABC* promoter of *Escherichia coli*: roles of TdcA and TdcR. *J Bacteriol* 176: 6214-20

Hasona A, Self WT, Ray RM, Shanmugam KT (1998) Molybdate-dependent transcription of *hyc* and *nar* operons of *Escherichia coli* requires MoeA protein and ModE-molybdate. *FEMS Microbiol Lett* 169: 111-6

Hengge-Aronis R (2002) Signal Transduction and Regulatory Mechanisms Involved in Control of the S (RpoS) Subunit of RNA Polymerase. *Microbiology and Molecular Biology Reviews* 66: 373-395

Ho H, Miu A, Alexander MK, Garcia NK, Oh A, Zilberleyb I, Reichelt M, Austin CD, Tam C, Shriver S, Hu H, Labadie SS, Liang J, Wang L, Wang J, Lu Y, Purkey HE, Quinn J, Franke Y, Clark K, Beresini, MH, Tan MW, Sellers, BD, Maurer, T, Koehler, MFT, Wecksler, AT, Kiefer, JR, Verma, V, Xu, Y, Nishiyama, M, Payandeh, J, Koth, CM (2018) Structural basis for dual-mode inhibition of the ABC transporter MsbA. *Nature* 557: 196-201

Hu Y, Benedik MJ, Wood TK (2012) Antitoxin DinJ influences the general stress response through transcript stabilizer CspE. *Environ Microbiol* 14: 669-79

Hu KH, Liu E, Dean K, Gingras M, DeGraff W, Trun NJ (1996) Overproduction of three genes leads to camphor resistance and chromosome condensation in *Escherichia coli*. *Genetics* 143: 1521-32

Huang YH, Ferrieres L, Clarke DJ (2006) The role of the Rcs phosphorelay in Enterobacteriaceae. *Res Microbiol* 157: 206-12

Hwang PM, Bishop RE, & Kay L E (2004) From The Cover: The integral membrane enzyme PagP alternates between two dynamically distinct states. *Proc. Natl. Acad. Sci*. 101:26: 9618–9623.

Ito A, May T, Kawata K, Okabe S (2008) Significance of *rpoS* during maturation of *Escherichia coli* biofilms. *Biotechnol Bioeng* 99: 1462-71

Jacobi A, Rossmann R, Bock A (1992) The *hyp* operon gene products are required for the maturation of catalytically active hydrogenase isoenzymes in *Escherichia coli*. *Arch Microbiol* 158: 444-51

Jacobs MHJ, Van Der Heide T, Tolner B, Driessen AJM, Konings WN (1995) Expression of the *gltP* gene of *Escherichia coli* in a glutamate transport-deficient mutant of Rhodobacter sphaeroides restores chemotaxis to glutamate. *Mol Microbiol* 18: 641-647

Jia W, El Zoeiby A, Petruzziello TN, Jayabalasingham B, Seyedirashti S, Bishop RE (2004) Lipid trafficking controls endotoxin acylation in outer membranes of *Escherichia coli*. *J Biol Chem* 279: 44966-75

Johnston D, Tavano C, Wickner S, Trun N (2006) Specificity of DNA binding and dimerization by CspE from *Escherichia coli*. *J Biol Chem* 281: 40208-15

Kang Y, Weber KD, Qiu Y, Kiley PJ, Blattner FR (2005) Genome-wide expression analysis indicates that FNR of *Escherichia coli K-12* regulates a large number of genes of unknown function. *J Bacteriol* 187: 1135-60

Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muniz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M,

Altman T, Kaipa P, Spaulding A, Pacheco J, Latendresse M, Fulcher C, Sarker M, Shearer AG, Mackie A, Paulsen I, Gunsalus RP et al. (2011) EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res* 39: D583-90

Kiley PJ, Beinert H (2003) The role of Fe-S proteins in sensing and regulation in bacteria. *Curr Opin Microbiol* 6: 181-5

Klein G, Lindner B, Brade H, Raina S (2011) Molecular basis of lipopolysaccharide heterogeneity in *Escherichia coli*: envelope stress-responsive regulators control the incorporation of glycoforms with a third 3-deoxy-alpha-D-manno-oct-2-ulosonic acid and rhamnose. *J Biol Chem* 286: 42787-807

Klein G, Raina S (2015) Regulated Control of the Assembly and Diversity of LPS by Noncoding sRNAs. *BioMed Res Int* 2015: 153561

Konovalova A, Perlman DH, Cowles CE, Silhavy TJ (2014) Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of beta-barrel proteins. *Proc Natl Acad Sci U S A* 111: E4350-8

Kulkarni MM (2011) Digital multiplexed gene expression analysis using the NanoString nCounter system. *Curr Protoc Mol Biol* Chapter 25: Unit25B.10

Lacour S, Landini P (2004) SigmaS-dependent gene expression at the onset of stationary phase in *Escherichia coli*: function of sigma S-dependent genes and identification of their promoter sequences. *J Bacteriol* 186: 7186-95

Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* 317: 41-72

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

Lubitz SP, Weiner JH (2003) The *Escherichia coli ynfEFGHI* operon encodes polypeptides which are paralogues of dimethyl sulfoxide reductase (DmsABC). *Arch Biochem Bi* 418: 205-16

Lima S, Guo MS, Chaba R, Gross CA, Sauer RT (2013) Dual molecular signals mediate the bacterial response to outer-membrane stress. *Science (New York, NY)* 340: 837-41

Majdalani N, Chen S, Murrow J, St John K, Gottesman S (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol* 39: 1382-94

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

Malinverni JC, Silhavy TJ (2009) An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc Natl Acad Sci USA* 106: 8009-14

Marguerat S, Bahler J (2010) RNA-seq: from technology to biology. *Cell Mol Life Sci: CMLS* 67: 569-79

Martinez-Antonio A, Collado-Vides J (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* 6: 482-9

May KL, Silhavy TJ (2018) The *Escherichia coli* Phospholipase PldA Regulates Outer Membrane Homeostasis via Lipid Signaling. *MBio* 9: 379-18.

McNicholas, PM, Chiang, RC, Robert, P (1998) Anaerobic regulation of the *Escherichia coli dmsABC* operon requires the molybdate-responsive regulator ModE. *Mol Microbiol 27*: 197–208.

Miller JH (1972) Experiments in Molecular Genetics. CSH Laboratory Press, Cold Spring Harbor, NY

Miller SI, Ernst RK, Bader MW (2005) LPS, TLR4 and infectious disease diversity. *Nat Rev Microbiol* 3: 36-46

Miller VL, MEKALANOS JJ (1988) A Novel Suicide Vector and Its Use in Construction of Insertion Mutations: Osmoregulation of Outer Membrane Proteins and Virulence Determinants in Vibrio cholerae Requires *toxR*. *J Bacteriol* 170:2575-2583

Navasa N, Rodriguez-Aparicio L, Ferrero MA, Monteagudo-Mera A, Martinez-Blanco H (2013) Polysialic and colanic acids metabolism in *Escherichia coli* K92 is regulated by RcsA and RcsB. *Biosci Rep* 33: 405–415

Nishimura Y, Tan IKP, Ohgami Y, Kohgami K, Kamihara T (1983) Induction of membrane-bound l-lactate dehydrogenase in *Escherichia coli* under

conditions of nitrate respiration, fumarate reduction and trimethylamine-N-oxide reduction. *FEMS microbiology letters* 17: 283-286

Nishino K, Honda T, Yamaguchi A (2005) Genome-wide analyses of *Escherichia coli* gene expression responsive to the BaeSR two-component regulatory system. *J Bacteriol* 187: 1763-72

Nitta T, Nagamitsu H, Murata M, Izu H, Yamada M (2000) Function of the sigma(E) regulon in dead-cell lysis in stationary-phase *Escherichia coli*. *J Bacteriol* 182: 5231-7

Nikaido H, Vaara M (1985) Molecular basis of bacterial outer membrane permeability. *Microbiological reviews* 49: 1-32

Ogasawara H, Shinohara S, Yamamoto K, Ishihama A (2012) Novel regulation targets of the metal-response BasS-BasR two-component system of *Escherichia coli*. *Microbiology (Reading, England)* 158: 1482-92

Ozenberger BA, Nahlik MS, McIntosh MA (1987) Genetic organization of multiple fep genes encoding ferric enterobactin transport functions in *Escherichia coli*. *J Bacteriol* 169: 3638-46

Pandey, L (2013) A NEW ROLE FOR PAGP IN SIGNAL TRANSDUCTION. (MSc.Thesis). Retrieved from https://macsphere.mcmaster.ca/bitstream/ 11375/ 13511/1/fulltext.pdf.

Pao SS, Paulsen IT, Saier MH, Jr. (1998) Major facilitator superfamily. *Microbiol Mol Biol Rev: MMBR* 62: 1-34

Park SJ, Chao G, Gunsalus RP (1997) Aerobic regulation of the *sucABCD* genes of *Escherichia coli*, which encode alpha-ketoglutarate dehydrogenase and succinyl coenzyme A synthetase: roles of ArcA, Fnr, and the upstream *sdhCDAB* promoter. *J Bacteriol* 179: 4138-42

Partridge JD, Browning DF, Xu M, Newnham LJ, Scott C, Roberts RE, Poole RK, Green J (2008) Characterization of the *Escherichia coli K-12 ydhYVWXUT* operon: regulation by FNR, NarL and NarP. *Microbiology (Reading, England)* 154: 608-18

Passalacqua KD, Varadarajan A, Weist C, Ondov BD, Byrd B, Read TD, Bergman NH (2012) Strand-specific RNA-seq reveals ordered patterns of sense and antisense transcription in Bacillus anthracis. *PloS one* 7: e43350

Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli K-12*. *Mol Genet Genomics* 272: 580-91

Picard F, Dressaire C, Girbal L, Cocaign-Bousquet M (2009) Examination of post-transcriptional regulations in prokaryotes by integrative biology. *C R Biol* 332: 958-73

Rahman M, Hasan MR, Oba T, Shimizu K (2006) Effect of *rpoS* gene knockout on the metabolism of *Escherichia coli* during exponential growth phase and early stationary phase based on gene expressions, enzyme activities and intracellular metabolite concentrations. *Biotechnol Bioeng* 94: 585-95

Raetz CR, Reynolds CM, Trent MS, Bishop RE (2007) Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* 76: 295-329

Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* 71: 635-700

Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordonez B, Gullane P, Irish J, Jurisica I, Kamel-Reid S (2011) mRNA transcript quantification in archival samples using multiplexed, color-coded probes. *BMC Biotechnol* 11: 46

Ruiz N, Falcone B, Kahne D, Silhavy TJ (2005) Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* 121: 307-17

Ruiz N, Silhavy TJ (2005) Sensing external stress: watchdogs of the *Escherichia coli* cell envelope. *Curr Opin Microbiol* 8: 122-6

Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ (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: 5537-42

Sambasivarao D, Weiner JH (1991) Differentiation of the multiple S- and N-oxide-reducing activities of *Escherichia coli. Current Microbiology* 23: 105-110

Sambasivarao D, Weiner JH (1991) Dimethyl sulfoxide reductase of *Escherichia coli*: an investigation of function and assembly by use of in vivo complementation. *J Bacteriol* 173: 5935-43

Salmon KA, Hung SP, Steffen NR, Krupp R, Baldi P, Hatfield GW, Gunsalus RP (2005) Global gene expression profiling in *Escherichia coli K12*: effects of oxygen availability and ArcA. *J Biol Chem* 280: 15084-96

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

Sambrook J, & Russell DW (2001) Molecular cloning: a laboratory manual 3rd edition. Cold Spring Harbour Laboratory Press, UK

Salvail H, Lanthier-Bourbonnais P, Sobota JM, Caza M, Benjamin JA, Mendieta ME, Lepine F, Dozois CM, Imlay J, Masse E (2010) A small RNA promotes siderophore production through transcriptional and metabolic remodeling. *Proc Natl Acad Sci U S A* 107: 15223-8

Sand O, Gingras M, Beck N, Hall C, Trun N (2003) Phenotypic characterization of overexpression or deletion of the *Escherichia coli crcA*, *cspE* and *crcB* genes. *Microbiology (Reading, England)* 149: 2107-17

Sawers G (2001) A novel mechanism controls anaerobic and catabolite regulation of the *Escherichia coli tdc* operon. *Mol Microbiol* 39: 1285-98

Schroeder A, Mueller O, Stoker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Grazow M and Ragg T (2006) The RIN: a TNA integrity number for assigning integrity values to RNA measurements. BMC *Mol Biol* 7:3-3

Sevin DC, Fuhrer T, Zamboni N, Sauer U (2017) Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in *Escherichia coli*. *Nat Methods* 14: 187-194

Silhavy T, Berman M, Enquist L (1984) Experiments with gene fusions. Cold Springs Harbor: *Cold Spring Harbor Laboratory Press*. U S A

Simons RW, Houman F, Kleckner N (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* 53: 85-96

Smith AE, Kim SH, Liu F, Jia W, Vinogradov E, Gyles CL, Bishop RE (2008) PagP activation in the outer membrane triggers R3 core oligosaccharide truncation in the cytoplasm of *Escherichia coli* O157:H7. *J Biol Chem* 283: 4332-43

Sherman DJ, Xie R, Taylor RJ, George AH, Okuda S, Foster PJ, Needleman DJ, Kahne D (2018) Lipopolysaccharide is transported to the cell surface by a membrane-to-membrane protein bridge. *Science (New York, NY)* 359: 798-801

Shprung T, Peleg A, Rosenfeld Y, Trieu-Cuot P, Shai Y (2012) Effect of PhoP-PhoQ activation by broad repertoire of antimicrobial peptides on bacterial resistance. *J Biol Chem* 287: 4544-51

Stanley NR, Sargent F, Buchanan G, Shi J, Stewart V, Palmer T, Berks BC (2002) Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol Microbiol* 43: 1005-21

Stroud RM, Kay LM, Dickerson RE (1972) The crystal and molecular structure of DIP-inhibited bovine trypsin at 2.7 Angstrom resolution. *Cold Spring Harb Symp Quantit Biol* 36: 125-40

Stockbridge RB, Kolmakova-Partensky L, Shane T, Koide A, Koide S, Miller C, Newstead S (2015) Crystal structures of a double-barrelled fluoride ion channel. *Nature* 525: 548-51

Stokes JM, MacNair CR, Ilyas B, French S, Cote JP, Bouwman C, Farha MA, Sieron AO, Whitfield C, Coombes BK, Brown ED (2017) Pentamidine sensitizes Gram-negative pathogens to antibiotics and overcomes acquired colistin resistance. *Nature microbiology* 2: 17028

Sturm A, Schierhorn A, Lindenstrauss U, Lilie H, Bruser T (2006) YcdB from *Escherichia coli* reveals a novel class of Tat-dependently translocated hemoproteins. *J Biol Chem* 281: 13972-8

Tam C, Missiakas D (2005) Changes in lipopolysaccharide structure induce the sigma(E)-dependent response of *Escherichia coli*. *Mol Microbiol* 55: 1403-12

Tang H, Rothery RA, Voss JE, Weiner JH (2011) Correct assembly of iron-sulfur cluster FS0 into *Escherichia coli* dimethyl sulfoxide reductase (DmsABC) is a prerequisite for molybdenum cofactor insertion. *J Biol Chem* 286: 15147-54

Tang H, Rothery RA, Weiner JH (2013) A variant conferring cofactor-dependent assembly of *Escherichia coli* dimethylsulfoxide reductase. *Biochim Biophys Acta* 1827: 730-7

Van Elsas JD, Semenov AV, Costa R, Trevors JT (2011) Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *The ISME journal* 5: 173-83

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10: 57-63

Walsh NP, Alba BM, Bose B, Gross CA, Sauer RT (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

Weiner JH, Rothery RA, Sambasivarao D, Trieber CA (1992) Molecular analysis of dimethylsulfoxide reductase: a complex iron-sulfur molybdoenzyme of *Escherichia coli. Bioch Biophys Acta* 1102: 1-18

Weiner JH, Shaw G, Turner RJ, Trieber CA (1993) The topology of the anchor subunit of dimethyl sulfoxide reductase of *Escherichia coli*. *J Biol Chem* 268: 3238-44

Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93: 93-101

Wu T, McCandlish AC, Gronenberg LS, Chng SS, Silhavy TJ, Kahne D (2006) Identification of a protein complex that assembles lipopolysaccharide in the

outer membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A*. 103: 11754-9

Xi H, Schneider BL, Reitzer L (2000) Purine catabolism in *Escherichia coli* and function of xanthine dehydrogenase in purine salvage. *J Bacteriol* 182: 5332-41

Yamada M, Talukder AA, Nitta T (1999) Characterization of the *ssnA* gene, which is involved in the decline of cell viability at the beginning of stationary phase in *Escherichia coli*. *J Bacteriol* 181: 1838-46

Yan A, Guan Z, Raetz CR (2007) An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in *Escherichia coli*. *J Biol Chem* 282: 36077-89

Zhang G, Baidin V, Pahil KS, Moison E, Tomasek D, Ramadoss NS, Chatterjee AK, McNamara CW, Young TS, Schultz PG, Meredith TC, Kahne D (2018) Cell-based screen for discovering lipopolysaccharide biogenesis inhibitors. *Proc Natl Acad Sci U S A* 115: 6834-6839

Zientz E, Six S, Unden G (1996) Identification of a third secondary carrier (DcuC) for anaerobic C4-dicarboxylate transport in *Escherichia coli:* roles of the three Dcu carriers in uptake and exchange. *J Bacteriol* 178: 7241-7

Zientz E, Janausch IG, Six S, Unden G (1999) Functioning of DcuC as the C4-dicarboxylate carrier during glucose fermentation by *Escherichia coli*. *J Bacteriol* 181: 3716-20