Characterization of novel envelope proteins and their relationship with the Cpx response and stress resistance in *Escherichia coli*

by

Justin Graydon Bishop

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
MICROBIOLOGY AND BIOTECHNOLOGY

Department of Biological Sciences
University of Alberta

Abstract

Stress is a major factor every organism needs to mitigate to survive. Environmental factors such as oxygen content, temperature, and alkalinity can all induce stress and prevent life from progressing. For *Escherichia coli* and other members of the mammalian gastrointestinal tract they need to survive fluctuations in pH, osmolarity, and nutrient availability all while fighting off the host immune system, bacteriophages, and other microbes. The bacterial envelope, the first layer of protection against external stress, is a physical barrier separating the external environment from the cell's cytosol and acts as a matrix for respiration, transport, and the proton motive force. If the envelope fails, cellular processes fail. Envelope homeostasis is maintained by a suite of envelope stress responses that sense and respond to perturbations through the regulation of various envelope processes. *E. coli* possess five major envelope stress responses that work together to mitigate stress at the various membrane layers. The Cpx stress response is a crucial component that helps maintain envelope homeostasis at the inner membrane and periplasmic space through the regulation of inner and transmembrane complexes, chaperones, and proteases. Regulon analysis reveals the Cpx response influences a broad set of cellular processes ranging from virulence to metabolism. The Cpx response's origins lie in the assembly of the conjugative pilus. Conjugation is one of three methods of horizontal gene transfer, the others being natural transformation and transduction by bacteriophages. Some bacteriophages can integrate into the host genome as prophages and have been shown to aid in resistance to various stress, influence host gene regulation, and provide a competitive advantage over other bacterial cells. In this study, our aim was to search for a link between prophage encoded genes in enteropathogenic *E. coli* strain E2348/69 and the Cpx response to determine if they contributed to mitigation of envelope stress. We found that E2348/69 contains active prophages with a narrow host range. We cross referenced the prophage regions in E2348/69 with a transcriptome

dataset collected during Cpx activation and located three putative genes (*yfgG, yfgH, and yfgI*) near, but not encoded by prophage region 10. Further examination of these genes revealed a putative CpxR binding site upstream of *yfgG* and gene expression analysis showed the expression of *yfgG*, and possibly *yfgH,* were influenced by the Cpx response. We found that these proteins are localized to the envelope and aid in resistance to alkaline, cobalt, and nutrient limitation in conjunction with the Cpx response. Absence of the outer membrane lipoprotein YfgH strongly induced the Cpx response in a CpxP dependent manner, a novel finding in the field. Transcript and protein analysis revealed an increase in cellular CpxP levels and periplasmic DegP levels. Together our data suggests these envelope localized proteins are key factors in envelope homeostasis mediated by the Cpx response.

Acknowledgments

I would like to take this opportunity to give a massive thank you to my supervisor Dr. Tracy Raivio for taking me in as an undergrad and then as a graduate student. Working in your lab has been an absolute pleasure over the years. Thank you for all your guidance and insight into this project. None of this would be possible without you. To my co-supervisor, Dr. Amit Bhavsar, you're insight and questions have been greatly appreciated to help shape this project. Troy, and the rest of the MBSU staff, thank you for teaching me all the ins and outs of qPCR. Your help has been invaluable to this project. To the Raivio lab members, you all are the best group of people to work and hang with. Thank you for the support, help, and mostly, all the laughs. The last few years would not have been the same without you. Lastly, I would like to thank my parents, sisters, and friends for the encouragement throughout this degree.

Now it's time to go surf… 

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A/E: Attaching and Effacing

Amk^R: Amikacin resistant

Amp^R: Ampicillin resistant

ATP: Adenosine triphosphate

BLAST: Basic Local Alignment Search Tool

Cam^R: Chloramphenicol resistant

CP: Cytoplasm

CPS: Counts per second, a measure of luminescence

CTD: C-terminal domain

EPEC: Enteropathogenic *Escherichia coli*

ESR: Envelope stress response

ETC: Electron transport chain

Fe/S: Iron sulphur core/cofactors

IM: Inner membrane

IPTG: Isopropyl-β-D- thiogalactopyranoside

Kan^R: Kanamycin resistant

LEE: Locus of enterocyte effacement

LPS: Lipopolysaccharide

MBP: Maltose binding protein MalE

MIC: Minimum inhibitory concentration

mRNA: messenger RNA

NADH: Nicotinamide adenine dinucleotide

NCBI: National Center for Biotechnology Information

NTD: N-terminal domain

OD_{600}: Optical density at 600nm

OM: Outer membrane

OMP: Outer membrane protein

OX: Overexpression

PAGE: Polyacrylamide gel electrophoresis

PCR: Polymerase chain reaction

PMF: Proton motive force

PP: Periplasm

qPCR: Quantitative PCR

RNAP: RNA polymerase

SDS: Sodium dodecyl sulfate

SEM: Standard error of the mean

sRNA: small RNA

Chapter 1: Introduction

Escherichia coli

Escherichia coli are species of Gram-negative bacteria that are among the leading global causes of antimicrobial resistance associated deaths (1). This motile, facultative anaerobe ranges in pathogenicity and hosts a diverse arsenal of resistance and virulence factors to aid in its survival. One of the most well-known groups of *E. coli* is Shiga-Toxin producing *E. coli* (STEC), also known as enterohemorrhagic *E. coli* (EHEC), and are a large factor in major food recalls and *E. coli* outbreaks (2). EHEC causes enterohemorrhagic gastroenteritis and in severe cases, hemolytic-uremic syndrome leading to blood lysis and kidney failure (3). This supercharged group produces a toxin, called Shiga toxin, that inhibits protein synthesis and is primarily responsible for EHEC's disease characteristics (3,4). Not all pathogenic *E. coli* carry the Shiga toxin, enteropathogenic *E. coli* (EPEC) lacks the toxin but is still a leading cause of infantile diarrhea (5). A common subset of virulence genes are responsible for *E. coli* caused gastroenteritis and are found on a pathogenicity island (PAI) called the Locus of Enterocyte Effacement (LEE) (6,7). The LEE encodes a Type III secretion system (T3SS) and various effectors that are responsible for the attachment and effacing (A/E) lesions seen on enterocytes during many *E. coli* infections (7). The hallmark of this phenotype is the formation of a pedestal-like structure causes by actin rearrangement induced by T3SS injected effectors (7). While STEC and EPEC infections are isolated to the gastrointestinal tract, *E. coli* infection can go beyond the gut and are caused by Extraintestinal pathogenic *E. coli* (ExPEC). Included in this group are avian pathogenic *E. coli* (APEC), uropathogenic *E. coli* (UPEC), and neonatal meningitis *E. coli* (NMEC), the latter causing bacteriemia induced meningitis in infants (8). This group uses a diverse subset of specialized genes on various PAIs that aid in its survival away from the gastrointestinal tract (8). For example, the P-pilus commonly found in UPEC allows attachment to urethral cells and the *iss* gene aids in resistance to serum (8).

While *E. coli* are best known for their adverse effects on humans, they actually play an important role as members of the mammalian intestinal microbiome. Commensal *E. coli* generally lack many of the virulence factors, such as those encoded on the LEE, that cause disease in humans but still possess a strong suite of defense mechanisms to help survive the harsh environment of

the gastrointestinal tract. Residents of the mammalian gut need to deal with large fluctuations in pH, nutrient availability, the host immune system, and other members of the microbiome. All members of the microbiome have adapted to thrive under these conditions and fill specific niches within the bacterial community. Residents of the gut also play an important role in protecting the host against opportunistic and invading pathogens. Probably the most well-known example is *E. coli* strain Nissle 1917 and its ability to help prevent infection (9). Nissle 1917, isolated from a solider during World War One who seemed impervious to intestinal disease in an area of high *Shigella* contamination, has adapted well to the environment of the gastrointestinal tract and combat with other bacteria (9). It has now been shown that Nissle 1917 possesses an arsenal of counter measures to fight invading pathogens such as antimicrobial peptides called microcins and modulation of the host immune system (9). Nissle 1917 has also adapted to the low iron availability in the gastrointestinal tract by the production of various siderophores (9). This in turn, has been shown to reduce colonization by enteric pathogen *Salmonella enterica,* limiting access to iron, a key signal for expression of virulence factors (10,11).

Beyond the native habitat of the gut, *E. coli* has become one of the most studied organisms and is considered a laboratory workhorse due to its rapid growth and easy to manipulate genome. It was among the first group of organisms to have their genome fully sequenced in 1997 and has been crucial in advancing the fields of microbiology and genetics (12). While considered well characterized, the function of many of *E. coli*'s genes remain unknown; as of 2019, around 1600 (35%) of *E. coli's* 4623 genes remain uncharacterized (13).

The Bacterial Envelope

The envelope is one of the key defenses a bacterium has to protect itself against the external environment. It acts as both a physical barrier and a framework for many of the cellular processes that have adapted to each organism's unique lifestyle and facilitates the transport of nutrients and waste in and out of the cell (14). Bacterial envelopes fall into three classifications: Gram-positive, Gram-negative, and the envelope of *Corynebacterineae* which includes *Mycobacterium tuberculosis.* All bacterial envelopes possess an inner phospholipid bilayer and peptidoglycan cell wall, but numerous differences exist between the three types. Gram-positives synthesize a thicker peptidoglycan layer with covalently attached surface-exposed teichoic acid

while Gram-negatives produce a thinner cell wall and lack teichoic acid (14). The most notable difference between Gram positive and negative envelopes is the presence of an outer membrane in Gram negatives which adds additional protection to the cell. The envelope of *Corynebacterineae* is quite unique in that it sits in-between the two others. This group of bacteria have covalently attached mycolic acid and a symmetrical outer membrane with unique porins (14). The Gram-negative envelope consists of three distinct parts: an outer lipid bilayer, an inner phospholipid bilayer, and a middle periplasmic space containing a thin peptidoglycan cell wall (15).

The outer membrane is the first point of contact that many cells have with other cells, the host immune system, and the remaining external environment (16). Structurally, it is an asymmetric lipid bilayer composed of inner phospholipid and outer lipopolysaccharide (LPS) leaflets that hold a collection of outer membrane proteins (OMP) and lipoproteins (16). The LPS layer is a crucial component in maintaining cell structure and helps create a densely packed selective barrier due to its amphipathic nature, negative charge, and saturated fatty acids providing and effective barrier to both hydrophobic and hydrophilic molecules (14,17). LPS is made up of three domains: a hydrophobic lipid A, a core oligosaccharide, and a highly diverse O-antigen, the latter used to classify various serogroups in Gram-negative pathogens (17). Lipid A forms the outer leaflet of the outer membrane and is the most conserved of the three components (17). It also can invoke a massive immune response when released in host organisms as it is present in many pathogens; if released in sudden large amounts the host can produce a lethal immune response known as toxic shock (17). The core oligosaccharide contains two parts, an inner well conserved core that attaches to Lipid A, and an outer core that provides a foundation for the O-antigen oligosaccharide (17). The O-antigen is a highly diverse polymer of repeating oligosaccharides that completes the LPS structure (17). Currently, 185 unique serotypes have been classified in *E. coli* alone, highlighting its diversity (18). Absence of the O-antigen can result in increased sensitivity to host immune responses and a drop in virulence (18). The O-antigen can also aid in immune suppression in UPEC and NMEC infections (18). Not all *E. coli* synthesize O-antigen, but still produce the Lipid A and core oligosaccharide (17). In these cases, the LPS is classified as rough, or lipooligosaccharide (17).

Outer membrane proteins make up about 70% of the cell surface and are transmembrane β-barrels that aid in the transport of molecules and ions, signalling, and outer membrane biogenesis (14,16). OmpF and OmpC, two porins involved in passive diffusion are among the most abundant proteins in the membrane with a combined estimate of 250,000 copies per cell (14). OMPs are also thought to cluster together rather than be spatially homogenous which could provide some rigidity to the outer membrane (16). Lipoproteins are proteins anchored in the outer membrane that contain an N-terminal cysteine residue where a lipid moiety is attached (14,16). Lipoproteins are massively abundant, estimated to be over a million proteins per cell in *E. coli* and are involved in a wide range of functions, many not yet known (14,16). So far, they have been identified to be involved in surface sensing, transport to the outer membrane, signalling, and structural integrity. NlpE and RcsF are both thought to monitor the outer membrane and periplasmic space for perturbations (19,20). RcsF is known to interact with the OMP BamA and is required to sense LPS defects (21,22). One of the most abundant lipoproteins in *E. coli* is Lpp and functions to anchor the outer membrane to the peptidoglycan layer by covalent interaction (16,23).

The periplasm is a small space between the inner and outer membrane that is densely packed with proteins and holds a thin peptidoglycan layer that adds rigidity to maintain cell shape and protect against changes in osmotic conditions (14). Peptidoglycan is composed of alternating units of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) glycan strands linked together by short peptides (24). Multiple layers of cross-linked glycan strands account for the rigidity of peptidoglycan and is the primary factor in overall shape of the bacterium (14). One of the most interesting phenomena associated with peptidoglycan is bacterial L-forms, a unique state in which the organism exists without a cell wall. The exact role of L-forms is not well understood although a recent study found *E. coli* strains capable of L-form switching in persistent urinary tract infections (25).

The periplasm also holds a variety of signalling proteins, proteases, chaperones, and envelope spanning complexes. Due to the proximity of the periplasm to the extracellular environment it is exposed to higher levels of stress than the cytoplasmic contents and thus the cell needs to monitor this space to ensure proteins are properly folded and translocated (26,27). This brings up

another issue, a lack of energetic molecules such as ATP, in the periplasm (27). Thus, all periplasmic biological processes have adapted to function independent of ATP but rather rely on binding affinity and free energy of folded conformations (27,28). Periplasmic chaperones either act to slow protein folding such as in preventing OMPs destined to the outer membrane from misfolding and aggregating or as a catalyst such as in the formation of disulfide bonds (26,27). For example, Skp aids in the stabilization of nascent OMP peptides, such as OmpA, during their transit to the outer membrane such as OmpA while the stress associated chaperone Spy stabilizes unfolded peptides and allows them to fold or refold while preventing aggregation of others (27). The periplasm is a strong oxidizing environment which promotes spontaneous bonds between thiol residues in proteins (29). Properly located disulfide bonds are an important component of maintaining the shape and function of some proteins, while aberrant bonds can lead to a reduced or loss of function (29) A group of inner membrane and periplasmic enzymes called the disulfide bond (Dsb) family aid in the oxidation and specificity of thiol bond formation by transferring the electrons form the thiol residues to the quinone pool (29).

Unlike the outer membrane, the inner membrane, also known as the cytoplasmic membrane, is an impermeable barrier to most molecules and acts as a substitute for organelles found in eukaryotes (14). The composition of the phospholipid inner membrane can vary across bacterial species and can change within a single cell to adapt to various environmental conditions (30). For example, the ratio of unsaturated to saturated fatty acids adjusts to maintain membrane fluidity during temperature swings (30). A large component of the inner membrane are the vast amount of integral membrane proteins, making up around 20-30% of the cell's total proteins (31). These proteins contain α-helix transmembrane domains and are involved in a variety of biological pathways from signalling to metabolism (32). Additionally, large envelope spanning complexes such as efflux pumps and Type III secretion systems are anchored at the inner membrane. All integral proteins require assistance when translocating and inserting into the inner membrane (32). Two transport systems exist to facilitate movement across and insertion into the inner membrane. The Sec and Tat systems (discussed later) are key components in envelope biogenesis but require large amounts of energy to accomplish their tasks. Due to its impermeability, the inner membrane allows the formation of the proton motive force (PMF), an electrochemical gradient of protons between the periplasm and cytosol formed by primary proton

pumps such as respiratory complexes coupled to electron movement (33). For example, the transfer of electrons by NADH dehydrogenase I pumps 3-4 protons across the inner membrane (34,35). The PMF consists of two components: a pH gradient and a transmembrane potential and is actively maintained by a multitude of cellular processes (33). The most well understood use of the PMF is to drive ATP synthesis, although other uses have been found such as in rotating the flagellar motor (36,37).

Protein Transport in the Envelope

All proteins localized to the envelope need to be either translocated across or inserted into the inner membrane. *E. coli* has several pathways to aid in protein movement and envelope biogenesis. The Sec and Tat pathways translocate proteins across the inner membrane, and the Lol and Bam pathways move lipoproteins and OMPs to the outer membrane (14). Specialized pathways also exist for surface exposed structures such as pili and LPS (14).

The Sec translocase is the most widely used translocation and insertion pathway in *E. coli* and moves unfold polypeptides across the inner membrane through the SecY channel (38,39). The highly conserved pathway consists of a translocation complex SecYEG, SecB chaperone, and the SecA ATPase (39,40), Additional auxiliary components SecDF, YajC, and YidC aid in later stages of translocation but their roles are not fully understood, although evidence suggests YidC plays a role in the insertion of certain inner membrane proteins (39). Depending on their N-terminal signal sequence, protein precursors are transported via two methods; a signal recognition particle (SRP) that is coupled to translation, or SecA mediated transport, which is the predominant pathway in *E. coli* (39). In SecA mediated transport, SecA binds to newly synthesized precursors in the cytoplasm and delivers them to the SecYEG complex, upon which it hydrolyzes ATP, and the protein is moved across the membrane (39,40). SecB aids in stabilizing unfolded proteins until translocation (40). After translocation, proteins either remain in the periplasm and are folded by chaperones or continue to the outer membrane via the Bam and Lol pathways.

While the Sec pathways supports transport of unfolded proteins, the Tat (Twin-arginine) translocation pathway moves pre-folded proteins into the periplasm. In *E. coli,* the Tat system

consists of membrane bound TatABCD and TatE (41). Upon binding by TatBC to a signal peptide, TatA is recruited to the complex thought to form a channel (41,42). No definitive mechanism of translocation exists for the Tat pathway as it is not well conserved across bacterial species, several of its members can be substitutes for each other and multiple complexes with various ratios of each protein have been observed (41–43). The number of targets for the Tat translocation system is abysmally small in *E. coli*, only around 30 proteins, but many of its substrates carry cofactors, such as Fe/S cores, specific metal ion insertions, or are hetero-oligomeric complex (41,42).

Protein transport to the outer membrane is mainly accomplished by two pathways: the β-barrel assembly machine (BAM), and the localization of lipoproteins (Lol) export pathway. After dissociation from the SecYEG complex, OMP peptides are bound by periplasmic chaperones SurA, FkpA, and Skp to mediate their transit across the periplasm to the Bam complex (44). The Bam complex consists of a central β-barrel BamA and four lipoproteins: BamB, BamC, BamD, and BamE (44). BamA contains 5 polypeptide transport-associated (POTRA) motifs that interact with BamBCDE and may play a role in the folding and insertion of OMP peptides into the outer membrane (44). The exact mechanism of OMP folding and insertion is not known, but one of the current models is known as lateral opening where the seam in BamA's β-barrel adopts an open conformation to provide a template for OMP folding, aided by the POTRA domains (45).

Prolipoproteins have a different fate after arrival into the periplasm as their journey to the outer membrane is mediated by the Lol pathway. The Lol pathway includes a combination of periplasmic chaperone LolA, outer membrane lipoprotein LolB, and an inner membrane ATP-binding cassette (ABC) transporter LolCDE (46). Additional inner membrane proteins, Lgt, LspA, and Lnt modify the lipoproteins before transit to the outer membrane in a process known as lipoprotein maturation (46). Prolipoproteins undergo maturation at the inner membrane where they are first acylated by Lgt with a diacyl lipid moiety at the terminal cysteine residue of a signal sequence known as the lipobox (46). The signal peptide is then cleaved by LspA and further lipid modification is catalyzed by Lnt (46). Whether the lipoproteins remain at the inner membrane or are transported to the outer membrane depends on the amino acid adjacent to the conserved cysteine residue where the presence of an aspartate residue signals localization of the

lipoprotein to the inner membrane (46). Around 90% of lipoproteins in *E. coli* are translocated to the outer membrane (46). After modification by Lnt, lipoproteins destined to the outer membrane are captured by LolE and extracted from the inner membrane by LolD ATP hydrolysis (46). The lipoprotein is then transferred to the outer membrane by the LolA chaperone and anchored into the outer membrane by LolB (46).

LPS synthesis takes place at the cytosolic side of the inner membrane and requires a multistep process to construct a complete LPS molecule and transport it to the outer membrane (17). LPS molecules are not substrates of the Sec of Tat pathways, but instead are translocated by the MsbA flippase to the periplasmic side of the inner membrane where any modifications, including O-antigen addition take place (17). The envelope spanning Lpt complex (LptA-G) mediates transfer across the periplasm and outer membrane (17). LptB2FG form an ABC transporter that provides energy by ATP hydrolysis to extract and move LPS to the outer membrane (17). It is thought that LPS is pushed along a bridge formed by LptCAD connecting the inner and outer membranes (17). Transport across the outer membrane and final insertion of LPS is mediated by a LptD β-barrel-like OMP and lipoprotein LptE (17).

Envelope Stress and Two Component Systems

All forms of life encounter various forms of environmental, mechanical, or internal stress that poses a serious threat to the survival of the organism. *E. coli*'s native habitat in the mammalian gastrointestinal tract is a ruthless environment where the cell is exposed to a range of stress that can severely damage the cell and impair biological processes. Fluctuations in pH or redox potential can induce misfolding of proteins as amino acid sidechains either are protonated, or deprotonated and reactive oxygen species are generated, resulting non-functional proteins and damaged membranes (47). Shifts in osmolarity can disrupt ion gradients and increase the differential pressure on the cell wall (47). The cell also must combat the hosts immune system and members of the native microbiota. For instance, human α-defensins, host produced antimicrobial peptides, can disrupt the bacterial membrane through pore formation, allowing leakage of cytoplasmic contents, and the destruction of the proton motive force and ion gradients (48,49). *E. coli* needs to mitigate these stresses to survive and colonize the mammalian gut. This

is accomplished by sensing and responding to perturbations in the envelope using several stress responses.

E. coli has 5 major envelope stress responses (ESR); Bae (Bacterial adaptive response), Cpx (Conjugative plasmid expression), Psp (Phage shock protein), Rcs (Regulator of capsule synthesis), and the Sigma E response. The Rcs and Sigma E response both aid in mitigation of outer membrane stress. The Rcs pathway responds to disturbances in the outer membrane, LPS, and peptidoglycan by regulation of capsule, biofilm, motility, and virulence genes while Sigma E responds to heat shock, altered LPS, and misfolded outer membrane proteins (OMP) though regulation of protein folding pathways (19,20,47). The Psp and Cpx responses are responsible for inner membrane and periplasmic homeostasis (20,47). The Psp response, originally identified in resistance to filamentous bacteriophage infection, monitors the inner membrane for severe damage such as organic solvent and ethanol exposure, mis-localized OMPs, and extreme heat shock that lead to a loss of proton motive force (PMF) resulting in upregulation of *psp* genes, although their exact effect on stress mitigation is unclear (47,50). The Cpx pathway responds to a diverse set of stimuli and its regulon is extensive, but its main role appears to be monitoring the inner membrane and periplasm for mis-folded proteins and appropriately regulating protein folding and degradation factors (20,51). The role of the Bae response is less understood, but the system has strong ties to the MdtABC efflux system and aids in zinc, ethanol, and indole resistance and is thought to aid in envelope homeostasis via efflux and regulation of the protein folding chaperone Spy (47,52). These systems do not exist independently of each other as there is overlap in their roles in envelope homeostasis and known cross talk between them. The periplasmic chaperone Spy is regulated by both the Cpx and Bae responses; DegP, a periplasmic endoprotease is regulated by Cpx and Sigma E (53–55). The small RNA *rprA* is dually regulated by the Cpx and Rcs systems (56,57). Additionally, transcription of *rpoE* has been shown to be repressed by CpxR (58,59). All the crosstalk between these responses indicates maintaining envelope homeostasis is a complex and tedious task that has been fine turned over the millennia.

Mechanistically, ESR's are signalling pathways that sense and respond to stimuli. One of the basic types of signalling pathways is the two-component systems (TCS). TCS use an inner membrane bound histidine kinase (HK) to sense stimuli and a cytoplasmic response regulator

(RR) to appropriately respond (60). Upon sensing an inducing cue directly or through auxiliary envelope proteins, the HK will auto-phosphorylate and transfer a phosphoryl group to its cognate RR (60). This addition of a phosphoryl group to the RR typically causes a conformational change, permitting binding to DNA (60). The HK characteristically possesses phosphatase activity and in the absence of any inducing cues the HK de-phosphorylates the RR to prevent DNA binding and maintains the system at a steady state (60). TCS are abundant across the Bacterial domain; *E. coli* alone has 32 TCS that range from regulation of anerobic respiration in the ArcA/B response, inorganic phosphate starvation with PhoR/B, to envelope homeostasis with the CpxR/A and BaeR/S (20,47,52,61,62).

The structure of HK's are highly conserved and typically contains periplasmic sensing, transmembrane, cytosolic sensing, and enzymatic domains (60). Signals are commonly sensed through periplasmic or cytoplasmic PAS-like (**P**ER-**A**RNT-**S**IM) domains, although it has been shown that some HK can detect signals through an α-helical domain (60,63). Upon detection of a signal, a HK will undergo a conformational change in its C-terminal HAMP (**h**istidine kinases, **a**denylyl cyclases, **m**ethylaccepting proteins, and other **p**rokaryotic signaling proteins) domain that exposes a conserved histidine reside in the enzymatic domain (60,64). As two HK dimerize, the histidine will cross-phosphorylate and upon interaction with the cognate RR, transfer the phosphate group to an aspartate residue in the receiver domain of the RR (60,64). Phosphorylation of the RR receiver domain induces a conformational change that permits a change in the output response, typically a modification in DNA binding affinity although binding to proteins and RNA have also been observed (60,64). Dephosphorylation of the RR depends on several factors such as prevalence of specific phosphatases, and RR intrinsic auto-dephosphorylation rates (60,64). In many cases, the HK will possess phosphatase activity to control the pathway output in the absence of inducing signals (64).

The Cpx Envelope Stress Response

The Cpx stress response is regulated by a canonical two component system that plays a major role in the maintenance of envelope homeostasis and detection of mis-folded periplasmic proteins in many Gram-negative bacteria (20,51). The pathway consists of an inner membrane bound histidine kinase CpxA, a cytoplasmic response regulator CpxR and auxiliary sensing

proteins CpxP and NlpE also aid in sensing and maintaining envelope homeostasis (20). CpxP is an inhibitor of the response, forming a bowl-like dimer that interacts with the periplasmic domain of CpxA (51,65,66). It is thought that CpxP monitors the periplasm by binding misfolded proteins, resulting in the de-repression of CpxA and activation of the system, although the exact method of binding misfolded proteins is unclear (20,65,66). The current models suggest binding mis-folded proteins results in the proteolysis of the protein-CpxP complex via DegP, alleviating repression of CpxA (20,67,68). Interestingly, CpxP is not required for many inducing signals of the Cpx response such as alkaline pH and NlpE overexpression (20,69).

Discovered when overexpression rescued LamB-LacZ-PhoA mediated toxicity, NlpE is an outer membrane lipoprotein that functions as a sensor for surface adhesion, copper stress, and the maturation and transport of lipoproteins (70–74). Deletion of NlpE has also been shown to reduce Cpx activation when DsbA mediated oxidative protein folding is absent and is required for Cpx detection of overexpressed OmpA (70,75). A previous structural study solving the crystal structure of NlpE hypothesized that a structural change in its N-terminal domain (NTD) allowed the protein to reach down and interact with CpxA via its C-terminal domains (CTD) but has since been contested as NlpE's CTD is expendable for Cpx activation while the NTD was found to directly interact with CpxA (20,70,76). Interestingly NlpE is not required for many Cpx activating cues such as alkaline pH, PapG overexpression, and EDTA (69). Overexpression of NlpE is a strong inducer of the response as elevated production results in mis-localization to the inner membrane, an inducing cue that has recently linked NlpE and the Cpx response to monitoring lipoprotein trafficking to the outer membrane (20,74,77–79).

The Cpx response was initially discovered to be a key component in the formation of the F-pilus in conjugation, but its role in multifaceted role in envelope homeostasis was accelerated when mutations in *cpxA* causing constitutive activation were found to rescue a mis-localized LamB-LacZ-PhoA fusion protein (20,80,81). The system is now known as one of the central pathways in sensing and mitigating the effects caused by envelope stress. Known inducers are alkaline pH, surface adhesion, high osmolarity, divalent cations, EDTA, and β-lactam antibiotics, furthermore, the Cpx response is required for resistance to many of these conditions (20,51,82). The CpxR regulon is extensive and has been linked to numerous pathways including adherence

and motility appendages, virulence mechanisms, metabolic pathways, lipoprotein biogenesis, other TCS, and the diguanylate cyclase *dgcZ*; potentially linking the Cpx pathway to a vast network of genes and phenotypes regulated by the secondary messenger molecule cyclic-di-GMP (20,51,83,84).

One of the more recent discoveries is the role of the Cpx response in colonization and pathogenesis (20,82). A collection of studies have linked the response to the regulation of virulence factors such as the LEE pathogenicity island and Type III secretion system, Type VI secretion system 2 in APEC, UPEC's P-pilus and α-hemolysin HlyA, and enterotoxigenic *E. coli's* colonization factor CS3 (73,85–89). Aside from pathogenesis, the Cpx response is also required for attachment and colonization in the gastrointestinal tract. Absence of Cpx drastically reduces the ability to colonize the mouse intestine and adhere to epithelial cells (90–92).

Horizontal Gene Transfer and Bacteriophages

In its early days, the Cpx response was linked to conjugal plasmid transfer when mutations in the pathway were observed to decreased DNA transfer via conjugation (81). Conjugal gene transfer is among the three main methods of horizontal gene transfer: the others being natural transformation and transduction via bacteriophages, commonly known as phages (93,94). Natural transformation is rare, only seen in about 1% of the characterized species but provides an explanation for gene transfer between species as conjugation and transduction as more specific and cross species transfer can be inefficient (93). Conjugation is the most studied form of horizontal gene transfer but in recent years there has been a resurgence in phage research (94). Phages are small virus-like particles that infect bacterial cells and exist in two types of replication cycles: lytic and temperate. Lytic phages are purely parasitic and behave similarly to eukaryotic viruses where they attach, inject their genetic material, and hijack host cell machinery to replicate resulting in cell lysis and release of mature phage particles (95). Unlike lytic phages, temperate phages can enter a phase called lysogeny, where the phage DNA temporarily integrates into the bacterial genome as an active prophage allowing vertical transmission to bacterial daughter cells (95). When cell viability decreases due to stress, such as by DNA damaging antibiotics, prophages can transition back into the lytic life cycle (95). The most well studied mechanism of prophage induction is Lambda phage's exploitation of the SOS pathway

through its C1 repressor, an inhibitor of lytic cycle genes (96). This repressor is recognized by the SOS protein RecA and undergoes autoproteolysis upon interaction with the ssDNA-RecA complex during DNA damage, de-repressing lytic cycle genes (96). Occasionally, prophages will lose their ability to excise and are permanently integrated into the bacterial chromosome as cryptic prophages.

Prophages and Lysogenic Conversion

Prophages have been known to carry genes that can provide a benefit to the bacterial cell. This phenomenon is called lysogenic conversion and can drastically increase the fitness of the host. For example, Shiga toxin, Cholera toxin, and Diphtheria toxin are all of phage origin and give the bacterium an advantage over the immune system (97). Aside from toxins, prophages can also contribute to stress resistance. Various studies have shown that the presence of prophages results in increased resistance to oxidative stress, high osmolarity, acid and alkaline stress, and serum. (98–103). Prophages have also been found to modulate the expression of host genes: the lambdoid transcription factor Cro has been linked to influencing virulence in enterohemorrhagic *E. coli*, a prophage in *Listeria monocytogenes* acts as a regulatory switch by disrupting the Com system required for macrophage escape, and the promoter of the TorCAD system that allows *E. coli* to use trimethylamine oxide (TMAO) as an alternate electron donor is replaced when the HK022 phage integrates into the genome (104–106). Phages can also modulate the bacterial envelope. Multiple phages have been linked to LPS modification, allowing evasion from the host immune system (98). For example, recently identified Kapi1 was found to modify the LPS O-antigen chain in the mouse commensal *E. coli* MP1 and *Pseudomonas aeruginosa* phage D3 converts its O-antigen from O5 to O16 (107–109). One of the most basic benefits of a lysogenized phage is protection from other invading phages, known as super infection immunity (98). This can either occur by expression of regulators or envelope proteins that repress invading phage genes or prevent phage entry (98). Bacteriophages HK97 and P22 encode predicted transmembrane proteins that provide protection against invading phages (99,110,111). Several *E. coli* phages encode an outer membrane lipoprotein that aids in superinfection immunity to FhuA-dependent phages by interacting with several OMPs (112,113).

The bacterial host's interaction with prophage is bi-directional and regulation can go both ways. For example, the lysis cassette of *E. coli's* cryptic prophage DLP12 is regulated by RpoE through direct binding upstream of the cassette and interaction with a phage anti-terminator that promotes complete transcription of the lysis genes (114). Additionally, host sRNAs have been found to regulate prophage encoded virulence factors and toxins (115,116). The Cpx and Rcs regulated sRNA RprA was found to possibly bind two prophage loci and PhoPQ regulated sRNA PinT blocks synthesis of horizontally acquired effectors SopE and SopE2 in *S. enterica* (117).

Objectives

With the intimate link between prophages, stress resistance, and cross-talk with host regulatory pathways, several questions arise; do the prophages in *E. coli* encode any genes with ties to the Cpx response, if so, are those genes involved in stress resistance, and does the Cpx response play a role in direct regulation of horizontally acquired genes? Thus, the theme of this study was to provide answers to these questions by exploring the relationship between prophages and the Cpx stress response in EPEC strain E2348/69. Our initial objectives were to classify the activity of prophages in E2348/69, determine if they encoded genes involved in stress adaptation, and explore if these genes had ties to the Cpx response. We identified three genes; *yfgG, yfgH,* and *yfgI* near a prophage boundary but determined they were native to *E. coli* and not prophage encoded. Characterization of these genes was still pursued based on their potential links to stress resistance and regulation by the Cpx response, which led to 4 additional sub-objectives for this study: 1) Determine if the Cpx response influenced the regulation of *yfgG*, *yfgH*, and *yfgI*. 2) Explore the role of YfgG, YfgH, and YfgI in stress resistance to known Cpx inducing conditions. 3) Explore the effect of overexpression and the absence of YfgG, YfgH, and YfgI on the Cpx pathway. 4) Determine the subcellular localization of YfgG, YfgH, and YfgI.

Chapter 2: Methods

Strains and Growth Conditions

All strains and plasmids used in this study are listed in Table 5. Primers used in this study are listed in Table 6. Unless specificized, all strains were grown in Lennox Broth (10g/L tryptone, 5g/L NaCl, 5g/L yeast extract) at 37℃ with 225rpm shaking or on agar at 37℃. Strains carrying a *cpxA24* mutation were grown at 30℃ in the presence of 3μg/mL amikacin (Amk) (118). When necessary, 30μg/mL or 50μg/mL kanamycin (Kan), 100μg/mL ampicillin (Amp), or 25μg/mL chloramphenicol (Cam) was added to the culture. Protein overexpression from pCA24N and pTrc99A based vectors was induced with 0.1mM isopropyl-β-D- thiogalactopyranoside (IPTG, Invitrogen).

Strain and Plasmid Construction

Gene knockout mutants were created using P1 transduction to transfer a mutant allele from the Keio collection into MC4100 (119,120). Overnight cultures of strains carrying the desired mutation were subcultured 1:50 in LB supplemented with 5mM $CaCl_2$, 0.2% glucose, and grown for 30 minutes at 37℃ with aeration. 100μL of P1vir lysate was added and the culture was grown until lysed. 100μL of chloroform was added, cells were vortexed and pelleted. The lysate was collected and filtered through a 0.45μm syringe filter and stored at 4℃. 200μL of overnight culture of recipient bacteria were pelleted and resuspended in 100μL of 10mM $MgSO_4$ and 5mM $CaCl_2$. 10-20μL of P1 lysate was added to the culture and incubated statically at 30℃ for 30 minutes. 1mL of LB supplemented with 10mM sodium citrate was added and cells were incubated statically at 37℃ for 30 minutes. The cells were pelleted and resuspended in 100μL 1M sodium citrate and the entire volume was plated on LB agar supplemented with 30μg/mL kanamycin and grown at 37℃ overnight. Colonies were patched onto LB agar supplemented with 50μg/mL kanamycin. Insertion of the kanamycin cassette was verified by PCR using flaking primers of the gene of interest or an upstream primer and the K1 primer that binds the resistance cassette (Table 6).

Clean deletion mutations were created using the pFLP2 FRT recombinase plasmid (121). The pFLP2 plasmid was transformed into the desired strain by either chemical transformation or

electroporation and recovered at 30℃ for at least 1 hour with aeration. Transformants were plated on LB agar supplemented with 100μg/mL ampicillin and grown overnight at 37℃. Colonies were patched onto LB agar and LB agar supplemented with 50μg/mL kanamycin to select for colonies that have lost the kanamycin cassette. Several kanamycin sensitive colonies were used to inoculate 2mL LB and were grown at 30℃ for 6 hours with aeration. Cultures were serially diluted and 100μL of 10^{-3} and 10^{-4} dilutions were plated onto LB agar without NaCl, supplemented with 5% sucrose and grown at room temperature for 2-3 days until colonies were visible. Sucrose resistance colonies were patches onto LB agar and LB agar with 100μg/mL ampicillin to select for colonies that has lost the pFLP2 plasmid. Removal of the kanamycin cassette was confirmed by PCR using flanking primers of the gene of interest on ampicillin sensitive colonies (Table 6).

Construction of Luminescent Reporters

Luminescent reporters *yfgG-lux* and *yfgH-lux* containing the promoter region for *yfgG* and *yfgH* were created using the pJW15 luminescent reporter plasmid as previously described (122). Primers containing XhoI/KpnI or EcoRI/BamHI cut sites were designed to amplify around +500 to -50 relative to the start codon. Promoter regions were amplified by PCR using Phusion High-fidelity DNA polymerase (ThermoFisher) following the manufactures protocol. Amplified promoter regions and pJW15 plasmid were double digested using XhoI/KpnI or EcoRI/BamHI Fast digest restriction endonucleases (Invitrogen) following the manufacturers protocol. Short fragments of DNA were removed using the QIAquick PCR cleanup kit (Qiagen) following the manufactures protocol. Promoter regions were ligated into pJW15 using T4 DNA ligase (Invitrogen) following the manufactures protocol and a modified incubation time and temperature of 16hrs at 16℃. Ligated plasmid was transformed into One Shot TOP10 chemically competent *E. coli* (Invitrogen) and plated on LB agar supplemented with 30μg/mL kanamycin. Inserts in transformants were verified by PCR using pNLP10 primers and DNA sequencing (Table 6).

Construction of His Tagged Protein Overexpression Vectors

YfgG, YfgH, and YfgI coding regions were amplified from the MC4100 chromosome using PCR using Phusion High-fidelity DNA polymerase (ThermoFisher) and primers containing

SacI/XbaI cut sites and a C-terminus 6xHis tag. pTrc99A plasmid DNA and amplified PCR product were double digested with SacI and XbaI Fast Digest restriction endonucleases (Invitrogen) following the manufacturers protocol and small DNA fragments were removed with the QIAquick PCR cleanup kit (Qiagen). Digested PCR product was ligated into digested pTrc99A plasmid using T4 DNA ligase (Invitrogen) following the manufactures protocol and a modified incubation time of 16 hours at 16℃. Ligated plasmid was transformed into One Shot TOP10 chemically competent *E. coli* (Invitrogen) and plated on LB agar supplemented with 100μg/mL ampicillin. Inserts were verified by PCR using pTrc99A_F and pTrc99A_R primers and DNA sequencing (Table S2). Protein overexpression in MC4100 was verified by SDS-PAGE and western blotting with α-His antibodies (Figure S 1).

Repair of pCA-yfgI by Q5 mutagenesis

Sequences of the ASKA overexpression vectors for YfgG, YfgH, and YfgI were verified by PCR, DNA sequencing, and multiple sequence alignment to the MC4100 (Accession: HG738867.1), BW25113 (Accession: CP009273), and E2348/69 (Accession: FM180568.1) genomes using Clustal Omega (123). A V145M mutation in pCA-yfgI was repaired using a Q5 Site-directed mutagenesis kit (New England Biolabs) and pCA_yfgI_Q5_F2 and pCA_yfgI_Q5_R2 primers following the manufactures protocol (Table S2). The correct reversion to MC4100 wild-type amino acid was verified by PCR and DNA sequencing using pCA24N_F and pCA24N_R primers (Table 6).

RNA-Seq Filtering and Prophage identification

The RNA-seq dataset was filtered for prophage associated genes to identify any link between prophages and the Cpx response. Prophage regions in E2348/69 (Accession no. NC_011601.1) were identified using the web server PHASTER (124,125). Prophage regions were cross referenced with the RNA-seq dataset to identify potential prophage genes whose regulation was influenced by the Cpx response. Prophage regions were annotated using BLAST. Prophage associated genes were further filtered by fold change, predicted function, or associated phenotypes, and read numbers. Genes with stress associated phenotypes and larger fold change or gene read numbers were given higher priority. Further protein function prediction was conducted using the PANDA server (126).

Comparison between E2348/69 prophage 10 and *Salmonella* phages SPC32H and Epsilon15 were done by BLAST and visualized using Kablammo (127,128). Prophages were annotated using the PHASTER output, BLAST, and NCBI's reference genomes (SPC32H Accession No: KC911856; Epsilon15 Accession No: NC_004775).

Minimum Inhibitory Concentration Assay

Overnight cultures of EPEC strain E2348/69 and K-12 strain MC4100 were subcultured 1:50 into 200 μL LB in a 96-well plate containing a 2-fold dilution series from 50μg/mL to 0.195μg/mL of mitomycin C dissolved in $MgSO_4$ (both E2348/69 and MC4100), or 512μg/mL to 0μg/mL of vancomycin (MC4100 only). Plates were incubated at 37℃ for 24hrs with aeration. MIC's were recorded both visually and by measuring optical density (OD_{600}) with a PerkinElmer Wallac Victor3 Plate Reader.

Prophage Spotting Assays

Overnight cultures of EPEC strain E2348/69 grown in LB supplemented with 1μg/mL mitomycin C at 37℃ for 24hrs with aeration were pelleted by centrifugation. The lysate was collected, and filtered with a 0.45μm syringe filter. A soft agar overlay was prepared by mixing 100μL *E. coli* culture with 3mL soft LB agar (0.7% agar) and poured onto LB agar and let solidify. A serial dilution of lysate was spotted on (10μL) onto the agar plate containing a overlays of various *E. coli* strain and incubated at 37℃ overnight. Strains used in a host range analysis were E2348/69, J96, DS17, MC4100, MG1655, BW25113, W3110, DH5α, GT115, SY327, and MP1.

Growth Assays

To explore phage activity in EPEC, overnight cultures of E2348/69 and MC4100 were subculture 1:20 into a 24 well plate with 1mL LB supplemented with increasing concentrations of mitomycin C (0μg/mL, 0.1μg/mL, 0.5μg/mL, 1.0μg/mL, 2.0μg/mL, 5μg/mL) and grown at 37℃. Every 2 hours, cultures were agitated with 225rpm shaking for 30 seconds before OD_{600} measurements were taken in PerkinElmer Wallac Victor^2 1420 plate reader.

To determine resistance to alkaline, osmotic, and cobalt stress, overnight cultures of MC4100 wild-type, *yfgG, yfgH, yfgI*, and *cpxR* single and double knockout mutants were subcultured 1:100 into 200µL LB media supplemented with: 100mM Tris-HCl at pH 7 or pH 8, 300mM NaCl, 0.75mM $CoCl_2$, or 0.75mM $CoCl_2$ with 100mM Tris-HCl at pH 7 in a 96-well plate. Cultures were grown for 16 hours with 237rpms of continuous shaking at 37℃ in an Epoch 2 Microplate Spectrophotometer (BioTek). Optical density was measured every 30min. For knockout complementation, media was supplemented with 100µg/mL ampicillin in addition to the above conditions for plasmid maintenance.

To determine the growth while under nutrient limitation, overnight cultures of Keio collection strains BW25113, *yfgG::kan*, *yfgH::kan, cpxR::kan* and MC4100 wild-type, *ΔyfgG*, *ΔyfgH,* and *ΔcpxR* strains were washed twice in PBS, and standardized to OD_{600} = 1.0. Cultures were diluted 1:10 into 180µL fresh M9 minimal media supplemented with either 0.4% glucose (Sigma), 0.4% malic acid (Sigma), or 0.4% succinic acid (Sigma) in a 96-well plate and grown for 48 hours with 237rpms of continuous shaking at 37℃ in an Epoch 2 Microplate Spectrophotometer (BioTek). Optical density was measured every hour.

To determine growth in the absence of siderophore transport, overnight cultures of wildtype MC4100, *yfgG, yfgH, yfgI*, and *exbB* single and double knockout mutants were standardized to an OD_{600} = 1.0 and serially diluted in PBS. 10µL of the 10^{-1}-10^{-7} dilutions were spotted onto M9 minimal media supplemented with 0.4% glucose (Sigma) and plates were incubated at 37℃ for 48 hours then for an additional 72 hours at room temperature. A growth control was spotted onto LB agar in a similar manner and grown overnight.

Luminescence Assays

Activity of the *yfgG* and *cpxP* promoters were assay by the pJW15 luminescent reporter as previously described (122). Strains carrying the reporter were grown overnight and subcultured 1:50 into LB supplemented with 30µg/mL kanamycin and grown to mid-log (OD_{600} = 0.4-0.6) at 37℃ with aeration. Strains carrying a *cpxA24* mutation were grown at 30℃. 200µL of culture

was transferred to a black 96-well plate and the optical density and luminescence in counts per second (CPS) were measured using a PerkinElmer Wallac Victor3 Plate Reader. CPS was standardized to OD_{600}.

To assay the activity of the promoters under various protein overexpression, strains carrying ASKA library overexpression vectors were grown in LB supplemented with 100μg/mL ampicillin and 0.1mM IPTG was added at mid-log to induce expression (129). Cultures were grown for an addition hour before optical density and luminescence was measured as described above.

Activity of pJW15-*yfgG* (*yfgG-lux*) in minimal media was determined by spin-down assay. 500μL of mid-log cultures were pelleted by centrifugation and resuspended in 500μL of LB or minimal media supplemented with 0.4% glucose, 0.4% malic acid, or 0.4% succinic acid. Activity of pJW25 (*cpxP-lux*) under cobalt and alkaline stress in *yfgG, yfgH*, and *yfgI* knockouts was determined by spin-down assay. 500μL of mid-log culture were pelleted and resuspended in 500μL LB supplemented with 100mM Tris HCl pH 7 or pH 8, or LB supplemented with 0.75mM $CoCl_2$.

For both *yfgG-lux* and *cpxP-lux,* 200μL of culture was transferred to a black 96-well plate and grown for 1 hour at 37℃ with aeration. Luminescence was measured as described above.

Gene Expression Assay

RNA-extraction and cDNA synthesis

Expression of *yfgG, yfgH,* and *yfgI* were assayed by RNA collection from MC4100, MP13, and E2348/69 wild-type and *cpxR* knockout strains carrying an IPTG inducible NlpE overexpression vector or vector control. Strains were subcultured 1:50 into LB supplemented with 100μg/mL ampicillin and grown to mid-log (OD_{600} = 0.4-0.6) at 37°C with aeration. NlpE expression was induced with 0.1mM IPTG for 1 hour before 1mL was collected for RNA isolation.

Gene expression experiments in the absence of YfgH were done by collecting RNA from wild-type MC4100*, yfgH, nlpE, cpxP,* and *cpxA* single and double knockout strains subcultured 1:50

into LB and grown to mid-log (OD_{600} = 0.4-0.6) at 37℃ with aeration. 1mL of culture was collected for RNA isolation.

The effect of alkaline pH on Cpx activation in the absence of YfgH was done by RNA collection from wild-type MC4100 and *ΔyfgH* knockout strains subcultured 1:50 into LB and grown to mid-log (OD_{600} = 0.4-0.6) at 37℃ with aeration. 1.5mL of culture was pelleted and resuspended in LB adjusted to pH 7 or pH 8.5 with HCl or NaOH and grown for an hour at 37℃ with aeration. 1mL of culture was collected for RNA isolation.

In all gene expression experiments, RNA was extracted and isolated using the Lucigen MasterPure Complete DNA and RNA purification kit (Epicentre Biotechnologies) following the manufactures protocol and stored at -80℃. RNA quality was verified by RIN using a 2100 Bioanalyzer System (Agilent Technologies). RNA concentrations were standardized to 500ng/μL and 1 μg of RNA of each sample was converted to DNA using Superscript 2 (Invitrogen). Briefly, 1μg of RNA was mixed with 4.95μL 200ng/μL random hexamer primers (ThermoFisher), 1μL 10mM deoxynucleoside triphosphates (dNTPs; Invitrogen), and nuclease free water to 10μL. The mixture was incubated in a thermocycler for 10min at 70℃ followed by 10min at 25℃ and chilled to 4℃. The RNA primer mixture was combined with 4μL 5x First Strand Buffer, 2μL 100mM DTT, 2.5μL Superscript 2 (200U/μL; Invitrogen) and 1.5μL nuclease free water for a final volume of 20μL. The following conditions were used to run the reaction: 10min at 25℃, 60min at 37℃, 60min at 42℃, 10min at 70℃, chilled to 4℃ and stored at -20℃.

Quantitative PCR

qPCR primers were designed using Primer Express 3.0 (ThermoFisher) and are listed in Table 6. 5μL of 2X QPCR Mastermix (*Dynamite*) was combined with 2.5μL 3.2μM stock solution of each primer and 500ng/mL cDNA template for a total of 10μL in a 96-well microtiter plate. The 2X QPCR Mastermix (*Dynamite*) used in this study is a proprietary mix developed and distributed by the Molecular Biology Service Unit (MBSU), in the department of Biological Science at the University of Alberta, Edmonton, Alberta, Canada. It contains Tris (pH 8.3), KCl,

$MgCl_2$, Glycerol, Tween 20, DMSO, dNTPs, ROX as a normalizing dye, SYBR Green (Molecular Probes) as the detection dye, and an antibody inhibited Taq polymerase. The plate was incubated at 95°C for 15 seconds, 60°C for 1 minute and repeated 40 times in a 7500 Fast Real-Time PCR system (Applied Biosystems). The Delta delta Ct ($2^{-\Delta\Delta Ct}$) method was used to determine relative gene expression using the number of cycles taken to reach a fluorescence threshold (cycle threshold [C_T]). Expression of *gyrA* was used as an endogenous control in all experiments to standardize the C_T*'s* between each sample and condition. Statistical significance was determined using a Student's T-Test or pairwise T-Test on log_2 transformed data to account for non-parametric data.

Membrane Fractionation by Sucrose Density Gradients

Cellular localization of YfgG, YfgH, and YfgI were determined by sucrose density gradients and western blotting using a modified protocol (130). Overnight cultures of strains carrying 6xHis YfgG, YfgH, or YfgI overexpression vectors were subcultured 1:100 in 250mL LB supplemented with 100μg/mL ampicillin and grown to an OD_{600} = 1.0 at 37℃ with 180 rpm shaking. Cells were pelleted by centrifugation using successive passes in 50mL conical tubes. The pellet was washed in 50mL 10mM Tris-HCl pH 7.5 and re-pelleted. Cells were resuspended in 10mL Tris-Sucrose (TS) buffer (0.75M sucrose in 10mM Tris-HCl pH 7.5) with 50μg/mL lysozyme and 2mM PMSF. 20mL of 1.65 mM EDTA was slowly added and cell suspensions were incubated for at least 10min on ice. Cell membranes were lysed using 2 passes through a French Pressure Cell Press at 20,000 PSI. 1mL of each sample was collected as a whole cell lysate. Cell debris was removed by centrifugation for 20min at 4℃. Total membranes were pelleted by ultracentrifugation at 38,000 rpm for 45min at 4℃ using a 50.2 Ti rotor (Beckman). Membranes were resuspended in 1mL TES buffer (1 vol TS buffer to 2 vol 1.65mM EDTA) by agitation with sterile plastic loops and pipetting up and down. Membranes were again pelleted by ultracentrifugation as described above and resuspended in 500μL of 25% (w/w) sucrose in 5mM EDTA. 100μL of sample was collected at this point as a total membrane fraction.

Membranes were separated using a six-step sucrose gradient (35-65% sucrose (w/w) in 5mM EDTA) and ultracentrifugation. 4mL of each solution was carefully layered in ultracentrifuge tubes and 400μL of membrane was added on top. Membranes were ultracentrifuged for 18hrs at

40,000rpm at 4℃ using a 50.2 Ti rotor (Beckman). 2mL fractions for a total of 12-13 fractions were pipetted off the top of each gradient and stored in 2mL cryotubes at -80℃. 30µL of each fraction was prepared with 30µL Laemmli buffer and visualized with SDS-PAGE and western blotting. Inner and outer membrane fractions were identified by CpxA and BamA localization respectively.

Periplasmic Protein Isolation

Periplasmic proteins were isolated as previously described (131). To assay His tagged YfgI protein levels in the periplasm, overnight cultures of MC4100 carrying either pTrc99A or pTrc-YfgI were subcultured 1:50 into LB supplemented with 100µg/mL ampicillin and grown to an OD_{600} ~1.0 at 37℃ with 225rpm shaking. To determine total periplasmic protein content in a *ΔyfgH* knockout; overnight cultures of wild-type, single and double knockouts of *yfgH, cpxR, and cpxP* were subcultured 1:50 into LB and grown to an OD_{600} = ~1.0 at 37℃ with shaking.

In all experiments, a whole cell sample was collected before periplasmic protein isolation. 1mL of culture was collected and pelleted by centrifugation and resuspended in 250µL 0.2M Tris-HCl pH 8.0. 250µL 0.2M Tris-HCl 1M sucrose pH 8.0, 2.5µL 0.1M EDTA, 7.5µL lysozyme (4mg/mL stock concentration), and 500µL dH_2O were added and cells were incubated on ice for 2 minutes. 20µL 1M $MgCl_2$ was added, and cells were incubated on ice for 30 minutes. Spheroplasts were pelleted at 14,000 rpm in a benchtop centrifuge for 10 minutes. Both the supernatants and cell pellets were collected as the periplasmic fraction and cell debris fraction respectively. Periplasmic proteins were concentrated using trichloroacetic acid (TCA) precipitation as described (132). 100µL 100% TCA was added to 1mL supernatant, vortexed and incubated overnight at 4℃. Proteins were pelleted by centrifugation at 10,000 x g at 4℃ for 15 minutes and washed once in ice-cold acetone and centrifugated again. Proteins were resuspended in 75-100µL dH_20.

Protein samples used in YfgI localization were immediately prepped in Laemmli buffer and visualized as described under Western Blotting. RNAP and CpxA localization were used as cytoplasmic, and membrane bound protein controls respectively. MalE localization was used as

periplasmic protein control. Periplasmic protein samples used in Δ*yfgH* experiments were quantified and standardized by BCA assay (Thermofisher) before being prepped with Laemmli buffer. Samples were run on 10% SDS-PAGE and stained using Colloidal Coomassie Blue as previously described (133). Gels were imaged using a ChemiDoc MP imager (Bio-Rad). Wild-type and Δ*yfgH* periplasmic lanes were cut out and sent for mass spectrometry protein analysis at the Alberta Proteomics and Mass Spectrometry Facility at the University of Alberta. Results were filtered by GO annotation for cellular localization and cross referenced with the EcoCyc database (134). Relative fold change was determined by change in PSM (Peptide Spectral Matches).

Western Blotting

To verify the expression of YfgG, YfgH, and YfgI from constructed overexpression vectors and to determine if leaky expression from the IPTG inducible promoter was detectible, overnight cultures were subcultured 1:50 into LB supplemented with 100μg/mL ampicillin and grown at 37℃ with aeration. Either 0.1mM IPTG or water was added to the cultures at mid-log and grown for an additional 30 minutes. 200μL of culture was standardized to OD_{600} = 2.0, pelleted by centrifugation, and washed 2-3 times in phosphate buffered saline (PBS) (8g/L NaCl, 0.2g/L KCl, 1.44g/L Na_2HPO_4, 0.24g/L KH_2PO_4). Pellets were resuspended in equal volumes of dH_2O and 2X Laemmli buffer (Sigma). Samples were boiled for 5 minutes and stored at -20℃.

Samples used to identify YfgG, YfgH, or YfgI were routinely run on 15% SDS-PAGE gels. Samples used to identify CpxP were run on 12% SDS-PAGE gels. Samples used to identify BamA, CpxA, or MalE were run on 10% SDS-Page gels. Samples were electrophoresed at 100V. Proteins were transferred to nitrocellulose membranes with a Trans Blot Semi-dry transfer system (Bio-Rad) at 15V for 25-30min using Towbin transfer buffer (192mM glycine, 25mM Tris, 20% methanol). Membranes were rinsed once and let soak for 5 minutes in Tris-buffered saline (TBS) (2.4g/L Tris-HCl, 0.56g/L Tris base, 8.8g/L NaCl). Membranes were incubated for 1 hour with agitation in 5% skim milk prepared in TBST (TBS with 0.1% Tween-20). Membranes were then rinsed twice in TBST and incubated with a primary antibody in 2% bovine serum albumin in TBST overnight at 4℃ with agitation. For verification of YfgG, YfgH, and YfgI expression, membranes were co-incubated with 1:5000 dilutions of α-His rabbit and α-

RNAP mouse polyclonal antibodies. For identifying proteins during membrane fractionations membranes were incubated with either 1:25000 α-BamA, 1:10000 (YfgG sample) or 1:5000 (YfgH and YfgI samples) α-MBP-CpxA, or 1:25000 α-His antibodies (Invitrogen). To identify proteins during periplasmic preparations, membranes were incubated with 1:5000 α-MBP-CpxA, 1:25000 α-His antibodies, 1:5000 α-MBP-CpxP, or 1:5000 α-RNAP antibodies (BioLegend). After incubation, membranes were briefly rinsed and washed 4 times in TBST for 5 minutes each with agitation. Membranes were then incubated with a 1:12500 dilution of LI-COR IRDye fluorescent secondary antibodies (Fisher Scientific) in 5% skim milk in TBST for 1 hour with agitation in an aluminium wrapped 50mL conical tube to avoid fluorophore bleaching. After incubation, membranes were rinsed once and washed 4 times for 5 minutes in TBST. Blots were imaged using a ChemiDoc imager (Bio-Rad).

Chapter 3: Results

Enteropathogenic *E. coli* contains active prophages

Analysis of EPEC strain E2348/69's genome (Accession number: NC_011601.1) using PHASTER revealed 9 unique prophages, 4 of which are predicted to be lambdoid (Figure 1; 124,125). The 9 unique prophages generally aligned to predicted prophages annotated by NCBI, although not all NCBI annotated prophages in the genome were identified by PHASTER and the prophage boundaries differed slightly, due in part to the algorithm used to identify prophage regions and the vast variability among bacteriophages.

The most studied method of prophage induction into the lytic cycle is through DNA damage and hijacking the signal off the host's SOS response (96). To test whether EPEC contained any active prophages, cultures were grown in the presence of subinhibitory concentrations of the DNA damaging agent mitomycin C (MMC). Increasing concentrations of MMC resulted in a drop in optical density as the cultures transitioned to stationary phase, suggesting that prophages were entering the lytic cycle and lysing the cells, leading to a drop in cell density (Figure 2). This drop in optical density was not observed in MC4100, as K-12 *E. coli* is only known to contain cryptic prophages (Figure 2A-B; 135). To further explore if the drop in cell density seen during MMC exposure was due to phage lysis, supernatant collected from overnight cultures grown in the presence of MMC were spotted in a dilution series onto a soft agar overlay to assay for plaques. Clearing was observed and individual plaques appeared at higher lysate dilutions indicating prophage were likely present in the supernatant, although further analysis is needed to definitely determine if phage particles were released. (Figure 2C). A host range analysis of several ExPEC, J96 and DS17 and commensal *E. coli* strains showed that the potential phages present in the E2348/69 lysate could only infect BW25113 and DH5α out of the 11 strains screened (Figure 2D). Not surprisingly, E2348/69 showed no signs of phage infection as prophages are known to provide infection immunity from invading phages (98). Interestingly, not all the K-12 strains screened were infected, indicating the various genetic differences may play a role in bacteriophage resistance.

Survey of the Cpx transcriptome in EPEC E2348/69 for prophage associated genes

A preliminary RNA-seq dataset outlining the CpxR transcriptome during Cpx activation by NlpE overexpression and in a *cpxR* knockout was filtered for differential expression of genes associated with prophage regions identified with PHASTER (124,125). Preference was geared towards genes predicted to be involved in stress resistance and/or associated with the bacterial envelope. Two membrane associated genes, *yfgG* a small predicted inner membrane protein, and *yfgH* a putative lipoprotein, were identified and have been linked with cobalt resistance and outer membrane integrity, respectively in previous studies (Table 1; Table 2; 136,137). A third gene, *yfgI*, immediately downstream of *yfgH* has been linked to nalidixic acid resistance but its predicted localization is cytosolic and membrane bound (Table 1; Table 2;138). Both *yfgG* and *yfgH* were upregulated between 2-2.5 fold during NlpE overexpression while *yfgI* only was found upregulated when NlpE overexpression was compared to NlpE overexpression in a *cpxR::kan* knockout (Table 1).

According to the PHASTER prediction, all three genes resided well inside the predicted prophage region 10 (PP10). Several other methods of prophage prediction showed contrasting results. ProphET's prediction for the boundary of PP10 was immediately upstream of *yfgG* while NCBI's prediction excluded *yfgG, yfgH,* and *yfgI* from the prophage entirely (Figure 3;139). One of PP10's closest relatives is the *Salmonella* Epsilon15 phage: 91% identity with 46% coverage, which is known to integrate into *guaA*, the gene immediately downstream of *yfgI* (140). In addition, *yfgG, yfgH*, and *yfgI* are all encoded on the *E. coli* K-12 genomes where PP10 is absent (Figure 3). Together this provides strong evidence that these three genes are not prophage encoded. Regardless, study of these genes was pursued due to their potential link to the Cpx response, predicted membrane association, and stress resistance phenotypes.

Expression of *yfgG* and *yfgH* are affected by the Cpx response

The promoter regions for *yfgG* and *yfgH* were analyzed for putative CpxR binding sites (Consensuses Binding Sequence: GTAAAN$_5$GTAAA) using PRODORIC's online webserver Virtual Footprint (141). Two putative CpxR binding sites were identified upstream of *yfgG*, suggesting direct regulation by CpxR (Figure 4). To assay whether *yfgG, yfgH*, and *yfgI* expression were influenced by the Cpx response and validate the RNA-seq data, RNA was

collected from E2348/69 wild-type or *cpxR* knockout strains carrying either an IPTG-inducible NlpE overexpression plasmid or empty vector that were grown to mid-log (OD_{600} = ~0.5) and induced for 30 minutes with 0.1mM IPTG. Changes in gene expression of *yfgG*, *yfgH*, *yfgI,* and the endogenous control *gyrA* were assayed by quantitative PCR (qPCR). Overexpression of NlpE resulted in a ~7.5-fold increase of *yfgG* and a 3-fold increase of *yfgH* transcripts while expression of *yfgI* did not significantly differ from the vector control (Figure 5A). In the absence of CpxR, transcript levels did not significantly differ from wild-type levels (Figure 5A).

Prophages have been known to alter gene expression, for example prophage encoded TorI negatively regulates the *torCAD* operon therefore altering the utilization of trimethylamine N-oxide (TMAO) as an electron acceptor (142). To investigate whether the presence of prophages, or other mobile genetic elements in EPEC affected the regulation of *yfgG, yfgH*, and *yfgI*, RNA was collected during NlpE overexpression and from a *cpxR* knockout in two non-pathogenic *E. coli* strains lacking PP10, MC4100, a typical K-12 laboratory strain and MP1, a recent mouse isolate found to host a prophage capable of LPS O-antigen modification (107). In both strains, NlpE overexpression induced upregulation of both *yfgG* and *yfgH,* while no change in transcript levels was seen for *yfgI,* following a similar trend seen in EPEC (Figure 5A-C). When CpxR was absent, transcript levels of all three genes were 2-fold lower than the control strain in both non-pathogens (Figure 5B,C). These data suggest that the presence of PP10 may not affect expression of *yfgG, yfgH,* and *yfgI* during activation of the Cpx response but when CpxR is absent, PP10 may affect expression levels. It must be noted, though, that major strain differences between MC4100 and E2348/69 exist, and further investigation is needed to determine whether or not PP10 affects expression of these genes. Since there was little difference between EPEC and MC4100 in terms of *yfgG* expression during Cpx activation, the effect of the Cpx response on *yfgG* regulation was further explored in MC4100 for ease of genetic manipulation.

To explore various signals that could activate expression of *yfgG* and *yfgH*, the upstream promoter regions of each were cloned into the pJW15 vector to create luminescent reporters *yfgG-lux* and *yfgH-lux* (122). An increase in *yfgG-lux* luminescence was observed during NlpE overexpression, similarly to the qPCR expression data, and in a constitutively active *cpxA24* mutant, providing further support for the conclusion that *yfgG* expression is induced when the

Cpx response is activated (Figure 5D,E;80). No change in luminescence from *yfgH-lux* was detected during NlpE OX and raw CPS (photon counts per second) values were very low, close to background levels of LB (Figure 5F; Figure S 2).

Sensitivity to various stressors is increased in the absence of YfgG, YfgH, and YfgI

The Cpx response has been shown to be required for resistance to alkaline and osmotic stress (66,143,144). To explore whether YfgG, YfgH, and YfgI contribute to alkaline pH and osmotic resistance, single and double knockouts of *yfgG, yfgH,* and *yfgI* along with *cpxR* were grown under neutral, alkaline (pH 8), and high salt (0.3M NaCl) conditions. When grown in LB at neutral pH, only an *yfgG cpxR* double knockout showed any signs of growth deficiencies (Figure 6). Complementation of YfgG restored growth to levels comparable with the wild-type strain. Growth under osmotic stress did not intensify the growth reduction in the *yfgG cpxR* double knockout and no other growth phenotypes were seen in the *yfgH* or *yfgI* knockouts except for an extended lag phase in the *yfgI cpxR* double knockout although overall growth was similar to wild-type levels (Figure 7). When grown under alkaline stress, both *yfgG cpxR* and *yfgI cpxR* double knockouts showed severe growth deficiencies compared to wild-type and the single knockouts (Figure 8). Complementation of *yfgG* and *yfgI* did not restore growth suggesting overexpression of these proteins may be toxic while under alkaline stress.

Overexpression of YfgG has been shown to aid in resistance to cobalt and nickel induced stress while the Cpx response has been linked to copper resistance (136). Cobalt, nickel, and copper are known to induce reactive oxygen species and impact the Fe/S cores present in certain enzymes (145,146). The link between cobalt, *yfgG* and the Cpx response was explored by growing knockouts of *yfgG* and *cpxR* in media supplemented with cobalt chloride (0.75mM). The single knockout of *yfgG* showed reduced growth in the presence of cobalt and complementation restored growth above and beyond wild-type levels, similarly to the published phenotype (Figure 9A;136). Interestingly, the *cpxR* knockout had better growth than the wild-type strain suggesting the Cpx response may slow growth in some manner under cobalt stress. Absence of *yfgG* appears to have a stronger effect than absence of *cpxR* as the double knockout more resembles growth of the single *yfgG* knockout. Complementation of YfgG in both the single and double knockout improved growth better than wild-type and single CpxR knockout levels indicating that YfgG is

solely responsible for this phenotype. Cobalt chloride was found to be mildly acidic in solution and when added to growth media the measured pH was ~6.0. To account for growth phenotypes attributed to acid stress, single and double knockouts of *yfgG* and *cpxR* were grown in buffered media (pH 7) supplemented with cobalt chloride (0.75mM). The growth phenotypes seen previously were reduced towards wild-type levels (Figure 9B). The single knockout of *yfgG* still had reduced growth but not as severe when compared to growth under cobalt stress in unbuffered media. The effect of cobalt stress was also explored in the absence of the lipoprotein YfgH in a similar manner. The single knockout of *yfgH* showed poor growth in the presence of cobalt (Figure 9C). Complementation of YfgH did not restore growth to wild-type levels, but rather appears to be toxic under these conditions. Deletion of *cpxR* in the absence of *yfgH* showed improved growth compared to the single *yfgH* knockout. Here, complementation had no effect on growth. These results implicate both *yfgG* and *yfgH* in cobalt resistance and could implicate the involvement of *yfgG* in acid resistance.

To assess the activity of the Cpx response under cobalt stress with or without YfgG, a luminescent reporter containing the promoter region of *cpxP*, *cpxP-lux*, was transformed into wild-type and *yfgG* knockout strains and subjected to cobalt stress (0.75mM) for 1 hour at mid-log. The presence of cobalt resulted in mild repression of the Cpx response as luminescent activity was lower in cells exposed to cobalt compared to those grown in LB (Figure 9D). The Cpx response was further repressed in the absence of YfgG. The absence of YfgG does decrease Cpx activity somewhat but the effect is more pronounced when cells are exposed to cobalt. These results suggests that the mechanism by which YfgG aids in cobalt resistance is independent of CpxR mediated regulation since absence of *yfgG* both decreases Cpx activity while also leaving cells more susceptible to cobalt. Additionally, repression of the Cpx response appears to result in increased resistance as a *cpxR* knockout of the response grows better than wild-type cells.

Recently the Cpx response has been implicated in the turnover of electron transport chain (ETC) proteins (147). Additionally, growth on malate and succinate, carbon sources that require the ETC, is severely delayed in the absence of CpxR (147). It was also shown that growth on succinate was found to induce the Cpx pathway (147). With succinate being an activating cue for

the Cpx pathway, and in the context of previous results (Figure 5), we hypothesized that expression of the *yfgG* promoter would be increased under this condition. Activity of the *yfgG* promoter was monitored via the *yfgG-lux* reporter plasmid during short term exposure to minimal media supplemented with glucose (0.4%), malic acid (0.4%), or succinic acid (0.4%). An increase in promoter activity was seen under all three carbon sources, not just malate and succinate, suggesting expression of *yfgG* under these conditions might primarily be influenced by nutrient limitation rather than the Cpx response, as glucose was not found to activate the response (Figure 10A;147). To explore the link between YfgG, YfgH, and growth on minimal media, wild-type BW25113 and MC4100 were grown alongside single knockouts of *cpxR, yfgG,* and *yfgH* in M9 minimal media supplemented with glucose (0.4%), malic acid (0.4%), or succinic acid (0.4%). Knockouts of either *yfgG* or *yfgH* in both BW25113 and MC4100 grew similar to wild-type with glucose as a carbon source (Figure 11; Figure 12). In BW25113, knockouts of *yfgG* and *yfgH* grew slightly slower than the wild-type strain when grown on malate and succinate. In MC4100, the knockout of *yfgG* showed no difference compared to the wild-type strain, while on the other hand, absence of YfgH showed severe growth defects similar to the *cpxR* knockout when grown solely on malate and succinate. This suggests that YfgH plays a key role in growth while the ETC is in high demand. YfgG may be involved in some manner during growth under nutrient limiting conditions, but its presence appears to be non-essential for growth.

A study screening the *E. coli* genome for genetic interactions identified an aggravating interaction between *yfgH* and *exbB*, a member to the TonB-ExbB-ExbD complex involved in siderophore and Vitamin B12 transport (137). Double knockouts of *yfgG, yfgH,* and *yfgI* were created with *exbB* and plated on minimal media supplemented with glucose (0.4%) to determine if low nutrient conditions impaired growth in the absence of iron sequestration. Slowed growth was seen in the single knockouts of *yfgH*, *yfgI*, and *exbB* after 48 hours on minimal media while only subtle difference in growth were seen on rich media (Figure 10B). After 120 hours these strains reached higher levels of growth but still slight lower than wild-type. Severe growth defects were seen in the *yfgH exbB* and *yfgI exbB* double knockouts on minima media. Even after 120 hours, growth was still low, especially in the *yfgI exbB* knockout. This could implicate both

YfgH and YfgI in the transport of siderophores and other TonB-dependent substrates under nutrient limited conditions.

Together these results show that the absence of YfgG, YfgH, and YfgI impairs the bacterium's ability to adjust to growth under stress. In certain cases, their growth phenotypes appear when Cpx mediated envelope homeostasis is missing, indicating absence of these proteins may induce some form of envelope stress that presents a serious challenge to the cells ability to mitigate disruptions in the envelope from external sources.

Absence of YfgH induces the Cpx and Rcs Envelope Stress Responses

The Cpx response is thought to use auxiliary proteins NlpE and CpxP to monitor the envelope (20). CpxP binds to misfolded proteins, allowing degradation by the endoprotease DegP, alleviating repression of CpxA and resulting in activation of the Cpx response. NlpE is thought to act as a sentinel for lipoprotein trafficking and reducing stress in the periplasm, yet its mechanism of activation is unknown (67,70). To explore whether the absence of YfgG, YfgH, and YfgI influence Cpx activity, *cpxP* expression was monitored via the *cpxP-lux* reporter in knockout mutants of the three genes. Removal of both YfgH and YfgI resulted in significantly higher Cpx activity than wild type while the Cpx response was slightly repressed in the absence of YfgG (Figure 13A).

To investigate the mechanism by which the Cpx response is being activated by the *yfgH* knockout, the mRNA levels of Cpx regulon members *cpxP, spy*, and *degP* were assayed by qPCR in *cpxA, cpxP,* and *nlpE* knockouts. Transcript levels of *cpxP, spy,* and *degP* were highly upregulated in the absence of YfgH (Figure 13B-D). Transcript levels of *spy* in particular were increased over 200-fold when YfgH was removed. Activation of the Cpx response in the absence of YfgH is also dependent on both CpxA and CpxP, but not NlpE as *cpxP, spy*, and *degP* expression levels decreased when either *cpxA* or *cpxP* were absent alongside YfgH, but a slight increase in *cpxP* levels was seen in the *nlpE yfgH* double knockout. Activation of the Cpx response is also not due to increased levels of *nlpE* expression as there was no drop in *cpxP* transcript in the *nlpE* knockout and *nlpE* transcripts were actually less in the absence of YfgH (Figure 13E). An increase in *cpxP* transcript levels were seen upon removal of *yfgH* in the *cpxA*

knockout background. A high amount of variability was found in the expression levels of *cpxP* between *yfgH cpxA* double knockout replicates (Figure S 3). Expression levels of *cpxP* in two *yfgH cpxA* knockout samples grouped at a level similar to expression levels in the single *cpxA* mutant while two samples grouped higher, although less than *yfgH* mutant levels. This variability could indicate the presence of suppressor mutations in the *yfgH cpxA* double knockout as the cell attempts to counteract the effect of the Δ*yfgH* mutation. To provide additional evidence of Cpx pathway activation, levels of Cpx regulon members *dsbA, ldtC* (formerly *ycfS*), *yebE,* and *yncJ* were assayed in the *yfgH* knockout (59,83). Levels of *ppiD,* which encodes a periplasmic chaperone, were also measured as evidence suggests CpxR doesn't influence regulation (59). A strong increase in all but *ppiD* transcript levels were seen in the absence of YfgH further implying an overall general activation of the Cpx response under this condition (Figure 13F).

To determine if absence of YfgH and YfgI limited the ability of the Cpx response to respond to stimulus, mid-log cells were exposed to alkaline stress. The removal of YfgI had no effect on the cells ability to sense alkaline stress as the difference in activation levels between the wild-type and knockout strains were similar to neutral conditions (Figure 14A-C). On the other hand, the YfgH knockout had no difference in Cpx activity between neutral and alkaline pH while wild-type cells saw an increase in Cpx activity upon growth in alkaline conditions. This suggests the Cpx response is either unable to sense alkaline stress or is fully saturated by the base-level activation in the absence of YfgH. To parse out whether this is due to a unique role in sensing pH or an overall signalling phenotype, the ability of the Cpx response to sense NlpE overexpression was assayed in a YfgH knockout. No significant difference in *cpxP-lux* activity was found between the vector control and NlpE overexpression in the *yfgH* knockout (Figure 14D). This suggests that YfgH may play a role in signaling during envelope stress and in its absence, the Cpx response is unable to respond accordingly.

To corroborate the lack of induction of a luminescent *cpxP-lux* reporter by alkaline pH in an *yfgH* null mutant, transcript levels were measured in cells exposed to alkaline stress in a similar manner. Similar to the luminescent data, there is no change in *cpxP* mRNA levels between neutral and alkaline conditions in an *yfgH* null background (Figure 15A). Interestingly, a drop in *cpxP* expression is seen when exposed to alkaline stress in a wild-type background, suggesting

repression of the pathway, a response opposite to what is known. This experiment was repeated a total of 4 times; 3 times with buffered LB pH 8 and once with unbuffered LB pH 8.5, isolating fresh RNA each time, and resulted in a similar drop in *cpxP* transcript levels upon exposure to alkaline stress. The transcription of additional genes known to either be regulated by CpxR or have differential regulation under alkaline stress were also measured to investigate this unexpected result. Transcript levels for *spy, ltdC,* and *dsbA* were all found to be upregulated by alkaline pH, indicating the Cpx response may be active, while expression of *degP* and *yebE* did not differ significantly (Figure 15B). Non-Cpx regulated *pyrC* was previously found to be upregulated under alkaline stress, but here we show *pyrC* is downregulated under alkaline stress, although experimental conditions were different between studies (148). Next, protein levels in an *yfgH* knockout were assayed to determine if a) deletion of *yfgH* resulted in an increase of CpxP protein levels in the cell and b) the drop in *cpxP* transcript under alkaline stress correlated with protein levels. In wild type cells, CpxP levels are low and difficult to detect by immunoblotting but an increase in CpxP protein levels were observed in an *yfgH* knockout and CpxP was not detectable in either a *cpxR yfgH* or *cpxP yfgH* double knockout (Figure 15C). When exposed to alkaline stress, increased CpxP levels are also observed compared to wild type levels although less than when *yfgH* is absent (Figure 15D). While hard to quantify, CpxP levels in an *yfgH* knockout background appear unchanged during alkaline stress. This suggests that, in a wildtype strain background, there could be some post-transcriptional regulation that is causing a divergence in mRNA and CpxP proteins levels during alkaline stress, possibly due to the sRNA CpxQ, known to influence CpxP protein levels (149,150). This divergence is absent in the *yfgH* mutant, and both *cpxP* transcript and CpxP protein levels remain the same regardless of pH. These finding support the conclusion that YfgH influences Cpx signaling.

A previous study showed that a knockout of *yfgH* resulted in morphological changes to the bacterial envelope when exposed to vancomycin, a phenotype that was exacerbated when lipoprotein synthesis was reduced suggesting a role for YfgH in outer membrane integrity (137). To determine if absence of YfgH increased *E. coli's* sensitivity to vancomycin, an *yfgH* knockout's growth in vancomycin was assayed by minimum inhibitory concentration (MIC) and growth in a subinhibitory concentration. No difference in MIC or growth was seen between wild type MC4100 and the *yfgH* knockout (Figure 16A-B). To further investigate whether YfgH is

involved in outer membrane integrity, the supernatant of wild-type, and knockouts of *yfgH, cpxR,* and *cpxP* were assayed for changes in protein content. Very low levels of protein were detected in the supernatants of wild-type, and *cpxR* and *yfgH* knockouts (Figure 16C). Absence of CpxP resulted in an increase the presence of proteins in the supernatant that was further increased when YfgH was removed. YfgH may play a role in outer membrane integrity even though growth is unaffected by vancomycin stress, as removal of the lipoprotein increases the amount of proteins detected in the periplasm, possibly due to a weakened outer membrane. Alternatively, this may be a product of an accumulation of misfolded proteins in the periplasm induced by removal of YfgH, thereby resulting in the activation of the Cpx response (Figure 13).

5 major stress responses dominate control over envelope homeostasis in *E. coli*, the Cpx response being one of them. The data present here suggests a disruption of envelope homeostasis by the deletion of *yfgH.* To determine if the absence of YfgH induced widespread envelope stress, transcript levels of *pspA, baeR, rcsA,* and *rpoE* were measured by qPCR. Levels of *rcsA* were seen to be upregulated about 7-fold suggesting activation of the Rcs ESR and possible disruption of the outer membrane and/or cell wall (Figure 17). Both *pspA* and *baeR* saw a mild increase in transcript levels although less than 2-fold.

Together, these data show that in the absence of YfgH, the Cpx response is highly active, resulting in upregulation of many members of the CpxR regulon and an increase in CpxP protein levels in the cell. This phenotype appears to be primarily depended on CpxP, but not NlpE for Cpx activation. In the absence of YfgH, the system can no longer respond to alkaline pH or NlpE overexpression. A knockout of YfgH also resulted in an increase of *rcsA* transcript, suggesting possible disruption to the outer membrane or cell wall, although no change in growth was observed under vancomycin stress. Additionally, no proteins were detected in *yfgH* supernatant unless CpxP was absent.

Deletion of *yfgH* leads to an accumulation of protein folding and degradation factors in the periplasm

The majority of Cpx inducing cues are thought to induce misfolding of proteins in the periplasm. To investigate whether a knockout of YfgH leads to an accumulation of proteins in the

periplasm, whole cell, and periplasmic contents of an *yfgH* knockout, along with *cpxP,* and *cpxR* knockouts were collected, standardized, and visualized by SDS-PAGE and colloidal Coomassie (133). In whole cell fractions, protein levels were relatively consistent between samples with only a slight increase in protein density in the *yfgH* knockout around 30 and 40kDa (Figure 18 left). Several differences were seen in the periplasmic protein fractions with differences in band intensities just above and below 50kDa but no overall accumulation was seen (Figure 18 right). Just below 50kDa, an increase in protein density is seen in the *yfgH* knockout and deceases in both the *yfgH cpxR* and *yfgH cpxP* double knockouts, indicating this protein may be influenced by the Cpx response. The wild-type and Δ*yfgH* lanes were excised and sent for whole lane protein analysis. Absence of *yfgH* resulted in an increase of membrane and periplasmic associated proteins in the periplasmic isolate, most notably DegP, PpiA, and DsbA (Table 3). Additionally, several proteins were only detected in either the wild-type or Δ*yfgH* samples. CpxP, Spy, Slt, LtdC, and RcsF were only detected in the absence of yfgH (Table 4). While OmpF and RseB were only detected in the wild-type sample. This indicates that absence of YfgH alters envelope homeostasis and may be inducing misfolding of proteins in the periplasm.

YfgG, YfgH, and YfgI are envelope localized proteins that activate the Cpx response

The primary role of the Cpx response is to maintain envelope homeostasis by monitoring protein folding in the periplasm. Overexpression of NlpE and OmpA is known to induce activation of the response and overexpression in general can induce misfolding for proteins due to clogged translocation and biosynthesis pathways (74,75). If YfgG, YfgH, and YfgI are indeed envelope localized proteins, overexpression could similarly induce the Cpx response. Additionally, the data presented so far indicates removal of these proteins induces envelope stress. To determine the impact of overexpression, cells containing YfgG, YfgH, and YfgI overexpression vectors were monitored for *cpxP* promoter activity via a luminescent reporter. Overexpression of both YfgH and YfgI were found to activate the Cpx response (Figure 19). A slight increase in Cpx activity was seen when YfgG was overexpressed although not significant.

Previous bioinformatics predictions place YfgG and YfgH in the cellular envelope while YfgI appears to be cytosolic (134). YfgG contains an α-helix transmembrane domain indicating inner membrane localization, while YfgH contains a lipoprotein signal sequence and is likely

transported to the outer membrane via the Lol pathway (134). Further analysis of the YfgI peptide sequence revealed a Sec/SPI signal sequence predicted by the SignalP 6.0 webserver, suggesting YfgI may be translocated into the periplasm (151). Little experimental evidence exists for the localization of these proteins. To address this issue and to determine their location in the cell, bacterial membranes from strains overexpressing His-tagged YfgG, YfgH, and YfgI were separated by a sucrose density gradient. Localization of CpxA and BamA were used to identify inner and outer membrane fractions respectively. YfgG and YfgH were identified in the inner and outer membrane fractions with peak band intensity corresponding with CpxA and BamA respectively (Figure 20A). YfgI had band intensity peaks corresponding with both CpxA and BamA suggesting it might be a periplasmic protein with possible anchoring or interaction with both the inner and outer membrane. To determine if YfgI is a soluble periplasmic protein, periplasmic content was collected from cells overexpressing YfgI. RNAP, CpxA and MalE were used as cytosolic, membrane and periplasmic controls respectively. A small amount of YfgI was detected in the periplasmic fraction (PP) while much of the protein was found in the cell debris fraction (CD) (Figure 20B). A small amount of RNAP was also detected in the PP fraction, indicating possible cytoplasmic contamination, indicating that any YfgI detected in the PP fraction may also be contamination. Thus, YfgI may not be a soluble periplasmic protein, but instead an inner membrane protein that extends into the periplasm, possibly interacting with the outer membrane.

Chapter 4: Discussion

The integrity of the bacterial envelope is essential to a cells survival as it is home to a multitude of fundamental pathways and the cell must do whatever it takes to mitigate disturbances in its envelope. Mobile genetic elements such as prophages have also been found to modify the composition of the envelope, providing protection against invading bacteriophages and the host immune system (98). The Cpx response, along with several other ESRs, maintains inner membrane and periplasmic homeostasis by controlling the expression of various envelope localized proteins, chaperones, proteases, and biosynthetic pathways (20,47). Transcriptome studies of the Cpx regulon show the response has a wide reaching effect on overall gene expression in the cell and many of the regulon members have yet to be fully understood (83). In this study, we found that EPEC strain E2348/69 possibly contains active prophages capable of infecting K-12 strains BW25113 and DH5α. We also identified *yfgG* as a new member of the CpxR regulon that encodes an envelope localized protein with links to stress resistance. The regulation of additional genes, *yfgH* and *yfgI,* may also be influenced by the Cpx envelope stress response and aid in resistance to various stress and nutrient limiting conditions. Our results suggest outer membrane localized YfgH is important for the maintenance of envelope homeostasis as the Cpx response is highly active in its absence, strongly increasing the level of periplasmic proteases and chaperones, possibly leaving the response blind to other causes of envelope stress.

Analysis of prophages in E2348/69

Horizontal gene transfer is one of the main drivers of bacterial evolution. Studies have identified many non-native genes to aid in response to stress (101,103,105,152). To date, no link between prophages and the Cpx response has been found. In this work we attempted to screen the Cpx regulon in EPEC strain E2348/69 for possible connections to prophage regions.

E2348/69 was found to contain a large collection of prophages. The activity status of each individual prophage was not identified in this study or any other study, but here we present evidence suggesting at least one prophage may be active as plaques were detected on an overlay with BW25113 (Figure 2C). Further analysis will need to be done to determine if the clearing seen was indeed due to the presence of a bacteriophage. Half of the prophages identified in

E2348/69 are predicted to be lambdoid in nature and likely depend on the canonical method of prophage induction through the SOS-response (Figure 1; 96). The basic premise of SOS-based induction is used by other types of phages, although other methods do exist (152). Various studies have highlighted non-SOS based induction such as through heat shock or compounds found in the human diet (152–155). Several prophages identified in E2348/69 are predicted to be similar to P2 and Epsilon15 phages (Figure 1). P2 is generally considered non-inducible and the mechanism of induction for Espilon15 has yet to be identified (140,156). It is plausible that these prophages are active but may not be induced by DNA damage. This also highlights another issue; it is possible that not all the prophages induced are able to infect BW25113. Phage host ranges can be highly specific, and it is rare that a single phage will infect more than one species as binding and infection depend on the phage tail fibers and bacterial cell binding sites (157). Here we show that out of the 11 strains screened, the E2348/69 lysate only produced plaques on 2 strains: BW25113 and DH5α (Figure 2D). Additional work involving the isolation, identification, and extending the host ranges of the possible phages present in the lysate needs to be done in order to draw any conclusions about individual specific phage activity and host range.

The genes of interest that appear to be influenced by Cpx activation were identified near E2348/69's prophage region 10 (PP10). PP10's sequence shares high similarity to several active *S. enterica* prophages, SPC32H and Epsilon15 being the most characterized (Figure S 4). PP10 encodes a repressor and anti-repressor almost identical to SPC32H's lytic cycle regulators, *ant* and *rep* genes (99% and 97% respectively), suggesting PP10's lytic-lysogenic regulation could be similar (158). Epsilon15 is a serotype-converting *Podoviridae*, sharing 91% identity with 46% coverage to PP10, and has been experimentally shown to integrate into *guaA*, the gene downstream of *yfgI* in *E. coli* strains missing prophage 10 (Figure S 4;140). Integration of Epsilon15 does not disrupt the coding sequence of *guaA* due to a small fragment of *guaA* encoded on the prophage that restores the coding sequence of the gene (140). This small fragment of *guaA* is also present in PP10, indicating PP10 may integrate in a similar manner (Figure S 5).

The possibility of PP10 integrating into *guaA* provides additional evidence against *yfgGHI* being prophage encoded, in addition to the presence of them on the K-12 genome. Furthermore, the

exact boundaries of PP10 differ depending on the prediction software used (Figure 3). PHASTER includes upstream genes *ppx, yfgF, yfgGHI,* and downstream genes *guaA* and *guaB*. ProphET includes *yfgGHI* and *guaAB* in the prophage while they are excluded in the NCBI annotation and original genome sequence and analysis for E2348/69 (160). Being the most abundant biological entity, bacteriophages are incredibly diverse and relatively few phages have been fully characterized (161–163). Current prophage prediction software generally use databases designed for bacterial genome annotation or methods that use phage gene density to make their predictions, both of which rely on previously characterized phage genes (124,125,139,164,165). This poses a problem since bacteriophage databases are limited compared to their inherent diversity, making it difficult to accurately predict prophage regions. For example, ProphET and PHAST (and PHASTER) use tRNA, a common insertion point to refine prophage boundaries, but this may not be the best method as seen with Epsilon15's integration into *guaA* (Figure S 5; 124,125,139). The problem at hand is somewhat paradoxical as the programs rely on a limited database, but the database is limited due to inaccuracies in the software. Regardless, developers are aware of the current issues and new and improved software is being developed to work around the limitations. In all cases, determining prophage boundaries cannot be purely based on bioinformatics, experimental evidence is needed to corroborate the predictions. This is true for PP10 as well and in order to determine its boundary, it needs to be determined whether it is cryptic or active and only then can PCR and sequencing based approaches be used to assay the boundary.

The Cpx response regulates inner membrane protein YfgG

The Cpx regulon is extensive and include a variety of characterized and uncharacterized genes (83). We identified three consecutive genes, *yfgGHI* in a preliminary RNA-seq dataset and have shown through qPCR that the expression of *yfgG* and *yfgH* are elevated during Cpx activation across three *E. coli* strains: E2348/69, MC4100, and MP13 (Figure 5). Upregulation of *yfgG* but not *yfgH* was also seen in a previous Cpx transcriptome study in MC4100 and E2348/69 (83). Presence of a CpxR binding site upstream of *yfgG* provides support for possible direct regulation by CpxR and removal of CpxR results in down regulation of *yfgGHI* in MC4100 and MP13 but interestingly not in E2348/69 (Figure S 2). Many factors could be attributing to this difference in expression as major genetic differences exist between the three strains. The most obvious

difference is the presence of the LEE pathogenicity island in E2348/69 but its absence in MP1, MP13's parent strain, and MC4100. As mentioned previously, this segment of DNA encodes the T3SS and unique collection of genes that account for the attaching and effacing phenotype seen in many pathogenic *E. coli* (7). The LEE encodes several regulators and chaperones that aid in coordinating the expression of virulence genes (7). The central LEE regulator Ler is required for A/E lesion formation and its own regulation is complex, regulated by many factors, including CpxR (7,166). A semi-recent study highlighted the extension of Ler beyond the LEE (167). They found that Ler influences the expression of over 50 extra-LEE genes in E2348/69, most notably *pagP,* involved in lipid A synthesis; RcsA, and a small subset of prophage encoded T3SS effector homologs (167). In addition to LEE, E2348/69 contains many mobile genetic elements including 10-13 prophages, and over 100 insertional elements (160). Any one of these could account for the observed difference in expression of *yfgGHI* in the absence of *cpxR* between E2348/69 and MC4100/MP1.

Genetically, E2348/69 and MP1 share a more recent ancestor than MC4100 (168). E2348/69 and MP1 are classified in the B2 phylogroup of *E. coli* while MC4100 is in group A (168). The B2 group is highly diverse and includes several virulent ExPEC strains as well as the commensal Nissle 1917 (169–171). Group A generally consists of non-pathogenic commensals lacking many of the associated virulence traits (169,172). The similarity of the regulation of *yfgGHI* in the absence of CpxR between MP1 and MC4100 provides some support for possible mobile genetic element regulation in E2348/69, such as prophages. In E2348/69 PP10 lies immediately downstream of *yfgGHI* while it is absent in MP1 and MC4100, and encodes several putative regulatory factors, making it a good candidate for further investigation.

The spatial organization of *yfgGHI* is sequential, with ~300bp of space between *yfgG*/*yfgH* and 15bp between *yfgH/yfgI* (Figure 4). The small gap between *yfgH/yfgI* suggests these genes are most likely transcribed in an operon while the larger gap between *yfgG/yfgH* potentially leaves room for a promoter. A luminescent reporter containing promoter region of *yfgG* showed increased activity during Cpx activation, while an *yfgH* reporter showed very low to zero luminescence, barely above background levels (Figure 5F; Figure S 2). There are two possible explanations for this result: 1) *yfgGHI* are transcribed in an operon, 2) activity levels of an *yfgH*

promoter are too low to detect by a luminescent reporter. The decreasing expression trend for *yfgGHI* is consistent with polarity in operonic transcription (173). Operonic polarity is when expression levels decrease for genes further from the promoter region and can be mediated by Rho-dependent terminators and CRP-cAMP (cyclic adenosine monophosphate), among others (173–175). The nature of expression levels and luminescent data suggest *yfgGH* could be transcribed as a single transcriptional unit but this most likely is not the case as a Rho-independent terminator, determined by the terminator prediction webtool ARNold, lies between the two genes (Figure 4;176). Additionally, a σ^{70} promoter has been predicted upstream of *yfgH* (177,178). This brings in the second explanation where transcriptional activity may be too low to detect by reporter assay. Two studies have found that absolute YfgH protein and mRNA levels are very low (179,180). While robust, the light produced from the luminescent based reporter may not be detectable due to low levels of expression, accounting for the lack of luminescence detection from the *yfgH-lux* reporter. The method used in this study to quantify mRNA transcript levels only measured relative expression with reference to a control group, not absolute mRNA levels. Further experiments could investigate the absolute levels of mRNA by qPCR, but such methods are not suitable for high throughput screening. Brighter luciferase reporters, such as NanoLuc, are available that could prove useful for monitoring activity for low promoter expression (181).

Inner membrane proteins YfgG and YfgI aid in resistance to stress

Exposure to alkaline pH induces many responses in the cell, such as increased generation of metabolic acid, change in expression of ion/proton transporters, and alteration of cell envelope structures (182). The Cpx response has been shown to aid in stress mitigation during alkaline stress and is thought to aid in abolishing misfolded proteins induced by alkaline conditions (66). Large shifts in pH can disrupt enzymatic activity, protein folding, and the PMF (33,182,183). Enzymatic activity and protein folding depends highly on the transfer of protons and electrons using the pKa of many amino acid side chains and shifts in pH can cause protonation or deprotonation that changes their enzymatic potential (183). The proton motive force is linked with pH in a complex way and can change dramatically when pH is shifted (33). Generally, bacteria aim to maintain the cytoplasm alkaline relative to external pH but may reverse this under severe stress to retain overall PMF (33). Bacteria counteract changes in pH by regulating

proton pumping machinery, thus counteracting the either high or low concentration of protons inside the cell (33). Under alkaline stress, the cell retains or imports protons to the cytoplasm (33). The periplasm is assumed to be at a pH similar to the external environment and as the pH deviates from neutral, periplasmic proteins increasingly become more susceptible to misfolding and reduced enzymatic activity. In this study we show that resistance to alkaline stress is reduced when YfgG and YfgI are removed in the absence of CpxR (Figure 8). The *yfgG cpxR* double knockout exhibits a decrease in growth under normal conditions, but the effect seen under alkaline stress is much larger. This additive phenotype seems counter intuitive as removal of CpxR resulted in decreased *yfgG* expression, and one would expect a similar phenotype between the single *cpxR* and double *yfgG cpxR* knockouts. This effect may rather be due to possible regulation of *yfgG* beyond the Cpx response, so that in the absence of CpxR expression of *yfgG* is still maintained at some level. Thus, the growth phenotype in alkaline pH of the single *cpxR* knockout is not entirely dependent on *yfgG,* but removal of both genes results in a more impaired phenotype. Interestingly, complementation did not restore growth in either the *yfgG cpxR* or *yfgI cpxR* knockouts while under alkaline stress but did restore growth of the *yfgG cpxR* double knockout under neutral conditions (Figure 6). This could indicate that overexpression of YfgG and YfgI may be toxic under alkaline conditions in the absence of CpxR as the physiological levels of each protein are not known. Additionally, we found overexpression of YfgI induces Cpx activation, which may account for the lack of complementation in the absence of CpxR as the cell may not be able to mitigate a possible aggregation of protein at the inner membrane (Figure 19).

CpxR is required for survival in alkaline conditions above pH ~8.8, but here we show that removal of either YfgG or YfgI at pH 8 exacerbates the effects (66). The roles YfgG and YfgI could play in alkaline resistance are numerous but its phenotype with CpxR is particularly interesting as not much is known about Cpx mediated resistance to alkaline stress and both CpxP and NlpE are not required to induce the response (69). One possible explanation for the Cpx response's function in alkaline resistance is its role in inner membrane protein turnover. The Cpx response has been implicated in the regulation and maintenance of inner membrane complexes, including transporters and metabolic complexes that contribute to the proton motive force (51). Without CpxR, the cell may not be able to maintain control over the pH and electrochemical

gradient across the inner membrane during alkaline stress as there could be an accumulation of non-functional transporters and metabolic proteins. YfgG and YfgI could potentially aid in the stabilization of these complexes or signalling although Cpx induction is seen in these knockouts when exposed to alkaline stress, suggesting it may not be the latter.

Previously, YfgG was identified as playing a role in resistance to cobalt and nickel, and overexpression aided in survival (136). Here we show that while involved in resistance to cobalt, the effect of removing YfgG is more pronounced in unbuffered media where the pH was mildly acidic, leading back to proton transporters (Figure 9A). Monitoring growth of an *yfgG* knockout under purely acidic conditions is necessary to confirm this and should be extended to YfgI due to its alkaline phenotype. Complementation in both unbuffered and buffered media supplemented with cobalt restored growth to wildtype levels indicating that indeed YfgG does play a role in cobalt resistance, but this role could be connected to inner membrane protein stability, as speculated with alkaline pH. A previous study found that AcrZ (formerly YbhT), a small, 49AA, α-helical transmembrane domain protein enhances resistance to certain antibiotics provided by the AcrAB-TolC efflux pump (184). AcrZ was found to directly interact the with a groove in the transmembrane domain of AcrB and is hypothesized to alter substrate specificity by inducing a conformational change in the drug binding pocket of AcrB (185,186). YfgG may serve a similar role as its length, at 63 AA, and structure is similar. We also show that deletion of *cpxR* results in better growth than wild-type cells under cobalt stress in unbuffered media (Figure 9A). The effect due to the absence of CpxR is likely caused by the acid conditions rather than cobalt as cells grew only slightly better in buffered media (Figure 9B). The Cpx response is not known to be involved in acid resistance and is in fact repressed under low pH (148). Under acidic conditions, cells need to actively pump protons out to maintain proper intracellular pH and PMF thus removal of the Cpx response may further alleviate the repression of the ETC complexes responsible for this activity (33).

Aside from the difference in acidity, exposure to cobalt primarily results in the disruption of Fe/S cores which are required for catalytic processes in many enzymes (145,187). The cell tightly controls the intracellular concentration of cobalt, and other divalent ions such as nickel and zinc. *E. coli* contains several metal transporters: importers CorA and ZupT, and exporter RcnAB, that

all are capable of transporting cobalt (188–191). Additionally, the TonB-ExbB-ExbD complex facilitates iron acquisition as well as that of Vitamin B12, also known as cobalamin, which contains a cobalt cofactor (192). When under cobalt stress, cells will upregulate iron uptake, Fe/S biogenesis machinery, and efflux complexes (187,193). YfgG could be aiding in any of these processes although no effect was seen when grown on minimal media in the absence of ExbB indicating it may not be linked to iron availability or transport (Figure 10B). The role of the Cpx response under cobalt stress has not been studied, but it is known to be activated by copper: which was identified to induce oxidative stress and disrupt Fe/S cores similarly to cobalt (51,194). We show that the Cpx response is not induced by cobalt but slightly repressed, and in the absence of YfgG, pathway activity decreases further under cobalt stress, suggesting the activation of the response is not required for adaptation to cobalt stress. This conclusion is further supported by the observation that growth of the *cpxR* knockout was similar or better than wild-type cells (Figure 9D). This may also implicate YfgG in some form of signalling under specific stress as no difference in Cpx activity was observed under alkaline stress in an *yfgG* null (Figure 14A).

Recently, the Cpx response has been implicated in turnover of respiratory complexes and it is required under non-fermentable carbon sources that enter the TCA cycle directly, inducing large demand for the ETC (147). Our data shows slowed growth under nutrient limitation in the absence of YfgI when glucose, a fermentable sugar, is the sole carbon source (Figure 10B). The growth defect seen with YfgI is further intensified when ExbB is removed. This could link YfgI with iron uptake, envelope complex stability, or the PMF. The TonB-ExbB-ExbD transducing system actively transports siderophores using the PMF harnessed by inner membrane complex ExbB-ExbD (192). If YfgI is involved in general envelope complex stability, removal of the protein could compromise critical ETC and nutrient transport complexes which could then be amplified when iron acquisition is impaired. This could be investigated by looking for direct interaction between YfgI and ExbB or other transporters. Under this hypothesis, removal of YfgI could cause a disruption in the periplasm and inner membrane, possibly activating the Cpx response. We show that removal of YfgI does indeed cause mild activation the Cpx response (Figure 13A). This most likely isn't a signalling phenotype as the Cpx response can still react to alkaline stress, although it could be signal specific (Figure 14C). YfgI has previously been

identified to aid in resistance to nalidixic acid by an unknown mechanism (138). Nalidixic acid is a quinolone antibiotic and targets DNA gyrase in the cytoplasm (195). Resistance mechanisms in Gram negatives include target site alteration, altered membrane permeability, and increased efflux (195). We found YfgI to be a secreted inner membrane protein, eliminating a role in target site alteration, and adding support for the possibility of aiding in the stability of envelope complexes (195). Absence of YfgG does induce a small growth phenotype in *E. coli* strain BW25113 but not in MC4100 (Figure 11). BW25113 is a more robust strain that carries a collection of cryptic prophages known to aid in resistance to stress and this could be the cause of this difference (103). Interestingly, expression of *yfgG* is increased under nutrient limitation, regardless of the carbon source (Figure 10A). This upregulation is likely CpxR independent as growth on minimal media containing glucose does not activate the response (147).

The exact role that YfgG and YfgI have in the envelope remains a mystery, but here we provide initial evidence for their role in alkaline and low nutrient conditions. These proteins may act to stabilize envelope complexes such as proton pumps and nutrient transporters, and removal of YfgG or YfgI could causes a change in activity or misfolding, but this is speculation at this point and needs to be verified experimentally (Figure 21). Further studies should focus on the relationship between YfgG, YfgI, the Cpx response and inner membrane transporters to determine if indeed absence of these envelope proteins affects stability and protein levels. Additional screening should be done to determine the conditions under which these genes are expressed. For example, the expression of *yfgG* and *yfgI* could be assayed in the absence of various inner membrane proteins or under alkaline stress. The Psp and Bae ESR's are also key components in the maintenance of envelope homeostasis, specifically the PMF and efflux, respectively, making them good candidates for possible interaction with YfgG and YfgI (47,50).

Absence of YfgH induces stress in the envelope

Lipoproteins have many diverse roles in the envelope and can be localized to the periplasmic side of the inner membrane and outer membrane or surface exposed (196). Lipoproteins are known for their role in LPS biogenesis, membrane integrity, sensing and signalling, and many other functions in the envelope (196). Here we identified YfgH to be an outer membrane lipoprotein that aids in resistance to cobalt stress and nutrient limitation. Further, we show that

removal of *yfgH* leads to a large induction of the Cpx response. Growth under alkaline stress does impact the cell somewhat, as there is a slight delay of growth in an *yfgH* knockout (Figure 8B). While this appears to be due to an activated Cpx response that is "primed" due to YfgH's absence, removal of *cpxR* from the *yfgH* background does not affect growth suggesting this phenotype may not depend on the Cpx response and its regulation of inner membrane proton pumps.

Our data shows that absence of YfgH impairs growth under cobalt stress at neutral pH and removal of CpxR almost restores growth to wild-type levels (Figure 9C). Several interweaved factors could cause this phenotype: activation of the Cpx response could be toxic to cells under cobalt stress, and absence of YfgH could cause misfolding of proteins rendering envelope processes ineffective. The transport of toxic compounds, metals, and nutrients is key for cells to survive stress and the absence of YfgH may be causing mis localization or misfolding of specific envelope complexes, causing an inability to properly efflux toxic molecules. The presence of irregular proteins could also lead to Cpx activation that causes dysregulation of transporters. So far, no cobalt transporters have been identified in the Cpx regulon, but it is known to repress inner membrane proteins, including respiratory complexes, thought to reduce accumulation of aberrant proteins at the inner membrane during envelope stress (51). Removal of CpxR could alleviate this repression, allowing the cell to efflux cobalt effectively. This could partially explain the growth deficiency in MC4100 when grown solely on succinate or malate (Figure 12). The succinate/malate:H^+ symporter, DctA, appeared in a Cpx transcriptome study and was found to be downregulated during Cpx activation (83). DctA uses the PMF to transport succinate, L and D-malate, and several other carbon compounds across the inner membrane, and cells devoid of this protein struggle to grow on these carbon sources (197). We show that removal of YfgH induces strong activation of the Cpx response, which may result in the repression of the DctA transporter by CpxR. This could then lead to slowed growth when malate or succinate are the sole source of carbon as the transport of these carbohydrates across the inner membrane may be impaired. This isn't a one way street but rather a balancing act because the Cpx response is also important for the maintenance of respiratory complexes and its absence results in delayed or no growth, but strong activation also reduces metabolic activity of succinate dehydrogenase (147). Absence of YfgH in BW25113 does hinder growth as its growth is similar to the *cpxR* knockout

on succinate and somewhat on malate. Several strain differences exist between MC4100 and BW25113, for example MC4100 is missing the e14 prophage element, found to aid in acid resistance, and could account for the difference in growth during nutrient limitation (103,198,199).

The Cpx response is thought to use CpxP and NlpE as monitors of the envelope, but few inducing signals actually require the presence of these auxiliary proteins, in fact no known signal is known to require CpxP (51). Here we highlight the first activating signal identified that requires CpxP to sense the stress-induction by the absence of YfgH. This is evident by the fact that the induction of the Cpx response upon *yfgH* deletion is negated in the absence of CpxP (Figure 13B-D). YfgH may not act as a repressor of the system as overexpression also activates the response. Perhaps YfgH over-expression leads to induction of the Cpx response in a similar way to how NlpE induces, through accumulation at the inner membrane (Figure 19;70). The specific interaction between YfgH and CpxP should be explored. All known inducing cues pass through CpxA, including the one demonstrated here (51). Detection of the signal is not dependent on the presence of NlpE either as no change was found when NlpE was removed. Additionally, we show that the pathway does not respond to stimulus, alkaline pH and NlpE overexpression, in the absence of YfgH (Figure 14). Whether this is due to possible YfgH involvement in signalling, saturation of induction, or depletion of CpxP in the periplasm is unknown, although the dependence on CpxP suggests an accumulation of misfolded proteins. Increases in *dsbA, degP,* and *spy* transcript levels also suggest there could be issues with protein folding. We assayed *yfgH* knockouts for their total and periplasmic protein content. We found elevated levels of proteins in the total cell and a specific increase in the periplasm, all of which appears dependent on CpxP and CpxR (Figure 18). This indicates that the absence of YfgH may not cause widespread aggregation of misfolded periplasmic proteins, but rather the increase in protein is in response to the knockout not the cause of it. A few bands in the periplasmic fractions caught our attention, specifically a band around 45-50 kDa increased in intensity in the absence of YfgH but not when both YfgH and CpxP/R were deleted. This band could be numerous proteins but logically DegP makes the most sense as it is a Cpx regulated, periplasmic 49.3 kDa protein that is upregulated in the absence of YfgH. Additionally, mass spectrophotometry protein analysis of the wild-type and Δ*yfgH* periplasmic lanes revealed that

DegP protein levels were found to be increased ~5 fold in the absence of YfgH, providing support that the band is DegP (Table 3). With elevated CpxP and DegP protein levels, CpxP could be degraded by DegP upon entry into the periplasm as it binds misfolded peptides, leaving CpxA in an activated state.

Transcript levels of *spy* and *degP* were slightly increased in the deletion mutants of *cpxA* and *cpxP* when induced by YfgH's absence and could signify induction of other ESR's beyond the Cpx response. The BaeRS and Sigma E ESR's are known to co-regulate these genes with the Cpx response in response to inner and outer membrane stress (53–55). Out of the 5 ESR's, only Cpx and Rcs were substantially activated in the absence of YfgH (Figure 17). Interestingly, no change was detected in *rpoE* transcript levels. CpxR is known to negatively regulate the Sigma E response and activation of the Cpx response represses *rpoE* expression (58,59). Our results show no change in *rpoE* expression levels between the wild type and *yfgH* knockout. Deciphering whether or not the absence of YfgH induces the Sigma E response is difficult as *degP* and *rpoE* are both CpxR and RpoE regulon members (55,58,59,200). Since regulation of *rpoE* by CpxR and RpoE are inverse of each other, there could be opposing effects between the two transcriptions factors when both stress responses are induced. This could produce an overall cancellation effect between CpxR and RpoE on *rpoE* transcription, resulting in no change in gene expression and could explain what we've observed here (Figure 17). Alternatively, the Sigma E response may still be inhibited as the periplasmic inhibitor, RseB, was only detected in the wild-type periplasmic isolates (Table 4). It is not known if RseB is degraded upon sensing a signal, but it is transcribed in an operon with *rpoE* and *rseA* and should be repressed when the Cpx response is activated, possibly explaining the lack of detection in the Δ*yfgH* mutant (47,201). No conclusions can be drawn from these data regarding the activation of the Sigma E pathway in response to the absence of YfgH, and further investigation is necessary to determine if the response is activated. Future studies should focus on the expression of RpoE regulon members that aren't influenced by the Cpx response, such as *surA,* in the absence of YfgH. Removal of *cpxR* would also eliminate any Cpx-dependent regulation.

The connection of the Rcs response in combination with the Cpx pathway in response to YfgH induced stress implies that stress is present at all areas of the envelope and could additionally

provide clues to the lipoprotein's function. The two ESRs possibly have some overlap in their inducing cues: the Rcs response can detect LPS and peptidoglycan defects, while the Cpx response is sensitive to lipoprotein trafficking (19,20). YfgH's role is most likely not involved with lipoprotein trafficking as deletion of NlpE did not result in a difference in Cpx induction. Furthermore, NlpE dependent activation of the Cpx response is independent of CpxP although there could be an unknown mechanism that has not be discovered yet (Figure 13B; 70). There is some evidence that YfgH may be involved with peptidoglycan synthesis. One of the highly upregulated genes in the *yfgH* knockout was *ldtC,* which encodes a transpeptidase that facilitates the linkage of peptidoglycan to the lipoprotein Lpp, thereby anchoring the outer membrane to the cell (202). Additionally, several peptidoglycan associated proteins were only detected in the absence of YfgH, including LdtC, Slt, and AmiA (Table 4). Interestingly, LdtB and LdtE were found at lower levels compared to wild type (Table 3; Table 4). The Cpx response has been linked to modifying peptidoglycan cross-linkages through upregulation of the LdtC and LdtD transpeptidase (203). YfgH may somehow aid in facilitating these reactions. An absence of YfgH may cause improper anchoring of Lpp to the peptidoglycan layer, or Cpx activation could be causing an imbalance in PG-modifying protein levels, both resulting in a weakened cell wall structure and compromising the outer membrane.

Alternately, YfgH may play a more general role in outer membrane stability as previously hypothesized, although this is not inconsistent with the speculated role with peptidoglycan as described above (137). Deletion of the *yfgH* increases susceptibility to low levels of vancomycin resulting in membrane protrusions and this phenotype was amplified with a knockdown of *lptD* leading to cell lysis, possibly linking YfgH to LPS synthesis (137). We show that absence of YfgH does not impact the MIC or growth in vancomycin, suggesting the previously reported morphological phenotype may not impact cell growth (Figure 16A-B). Absence of YfgH does result in an increase of extracellular proteins in the supernatant when combined with a *cpxP* knockout, a phenotype seen previously in the absence of DegP that is intensified with heat shock (Figure 16C; 204). This indicates impaired membrane integrity could lead to a leaky membrane when periplasmic proteolysis is reduced. The same study that linked YfgH to outer membrane integrity also found that absence of YfgH decreases OmpA levels (137). OmpA and other OMPs are vital components for transport, envelope biogenesis, and the integrity of the outer membrane

(205). Whether or not this is the sole reason an *yfgH* knockout exhibits a loss of membrane integrity is unknown, but it is clear that absence of YfgH causes a disturbance in the envelope as the periplasmic proteins content shows a large disparity between wild-type and Δ*yfgH* cells (Table 3; Table 4). Several Cpx regulated protein folding and degradation factors are seen to be increased or solely detected in the absence of YfgH, namely CpxP, Spy, DegP, DsbA, and PpiA, providing evidence of that protein misfolding is taking place in the periplasm. Ag43, an autotransporter involved in self-recognition and cell aggregation, was found to have the highest relative increase of 12-fold between wild-type and the Δ*yfgH* mutant (Table 3;206). The Rcs response has been shown to repress motility and other surface structures such as Ag43, but here we see an increase of protein even though our results suggest the Rcs response may be active (207). Removal of YfgH may be causing either miscommunication or misdirection in the envelope as ESRs try to restore homeostasis.

In our analysis of *cpxP* transcript levels to corroborate the alkaline blind phenotype we noticed an interesting response by the Cpx pathway to alkaline stress. As seen with the luminescent data, *cpxP* transcripts remained similar between neutral and alkaline conditions in the absence of YfgH, but in wild-type cells a reduction of *cpxP* transcript was measured even though protein levels and *cpxP-lux* activity were increased (Figure 14B; Figure 15). This is contradictory to not only the data in this study, but other published work as well (66). Additionally, only *cpxP* transcripts are reduced as *spy, ldtC, and dsbA* transcripts are increased, indicating the Cpx pathway is active, although no change was seen in *degP* levels. No previous experiments have been done to assay the Cpx's response to alkaline stress at the mRNA level; promoter activity has previously only been assayed through reporter assays with luciferase or *lacZ*. A previous microarray has looked at the total cellular response to alkaline stress and found that *cpxP* transcripts were upregulated ~2-fold (148). Identified in the same study, *pyrC* was found to be up regulated ~ 3 fold and was included as a Cpx-independent control in our experiments (148). Our results show downregulation of *pyrC* upon induction of alkaline stress. This could in fact be due to differences in study design. Maurer et al. collected RNA from cells after long term growth under alkaline stress whereas we only exposed our cells to alkaline stress for an hour at mid-log (148). It is possible that the role of the Cpx response in alkaline resistance is growth phase

dependent and the response could participate in various aspects over time, similar to its role in biofilm regulation (51,208).

Regardless of the difference between studies, our data shows a discrepancy between promoter activity, mRNA levels, and protein levels. Since the early days of transcriptomics, studies have highlighted some discrepancies between mRNA and protein levels due to mechanisms of post-translational regulation (209,210). We might be encountering that phenomenon here. CpxP protein levels are known to be modulated by the recently identified sRNA CpxQ, though current data suggests mRNA is unaffected and CpxQ only decreases protein levels (150). At this point in time, little is known about CpxQ and it may be possible that its activity is variable and could stabilize mRNA upon sensing a specific cue. Other sRNAs may also play a role in modulating CpxP's transcription or CpxQ's stability. The discrepancy between promoter activity and mRNA levels could be due to transcriptional attenuation where additional regulatory steps are taken once transcription begins (211). This method of regulation could be excluded by a promoter luciferase fusions as only a few dozen nucleotides downstream from the start codon are cloned into the reporter. Several classes of transcriptional attenuation exist and can rely on a 5' terminator sequence in a leader peptide, interaction with a RNA binding protein, or a riboswitch that can detect changes in physiological conditions (211). Additional considerations to factor in are the time exposed to the condition, translational efficiency of the transcript, and protein half-lives (209,210). In this specific case, perhaps the need for CpxQ is low under alkaline conditions while the need for CpxP protein remains high, thus the stability of mRNA is increased allowing prolonged translation. This could allow the overall transcript levels to decrease while maintaining protein levels and could be explored via RNA stability assay and expression levels of CpxQ.

The role of YfgH has yet to be uncovered but the data presented in this study shows that the outer membrane lipoprotein plays a crucial part in the maintenance of envelope homeostasis (Figure 21). YfgH may be involved in outer membrane integrity, or LPS and peptidoglycan synthesis. Removal of YfgH resulted in the induction of two ESRs and an increase in periplasmic chaperones, proteases, and peptidoglycan modifying proteins. Future studies could explore if YfgH has any further links to peptidoglycan or LPS biosynthesis and whether or not the Cpx

response is involved. Furthermore, the relationship between YfgH and the Rcs response should be explored to determine how the cell is modulating the outer membrane in the absence of YfgH. This could provide key insights into the function of YfgH. Due to their genetic organization, it is tempting to speculate on the relationship between YfgG, YfgH, and YfgI, and if any interaction between the proteins takes place as removal of these proteins share some similar phenotypes.

Figures and Tables

A

GenBank Annotation	PHASTER Annotation	Type
PP1	N/A	Lambdoid
PP2	1	Lambdoid
PP3	2	P2
N/A	3 (sits inside 4)	Lambdoid
PP4	4 (3 sits inside 4)	Lambdoid
PP5	5	Lambdoid
PP6	6	Lambdoid
PP7	7	P2
PP8	8	P22
PP9	9	Mu
PP10	10	Epsilon 15
N/A	11	Epsilon 15
PP11	N/A	P2
PP12	N/A	P4
PP13	N/A	NA

B

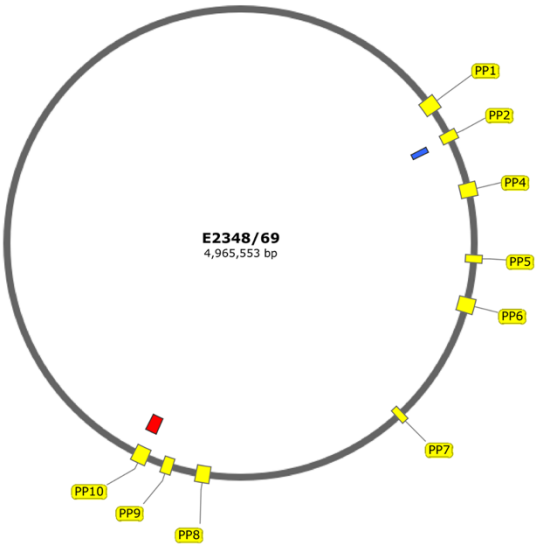
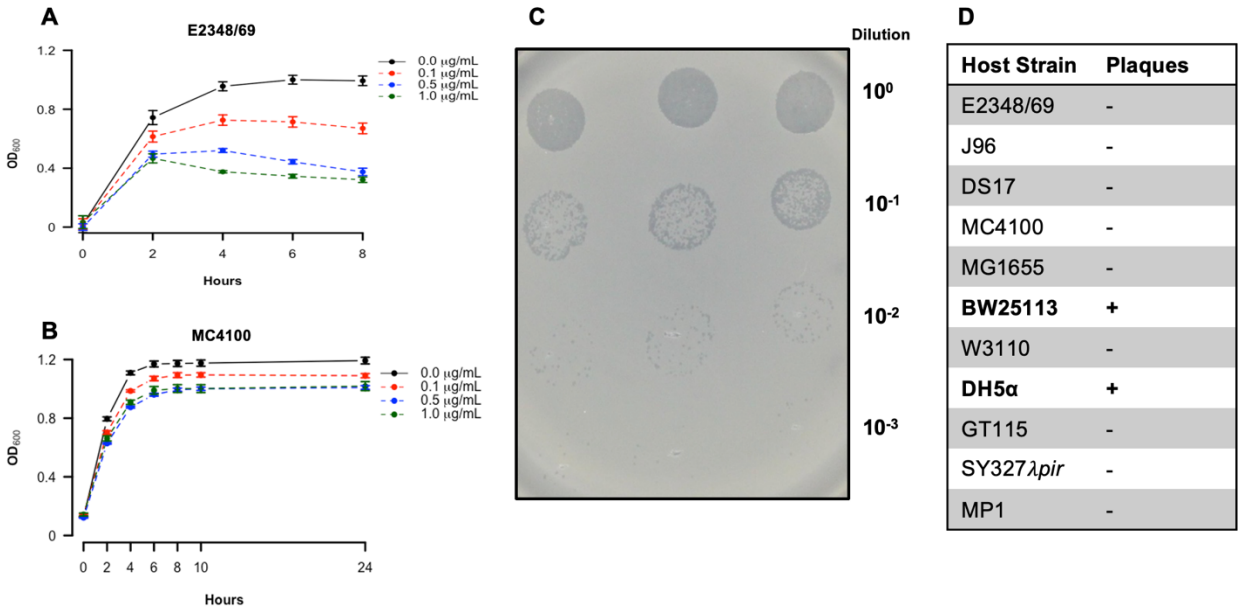


Figure 1: Predicted prophage regions on EPEC strain E2348/69

A) Prophage regions on E2348/69's genome (Accession no: NC_011601.1) were predicted using PHASTER and compared to the NCBI GenBank genome (124,125). Phage type was determined using a combination of the GenBank entry and BLAST with NCBI's viral database (128). **B)** Location of PHASTER predicted prophages on E2348/69's genome. Prophages 3 and 11 are represented in blue and red respectively.



Host Strain	Plaques
E2348/69	-
J96	-
DS17	-
MC4100	-
MG1655	-
BW25113	**+**
W3110	-
DH5α	**+**
GT115	-
SY327*λpir*	-
MP1	-

Figure 2: E2348/69 may contain active prophages capable of cell lysis

Growth of E2348/69 **(A)** and MC4100 **(B)** under DNA damaging conditions. Overnight cultures were subcultured 1:20 into LB supplemented with increasing concentrations of mitomycin C in a 24-well plate and grown at 37°C with shaking. OD_{600} measurements were taken every 2 hrs. Error bars represent SEM (n=3). **C)** Mitomycin C induced lysis results in the release of active phage particles. Supernatant from E2348/69 cultures grown overnight at 37°C in the presence of 1µg/mL mitomycin C was spotted in a dilution series on a soft agar overlay of BW25113 and incubated overnight at 37°C **D)** Host range analysis of E2348/69 lysate. Supernatant from E2348/69 cultures grown overnight in the presence of 1µg/mL mitomycin C was spotted on s soft agar overlay of various *E coli* strains and incubated overnight at 37°C. Bolded strains indicate the presence of plaques.

Table 1: Genes identified in a preliminary RNA-seq dataset to be differentially regulated by the Cpx response

Gene	Fold Change		Up/Dn Reg.	PANDA Prediction[126]/Known Function
	NlpE/ WT	Δ*cpxR* NlpE/ NlpE		
yfgG	2.288	-3.32	Up	membrane bound / resistance to cobalt and nickel stress[136]
yfgH	2.47	-3.49	Up	response to toxic substance, membrane bound / OM integrity[137]
yfgI	n.d	-2.71	Up	response to stress, cytosolic, membrane bound / nalidixic acid resistance[138]

n.d: no data

Table 2: Protein features of YfgG, YfgH, and YfgI

Protein	Length (AA)	Molecular weight (kDa)	pI[1]	Transmembrane domains[2]	Signal Sequence[3]
YfgG	63	7.46	11.89	1	Other
YfgH	172	17.71	8.67	-	Sec/SPII (Lipoprotein)
YfgI	179	20.23	6.23	-	Sec/SPI

1: Isoelectric point determined using the pI/MW tool on the ExPASy server (212).

2: Transmembrane domains determined using DeepTMHMM (213).

3: Signal sequences determined using SignalP 6.0 (151).

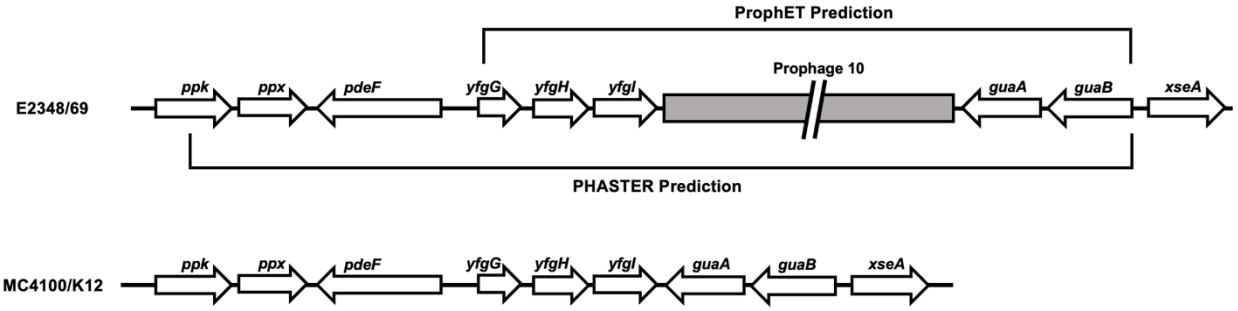


Figure 3: The boundary of prophage 10 differs based on prediction software

Prophage regions were annotated using NCBI's MC4100 (Accession No: HG738867) and E2348/69 (Accession no: NC_011601.1) reference genomes, PHASTER, and ProphET (124,125,214). Grey region represents prophage 10 region on E2348/69 annotated genome. Gene sizes are not to scale.

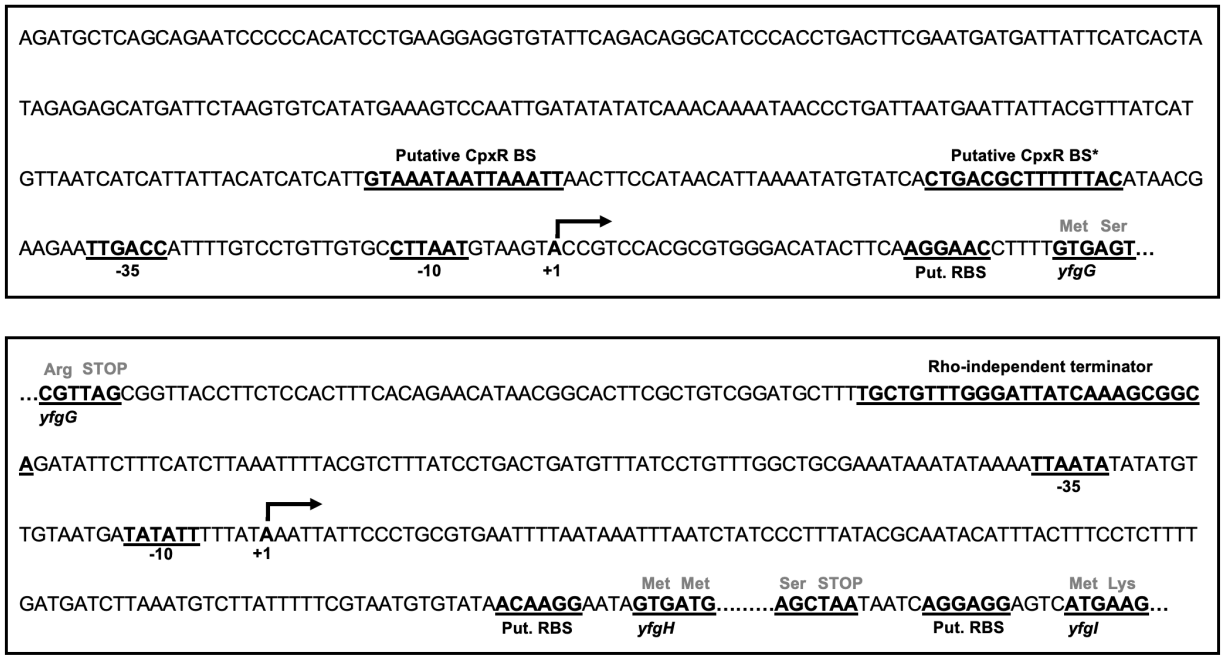


Figure 4: Analysis of *yfgG* and *yfgH* promoter regions

The regions upstream of *yfgG* **(top)** and *yfgH* **(bottom)** were analyzed for CpxR binding sites using the online webserver Virtual Footprint (PRODORIC; 141) and by visual inspection for the CpxR consensus sequence GTAAAN$_5$GTAAA. Transcriptional start sites and σ^{70} boxes (-10 and -35) were mined from the EcoCyc database and subsequent studies (134,215). Terminator sequences were predicted using ARNold (176). Amino acids denoted in grey. * Denotes binding on reverse strand.

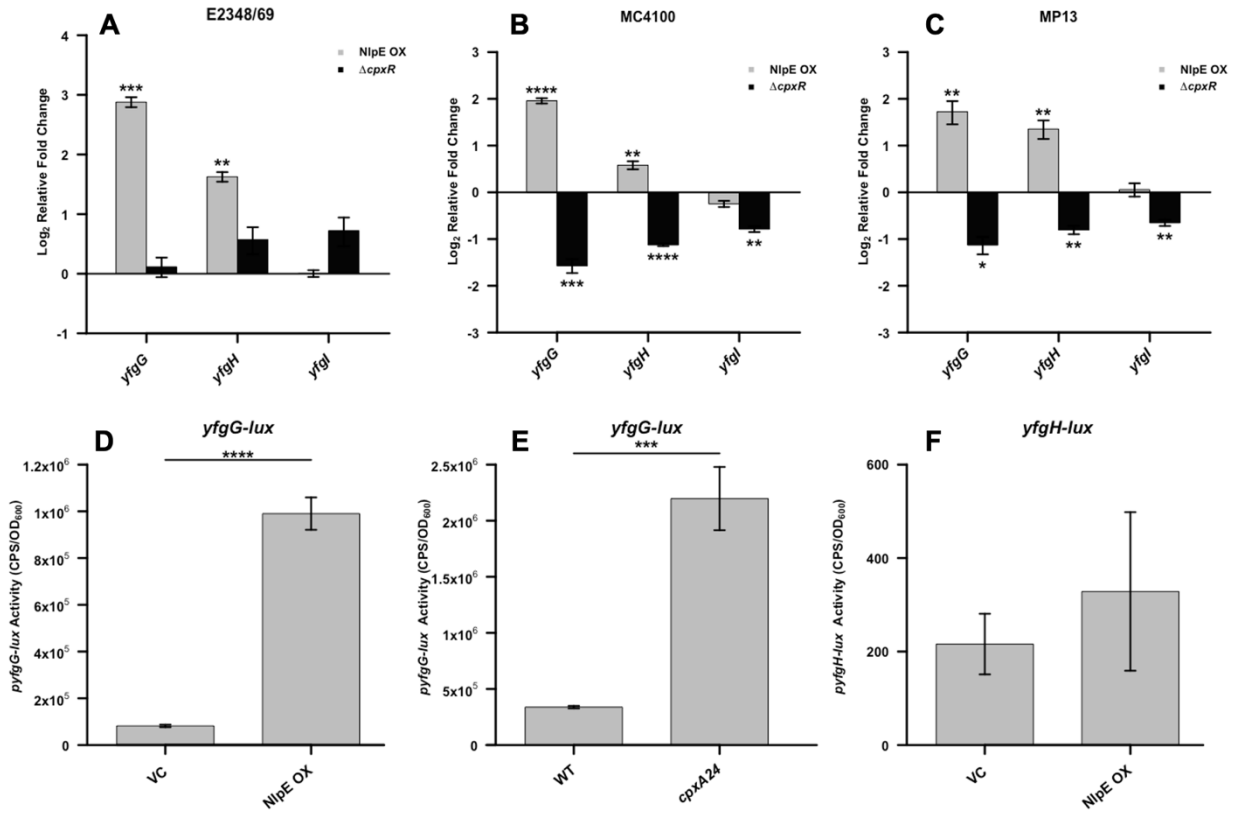


Figure 5: The Cpx response regulates *yfgG* and *yfgH* across multiple *E. coli* strains. **A-C)** RNA was collected from **(A)** E2348/69, **(B)** MC4100, and **(C)** MP13 wildtype and *cpxR* knockout (black bars) strains carrying either an NlpE overexpression (grey bars) or empty vector after 30min induction with 0.1mM IPTG at mid-log. Transcript levels of *yfgG, yfgH,* and *yfgI* were determined by qPCR. Fold change is relative to the respective wild-type stain carrying an empty vector **D)** MC4100 strains transformed with the *yfgG-lux* reporter and either an NlpE overexpression or empty vector were subcultured 1:50 into LB media and grown to mid-log. 0.1mM IPTG was added, and cells were grown for an additional hour measurement. **E)** Wildtype MC4100 or a *cpxA24* mutant carrying the *yfgG-lux* reporter and were grown to mid-log. **F)** MC4100 strains transformed with a luminescent report containing the *yfgH-lux* reporter and either an NlpE overexpression or empty vector were subcultured 1:50 into LB media and grown to mid-log. 0.1mM IPTG was added, and cells were grown for an additional hour before measurement. **D-F)** In all cases, cells were grown at 37°C with 225 rpms shaking. 200µL culture was transferred to a black 96-well plate where CPS and OD_{600} were measured. CPS was standardized to optical density **A-F)** Error bars represent SEM (n=3). Asterisks (*) indicate

statistical significance (*: $p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test).

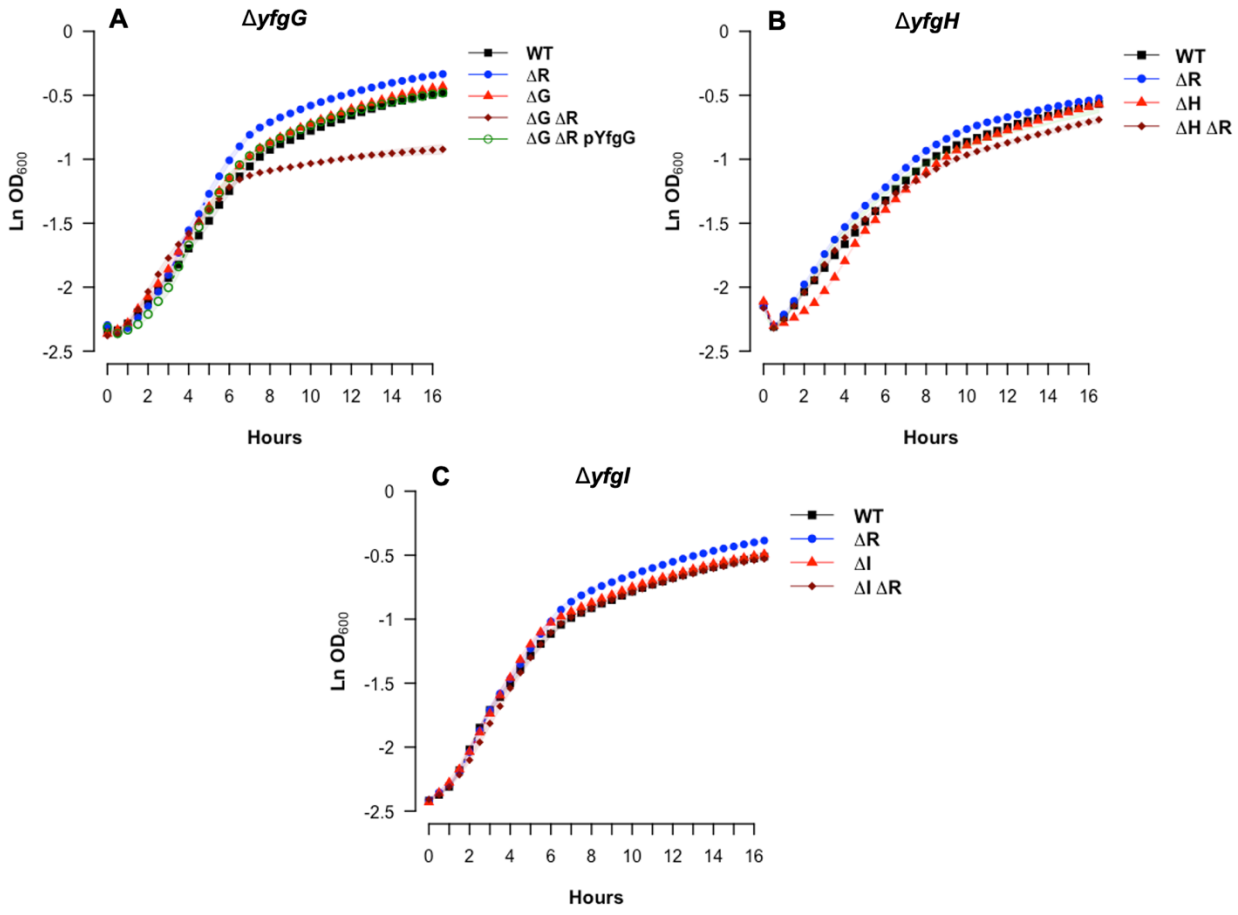


Figure 6: A dual knockout of *yfgG* and *cpxR* impairs growth under neutral conditions. MC4100 wildtype, single and dual knockouts of *yfgG* (**A**), *yfgH* (**B**), *yfgI* (**C**) and *cpxR* were subcultured in LB supplemented with 100mM Tris-HCl pH 7 and grown at 37°C for 16 hours in a 96-well plate with 237rpm orbital shaking. OD_{600} readings were taken every 30min. Shaded regions represent SEM (n=3). This is a representation of at least 3 independent experiments.

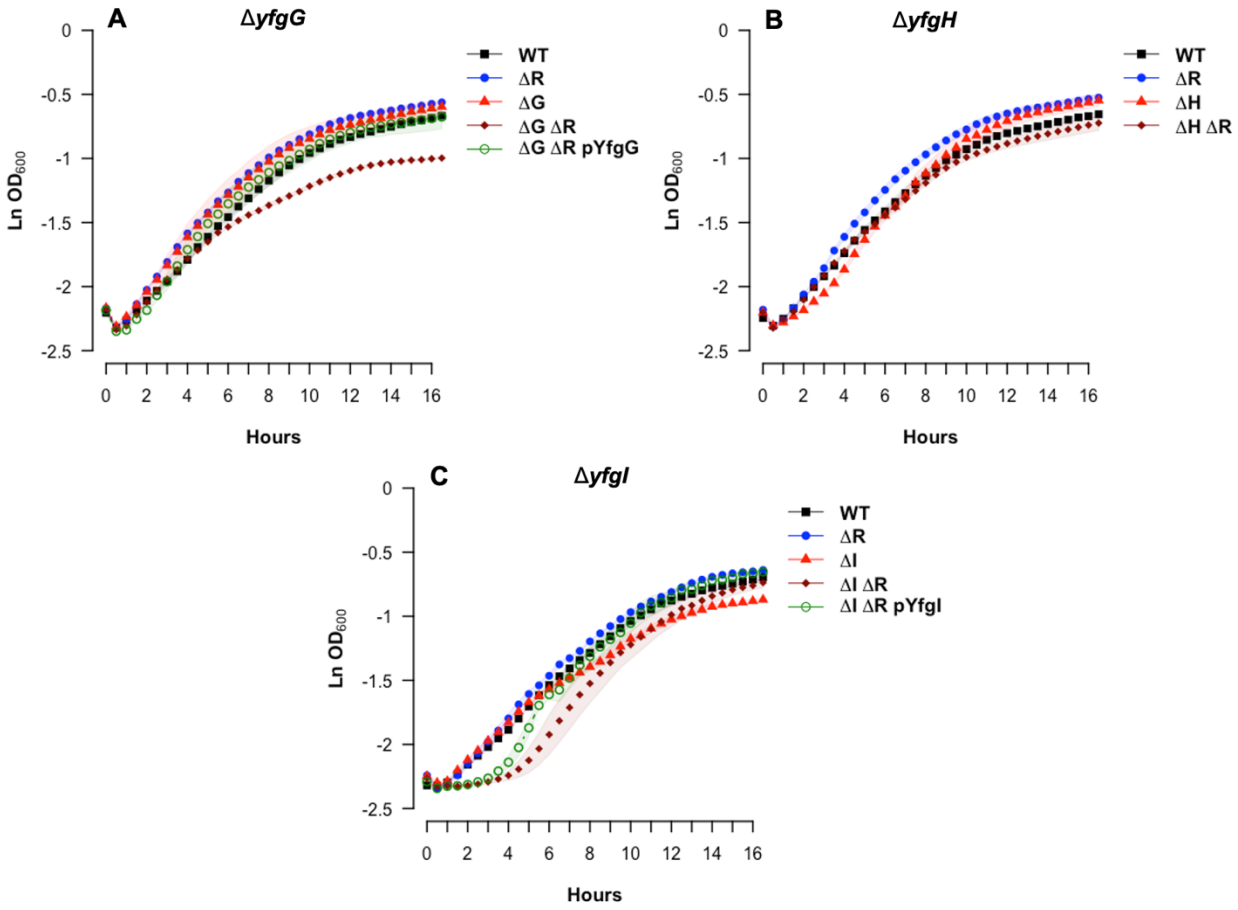


Figure 7: Absence of YfgG, YfgH, or YfgI does not affect growth in osmotic conditions. MC4100 wildtype, single and dual knockouts of *yfgG* (**A**), *yfgH* (**B**), *yfgI* (**C**) and *cpxR* were subcultured in LB supplemented with 300mM NaCl and grown at 37°C for 16 hours in a 96-well plate with 237rpm orbital shaking. OD_{600} readings were taken every 30min. Shaded regions represent SEM (n=3). This is a representation of at least 2 independent experiments.

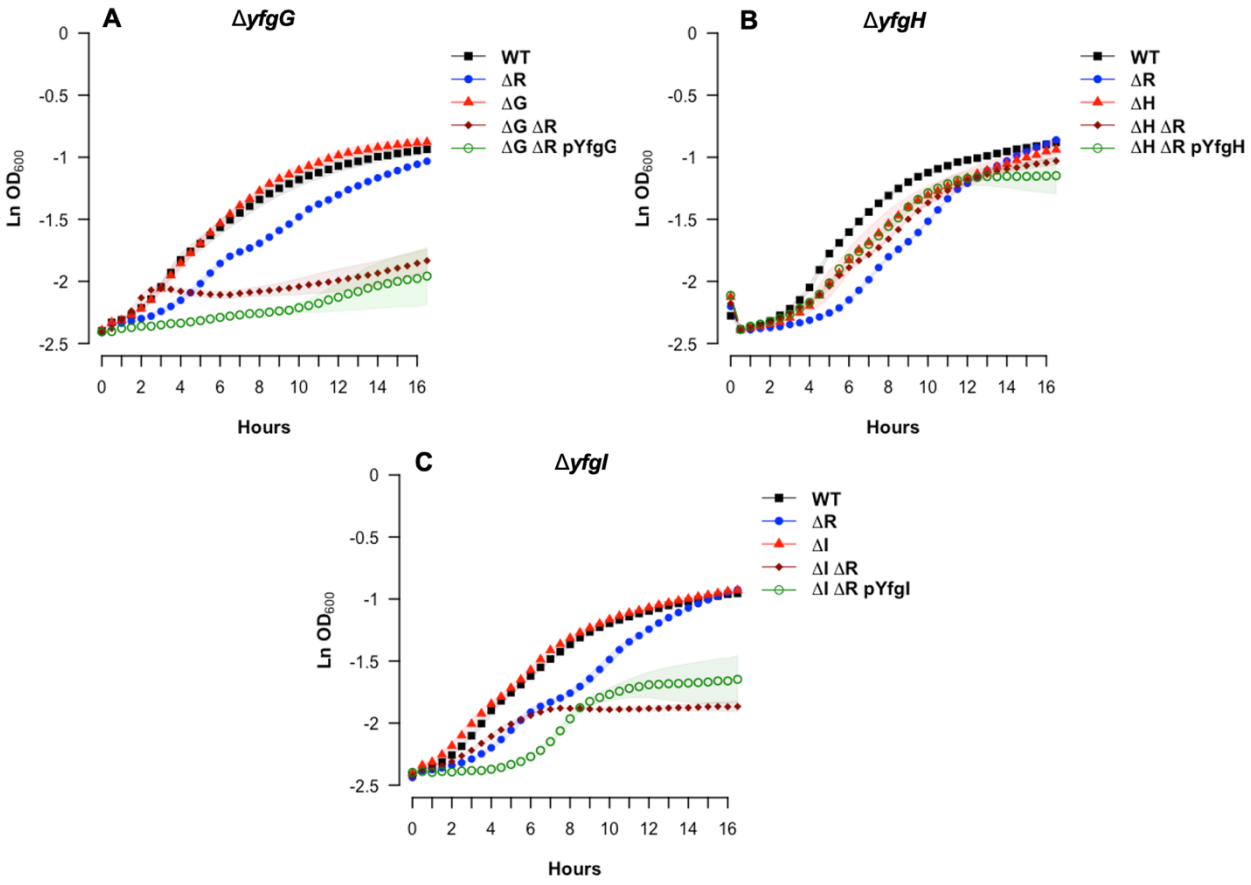


Figure 8: Severe growth deficiencies are present under alkaline stress in the absence of *yfgG* or *yfgI* in a *cpxR* knockouts

MC4100 wildtype, single and dual knockouts of *yfgG* **(A)**, *yfgH* **(B)**, *yfgI* **(C)** and *cpxR* were subcultured in LB supplemented with 100mM Tris-HCl pH 8, 0.1mM IPTG at 37°C for 16 hours in a 96-well plate with 237rpm orbital shaking. OD_{600} readings were taken every 30min. Shaded regions represent SEM (n=3). This is a representation of at least 3 independent experiments.

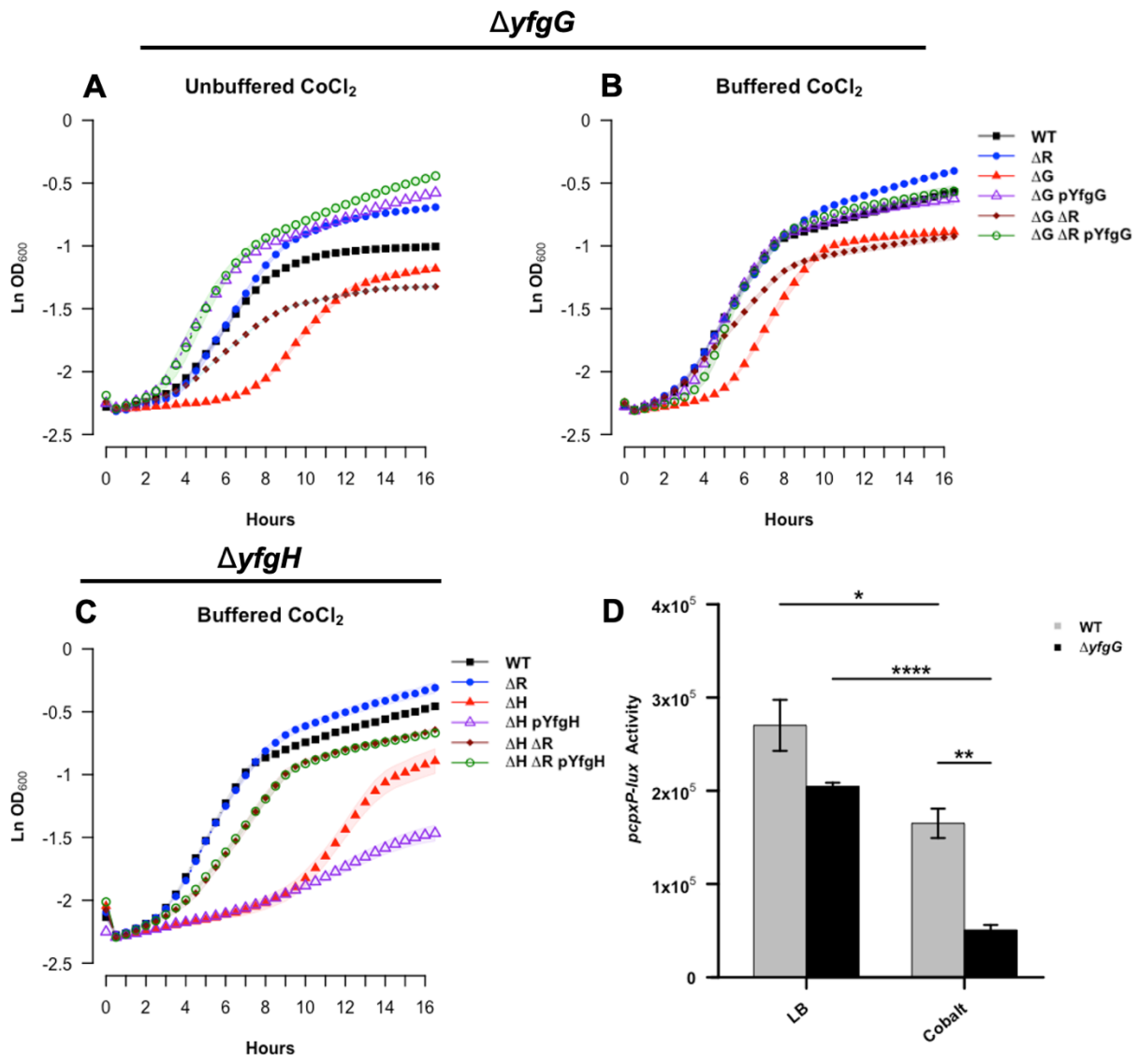


Figure 9: YfgG and YfgH contribute to cobalt resistance

Wildtype MC4100 and knockout mutants (**A,B:** Δ*cpxR,* Δ*yfgG,* Δ*yfgG cpxR::kan*; **C:** Δ*cpxR,* Δ*yfgH,* Δ*yfgH cpxR::kan***)** were subcultured in LB supplemented with 0.75mM $CoCl_2$ with **(B,C)** or without **(A)** 100mM Tris HCl pH 7 and grown at 37°C for 16 hours in a 96-well plate with 237rpm orbital shaking. OD_{600} readings were taken every 30min. Shaded regions represent SEM (n=3). **D) Absence of YfgG decreases Cpx activity in the presence of cobalt.** Wildtype MC4100 and Δ*yfgG* mid-log cells carrying the *cpxP-lux* reporter were pelleted and resuspended in LB supplemented with 0.75mM $CoCl_2$. 200uL of culture was transferred to a black 96-well plate and incubated for an hour at 37°C with 225rpm shaking. Luminescence (CPS) was measured and standardized to OD_{600}. Error bars represent SEM (n=3). Asterisks (*) indicate

statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test). **A-C)** This is a representation of at least 2 independent experiments.

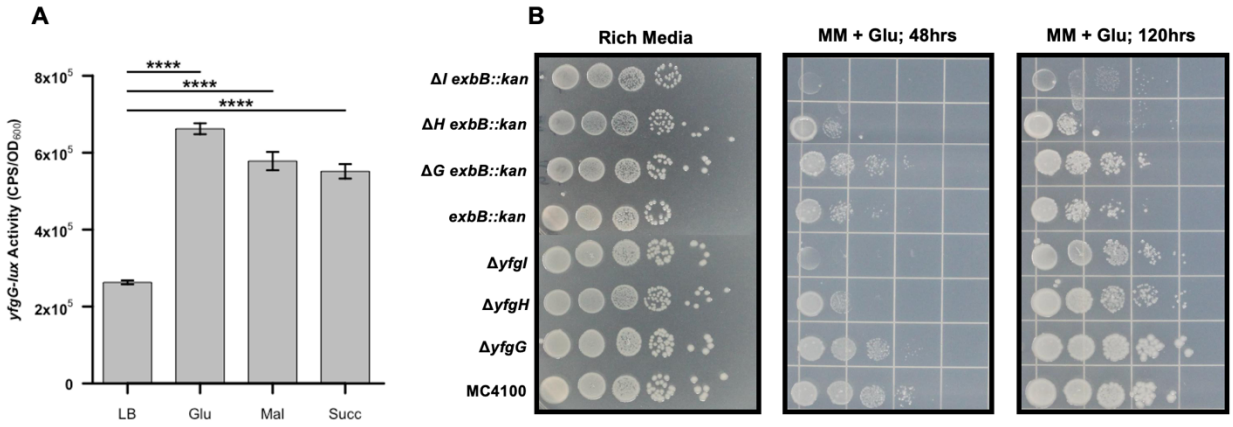


Figure 10: Nutrient limiting conditions activate expression of *yfgG*

A) MC4100 cells at mid-log carrying the *yfgG-lux* reporter were pelleted, washed twice with PBS, and resuspended in M9 minimal media supplemented with various carbon sources (0.4% glucose [Glu], 0.4% succinic acid [Succ], 0.4% malic acid [Mal]). 200µL culture was transferred to a black 96-well plate and grown for an hour at 37 °C with 225 rpm shaking. Luminescence (CPS) was measured and standardized to optical density (OD_{600}). Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test). This is a single stand-alone experiment and should be repeated. **B) Growth in minimal media is limited in the absence of YfgH, YfgI, and ExbB.** Overnight cultures of wildtype MC4100 and *yfgG* (ΔG), *yfgH* (ΔH), *yfgI* (ΔI), and *exbB* knockouts were standardized to $OD_{600} = 1.0$ and 10µL were spotted in a serial dilution series on M9 minimal media agar supplemented with 0.4% glucose. Plates were grown for 48 hours at 37°C then for an additional 72 hours at room temperature.

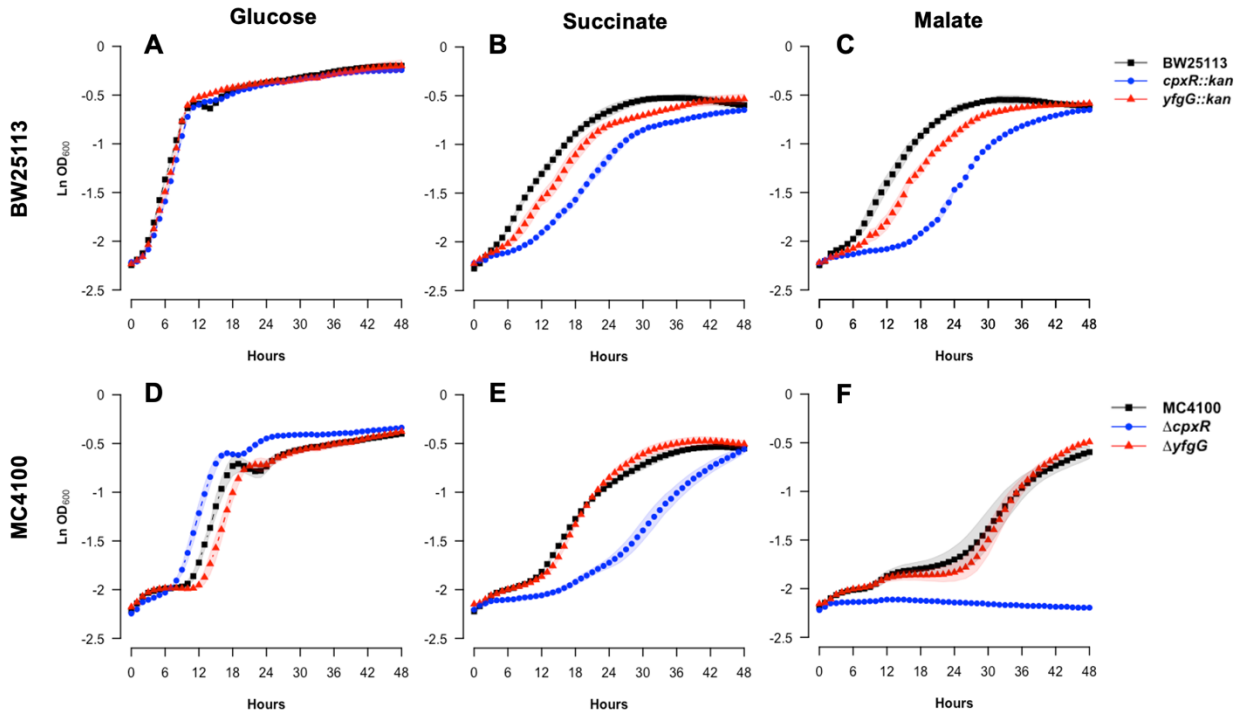


Figure 11: Absence of YfgG has little effect on growth under nutrient limitation

Overnight cultures of wildtype, Δ*cpxR* (blue), and Δ*yfgG* (red) knockouts of *E. coli* strains BW25113 **(A-C)** and MC4100 **(D-F)** were washed twice in PBS, standardized to OD_{600} = 1.0 and subcultured 1:10 in 200µL M9 minimal media supplemented with various carbon sources (0.4% glucose, 0.4% succinic acid, 0.4% malic acid) in a 96-well plate. Cultures were grown for 48 hours at 37°C with 237rpm orbital shaking. OD_{600} readings were taken every hour. Shaded regions represent SEM (n=3). This is a representation of at least 2 independent experiments.

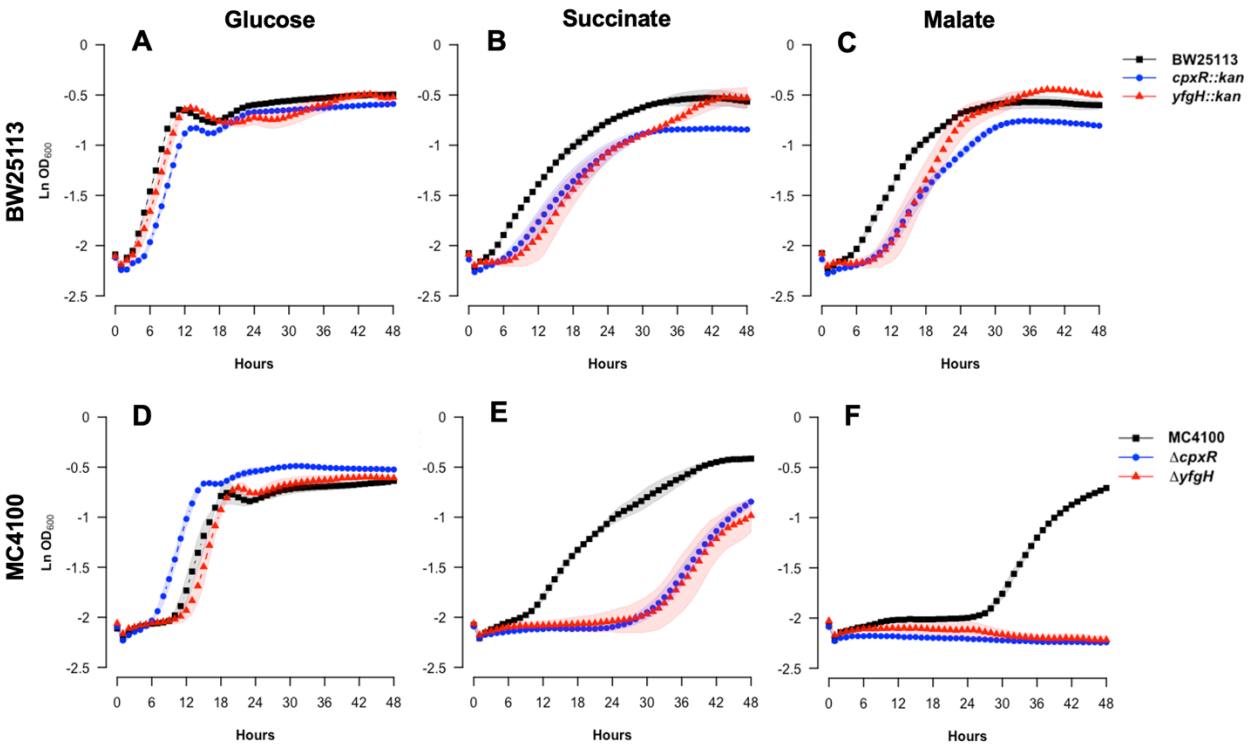


Figure 12: Absence of YfgH severely impacts growth on succinate and malate in MC4100

Overnight cultures of wildtype, Δ*cpxR* (blue), and Δ*yfgH* (red) knockouts of *E. coli* strains BW25113 **(A-C)** and MC4100 **(D-F)** were washed twice in PBS, standardized to OD_{600} = 1.0 and subcultured 1:10 in 200µL M9 minimal media supplemented with various carbon sources (0.4% glucose, 0.4% succinic acid, 0.4% malic acid) in a 96-well plate. Cultures were grown for 48 hours at 37°C with 237rpm orbital shaking. OD_{600} readings were taken every hour. Shaded regions represent SEM (n=3). This is a single stand-alone experiment and should be repeated.

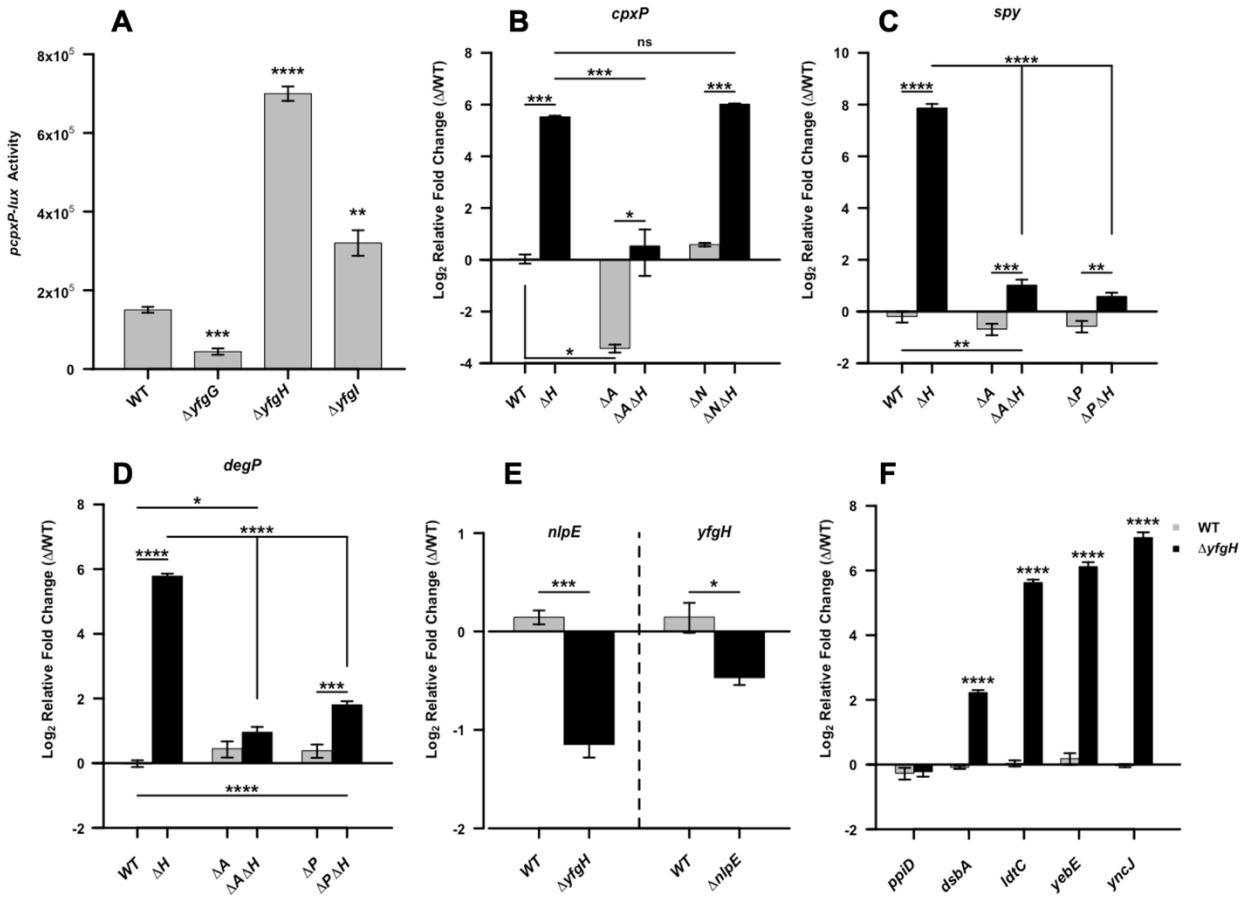


Figure 13: The Cpx stress response is activated in the absence of YfgH in a CpxA and CpxP dependent manner

A) MC4100 knockout strains of *yfgG, yfgH,* and *yfgI* carrying the *cpxP-lux* reporter were grown to mid-log, at which point 200µL culture was transferred to a black 96-well plate and grown for an hour at 37 degrees with 225 rpm shaking. Luminescence (CPS) were measured and standardized to optical density (OD_{600}). **B-D)** Relative gene expression of *cpxP* (**B**)*, spy* (**C**)*,* and *degP* (**D)** were determined by mRNA transcript levels measured by qPCR on wild-type MC4100, *yfgH* (ΔH)*, cpxA* (ΔA)*,* and *cpxP* (ΔP) knockout strains at mid-log. **E)** Expression levels of *nlpE* and *yfgH* determined by mRNA transcript levels measured by qPCR in wild-type MC4100 and *nlpE, or* yfgH knockouts at mid-log. **F)** Expression levels of other Cpx regulated genes (*dsbA, ldtC, yebE, yncJ*) and *ppiD* in $yfgH^{+}/yfgH^{-}$ strains determined by mRNA transcript levels measured by qPCR in wild-type MC4100 and *yfgH* knockouts at mid-log. **B-E)** Target genes for measured transcripts are indicated at the top of each chart. An internal control of *gyrA* was used to standardize expression levels between replicates and samples. Black bars represent Δ*yfgH*

single or double knockouts. Fold change is relative to wild-type (WT) expression levels for each respective gene indicated at the top of each chart (**B-E**) or x-axis (**F**). **A-F)** Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test).

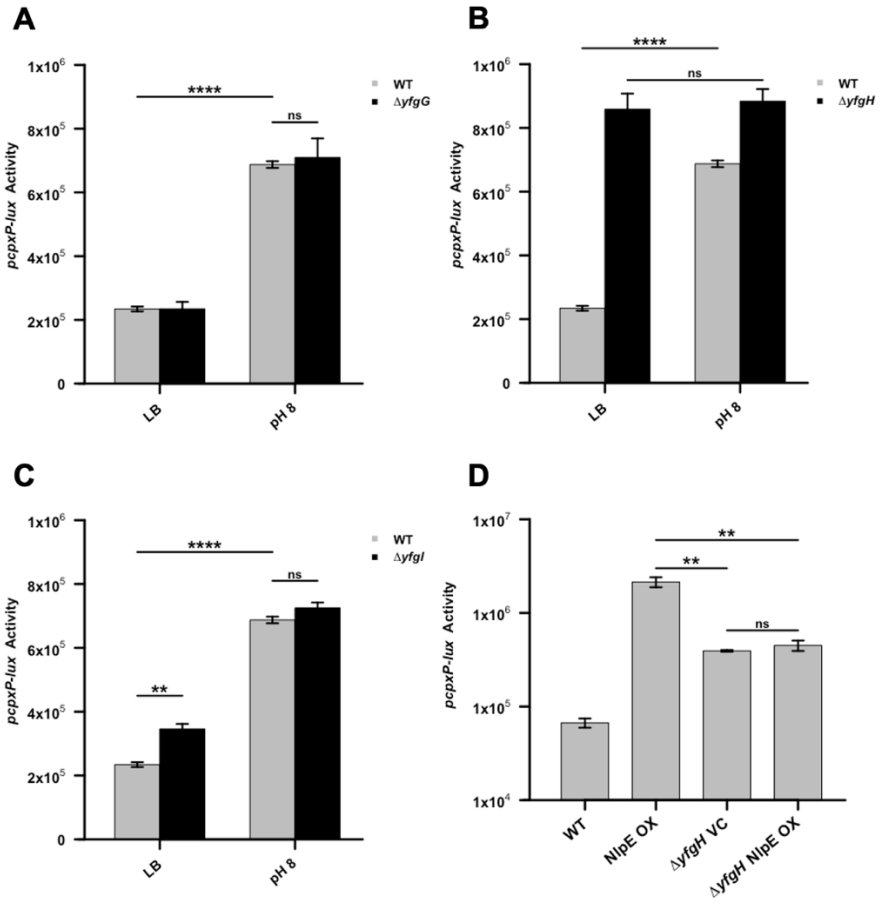


Figure 14: The Cpx pathways is unresponsive to stimuli in the absence of YfgH

A-C) Wildtype MC4100 and Δ*yfgG, ΔyfgH, and* Δ*yfgI* knockouts carrying the *cpxP-lux* reporter were grown to mid-log, pelleted, and resuspended in LB media supplemented with 100mM Tris-HCl pH 7 or pH 8. **D)** Wild-type MC4100 and Δ*yfgH* cells at mid-log carrying the *cpxP-lux*

reporter and a NlpE overexpression or vectors control were induced for 1 hour with 0.1mM IPTG. **A-D)** 200µL culture was transferred to a black 96-well plate and grown for an hour at 37°C with 225 rpm shaking. Luminescence (CPS) was measured and standardized to optical density (OD_{600}). Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test). This is a representation of at least 2 independent experiments.

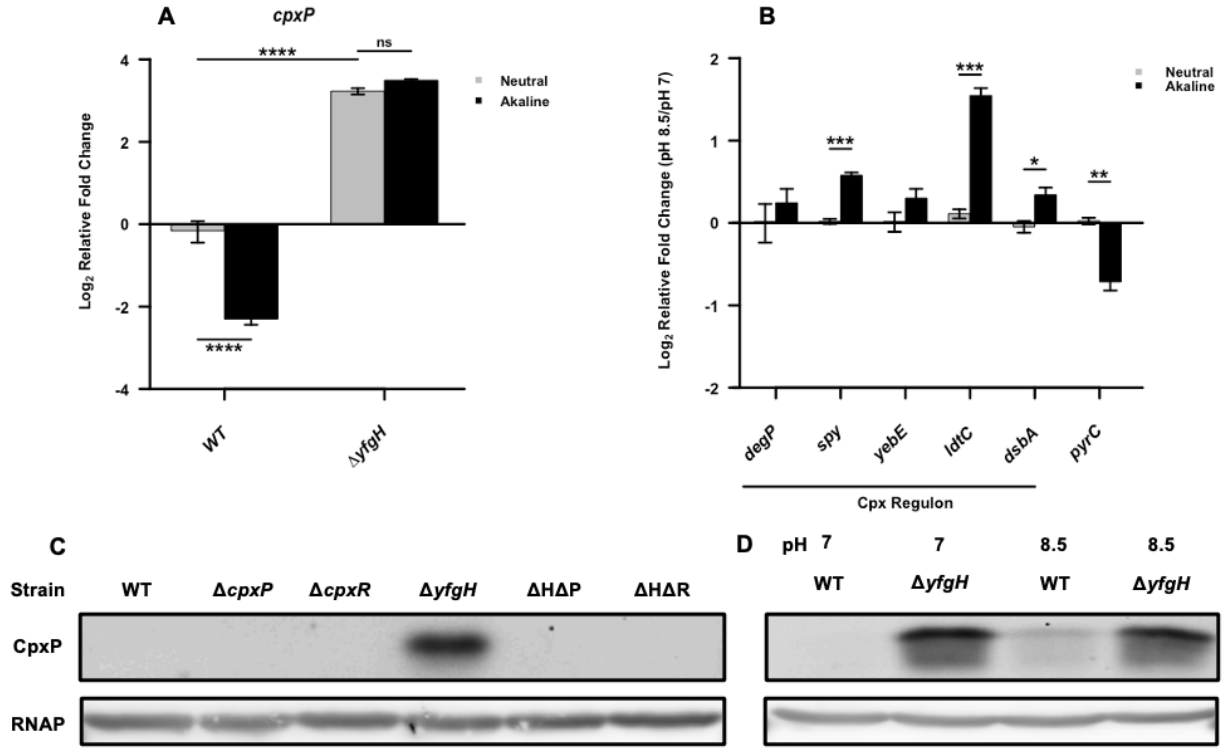


Figure 15: A-B) Transcript levels of *cpxP* are decreased during alkaline stress but unchanged with the removal of YfgH

A) Change in *cpxP* gene expression was determined by *cpxP* mRNA transcripts measured by qPCR in wildtype MC4100 and Δ*yfgH* cells exposed to alkaline stress (pH 8.5) for one hour. This is a representation of 4 independent experiments. **B)** Gene expression determined by mRNA transcript levels measured by qPCR of Cpx regulon members (*degP, spy, yebE, ldtC, dsbA*) and non-Cpx member *pyrC* after one hour exposure to alkaline stress (pH 8.5) in wild-type MC4100.

A-B) An internal control of *gyrA* was used to standardize expression levels between replicates and samples. Fold change is relative to wild-type (WT) at neutral pH. Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test). **C-D) CpxP protein levels are increased in the absence of YfgH. C)** Mid-log cells of wildtype MC4100 and *yfgH, cpxP, and cpxR* knockouts standardized to OD_{600} = 2.0 and 20µL were prepped for analysis with 2x Laemmli buffer. **D)** Wildtype MC4100 and *yfgH* knockouts were grown to mid-log, pelleted, resuspended in LB supplemented with 100mM Tris-HCl at pH 7 or pH 8 and grown for an additional hour at 37℃ with 225 rpms shaking. Cells were collected, standardized to OD_{600} = 2.0 and 20µL were prepped for analysis with 2x Laemmli buffer. **A-B)** Protein samples were run on 12% SDS-PAGE and visualized with α-MBP-CpxP and α-RNAP primary antibodies and IRDye fluorescent secondary antibodies (LI-COR). This is a representation of at least three independent experiments.

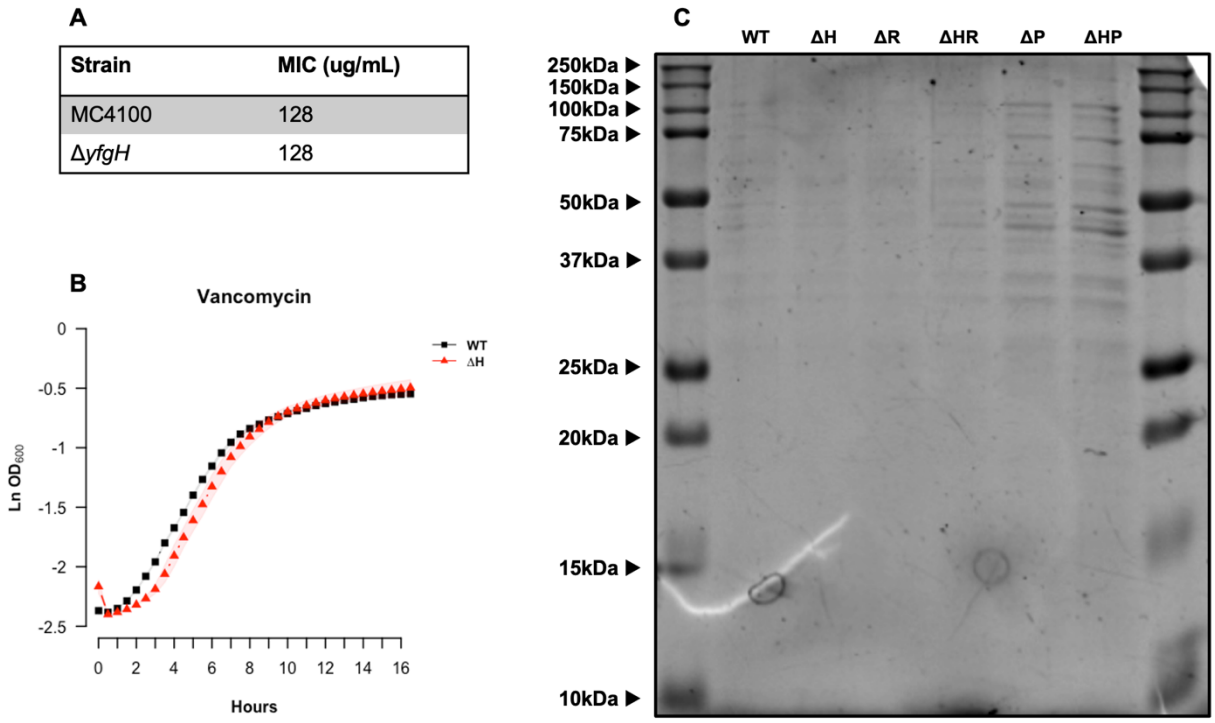


Figure 16: Absence of YfgH does not impact growth on vancomycin

A) MC4100 and Δ*yfgH* were subcultured 1:50 into 200µL LB supplemented with 100 µg/mL ampicillin and various concentrations of vancomycin in a 96-well plate. Cultures were grown for 24 hours at 37°C. MIC was determined both visually and by optical density. This experiment was in biological triplicate (n=3). **(B)** MC4100 and Δ*yfgH* were subcultured 1:100 into 200µL LB supplemented with 100 µg/mL ampicillin and 32 µg/mL vancomycin and grown at 37°C for 16 hours in a 96-well plate with 237rpm orbital shaking. OD_{600} readings were taken every 30min. Shaded regions represent SEM (n=3). These are single stand-alone experiments and need to be replicated. **C) Absence of YfgH causes a leaky membrane in the absence of CpxP.** Extracellular protein content was collected from the supernatant of wild-type MC4100 (WT), *yfgH* (ΔH), *cpxP* (ΔP) and *cpxR* (ΔR) knockouts grown to late log (OD_{600} = ~1.0) as described under Methods. Protein concentrations were measured by BCA Assay and standardized to the lowest concentration. Protein samples were run on 12% SDS-Page and stained with colloidal Coomassie to visualize the bands. This is a representation of at least two independent experiments.

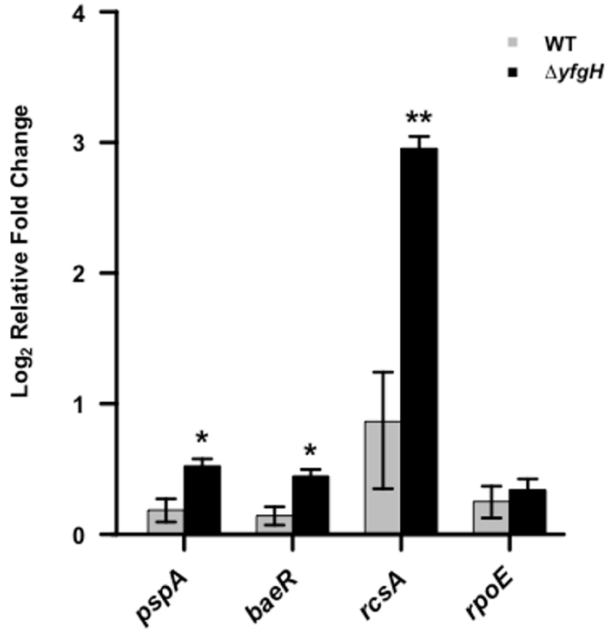


Figure 17 Absence of YfgH activates the Rcs response

Transcript levels of four envelope stress responses (Psp, Bae, Rcs, Sigma E) were assayed by qPCR in wild-type MC4100 and Δ*yfgH* knockout strains. An internal control of *gyrA* was used to standardize expression levels between replicates and samples. Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p < 0.05$; **: $0.01 > p > 0.001$, students T-test).

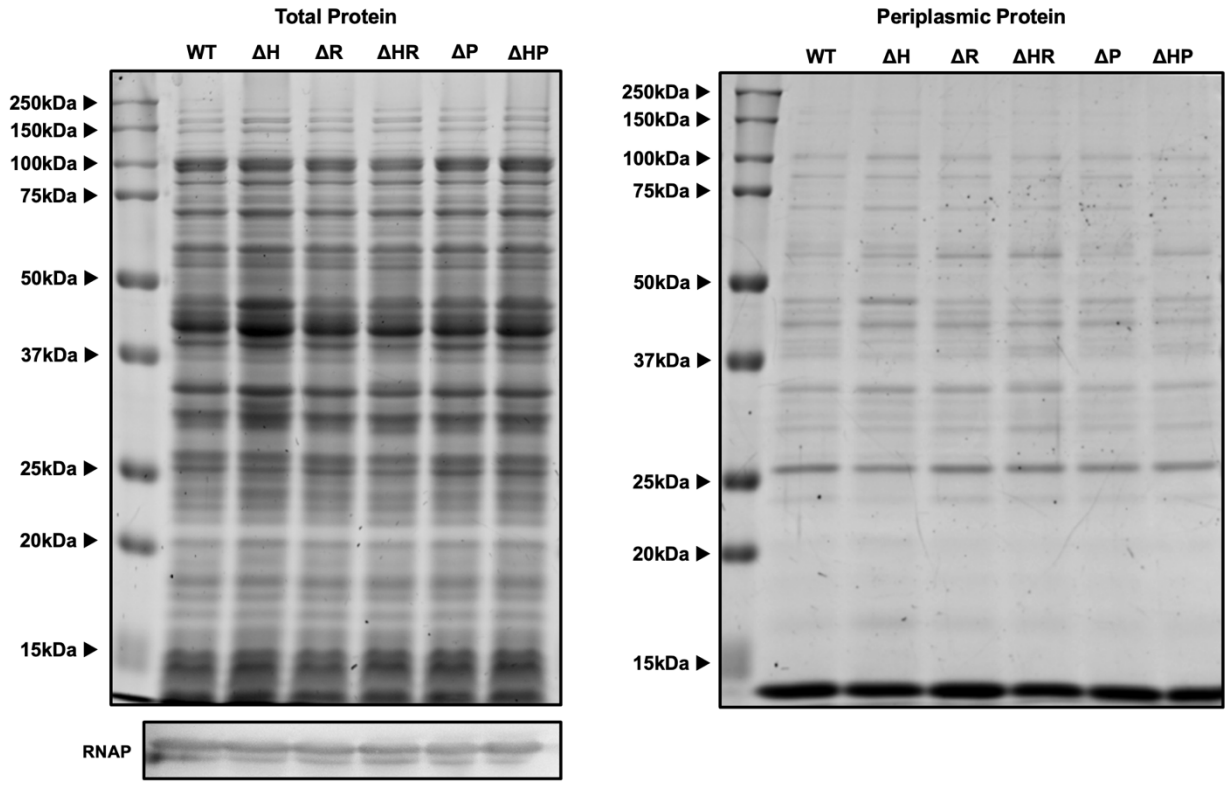


Figure 18: An increase in protein levels due to absence of YfgH is dependent on the Cpx response

Total (Left) and periplasmic protein (Right) content were collected from wild-type MC4100 (WT), *yfgH* (ΔH), *cpxP* (ΔP) and *cpxR* (ΔR) knockouts grown to late log at 37°C (OD_{600} = ~1.0) as described under Methods. Periplasmic protein concentrations were measured by BCA Assay and standardized to the lowest concentration. Protein samples were run on 12% SDS-Page and stained with colloidal Coomassie to visualize the bands. RNAP levels were assayed by western blot to verify equal protein levels for total whole cell protein content. This is a representation of three independent experiments

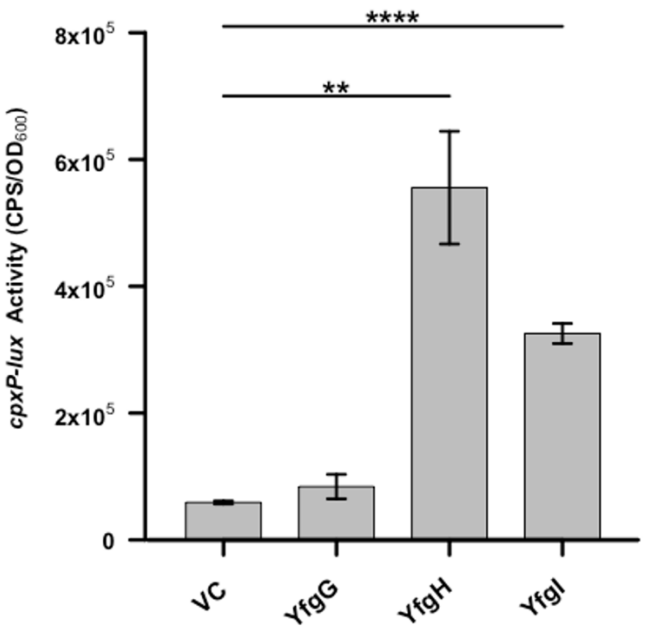


Figure 19: Overexpression of YfgH and YfgI induces the Cpx response

Wild-type MC4100 carrying the *cpxP-lux* reporter and either an NlpE overexpression or empty vector were subcultured 1:50 into LB media and grown to mid-log at 37°C. 0.1mM IPTG was added, and cells were grown for an additional hour before 200µL of culture was transferred to a black 96-well plate where CPS and OD_{600} was measured. CPS was standardized to optical density. Error bars represent SEM (n=3). Asterisks (*) indicate statistical significance (*:$p <$ 0.05; **: $0.01 > p > 0.001$, ***: $0.001 > p > 0.0001$, ****: $0.0001 > p$, students T-test).

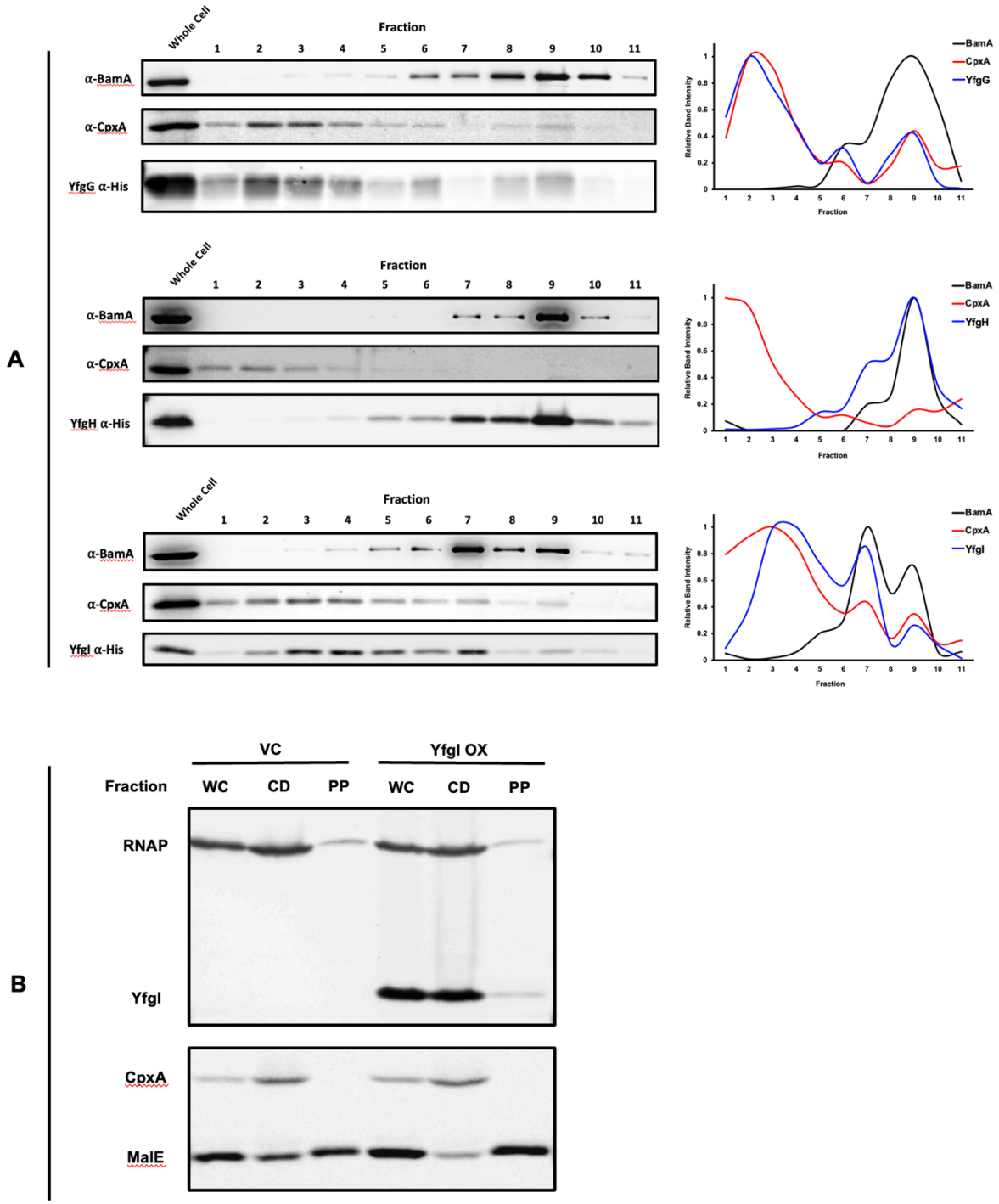


Figure 20: YfgG, YfgH, and YfgI are envelope localized proteins

MC4100 strains carrying IPTG inducible expression vectors of His-tagged YfgG, YfgH, and YfgI were subcultured 1:100 in 250mL LB with 100µg/mL ampicillin and grown at 37°C with ~180rpm shaking to an OD_{600} = ~1.0. Cultures were collected by centrifugation and membranes fractions were isolated using a sucrose density gradient (130). Fractions were run on 10% and

15% SDS-PAGE and visualized through immunoblotting with α-MBP-CpxA, α-BamA, α-His primary and IRDye fluorescent secondary antibodies (LI-COR). Band quantification was analyzed with ImageJ. **B)** MC4100 strains carrying a vector control or an IPTG inducible expression vector of His-tagged YfgI were subcultured 1:100 in 5mL LB with 100μg/mL ampicillin and grown at 37℃ with ~180rpm shaking to an OD_{600} = ~1.0. Cultures were collected by centrifugation and whole cell (WC), cell debris (CD) and periplasmic protein (PP) fractions were collected (131). Fractions were run on 10% and 15% SDS-PAGE and visualized through immunoblotting with α-MBP-CpxA, α-RNAP, α-His primary and IRDye fluorescent secondary antibodies (LI-COR).

Table 3: Top 10 increased or decreased proteins identified between periplasmic contents isolated from wild-type MC4100 and a Δ*yfgH* mutant.

Protein	Description	MW (kDa)	WT PSM	Δ*yfgH* PSM	Fold Change	Localization
Increased						
Ag43	Antigen 43, Surface autotransporter	106.8	2	24	**12.00**	PP/OM/EX
DegP	Periplasmic serine endoprotease	49.3	59	322	**5.46**	PP
PqqL	Putative zinc protease	104.6	21	104	**4.95**	PP/OM
OmpX	Outer membrane protein X	18.6	6	20	**3.33**	OM
YajG	Uncharacterized lipoprotein	20.9	3	9	**3.00**	OM
DsbA	Thiol:disulfide interchange protein	23.1	13	31	**2.38**	PP
PpiA	Peptidyl-prolyl cis-trans isomerase A	20.4	8	18	**2.25**	PP
OpgD	Glucans biosynthesis protein D	62.7	9	20	**2.22**	PP
OmpT	Outer Membrane protease	35.5	8	17	**2.13**	OM
PepN	Aminopeptidase N	98.9	4	8	**2.00**	IN
Decreased						
MalM	Maltose operon periplasmic protein	31.9	16	9	**0.56**	PP
YdeN	Uncharacterized sulfatase	62.8	18	10	**0.56**	PP
CpoB	Cell division coordinator	28.2	20	10	**0.50**	PP
LtdE	Probable L,D-transpeptidase	36.1	4	2	**0.50**	PP

SurA	Chaperone	47.3	47	22	**0.47**	PP
DsbC	Thiol:disulfide interchange protein	25.6	7	3	**0.43**	PP
LamB	Maltoporin	49.9	21	9	**0.43**	OM
GlpQ	Glycerophosphodiester	40.8	86	27	**0.31**	PP
Tsx	Nucleoside-specific channel-forming protein	33.6	20	6	**0.30**	OM
YhjJ	Peptidase M16 family protein	55.5	19	4	**0.21**	PP

PSM: Peptide Spectral Matches. The number of PSM's is the total number of identified peptide spectra matched for the protein and are roughly proportional to protein abundance. PSM cannot be used to compare different proteins. MW: molecular weight. Localization prediction based on GO (cellular component) annotation cross referenced with the EcoCyc database (134,216,217). EC: Extracellular; OM: Outer Membrane; PP: Periplasm; IM: Inner Membrane

Table 4: Proteins only found in the periplasmic content of either wild-type MC4100 or a Δ*yfgH* mutant.

Protein	**Description**	**MW (kDa)**	**Localization**
Found only in Δ*yfgH*			
BamB	Outer membrane protein assembly factor	41.9	OM
CpxP	Periplasmic protein	19	PP
Slt	Soluble lytic murein transglycosylase	73.3	PP
Spy	Periplasmic chaperone	18.2	PP
LtdC	Probable L,D-transpeptidase	34.6	PP
YebE	Inner membrane protein	23.7	IM
AcrF	Multidrug export protein	111.4	IM
AmiA	N-acetylmuramoyl-L-alanine amidase	31.4	PP
AmpC	Beta-lactamase	41.5	PP
ArgT	Lysine/arginine/ornithine-binding periplasmic protein	28	PP
CysP	Thiosulfate-binding protein	37.6	PP
DsbG	Thiol:disulfide interchange protein	27.5	PP
EnvC	Murein hydrolase activator	46.6	PP

NlpD	Murein hydrolase activator	40.1	PP
PspE	Thiosulfate sulfurtransferase	11.5	PP
RcsF	Outer membrane lipoprotein	14.2	PP /OM
Found only in wild-type			
OmpF	Outer membrane porin F	39.3	OM
RseB	Sigma-E factor regulatory protein	35.7	PP
BtuB	Vitamin B12 transporter	68.4	OM
FtsP	Cell division protein	51.8	PP
GudP	Probable glucarate transporter	49.1	IM
HybC	Hydrogenase-2 large chain	62.5	IM
OpgB	Phosphoglycerol transferase I	85.4	IM
NapA	Periplasmic nitrate reductase	93	PP
NarZ	Respiratory nitrate reductase 2 alpha chain	140.1	IM
SapA	Probable ABC transporter periplasmic-binding protein	61.5	PP
YgiS	Probable deoxycholate-binding periplasmic protein	60.7	PP
YnjE	Thiosulfate sulfurtransferase	48.2	PP

MW: molecular weight. Localization prediction based on GO (cellular component) annotation cross referenced with the EcoCyc database (134,216,217). EC: Extracellular; OM: Outer Membrane; PP: Periplasm; IM: Inner Membrane

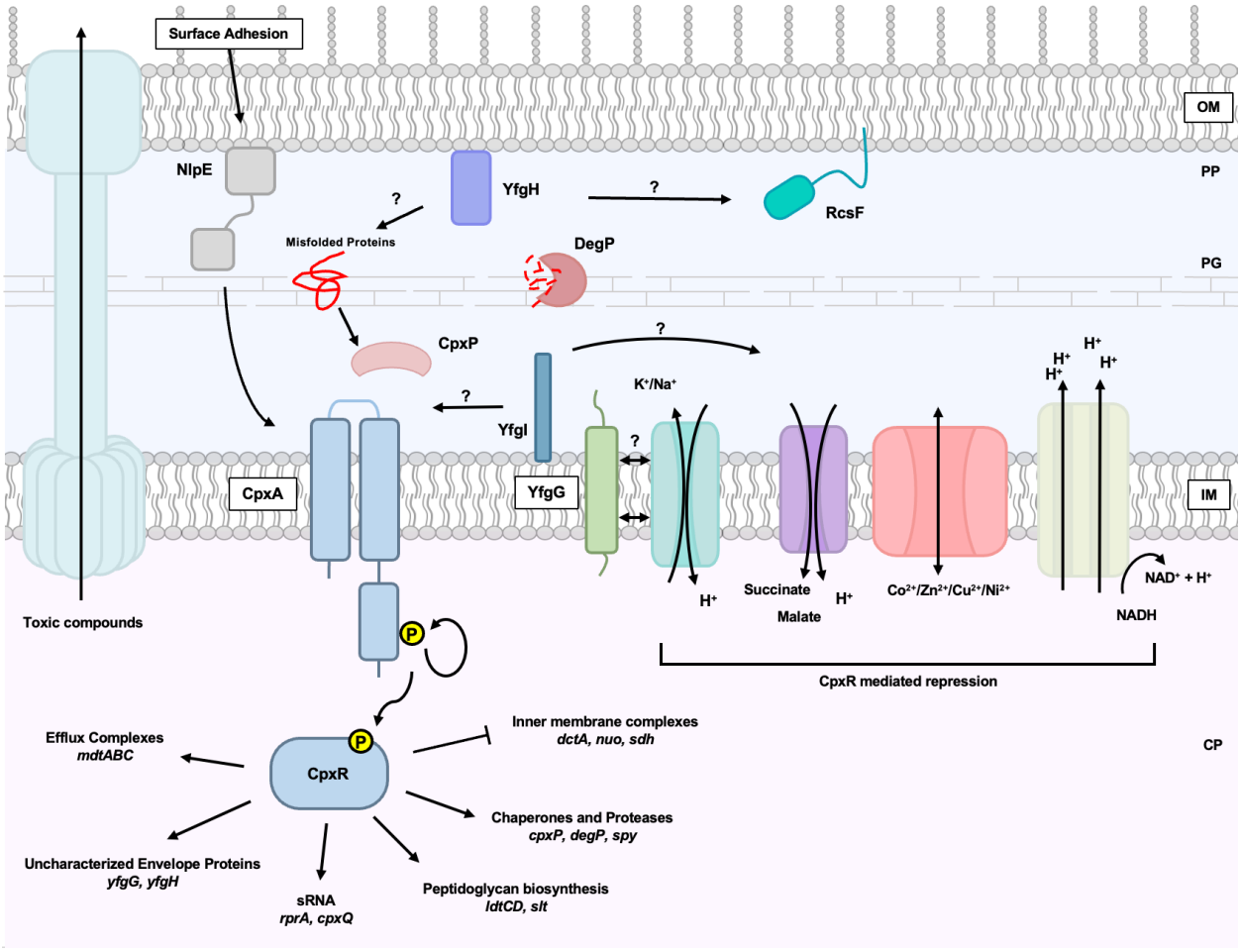


Figure 21: Model representing possible roles and interactions of YfgG, YfgH, and YfgI in the bacterial envelope.

Envelope localized YfgG, YfgH, and YfgI may play several roles in envelope homeostasis. Inner membrane proteins YfgG and YfgI may stabilize or modulate inner membrane proteins that transport protons, cations, and other molecules across the inner membrane. The Cpx response helps maintain homeostasis by regulation of inner membrane complexes, and periplasmic chaperones and proteases. Without either the Cpx response or YfgG/YfgI, these inner membrane proteins may become unstable, prone to misfolding or aggregation. The outer membrane lipoprotein YfgH could be involved in LPS and peptidoglycan synthesis, or general outer membrane integrity and may interact with the Cpx and Rcs responses. Removal of YfgH disrupts envelope homeostasis and strongly activates the Cpx response via CpxP rendering the response blind to additional envelope stress. This results in an increase of chaperones, proteases, and peptidoglycan modifying proteins in the periplasm, possibly due to misfolded proteins in the

periplasm. Cpx activation by removal of YfgH could result in the downregulation of various inner membrane complexes, such as carbohydrate and metal transporters.

Supplementary Tables and Figures

Table 5: List of strains and plasmids used in this study

Strain or Plasmid	Description	Source
MC4100	F-, [*araD*139]B/r, Δ(*argF-lac*)169, λ-, e14-, *flhD*5301, Δ(*fruK-yeiR*)725(*fruA*25), *relA*1, *rpsL*150(StrR), *rbsR*22, Δ(*fimB-fimE*)632(*::IS1*), *deoC1*	218
BW25113	*F*-, Δ(*araD-araB*)567, *lacI*+, Δ*lacZ*4787(*::rrnB-3*), λ-, (202, 258) *rph-1*, Δ(*rhaD-rhaB*)568, *hsdR*514	219
E2348/69	Prototypical EPEC O127:H6	220
ALN88	E2348/69 *cpxR::kan*	221
TR10	MC4100 *cpxA24*	222
J96	Uropathogenic O4:K6 *E. coli* isolated from pyelonephritis patient	223
DS17	Uropathogenic K5 *E. coli* isolated from pyelonephritis patient	224
MG1655	*F-, l-, rph-1*	225
W3110	*F- λ-* IN(*rrnD-rrnE*)1 *rph-1*	226
DH5α	F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1	Invitrogen
GT115	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) *φ80lacZ*ΔM15 Δ*lacX74 nupG recA1 araD*139 Δ(*ara-leu*)7697 *galE*15 *galK*16 *rpsL(StrA) endA1* Δ*dcm uidA*(Δ*MluI*)*::pir*-116 Δ*sbcC-sbcD*	InvivoGen
SY327 *λpir*	F^-, *araD*, Δ(*lac pro*), *argE*(Am), *recA56*, Rif^R, *gyrA* *λpir*	227

VT138	MC4100 Δ*cpxR*	Valeria Tsviklist, unpublished
JW3883	BW25113 *cpxR::kan*	219
JW3882	BW25113 *cpxA::kan*	219
JW5558	BW25113 *cpxP::kan*	219
JW5399	BW25113 *yfgG::kan*	219
JW5400	BW25113 *yfgH::kan*	219
JW2490	BW25113 *yfgI::kan*	219
MP1	*E. coli* isolated from laboratory mouse feces	168
MP13	MP1 att_{λ}::pAS07	168
KP16	MP13 *cpxR::kan*	Kat Pick, unpublished
KP116	MP13 pTrc99A	Kat Pick, unpublished
JGB50	MC4100 Δ*yfgH*	This Study
JGB99	MC4100 pTrc99A	This Study
JGB100	MC4100 pTrc-nlpE-wt-3	This Study
JGB101	VT138 pTrc99A	This Study
JGB102	VT138 pTrc-nlpE-wt3	This Study
JGB142	MC4100 pJW15-yfgG pTrc99A	This Study
JGB143	MC4100 pJW15-yfgG pTrc-nlpE-wt3	This Study
JGB144	VT138 pJW15-yfgG pTrc99A	This Study
JGB145	VT138 pJW15-yfgG pTrc-nlpE-wt3	This Study
JGB163	TR10 pJW15-yfgG	This Study
JGB124	MC4100 Δ*yfgG* pTrc99A	This Study
JGB125	MC4100 Δ*yfgG cpxR::kan* pTrc99A	This Study
JGB126	MC4100 Δ*yfgG cpxR::kan* pTrc-yfgG	This Study
JGB130	MC4100 Δ*yfgG* pTrc-yfgG	This Study
JGB132	MC4100 Δy*fgH* pTrc99A	This Study

JGB133	MC4100 Δ*yfgH* pTrc-yfgH	This Study
JGB135	MC4100 Δ*yfgH cpxR::kan* pTrc99A	This Study
JGB136	MC4100 Δ*yfgH cpxR::kan* pTrc-yfgH	This Study
JGB138	MC4100 Δ*yfgI* pTrc99A	This Study
JGB139	MC4100 Δ*yfgI* pTrc-yfgI	This Study
JGB140	MC4100 Δ*yfgI cpxR::kan* pTrc99A	This Study
JGB141	MC4100 Δ*yfgI cpxR::kan* pTrc-yfgI	This Study
JGB183	MC4100 Δ*yfgG* pJW25	This Study
JGB184	MC4100 Δ*yfgH* pJW25	This Study
JGB184	MC4100 Δ*yfgI* pJW25	This Study
JGB189	MC4100 pJW25	This Study
JGB199	MC4100 pJW25 pCA24N	This Study
JGB200	MC4100 pJW25 pNlpE	This Study
JGB201	MC4100 Δ*yfgH* pJW25 pCA24N	This Study
JGB202	MC4100 Δ*yfgH* pJW25 pNlpE	This Study
JGB204	MC4100 pJW25 pCA-yfgG	This Study
JGB205	MC4100 pJW25 pCA-yfgH	This Study
JGB206	MC4100 pJW25 pYfgI	This Study
JGB209	MC4100 *exbB::kan*	This Study
JGB210	MC4100 Δ*yfgG exbB::kan*	This Study
JGB211	MC4100 Δ*yfgH exbB::kan*	This Study
JGB212	MC4100 Δ*yfgI exbB::kan*	This Study
JGB216	MC4100 *ΔnlpE*	This Study
JGB218	MC4100 *ΔyfgH ΔnlpE*	This Study
JGB224	E2348/69 pTrc99A	This Study
JGB225	E2348/69 pTrc-nlpE-wt3	This Study
JGB226	ALN88 pTrc99A	This Study
JGB227	MP13 pTrc-nlpE-wt3	This Study
JGB228	KP16 pTrc99A	This Study
JGB229	MC4100 *cpxP::kan*	This Study

JGB230	MC4100 *ΔyfgH cpxP::kan*	This Study
JGB231	MC4100 *cpxA::kan*	This Study
JGB232	MC4100 *ΔyfgH cpxA::kan*	This Study
pJW25	pJW15 luminescent reporter with *cpxP* promoter insert; Kan^R	122
pJW15-yfgG-41	pJW15 luminescent reporter with *yfgG* promoter insert; Kan^R	This study
pTrc99A	IPTG inducible overexpression vector; Amp^R	228
pTrc-nlpE-WT3	IPTG inducible overexpression vector with C-terminal 6xHis tagged nlpE insert; Amp^R	147
pTrc-yfgG	IPTG inducible overexpression vector with C-terminal 6xHis tagged yfgG insert; Amp^R	This study
pTrc-yfgH	IPTG inducible overexpression vector with C-terminal 6xHis tagged yfgH insert; Amp^R	This study
pTrc-yfgI	IPTG inducible overexpression vector with C-terminal 6xHis tagged yfgI insert; Amp^R	This study
pFLP2	Heat dependent FRT recombinase, Amp^R	121
pCA-24N	Empty vector control for ASKA library containing a ITPG inducible promoter; Cam^R	129
pNlpE	ITPG-inducible NlpE overexpression vector from the ASKA library with a corrected E128D mutation; Cam^R	129
pCA-yfgG	ITPG-inducible YfgG overexpression vector from the ASKA library; Cam^R	129
pCA-yfgH	ITPG-inducible YfgH overexpression vector from the ASKA library; Cam^R	129
pYfgI	ITPG-inducible YfgI overexpression vector from the ASKA library with a corrected V145M mutation; Cam^R	129, This Study

Table 6: List of Primers used in this study

Name	5'-Sequence-3'
yfgG_OX_Sac1_F	AATGAGCTCTGGGACATACTTCAAGGAACCTTTTGTG
yfgG_6xHis_Xba1_R	AATTCTAGACTAGTGGTGATGGTGATGATGACGTTGTACCGGTG ATTCGACG
yfgH_OX_Sac1_F	AATGAGCTCCGTAATGTGTATAACAAGGAATAGTG
yfgH_6xHis_Xba1_R	AATTCTAGATTAGTGGTGATGGTGATGATGGCTCTTTTCAGGAC ATTTGGTA
yfgI_OX_Sac1_F	TCGGAGCTCAGCTAATAATCAGGAGGAGTCATG
yfgI_6xHis_Xba1_R	AATTCTAGATTAGTGGTGATGGTGATGATGATTAAACCAGCTGT CCGATTTG
pTrc99A_F	GTTCTGGCAAATATTCTGAAA
pTrc99A_R	ATTTAATCTGTATCAGGCTGA
YfgG_MC4100_Pm_Xho1_F	ATTCTCGAGATGCGCGAAAGTATTAATGTGGG
yfgG_MC4100_PM_Kpn1_R	ATAGGTACCCAGTACGATACGGGTCATGCG
yfgH_pm_V3_EcoR1_F	ATAGAATTCATCCACACTCTCCGTCG
yfgH_pm_V3_BamH1_R	ATAGGATCCTGCCACAGGCAGAAGAC
pNlp10_F	GCTTCCCAACCTTACCAGAG
pNlp10_R	CACCAACAATTAATGGATTGCA
yfgG_P1_F	CCACCTGACTTCGAATGATG
yfgG_P1_R	AGATGAAAGAATATCTGCCGC
yfgH_P1_F	CCTTTATACGCAATACATTTACTTTCCT
yfgH_P1_R	TCTCTTGCTGGATTTCCCT
yfgI_P1_F	GGCACCAAAGTGTATACCTCTAC
yfgI_P1_R	GAACAAAACCCTCTGTTACTACAG
cpxP_P1_F	GGAAGTCAGCTCTCGGTCATC
cpxA_P1_F	GACTGACCAGCCGTCCATAAG
cpxR_P1_F	AACTATGCGCATCATTTGCTCC
exbB_P1_F	AGAATTGGCATAAAGCAG
K1 Rev	CAGTCATAGCCGAATAGCCT
pCA_yfgI_Q5_F2	CAAACCGATGTGGTGCAATCAGAAG
pCA_yfgI_Q5_R2	CGCAGCTTTGTGCTTCAGTTCC
pCA24N_F	AGGCCCTTTCGTCTTCACCTC
pCA24N_R	AGGCAGATCGTCAGTCAGTCA
yfgG_qPCR_F	TCGCATGACCCGTATCGTAC
yfgG_qPCR_R	CTTTTTTGCTCTGATGGTGCTG
yfgH_qPCR_F	AGGAAGGCACCAAAGTGTATACCT
yfgH_qPCR_R	ACGCGTTTCGTTATACGTCGT
yfgI_qPCR_F	AGGAACTGAAGCACAAAGCTGC
yfgI_qPCR_R	CACGCTTTTCATCAGATCGC

cpxP_qPCR_F	GCAAAAGTCCGCAACCAAA
cpxP_qPCR_R	TCGTTGTTGATGTTTCTCGTTTAAA
GyrACFwdCpxQqPCR	CACCGCAACGCAAAACG
GyrACRevCpxQqPCr	TGATGGAGATAACCCCTTTCGT
gyrA_EPEC_qPCR_F	AAGGCGATAAAGTCGTCTCTCTGA
nlpE_qPCR_F	ACCTTCCTCCTTCGCTTCCT
nlpE_qPCR_R	TTCACCTTTGCTGTCGGTTAATAC
degP_qCPR_F	GGCGAAAGCGATGAAAGTTG
degP_qPCR_R	CAAAGCTGCTGATCGGCTTAC
dsbA_qPCR_F	GCGTACAGAAAACCCAGACCAT
dsbA_qPCR_R	GCCGCGTCGTACTCTTCA
ldtC_qPCR_F	TCATCAACACTCCGATAAAAGTCTCT
ldtC_qPCR_R	TCTCTGACAGCGGCTGATGT
ppiD_qPCR_F	CGCCGGGCATCAACTC
ppiD_qPCR_R	CTGAACATCTGCCAACGGTTT
spy_qPCR_F	CAGATCGCGGAAATCATGA
spy_qPCR_R	AAGGTATCGCTGGCAATGATG
yebE_qPCR_F	GCGTACAGCACGTTTGATCCT
yebE_qPCR_R	TGCCGCACGTTCTTTGG
yncJ_qPCR_F	GTGGATCATCAGCTGCGTCAT
yncJ_qPCR_R	CAAACCTTCCGCCGACTCT
tnaA_qPCR_F	CAGCGTGAAGCAGAATACAAAGAC
tnaA_qPCR_R	TCGCCAGCATATCGGCATA
baeR_qPCR_F	CGCGCGTCAAAACCATTT
baeR_qpCR_R	AATCAACGGGCTTTCAGCAT
pspA_qPCR_F	CGCGCGTCAAAACCATTT
pspA_qPCR_R	AATCAACGGGCTTTCAGCAT
rcsA_qPCR_F	TCGACGATATCCTTGGCGATA
rcsA_qPCR_R	CCATCCACATTCGCAACATACT
rpoE_qPCR_F	ATGTTGTCAGAAGAACTGAGACAGATA
rpoE_qPCR_R	CTTCATAGCTCAGGCCATCCA

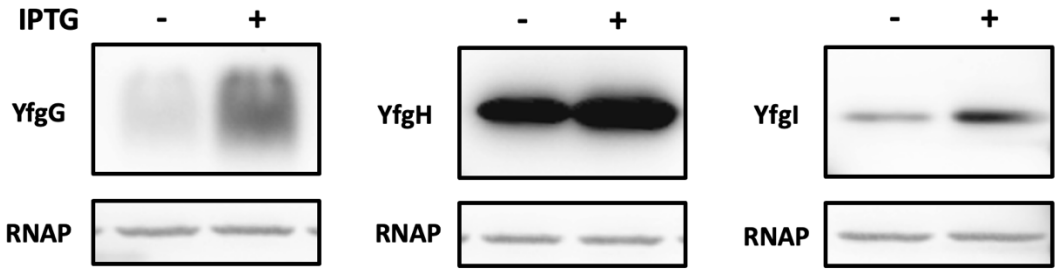


Figure S 1: Verification of His-tagged YfgG, YfgH, and YfgI leaky and IPTG induced expression

MC4100 strains carrying IPTG inducible expression vectors of YfgG, YfgH, and YfgI were subcultured 1:50 in 2mL LB with 100μg/mL ampicillin and grown at 37°C with 225rpm shaking to mid-log. Cultures were collected by centrifugation and standardized to OD_{600} = 2.0. Protein levels were run on 15% SDS-PAGE and visualized through immunoblotting with α-His, and α-RNAP primary and IRDye fluorescent secondary antibodies (LI-COR). RNAP levels were used as loading controls.

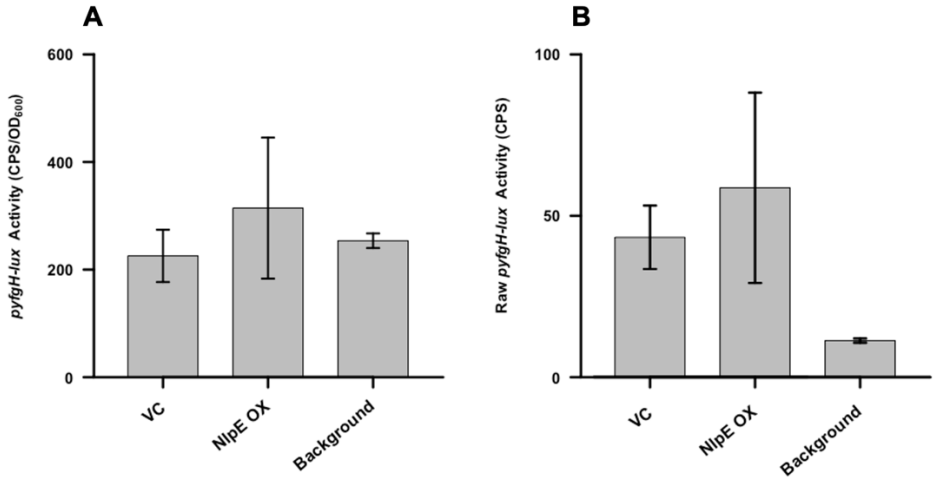


Figure S 2: The *yfgH-lux* reporter is barely active

A) LB unadjusted luminescent activity and **B)** Raw CPS (photon counts per second) for the *yfgH-lux* reporter. MC4100 carrying either the NlpE overexpression vector pTrc-nlpE-wt3 or the empty vector pTrc99A were subculture 1:50 in LB media and grown at 37°C to mid-log at which point 0.1mM IPTG was added. Cells were grown for an additional 30min where 200µL of culture was transferred to a black 96-well plate and CPS and OD_{600} was measured. CPS was standardized to optical density. Error bars represent SEM (n=3). This is a representation of at least 3 independent *yfgH-lux* reporter constructs.

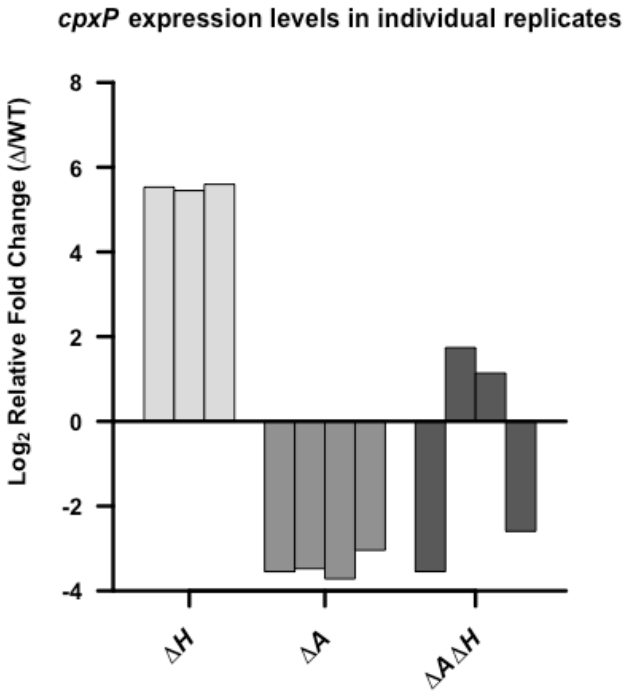


Figure S 3: Relative expression levels of *cpxP* in individual replicates of *yfgH* and *cpxA* knockouts.

Transcript levels of *cpxP* were measured by qPCR on wild-type MC4100, *yfgH,* and *cpxA* knockout strains at mid-log. Each bar represents an individual biological replicate. An internal control of *gyrA* was used to standardize expression levels between replicates and samples. Fold change is relative to wildtype *cpxP* expression levels. This figure represents single replicates from Figure 13B.

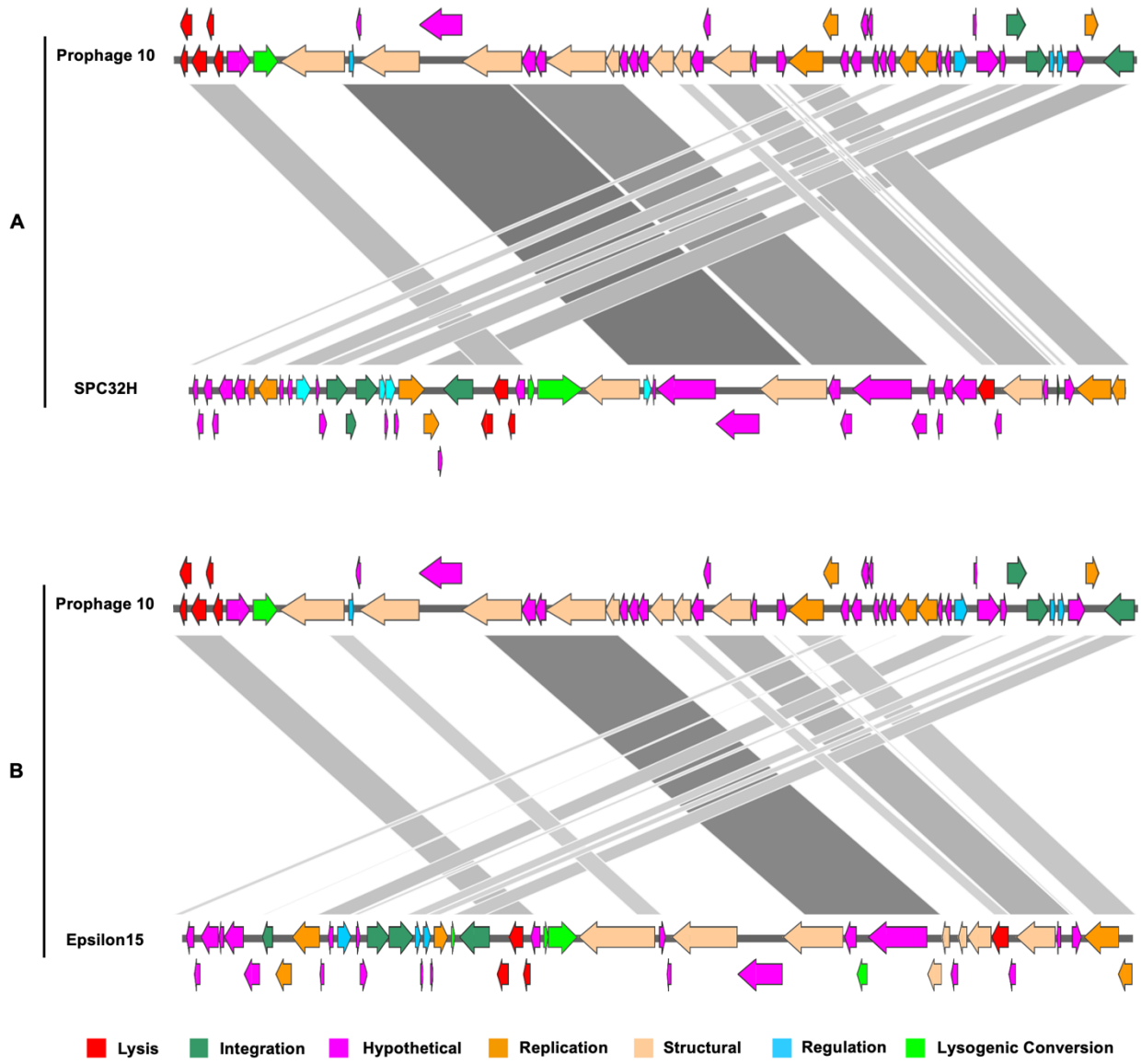


Figure S 4: Genome homology between E2348/69 prophage 10 and *Salmonella* phages SPC32H and Epsilon15

E2348/69 prophage 10 and **(A)** SPC32H and **(B)** Epsilon15 were annotated using PHASTER, BLAST, and GenBank reference genomes (SPC32H Accession No: KC911856; Epsilon15 Accession No: NC_004775) (124,125,128). Alignments were visualized with Kablammo (127). Darker lines represent regions with higher sequence similarity.

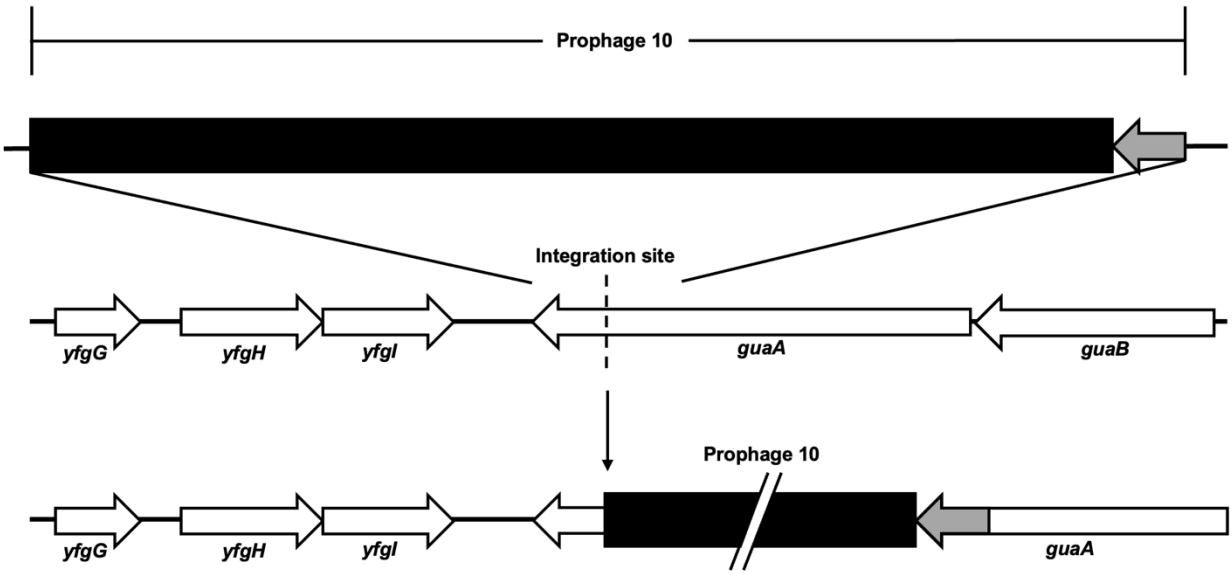


Figure S 5: Putative integration site for E2348/69 prophage 10

Representation of the putative integration site for prophage 10 into *guaA* in E2348/69. Complementary prophage 10 encoded *guaA* fragment is represented in grey and the rest of prophage 10 is in black.

References

1. Murray CJ, Shunji Ikuta K, Sharara F, *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 2022;**399**:629–55
2. Page ET. Trends in Food Recalls: 2004-13. EBI-191, U.S. Department of Agriculture, Economic Research Service, 2018
3. Nguyen Y, Sperandio V. Enterohemorrhagic *E. coli* (EHEC) pathogenesis. Frontiers in Cellular and Infection Microbiology 2012;**2**:90
4. Melton-Celsa AR. Shiga Toxin (Stx) Classification, Structure, and Function. Microbiology Spectrum 2014;**2**(4)
5. Hu J, Torres AG. Enteropathogenic *Escherichia coli*: foe or innocent bystander? Clinical Microbiology and Infection 2015;**21**(8):729–34.
6. Clarke SC, Haigh RD, Freestone PPE, Williams PH. Virulence of enteropathogenic *Escherichia coli*, a global pathogen. Clinical Microbiology Reviews 2003;**16**(3):365–78.
7. Franzin FM, Sircili MP. Locus of Enterocyte Effacement: A Pathogenicity Island Involved in the Virulence of Enteropathogenic and Enterohemorragic *Escherichia coli* Subjected to a Complex Network of Gene Regulation. BioMed Research International 2015;**2015:**534738
8. Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, *et al.* Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. Gut Pathogens 2019;**11**(10)
9. Sonnenborn U. Escherichia coli strain Nissle 1917—from bench to bedside and back: history of a special *Escherichia coli* strain with probiotic properties. FEMS Microbiology Letters 2016;**363**(19):fnw212
10. Deriu E, Liu JZ, Pezeshki M, *et al.* Probiotic Bacteria Reduce *Salmonella* Typhimurium Intestinal Colonization by Competing for Iron. Cell Host Microbe 2013;**14**(1):26-37
11. Tan Z, Lu P, Adewole D, Diarra MS, Gong J, Yang C. Iron requirement in the infection of *Salmonella* and its relevance to poultry health. Journal of Applied Poultry Research 2021;**30**(1):100101.
12. Blattner FR, Plunkett G, Bloch CA, *et al.* The Complete Genome Sequence of *Escherichia coli* K-12. Science 1997;**277**(5331):1453–62.

13. Ghatak S, King ZA, Sastry A, Palsson BO. The y-ome defines the 35% of *Escherichia coli* genes that lack experimental evidence of function. Nucleic Acids Research 2019;**47**(5):2446–54.
14. Silhavy TJ, Kahne D, Walker S. The Bacterial Cell Envelope. Cold Spring Harbor Perspectives in Biology 2010;**2**(5):a000414.
15. Glauert AM, Thornley MJ. The topography of the bacterial cell wall. 2003;**23**:159–98.
16. Sun J, Rutherford ST, Silhavy TJ, Huang KC. Physical properties of the bacterial outer membrane. Nature Reviews Microbiology 2021 20:4 2021;**20**(4):236–48.
17. Bertani B, Ruiz N. Function and biogenesis of lipopolysaccharides. EcoSal Plus 2018;**8**(1).
18. Liu B, Furevi A, Perepelov A v., *et al.* Structure and genetics of *Escherichia coli* O antigens. FEMS Microbiology Reviews 2020;**44**(6):655–83.
19. Meng J, Young G, Chen J. The Rcs System in Enterobacteriaceae: Envelope Stress Responses and Virulence Regulation. Frontiers in Microbiology 2021;**12:**627104
20. Hews CL, Cho T, Rowley G, Raivio TL. Maintaining Integrity Under Stress: Envelope Stress Response Regulation of Pathogenesis in Gram-Negative Bacteria. Frontiers in Cellular and Infection Microbiology 2019;**9**:313
21. Konovalova A, Mitchell AM, Silhavy TJ. A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. Elife 2016;**5**:e15276
22. Konovalova A, Perlman DH, Cowles CE, Silhavy TJ. Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins. Proc Natl Acad Sci U S A 2014;**111**(41):E4350–8.
23. Braun V. Covalent lipoprotein from the outer membrane of *Escherichia coli.* Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes 1975;**415**(3):335–77.
24. Vollmer W, Blanot D, de Pedro MA. Peptidoglycan structure and architecture. FEMS Microbiology Reviews 2008;**32**(2):149–67.
25. Mickiewicz KM, Kawai Y, Drage L, *et al.* Possible role of L-form switching in recurrent urinary tract infection. Nature Communications 2019;**10**(1):1–9.
26. Miot M, Betton JM. Protein quality control in the bacterial periplasm. Microbial Cell Factories 2004;**3**(1):1–13.

27. Stull F, Betton J-M, Bardwell JCA. Periplasmic Chaperones and Prolyl Isomerases. EcoSal Plus 2018;**8**(1).
28. Moon CP, Zaccai NR, Fleming PJ, Gessmann D, Fleming KG. Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. Proc Natl Acad Sci U S A 2013;**110**(11):4285–90.
29. Manta B, Boyd D, Berkmen M. Disulfide Bond Formation in the Periplasm of *Escherichia coli*. EcoSal Plus 2019;**8**(2).
30. Sohlenkamp C, Geiger O. Bacterial membrane lipids: diversity in structures and pathways. FEMS Microbiology Reviews 2016;**40**(1):133–59.
31. Wallin E, von Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Science 1998;**7**(4):1029–38.
32. Dalbey RE, Wang P, Kuhn A. Assembly of Bacterial Inner Membrane Proteins. Annual Reviews of Biochemistry. 2011;**80**:161–87.
33. Krulwich TA, Sachs G, Padan E. Molecular aspects of bacterial pH sensing and homeostasis. Nature Reviews Microbiology 2011;**9**(5):330-343.
34. Wikström M, Hummer G. Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications. Proc Natl Acad Sci U S A 2012;**109**(12):4431-4436
35. Bogachev A v., Murtazina RA, Skulachev VP. H+/e- stoichiometry for NADH dehydrogenase I and dimethyl sulfoxide reductase in anaerobically grown *Escherichia coli* cells. Journal of Bacteriology 1996;**178**(21):6233–7.
36. Maloney PC, Kashket ER, Wilson TH. A Protonmotive Force Drives ATP Synthesis in Bacteria. Proc Natl Acad Sci U S A 1974;**71**(10):3896-3900.
37. Biquet-Bisquert A, Labesse G, Pedaci F, Nord AL. The Dynamic Ion Motive Force Powering the Bacterial Flagellar Motor. Frontiers in Microbiology 2021;**12**:659464
38. Dalbey RE, Kuhn A. Protein Traffic in Gram-negative bacteria – how exported and secreted proteins find their way. FEMS Microbiology Reviews 2012;**36**(6):1023–45.
39. Cranford-Smith T, Huber D. The way is the goal: how SecA transports proteins across the cytoplasmic membrane in bacteria. FEMS Microbiology Letters 2018;**365**(11):fny093.
40. Crane JM, Randall LL. The Sec System: Protein Export in *Escherichia coli*. EcoSal Plus 2017;**7**(2).

41. Frain KM, Robinson C, van Dijl JM. Transport of Folded Proteins by the Tat System. The Protein Journal 2019;**38**(4):377-388.
42. Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway. Nature Reviews Microbiology 2012 10:7 2012;**10**(7):483–96.
43. Lee PA, Tullman-Ercek D, Georgiou G. The Bacterial Twin-Arginine Translocation Pathway. Annual Reviews of Microbiology 2006;**60**:373-95.
44. Plummer AM, Fleming KG. From Chaperones to the Membrane with a BAM! Trends in Biochemical Sciences 2016;**41**(10):872–82.
45. Tomasek D, Kahne D. The assembly of β-barrel outer membrane proteins. Current Opinion in Microbiology 2021;**60**:16–23.
46. Grabowicz M. Lipoproteins and Their Trafficking to the Outer Membrane. EcoSal Plus 2019;**8**(2).
47. Mitchell AM, Silhavy TJ. Envelope stress responses: balancing damage repair and toxicity. Nature Reviews Microbiology 2019;**17**:417–28.
48. Lehrer RI, Barton A, Daher KA, Harwig SSL, Ganz T, Selsted ME. Interaction of human defensins with *Escherichia coli*. Mechanism of bactericidal activity. Journal of Clinical Investigation 1989;**84**(2):553–61.
49. Lehrer RI, Lu W. α-Defensins in human innate immunity. Immunological Reviews 2012;**245**(1):84–112.
50. Brissette JL, Russel M, Weiner L, Model P. Phage shock protein, a stress protein of *Escherichia coli*. Proc Natl Acad Sci U S A 1990;**87**(3):862-6.
51. Raivio TL. Everything old is new again: An update on current research on the Cpx envelope stress response. Biochimica et Biophysica Acta - Molecular Cell Research 2014;**1843**(8):1529–41.
52. Leblanc SKD, Oates CW, Raivio TL. Characterization of the Induction and Cellular Role of the BaeSR Two-Component Envelope Stress Response of *Escherichia coli*. Journal of Bacteriology 2011;**193**(13):3367-75.
53. Raivio TL, Laird MW, Joly JC, Silhavy TJ. Tethering of CpxP to the inner membrane prevents spheroplast induction of the Cpx envelope stress response. Molecular Microbiology 2000;**37**(5):1186–97.

54. Raffa RG, Raivio TL. A third envelope stress signal transduction pathway in *Escherichia coli*. Molecular Microbiology 2002;**45**(6):1599–611.
55. Danese PN, Snyder WB, Cosma CL, Davis LJB, Silhavy TJ. The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. Genes & Development 1995;**9**(4):387–98.
56. Majdalani N, Hernandez D, Gottesman S. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. Molecular Microbiology 2002;**46**(3):813–26.
57. Vogt SL, Evans AD, Guest RL, Raivio TL. The Cpx envelope stress response regulates and is regulated by small noncoding RNAs. Journal of Bacteriology 2014;**196**(24):4229–38.
58. de Wulf P, McGuire AM, Liu X, Lin ECC. Genome-wide profiling of promoter recognition by the two-component response regulator CpxR-P in *Escherichia coli*. Journal of Biological Chemistry 2002;**277**(29):26652–61.
59. Price NL, Raivio TL. Characterization of the Cpx regulon in *Escherichia coli* strain MC4100. Journal of Bacteriology 2009;**191**(6):1798–815.
60. Zschiedrich CP, Keidel V, Szurmant H. Molecular Mechanisms of Two-Component Signal Transduction. Journal of Molecular Biology 2016;**428**(19):3752–75.
61. Brown AN, Anderson MT, Bachman MA, Mobley HLT. The ArcAB Two-Component System: Function in Metabolism, Redox Control, and Infection. Microbiology and Molecular Biology Reviews 2022. **86**(2):e0011021
62. Santos-Beneit F. The Pho regulon: A huge regulatory network in bacteria. Frontiers in Microbiology 2015;**6**(APR):402.
63. Milburn M v., Privé GG, Milligan DL, *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. Science 1991;**254**(5036):1342–7.
64. Buelow DR, Raivio TL. Three (and more) component regulatory systems - Auxiliary regulators of bacterial histidine kinases. Molecular Microbiology 2010;**75**(3):547–66.
65. Raivio TL, Popkin DL, Silhavy TJ. The Cpx Envelope Stress Response Is Controlled by Amplification and Feedback Inhibition. Journal of Bacteriology 1999;**181**(17):5263-72.

66. Danese PN, Silhavy TJ. CpxP, a Stress-Combative Member of the Cpx Regulon. Journal of Bacteriology 1998; **180**(4):831-9
67. Isaac DD, Pinkner JS, Hultgren SJ, Silhavy TJ. The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. Proc Natl Acad Sci U S A 2005;**102**(49):17775–9.
68. Tschauner K, Hörnschemeyer P, Müller VS, Hunke S. Dynamic Interaction between the CpxA Sensor Kinase and the Periplasmic Accessory Protein CpxP Mediates Signal Recognition in *E. coli*. PLoS ONE 2014;**9**(9):e107383.
69. DiGiuseppe PA, Silhavy TJ. Signal detection and target gene induction by the CpxRA two-component system. Journal of Bacteriology 2003;**185**(8):2432–40.
70. Delhaye A, Laloux G, Collet JF. The Lipoprotein NlpE Is a Cpx Sensor That Serves as a Sentinel for Protein Sorting and Folding Defects in the *Escherichia coli* Envelope. Journal of Bacteriology 2019;**201**(10):e00611-18.
71. May KL, Lehman KM, Mitchell AM, Grabowicz M. A stress response monitoring lipoprotein trafficking to the outer membrane. MBio 2019;**10**(3):e00618-19.
72. Otto K, Silhavy TJ. Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. Proc Natl Acad Sci U S A 2002;**99**(4):2287–92.
73. Shimizu T, Ichimura K, Noda M. The surface sensor NlpE of enterohemorrhagic *Escherichia coli* contributes to regulation of the type III secretion system and flagella by the Cpx response to adhesion. Infection and Immunity 2016;**84**(2):537–49.
74. Snyder WB, Davis LJB, Danese PN, Cosma CL, Silhavy TJ. Overproduction of NlpE, a New Outer Membrane Lipoprotein, Suppresses the Toxicity of Periplasmic LacZ by Activation of the Cpx Signal Transduction Pathway. Journal of Bacteriology 1995;**177(15)**:4216-23
75. Wang J. Molecular Mechanisms Used by Auxiliary Regulator NlpE to Signal the Cpx Envelope Stress Response in *Escherichia coli.* University of Alberta, 2017.
76. Hirano Y, Hossain MM, Takeda K, Tokuda H, Miki K. Structural Studies of the Cpx Pathway Activator NlpE on the Outer Membrane of *Escherichia coli*. Structure 2007;**15**(8):963–76.
77. Delhaye A, Collet JF, Laloux G. Fine-tuning of the Cpx envelope stress response is required for cell wall homeostasis in *Escherichia coli*. MBio 2016;**7**(1):e00047-16

78. Miyadai H, Tanaka-Masuda K, Matsuyama SI, Tokuda H. Effects of Lipoprotein Overproduction on the Induction of DegP (HtrA) Involved in Quality Control in the *Escherichia coli* Periplasm. Journal of Biological Chemistry 2004;**279**(38):39807–13.
79. May KL, Lehman KM, Mitchell AM, Grabowicz M. A stress response monitoring lipoprotein trafficking to the outer membrane. MBio 2019;**10**(3):e00618-19.
80. Cosma CL, Danese PN, Carlson JH, Silhavy TJ, Snyder WB. Mutational activation of the Cpx signal transduction pathway of *Escherichia coli* suppresses the toxicity conferred by certain envelope-associated stresses. Molecular Microbiology 1995;**18**(3):491-505
81. Mcewen J, Silverman P. Chromosomal mutations of *Escherichia coli* that alter expression of conjugative plasmid functions. Proc Natl Acad Sci U S A. 1980;**77**(1):513-7
82. Vogt SL, Raivio TL. Just scratching the surface: An expanding view of the Cpx envelope stress response. FEMS Microbiology Letters 2012;**326**(1):2–11.
83. Raivio TL, Leblanc SKD, Price NL. The *Escherichia coli* Cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. Journal of Bacteriology 2013;**195**(12):2755–67.
84. Jenal U, Reinders A, Lori C. Cyclic di-GMP: Second messenger extraordinaire. Nature Reviews Microbiology 2017;**15**(5):271–84.
85. Yi Z, Wang D, Xin S, *et al.* The CpxR regulates type VI secretion system 2 expression and facilitates the interbacterial competition activity and virulence of avian pathogenic *Escherichia coli*. Veterinary Research 2019;**50**(1):1–12.
86. Nagamatsu K, Hannan TJ, Guest RL, *et al.* Dysregulation of *Escherichia coli* α-hemolysin expression alters the course of acute and persistent urinary tract infection. Proc Natl Acad Sci U S A 2015;**112**(8):E871–80.
87. Ares MA, Abundes-Gallegos J, Panunzi LG, *et al.* The Coli surface antigen CS3 of enterotoxigenic *Escherichia coli* is differentially regulated by H-NS, CRP, and CpxRA global regulators. Frontiers in Microbiology 2019;**10**:1685
88. MacRitchie DM, Ward JD, Nevesinjac AZ, Raivio TL. Activation of the Cpx envelope stress response down-regulates expression of several locus of enterocyte effacement-encoded genes in enteropathogenic *Escherichia coli*. Infection and Immunity 2008;**76**(4):1465–75.

89. Jones CH, Danese PN, Pinkner JS, Silhavy TJ, Hultgren SJ. The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. EMBO J 1997;**16**(21):6394–406.
90. Lasaro M, Liu Z, Bishar R, *et al. Escherichia coli* isolate for studying colonization of the mouse intestine and its application to two-component signaling knockouts. Journal of Bacteriology 2014;**196**(9):1723–32.
91. Fujimoto M, Goto R, Haneda T, Okada N, Miki T. *Salmonella enterica* serovar Typhimurium CpxRA two-component system contributes to gut colonization in Salmonella induced colitis. Infection and Immunity 2018;**86**(7):e00280-18
92. Giannakopoulou N, Mendis N, Zhu L, Gruenheid S, Faucher SP, le Moual H. The virulence effect of CpxRA in *Citrobacter rodentium* is independent of the auxiliary proteins NlpE and CpxP. Frontiers in Cellular and Infection Microbiology 2018;**8**:320
93. Thomas CM, Nielsen KM. Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. Nature Reviews Microbiology 2005 3:9 2005;**3**(9):711–21.
94. von Wintersdorff CJH, Penders J, van Niekerk JM, *et al.* Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Frontiers in Microbiology 2016;**7**:173.
95. Fortier LC, Sekulovic O. Importance of prophages to evolution and virulence of bacterial pathogens. Virulence 2013;**4**(5):354–65.
96. Casjens SR, Hendrix RW. Bacteriophage lambda: Early pioneer and still relevant. Virology 2015;**479–480**:310–30.
97. Wagner PL, Waldor MK. Bacteriophage control of bacterial virulence. Infection and Immunity 2002;**70**(8):3985–93.
98. Cumby N, Davidson AR, Maxwell KL. The moron comes of age. 2012. Bacteriophage **2**(4): 225–228
99. Cumby N, Edwards AM, Davidson AR, Maxwell KL. The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. Journal of Bacteriology 2012;**194**(18):5012–9.
100. Yang J, Russell TW, Hocking DM, *et al.* Control of acid resistance pathways of enterohemorrhagic *Escherichia coli* strain EDl933 by PsrB, a prophage-encoded AraC-like regulator. Infection and Immunity 2015;**83**(1):346–53.

101. Li D, Tang F, Xue F, *et al.* Prophage phiv142-3 enhances the colonization and resistance to environmental stresses of avian pathogenic *Escherichia coli.* Veterinary Microbiology 2018;**218**:70–7.
102. Barondess JJ, Beckwfth J. A bacterial virulence determinant encoded by lysogenic coliphage λ. Nature 1990 346:6287 1990;**346**(6287):871–4.
103. Wang X, Kim Y, Ma Q, *et al.* Cryptic prophages help bacteria cope with adverse environments. Nature Communications 2010 1:1 2010;**1**(1):1–9.
104. Carey JN, Mettert EL, Fishman-Engel DR, Roggiani M, Kiley PJ, Goulian M. Phage integration alters the respiratory strategy of its host. Elife 2019;**8**:e49081
105. Rabinovich L, Sigal N, Borovok I, Nir-Paz R, Herskovits AA. Prophage excision activates listeria competence genes that promote phagosomal escape and virulence. Cell 2012;**150**(4):792–802.
106. Hernandez-Doria JD, Sperandio V. Bacteriophage Transcription Factor Cro Regulates Virulence Gene Expression in Enterohemorrhagic *Escherichia coli*. Cell Host and Microbe 2018;**23**(5):607-617.e6.
107. Pick K, Ju T, Willing BP, Raivio TL. Isolation and Characterization of a Novel Temperate *Escherichia coli* Bacteriophage, Kapi1, Which Modifies the O-Antigen and Contributes to the Competitiveness of Its Host during Colonization of the Murine Gastrointestinal Tract. MBio 2022;**13**(1):e02085-21
108. Newton GJ, Daniels C, Burrows LL, Kropinski AM, Clarke AJ, Lam JS. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. Molecular Microbiology 2001;**39**(5):1237–47.
109. Kuzio J, Kropinski AM. O-antigen conversion in *Pseudomonas aeruginosa* PAO1 by bacteriophage D3. Journal of Bacteriology 1983;**155**(1):203-212.
110. Susskind MM, Botstein D, Wright A. Superinfection exclusion by P22 prophage in lysogens of *Salmonella typhimurium*: III. Failure of superinfecting phage DNA to enter sieA+ lysogens. Virology 1974;**62**(2):350–66.
111. Hofer B, Ruge M, Dreiseikelmann B. The superinfection exclusion gene (sieA) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. Journal of Bacteriology 1995;**177**(11):3080–6.

112. Arguijo-Hernández ES, Hernandez-Sanchez J, Briones-Peña SJ, *et al.* Cor interacts with outer membrane proteins to exclude FhuA-dependent phages. Archives of Virology 2018;**163**(11):2959–69.
113. Uc-Mass A, Loeza EJ, de La Garza M, Guarneros G, Hernández-Sánchez J, Kameyama L. An orthologue of the cor gene is involved in the exclusion of temperate lambdoid phages. Evidence that Cor inactivates FhuA receptor functions. Virology 2004;**329**(2):425–33.
114. Rueggeberg KG, Toba FA, Bird JG, Franck N, Thompson MG, Hay AG. The lysis cassette of DLP12 defective prophage is regulated by RpoE. Microbiology (United Kingdom) 2015;**161**(8):1683–93.
115. Altuvia S, Storz G, Papenfort K. Cross-Regulation between Bacteria and Phages at a Posttranscriptional Level. Microbiology Spectrum 2018;**6**(4).
116. Fröhlich KS, Papenfort K. Interplay of regulatory RNAs and mobile genetic elements in enteric pathogens. Molecular Microbiology 2016;**101**(5):701–13.
117. Papenfort K, Espinosa E, Casadesús J, Vogel J. Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*. Proc Natl Acad Sci U S A 2015;**112**(34):E4772–81.
118. Raivio TL, Popkin DL, Silhavy TJ. The Cpx Envelope Stress Response Is Controlled by Amplification and Feedback Inhibition. Journal of Bacteriology 1999;**181**(17):5263-72.
119. Baba T, Ara T, Hasegawa M, *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. Molecular Systems Biology 2006;**2**:2006.0008
120. Silhavy T, Berman M, Enquist L. Experiments with gene fusions.1984. Cold Spring Harbor Laboratory.
121. Hoang TT, Karkhoff-Schweizer RR, Kutchma AJ, Schweizer HP. A broad-host-range F1p-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: Application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene 1998;**212**(1):77–86.
122. Wong JL, Vogt SL, Raivio TL. Using reporter genes and the *Escherichia coli* ASKA overexpression library in screens for regulators of the gram negative envelope stress response. Methods in Molecular Biology 2013;**966**:337–57.

123. Sievers F, Wilm A, Dineen D, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 2011;**7**:539.
124. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A Fast Phage Search Tool. Nucleic Acids Research 2011;**39**:W347-52.
125. Arndt D, Grant JR, Marcu A, *et al.* PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Research 2016;**44**:W16-W21.
126. Wang Z, Zhao C, Wang Y, Sun Z, Wang N. PANDA: Protein function prediction using domain architecture and affinity propagation. Scientific Reports 2018 8:1 2018;**8**(1):1–10.
127. Wintersinger JA, Wasmuth JD. Kablammo: an interactive, web-based BLAST results visualizer. Bioinformatics 2015;**31**(8):1305–6.
128. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology 1990;**215**(3):403–10.
129. Kitagawa M, Ara T, Arifuzzaman M, *et al.* Complete set of ORF clones of Escherichia coli ASKA library (A Complete Set of *E. coli* K-12 ORF Archive): Unique Resources for Biological Research. DNA Research 2005;**12**:291–9.
130. Dunstan RA, Hay ID, Lithgow T. Defining membrane protein localization by isopycnic density gradients. Methods in Molecular Biology 2017;**1615**:81–6.
131. Danese PN, Silhavy TJ. The sigma(E) and the Cpx signal transduction systems control the synthesis of periplasmic protein-folding enzymes in *Escherichia coli*. Genes and Development. 1997;**11**(9): 1183-93
132. Koontz L. TCA precipitation. Methods in Enzymology 2014;**541**:3–10.
133. Arndt C, Koristka S, Feldmann A, Bartsch H, Bachmann M. Coomassie-Brilliant Blue Staining of Polyacrylamide Gels. Methods in Molecular Biology 2012;**869**:465–9.
134. Keseler IM, Gama-Castro S, Mackie A, *et al.* The EcoCyc Database in 2021. Frontiers in Microbiology 2021;**12**:711077.
135. Ramisetty BCM, Sudhakari PA. Bacterial "grounded" prophages: Hotspots for genetic renovation and innovation. Frontiers in Genetics 2019;**10**(FEB):65.
136. Mutalik VK, Novichkov PS, Price MN, *et al.* Dual-barcoded shotgun expression library sequencing for high-throughput characterization of functional traits in bacteria. Nature Communications 2019;**10**(1):308.

137. Babu M, Díaz-Mejía JJ, Vlasblom J, *et al.* Genetic interaction maps in *Escherichia coli* reveal functional crosstalk among cell envelope biogenesis pathways. PLoS Genetics 2011;**7**(11):e1002377
138. Škunca N, Bošnjak M, Kriško A, *et al.* Phyletic Profiling with Cliques of Orthologs Is Enhanced by Signatures of Paralogy Relationships. PLoS Computational Biology 2013;**9**(1):e1002852
139. Reis-Cunha JL, Bartholomeu DC, Manson AL, Earl AM, Cerqueira GC. ProphET, prophage estimation tool: A stand-alone prophage sequence prediction tool with self-updating reference database. PLOS ONE 2019;**14**(10):e0223364.
140. Kropinski AM, Kovalyova I v., Billington SJ, *et al.* The genome of ε15, a serotype-converting, Group E1 *Salmonella enterica*-specific bacteriophage. Virology 2007;**369**(2):234–44.
141. Dudek CA, Jahn D. PRODORIC: state-of-the-art database of prokaryotic gene regulation. Nucleic Acids Research 2022;**50**(D1):D295–302.
142. Ansaldi M, Théraulaz L, Méjean V. TorI, a response regulator inhibitor of phage origin in *Escherichia coli*. Proc Natl Acad Sci U S A 2004;**101**(25):9423-8.
143. Jubelin G, Vianney A, Beloin C, *et al.* CpxR/OmpR Interplay Regulates Curli Gene Expression in Response to Osmolarity in *Escherichia coli*. Journal of Bacteriology 2005;**187**(6):2038-2049.
144. Prigent-Combaret C, Brombacher E, Vidal O, *et al.* Complex Regulatory Network Controls Initial Adhesion and Biofilm Formation in *Escherichia coli* via Regulation of the csgD Gene. Journal of Bacteriology 2001;**183**(24):7213-23.
145. Ranquet C, Ollagnier-de-Choudens S, Loiseau L, Barras F, Fontecave M. Cobalt stress in *Escherichia coli*: The effect on the iron-sulfur proteins. Journal of Biological Chemistry 2007;**282**(42):30442–51.
146. Majtan T, Frerman FE, Kraus JP. Effect of cobalt on *Escherichia coli* metabolism and metalloporphyrin formation. Biometals 2011;**24**(2):335-347.
147. Tsviklist V, Guest RL, Raivio TL. The Cpx Stress Response Regulates Turnover of Respiratory Chain Proteins at the Inner Membrane of *Escherichia coli*. Frontiers in Microbiology 2022;**12**:732288.

148. Maurer LM, Yohannes E, Bondurant SS, Radmacher M, Slonczewski JL. pH Regulates Genes for Flagellar Motility, Catabolism, and Oxidative Stress in *Escherichia coli* K-12. Journal of Bacteriology 2005;**187**(1):304-319.
149. Chao Y, Vogel J. A 3′ UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. Molecular Cell 2016;**61**(3):352–63.
150. Grabowicz M, Koren D, Silhavy TJ. The Cpxq sRNA negatively regulates skp to prevent mistargeting of β-barrel outer membrane proteins into the cytoplasmic membrane. MBio 2016;**7**(2):e00312-16
151. Teufel F, Almagro Armenteros JJ, Johansen AR, *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. Nature Biotechnology 2022 2022:1–3.
152. Golais F, Hollý J, Vítkovská J. Coevolution of bacteria and their viruses. Folia Microbiologica 2013;**58**(3):177–86.
153. Schröder W, van de Putte P. Genetic study of prophage excision with a temperature inducible mutant of Mu-1. Molecular and General Genetics MGG 1974 130:2 1974;**130**(2):99–104.
154. Jancheva M, Böttcher T. A Metabolite of *Pseudomonas* Triggers Prophage-Selective Lysogenic to Lytic Conversion in *Staphylococcus aureus*. Journal of the American Chemical Society 2021;**143**(22):8344–51.
155. Boling L, Cuevas DA, Grasis JA, *et al.* Dietary prophage inducers and antimicrobials: toward landscaping the human gut microbiome. Gut Microbes 2020;**11**(4):721–34.
156. Christie GE, Calendar R. Bacteriophage P2. Bacteriophage 2016;**6**(1):e1145782.
157. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption. FEMS Microbiology Letters 2016;**363**(4):2:fnw002
158. Kim M, Ryu S. Antirepression System Associated with the Life Cycle Switch in the Temperate Podoviridae Phage SPC32H. Journal of Virology 2013;**87**(21):11775-86.
159. Lerouge I, Vanderleyden J. O-antigen structural variation: mechanisms and possible roles in animal/plant–microbe interactions. FEMS Microbiology Reviews 2002;**26**(1):17–47.
160. Iguchi A, Thomson NR, Ogura Y, *et al.* Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. Journal of Bacteriology 2009;**91**(1):347–54.

161. Dion MB, Oechslin F, Moineau S. Phage diversity, genomics and phylogeny. Nature Reviews Microbiology 2020 18:3 2020;**18**(3):125–38.
162. Grose JH, Casjens SR. Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family Enterobacteriaceae. Virology 2014;**468–470**:421–43.
163. Pope WH, Bowman CA, Russell DA, *et al.* Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife 2015;**4**:e06416
164. Ecale Zhou CL, Malfatti S, Kimbrel J, *et al.* multiPhATE: bioinformatics pipeline for functional annotation of phage isolates. Bioinformatics 2019;**35**(21):4402–4.
165. Salzberg SL, Deicher AL, Kasif S, White O. Microbial gene identification using interpolated Markov models. Nucleic Acids Research 1998;**26**(2):544–8.
166. de la Cruz MA, Morgan JK, Ares MA, Yáñez-Santos JA, Riordan JT, Girón JA. The Two-Component System CpxRA Negatively Regulates the Locus of Enterocyte Effacement of Enterohemorrhagic *Escherichia coli* Involving σ32 and Lon protease. Frontiers in Cellular and Infection Microbiology 2016;**6**(FEB):11.
167. Bingle LEH, Constantinidou C, Shaw RK, *et al.* Microarray Analysis of the Ler Regulon in Enteropathogenic and Enterohaemorrhagic *Escherichia coli* Strains. PLoS ONE 2014;**9**(1):e80160.
168. Lasaro M, Liu Z, Bishar R, *et al.* Escherichia coli isolate for studying colonization of the mouse intestine and its application to two-component signaling knockouts. Journal of Bacteriology 2014;**196**(9):1723–32.
169. Picard B, Garcia JS, Gouriou S, *et al.* The Link between Phylogeny and Virulence in *Escherichia coli* Extraintestinal Infection. Infection and Immunity 1999;**67**(2):546-53.
170. Boyd EF, Hartl DL. Chromosomal Regions Specific to Pathogenic Isolates of *Escherichia coli* Have a Phylogenetically Clustered Distribution. Journal of Bacteriology 1998;**180**(5):1159-65.
171. Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of pathogenic *Escherichia coli*. Nature Reviews Microbiology 2020 19:1 2020;**19**(1):37–54.
172. Tenaillon O, Skurnik D, Picard B, Denamur E. The population genetics of commensal *Escherichia coli*. Nature Reviews Microbiology 2010 8:3 2010;**8**(3):207–17.

173. Kovács K, Hurst LD, Papp B. Stochasticity in Protein Levels Drives Colinearity of Gene Order in Metabolic Operons of *Escherichia coli*. PLoS Biology 2009;**7**(5):e1000115.
174. de Smit MH, Verlaan PWG, van Duin J, Pleij CWA. Intracistronic transcriptional polarity enhances translational repression: a new role for Rho. Molecular Microbiology 2008;**69**(5):1278–89.
175. Ullmannt A, Josepht E, Danchint A. Cyclic AMP as a modulator of polarity in polycistronic transcriptional units (positive regulation/rho factor/lactose and galactose operons/catabolite repression). Biochemistry 1979;**76**(7):3194–7.
176. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D. ARNold: a web tool for the prediction of Rho-independent transcription terminators. RNA Biology 2011;**8**(1):11–3.
177. Huerta AM, Collado-Vides J. Sigma70 Promoters in *Escherichia coli*: Specific Transcription in Dense Regions of Overlapping Promoter-like Signals. Journal of Molecular Biology 2003;**333**(2):261–78.
178. Salgado H, Peralta-Gil M, Gama-Castro S, *et al.* RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Research 2013;**41**(D1):D203–13.
179. Gorochowski TE, Chelysheva I, Eriksen M, Nair P, Pedersen S, Ignatova Z. Absolute quantification of translational regulation and burden using combined sequencing approaches. Molecular Systems Biology 2019;**15**(5):e8719.
180. Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 2014;**157**(3):624-35.
181. Hall MP, Unch J, Binkowski BF, *et al.* Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chemical Biology 2012;**7**(11):1848–57.
182. Padan E, Bibi E, Ito M, Krulwich TA. Alkaline pH Homeostasis in Bacteria: New Insights. Biochimica et Biophysica Acta 2005;**1717**(2):67-88.
183. di Russo N v., Estrin DA, Martí MA, Roitberg AE. pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pKas: The Case of Nitrophorin 4. PLOS Computational Biology 2012;**8**(11):e1002761.

184. Hobbs EC, Yin X, Paul BJ, Astarita JL, Storz G. Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. Proc Natl Acad Sci U S A 2012;**109**(41):16696–701.
185. Du D, Wang Z, James NR, *et al.* Structure of the AcrAB-TolC multidrug efflux pump. Nature 2014;**509**(7501):512-5.
186. Du D, Neuberger A, Orr MW, *et al.* Interactions of a Bacterial RND Transporter with a Transmembrane Small Protein in a Lipid Environment. Structure 2020;**28**(6):625-634.e6.
187. Thorgersen MP, Downs DM. Cobalt Targets Multiple Metabolic Processes in *Salmonella enterica*. Journal of Bacteriology 2007;**189**(21):7774-81.
188. Rodrigue A, Effantin G, Mandrand-Berthelot MA. Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli. Journal of Bacteriology 2005;**187**(8):2912–6.
189. Niegowski D, Eshaghi S. The CorA family: structure and function revisited. Cellular and Molecular Life Sciences 2007;**64**(19–20):2564–74.
190. Grass G, Franke S, Taudte N, *et al.* The metal permease ZupT from *Escherichia coli* is a transporter with a broad substrate spectrum. Journal of Bacteriology 2005;**187**(5):1604–11.
191. Blériot C, Effantin G, Lagarde F, Mandrand-Berthelot MA, Rodrigue A. RcnB is a periplasmic protein essential for maintaining intracellular Ni and Co concentrations in *Escherichia coli*. Journal of Bacteriology 2011;**193**(15):3785–93.
192. Porcheron G, Garénaux A, Proulx J, Sabri M, Dozois CM. Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: Correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. Frontiers in Cellular and Infection Microbiology 2013;**3**(DEC):90.
193. Fantino JR, Py B, Fontecave M, Barras F. A genetic analysis of the response of *Escherichia coli* to cobalt stress. Environmental Microbiology 2010;**12**(10):2846–57.
194. Argüello JM, Raimunda D, Padilla-Benavides T. Mechanisms of copper homeostasis in bacteria. Frontiers in Cellular and Infection Microbiology 2013;**4**(NOV):73.
195. Hooper DC, Jacoby GA. Mechanisms of drug resistance: quinolone resistance. Ann. N. Y. Acad. Sci. 2015;**1354**(1):12–31.

196. Braun V, Hantke K. Lipoproteins: Structure, Function, Biosynthesis. Subcellular Biochemistry 2019;**92**:39–77.
197. Davies SJ, Golby P, Omrani D, *et al.* Inactivation and Regulation of the Aerobic C4-Dicarboxylate Transport (dctA) Gene of *Escherichia coli*. Journal of Bacteriology 1999;**181**(18):5624-5635.
198. Grenier F, Matteau D, Baby V, Rodrigue S. Complete Genome Sequence of *Escherichia coli* BW25113. Genome Announcements 2014;**2**(5):1038–52.
199. Ferenci T, Zhou Z, Betteridge T, *et al.* Genomic Sequencing Reveals Regulatory Mutations and Recombinational Events in the Widely Used MC4100 Lineage of *Escherichia coli* K-12. Journal of Bacteriology 2009;**191**(12):4025-9.
200. Rouviere PE, de las Penas A, Mecsas J, Lu CZ, Rudd KE, Gross CA. rpoE, the gene encoding the second heat-shock sigma factor, sigma E, in *Escherichia coli*. The EMBO Journal 1995;**14**(5):1032-1042.
201. Missiakas D, Mayer MP, Lemaire M, Georgopoulos C, Raina S. Modulation of the *Escherichia coli* sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. Molecular Microbiology 1997;**24**(2):355–71.
202. Pazos M, Peters K. Peptidoglycan. Subcellular Biochemistry 2019;**92**:127–68.
203. Bernal-Cabas M, Ayala JA, Raivio TL. The Cpx Envelope Stress Response Modifies Peptidoglycan Cross-Linking via the l,d-Transpeptidase LdtD and the Novel Protein YgaU. Journal of Bacteriology 2015;**197**(3):603-14.
204. Zhang S, Cheng Y, Ma J, Wang Y, Chang Z, Fu X. Degp degrades a wide range of substrate proteins in *Escherichia coli* under stress conditions. Biochemical Journal 2019;**476**(23):3549–64.
205. Delcour AH. Outer membrane permeability and antibiotic resistance. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 2009;**1794**(5):808–16.
206. Danese PN, Pratt LA, Dove SL, Kolter R. The outer membrane protein, Antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. Molecular Microbiology 2000;**37**(2):424–32.
207. Ferrières L, Clarke DJ. The RcsC sensor kinase is required for normal biofilm formation in *Escherichia coli* K-12 and controls the expression of a regulon in response to growth on a solid surface. Molecular Microbiology 2003;**50**(5):1665–82.

208. Markova JA, Anganova E v., Turskaya AL, Bybin VA, Savilov ED. Regulation of *Escherichia coli* Biofilm Formation (Review). Applied Biochemistry and Microbiology 2018;**54**(1):1-11.
209. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell 2016;**165**(3):535–50
210. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics 2012 13:4 2012;**13**(4):227–32.
211. Turnbough CL. Regulation of Bacterial Gene Expression by Transcription Attenuation. Microbiology and Molecular Biology Reviews 2019;**83**(3):e00019-19.
212. Walker JM, Gasteiger E, Hoogland C, *et al.* Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook 2005:571–607.
213. Hallgren J, Tsirigos KD, Damgaard Pedersen M, *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. BioRxiv 2022:2022.04.08.487609.
214. Reis-Cunha JL, Bartholomeu DC, Manson AL, Earl AM, Cerqueira GC. ProphET, prophage estimation tool: A stand-alone prophage sequence prediction tool with self-updating reference database. PLOS ONE 2019;**14**(10):e0223364.
215. Huerta AM, Collado-Vides J. Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. Journal of Molecular Biology 2003;**333**(2):261–78.
216. Carbon S, Douglass E, Good BM, *et al.* The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research 2021;**49**(D1):D325–34.
217. Ashburner M, Ball CA, Blake JA, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics 2000;**25**(1):25–9.
218. Casadaban MJ. Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. Journal of Molecular Biology 1976;**104**(3):541–55.
219. Baba T, Ara T, Hasegawa M, *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. Molecular Systems Biology 2006;**2**:2006.0008.

220. Levine MM, Nalin DR, Hornick RB, *et al. Escherichia coli* Strains That Cause Diarrhœa But Do Not Produce Heat-Labile Or Heat-Stable Enterotoxins And Are Non-Invasive. The Lancet 1978;**311**(8074):1119–22.
221. Nevesinjac AZ, Raivio TL. The Cpx envelope stress response affects expression of the type IV bundle-forming pili of enteropathogenic *Escherichia coli.* Journal of Bacteriology 2005;**187**(2):672–86.
222. Raivio TL, Silhavy TJ. Transduction of envelope stress in *Escherichia coli* by the Cpx two-component system. Journal of Bacteriology 1997;**179**(24):7724-33.
223. Hull RA, Gill RE, Hsu P, Minshew BH, Falkow S. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. Infection and Immunity 1981;**33**(3):933-8.
224. Roberts JA, Marklund BI, Ilver D, *et al.* The Gal(α1-4)Gal-specific tip adhesin of *Escherichia coli* P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. Proc Natl Acad Sci U S A 1994;**91**(25):11889–93.
225. Jensen KF. The *Escherichia coli* K-12 “wild types” W3110 and MG1655 have an rph frameshift mutation that leads to pyrimidine starvation due to low pyrE expression levels. Journal of Bacteriology 1993;**175**(11):3401–7.
226. Bachmann BJ. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriological Reviews 1972;**36**(4):525-557.
227. Miller VL, Mekalanos JJ. 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. Journal of Bacteriology 1988;**170**(6):2575-83.
228. Amann E, Ochs B, Abel KJ. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. Gene 1988;**69**(2):301–15.