

Electronic Thesis and Dissertation Repository

10-1-2021 1:00 PM

Disruption of Insertion Sequence 200 (IS200) leads to a premature induction of the Cysteine regulon in *Salmonella Typhimurium*

Naomi-Jean Q. Scherba, *The University of Western Ontario*

Supervisor: Haniford, David B., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry

© Naomi-Jean Q. Scherba 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Bacteria Commons](#)

Recommended Citation

Scherba, Naomi-Jean Q., "Disruption of Insertion Sequence 200 (IS200) leads to a premature induction of the Cysteine regulon in *Salmonella Typhimurium*" (2021). *Electronic Thesis and Dissertation Repository*. 8158.

<https://ir.lib.uwo.ca/etd/8158>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Salmonella Typhimurium is a leading contributor to non-typhoid diseases with complex regulatory networks that are key to understanding its pathogenicity and virulence. I explore the role of the sRNA IS200, where its deletion in *Salmonella* leads to premature induction of the Cys regulon. The premature induction in Δ IS200 was validated through qRT-PCR and GFP transcriptional fusions and occurs at late exponential to early stationary phase. In addition, Δ IS200 leads to increased sensitivity to oxidative stress. The focus of this work has been to understand how Δ IS200 leads to premature Cys regulon induction, and I present three models: the futile import/export, the metabolic burden, and LrhA deprivation models. I conclude that the third model, which proposes a novel function for the global regulator LrhA, provides a promising direction for future research. Understanding this problem will help uncover how IS200 has integrated itself into the metabolic machinery of a medically important pathogen.

Keywords

Salmonella Typhimurium, cysteine biosynthesis, insertion sequence 200, sRNA, LrhA

Summary for Lay Audience

Pathogenesis and virulence are what allow bacteria to infect, persist and spread within populations. Proteins are functioning units in the cell to carry out jobs, and these proteins are regulated at multiple steps throughout their production. For example, transcription is the conversion of reading a DNA template or gene and copy it to messenger RNA (mRNA), and translation is turning that mRNA into a protein. Small RNAs (sRNA) will regulate post-transcriptionally, and the sRNA of interest in this project is the insertion sequence 200 (IS200).

Salmonella Typhimurium is a pathogenetic bacterium and a significant contributor to non-typhoid gastrointestinal diseases. Our lab has previously discovered that IS200 can regulate genes in its host *Salmonella*. When IS200 is deleted from the *Salmonella* strain, it causes the expression of the pathways responsible for pathogenesis to increase. Another consequence of deleting IS200 is that a pathway for creating cysteine, an important protein building block, is turned on earlier than expected. In this project, I validate and characterize this premature turn-on of the cysteine pathway in an IS200 deletion strain. When IS200 is removed, I show that the *Salmonella* strain is at a disadvantage when exposed to oxidative stress, which comes from harmful reactive oxygen molecules.

In addition to validating and characterizing the turning on of the cysteine pathway in *Salmonella* with IS200 deleted, I developed and tested three models to explain how IS200 might be causing this premature turn-on. Through a series of experiments, I show that the most promising of the three models I present involved a protein called LrhA. This protein regulates pathways involved in pathogenesis by preventing their expression, and I provide evidence that LrhA could be involved in regulating the cysteine pathway. My results imply that this is a novel function of LrhA and in *Salmonella* with IS200 deleted, LrhA could influence the cysteine pathway via direct and indirect regulation.

Co-Authorship Statement

Dr. David Haniford and Ryan Trussler constructed all strains except for DBH671 and DBH672 used in the experiments. In addition, Ryan Trussler cloned the functional pACYC184 *cysD* and *ydjN* GFP plasmids and the pBAD-LrhA-FLAG plasmid.

Acknowledgements

First and foremost, I would like to thank Dr. David Haniford for his guidance, encouragement, and support during this project. Thank you for being a great supervisor and pushing me to think critically and question beyond initial findings.

I would also like to thank Ryan Trussler for his vast knowledge, troubleshooting expertise and for being a soundboard for ideas.

I want to thank my advisory committee Dr. David Edgell and Dr. Brian Shilton, for their guidance and feedback.

Finally, I would like to thank my mom, whose support throughout my life has led me to where I am today, and to Randal, whose patience and love got me through a pandemic. Thank you to all my friends and family who have supported me throughout this adventure for the past two years.

Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Co-Authorship Statement.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	ix
List of Supplemental Tables.....	xi
List of Supplemental Figures.....	xii
List of Abbreviations.....	xiii
1 Introduction.....	1
1.1 <i>Salmonella</i> Typhimurium pathogenicity.....	1
1.2 Sulphur assimilation and cysteine metabolism.....	4
1.3 Small non-coding RNA and IS200.....	8
1.4 Objective and models.....	10
2 Materials and Methods.....	13
2.1 Strains, plasmids and growth conditions.....	13
2.2 RNA isolation and qRT-PCR.....	14
2.3 Reverse-Transcription PCR.....	14
2.4 Oxidative stress.....	15
2.5 Chromatin immunoprecipitation.....	15
2.6 Growth curves.....	16
2.7 Western blot.....	17
2.8 Glutathione determination.....	17

3	Results.....	19
3.1	Validating the premature induction of the Cys regulon in <i>Salmonella</i> SL1344 with a disruption of IS200	19
3.2	Providing exogenous cystine or disruption of CysB represses the Cys regulon induction in the Δ IS200 strain.....	23
3.3	Cysteine biosynthesis and oxidative stress	26
3.4	Model 1 – the futile import/export model.....	28
3.5	Model 2 – the metabolic burden model	37
3.6	Model 3 – the LrhA deprivation model	45
4	Discussion.....	51
4.1	Δ IS200 leads to a premature induction of the Cys regulon	52
4.2	IS200 is predicted to be a critical component of <i>Salmonella</i> fitness during macrophage invasion.....	53
4.3	Upregulation in the flagellar pathway contributes to <i>fliY</i> expression in Δ IS200.....	54
4.4	Preliminary evidence points to a connection between the flagellar pathway and cysteine metabolism.....	57
4.5	LrhA binds to multiple Cys regulon genes	59
4.6	LrhA could extend its control over the Cys regulon through multiple mechanisms	60
5	Conclusion	61
	Reference	63
	Appendix.....	71
	Supplemental Figures.....	71
	Supplemental Tables.....	77
	Curriculum Vitae	91

List of Tables

Table 1. Normalized signal of LrhA-DNA targets.	48
--	----

List of Figures

Figure 1. Overview of flagellar and invasion pathways.	3
Figure 2. Sulphur assimilation pathways and cystine transport, Cys regulon regulation, and FliA operon.	7
Figure 3. Simplified IS200 element.	9
Figure 4. Removal of 5' end of IS200 leads to Cys regulon induction at LE phase.	20
Figure 5. GFP monitoring of <i>cysD</i> and <i>ydjN</i> in <i>Salmonella</i> SL1344 WT and Δ IS200 strains grown in rich media.	22
Figure 6. Impact of adding exogenous cystine on the expression of select Cys regulon genes at LE phase.	24
Figure 7. Impact of <i>cysB</i> deletion on the Cys regulon induction during growth.	26
Figure 8. <i>Salmonella</i> WT and Δ IS200 exposure to oxidative stress.	28
Figure 9. The futile import/export model.	29
Figure 10. Transcript mapping downstream of the FliA operon.	30
Figure 11. Transcript mapping downstream of <i>fliY</i>	31
Figure 12. Western blot analysis to measure FliY and YecC levels.	32
Figure 13. Impact of <i>fliY</i> deletion on Cys regulon induction.	34
Figure 14. GFP monitoring of <i>cysD</i> and <i>ydjN</i> in WT and Δ IS200 strains containing Δ <i>fliY</i> grown in rich media.	36
Figure 15. The metabolic burden model.	37
Figure 16. Growth of <i>Salmonella</i> SL1344 WT and Δ IS200 in rich media.	38
Figure 17. Glutathione levels in <i>Salmonella</i> SL1344 WT and Δ IS200 with and without Δ <i>hilA</i> at LE phase.	40
Figure 18. Removal of <i>hilA</i> on <i>cysD</i> and <i>ydjN</i> expression via GFP monitoring in <i>Salmonella</i> SL1344 WT and Δ IS200 strains grown in rich media.	42
Figure 19. Removal of <i>flhDC</i> on <i>cysD</i> and <i>ydjN</i> expression via GFP monitoring in <i>Salmonella</i> SL1344 WT and Δ IS200 strains in rich media.	44
Figure 20. The LrhA deprivation model.	46
Figure 21. ChIP analysis of <i>Salmonella</i> SL1344 WT showing LrhA binding to select Cys regulon genes at EE, ME, and ES phase.	48

Figure 22. GFP monitoring of *cysD* and *yjN* expression in *Salmonella* SL1344 WT in the presence of LrhA overexpression. 50

List of Supplemental Tables

Supplemental Table 1. <i>Salmonella</i> strains used in this study.	77
Supplemental Table 2. Plasmids used in this study.	77
Supplemental Table 3. Oligonucleotides used in this study.	78
Supplemental Table 4. 165 annotated hits from unpublished RNA-seq data at early exponential phase.	80
Supplemental Table 5. 63 annotated hits from unpublished RNA-Seq data at mid-exponential phase.	85
Supplemental Table 6. 95 annotated hits from unpublished RNA-Seq data at late exponential phase.	87

List of Supplemental Figures

Supplemental Figure 1. Schematic overview of pACYC184 and pDH1089.....	71
Supplemental Figure 2. Growth curve analysis of <i>Salmonella</i> SL1344 WT.....	71
Supplemental Figure 3. Growth curves of <i>cysD</i> and <i>ydjN</i> GFP transcriptional fusions in <i>Salmonella</i> WT and Δ IS200.....	72
Supplemental Figure 4. Growth curve for <i>cysD</i> and <i>ydjN</i> GFP transcriptional fusions in Δ <i>fliY</i> backgrounds.	73
Supplemental Figure 5. Growth curves for <i>cysD</i> and <i>ydjN</i> GFP transcriptional fusions in Δ <i>hilA</i> and Δ <i>flhDC</i> backgrounds.	74
Supplemental Figure 6. Growth curve for overexpression of LrhA with <i>cysD</i> and <i>ydjN</i> GFP transcriptional fusions in <i>Salmonella</i> SL1344 WT.....	75
Supplemental Figure 7. γ -glutamyl cycle as a source of cysteine.....	76

List of Abbreviations

ABC	ATP-binding cassette
BSA	bovine serum albumin
CFU	colony-forming unit
DTNB	5,5'-dithiobis(2-nitro-benzoic acid)
EDTA	ethylenediaminetetraacetic acid
GFP	green fluorescent protein
HTH	helix-turn-helix
LTTR	LysR-type transcriptional regulator
NAS	N-acetyl serine
NBD	nucleotide-binding domain
OAS	O-acetyl serine
PBS	phosphate-buffered saline
PVDF	polyvinylidene difluoride
ROS	reactive oxygen species
SCV	<i>Salmonella</i> -containing vacuole
SDS	sodium dodecyl sulphate
SPI	<i>Salmonella</i> pathogenicity island
T3SS	type 3 secretion system
TAE	tris acetic acid EDTA
TBS	tris-buffered saline
TBST	tris-buffered saline Tween 20
TE	tris EDTA
TMD	transmembrane domain
TND	5-thio-2-nitrobenzoic acid

1 Introduction

1.1 *Salmonella* Typhimurium pathogenicity

Non-typhoid salmonella is a significant contributor to foodborne gastrointestinal diseases worldwide; a large portion of these cases are caused by *Salmonella enterica* subsp. *enterica* serovar Typhimurium (hereafter *Salmonella*). A global burden of disease report stated that in 2017 there were 95.1 million cases of enterocolitis (infection and inflammation of the digestive and colon), 535 000 cases of disease, and over 50 thousand deaths due to non-typhoid *Salmonella*¹. The largest population susceptible to *Salmonella* is South Africa due to a lack of clean water and malnutrition. Additionally, individuals with HIV, malaria, and sickle cell disease are at high risk¹.

Salmonella must bypass the many facets of the mammalian immune system. *Salmonella* can spread through food- and waterborne sources and from animals to people, and people to people—the prevalence of disease cases linked to *Salmonella* increases in warmer climates. *Salmonella* will invade the epithelial layer once exposed to the gut and digestive system by either passive diffusion facilitated through dendritic cells (phagocytotic accessory cells to the mammalian immune system) or active transport^{2,3}. Active transport primarily depends on the Type 3 Secretion System (T3SS) encoded in the *Salmonella* Pathogenicity Island 1 (SPI-1) and causes a ‘trigger’ mechanism for the host cell to engulf the bacterium^{4,5}. There are five different SPIs in *Salmonella* with more discovered in the typhoid strains^{6,7}. SPI-1 is involved in the invasion process⁸, SPI-2 encodes a secondary T3SS essential for survival post-infection⁹, SPI-3 is involved in intramacrophage survival³, SPI-4 is involved in epithelial adhesion¹⁰, and SPI-5 encodes for proteins required by SPI-1 and SPI-2^{11,12}. SPI-3, SPI-4, and SPI-5 are less characterized, and their functions and regulation are not entirely understood. Once *Salmonellae* are passed the epithelial layer, they are phagocytized by macrophages and form *Salmonella* containing vacuoles (SCV) which allow them to evade degradation and fusion with lysosomes through the expression of effector proteins^{3,5,12,13}. *Salmonella* will express and eject these effector proteins via a secondary T3SS encoded in SPI-2 and inject these effector proteins into the host cytoplasm from within the SCV. While they express SPI-2 encoded genes, which have been found to be essential in murine models for macrophage survival¹⁴, they can proliferate

within the SCV⁸. The macrophage will then undergo apoptosis, exposing the newly proliferated *Salmonella* and allowing them to reinfect the epithelial layer or other macrophages leading to the persistence and spread of infection.

In addition to the T3SSs encoded within SPI-1 and SPI-2, motility is a key determinant in successful host invasion and proliferation^{15–18}. Achouri *et al.* show that non-motile *Salmonella* has significantly lower uptake in macrophages as both motility and contact frequency are shown to be key factors¹⁹. The reduced uptake in macrophages was found in both $\Delta motAB$ (non-motile where flagella are produced but paralyzed) and $\Delta fliOPQR$ (non-motile where it lacks flagella); therefore, the presence of flagella is not as important as the ability to be motile¹⁹. The researchers have shown that even a cell biased towards the tumbling motion during swimming had a lower frequency of infection, whereas a cell biased towards running, or smooth swimming, caused an increased uptake of *Salmonella*. Centrifugation restored *Salmonella* uptake in macrophages for both $\Delta fliOPQR$ and $\Delta motAB$ non-motile strains but did not alter the uptake of cells that were biased for tumbling¹⁹. Flagella are also important in establishing biofilm formation and biofilm formation is a key determinant in persistent infection^{12,20,21}. Biofilm production occurs in an organized manner through the initial attachment, formation of microcolonies, and maturation. Flagella are involved in the initial adhesion and generate the force for cell contact. In some bacteria, flagella are crucial for biofilm maturation^{20,22}.

Given the importance of SPI-1 and flagellar pathways in infectivity and virulence, it is unsurprising that these pathways involve complex regulatory cascades. The T3SS encoded in SPI-1 is controlled in a feed-forward manner where there is a hierarchy of transcription factors that control the expression of downstream effector genes. For example, HilD is the transcription factor at the top of this cascade²³. It receives sensory inputs from its environment to activate HilA (a downstream transcription factor), and from there, HilA activates InvF, which encodes several effector proteins involved in the T3SS (Figure 1)^{24,25}. Similarly, flagellar production is regulated in a feed-forward manner with the heterohexamer FlhD₄C₂ transcription factor, a class 1 gene product, at the top of the hierarchy and the master regulator^{15,26–29}. It will then proceed to activate downstream class 2 genes. An important class 2 gene product is FliA, or sigma 28 factor, which is responsible for activating class 3 flagellar genes in addition to its operon^{30–32}.

fliA is grouped with *fliZ* in its class 2 operon, the latter encoding the FliZ protein that post-translationally controls HilD protein activity and thus providing one of the several links between invasion and flagellar pathways^{33–35}. An additional link between the two can be seen through the activation of *flhDC* transcription by HilD (Figure 1)^{33,35,36}.

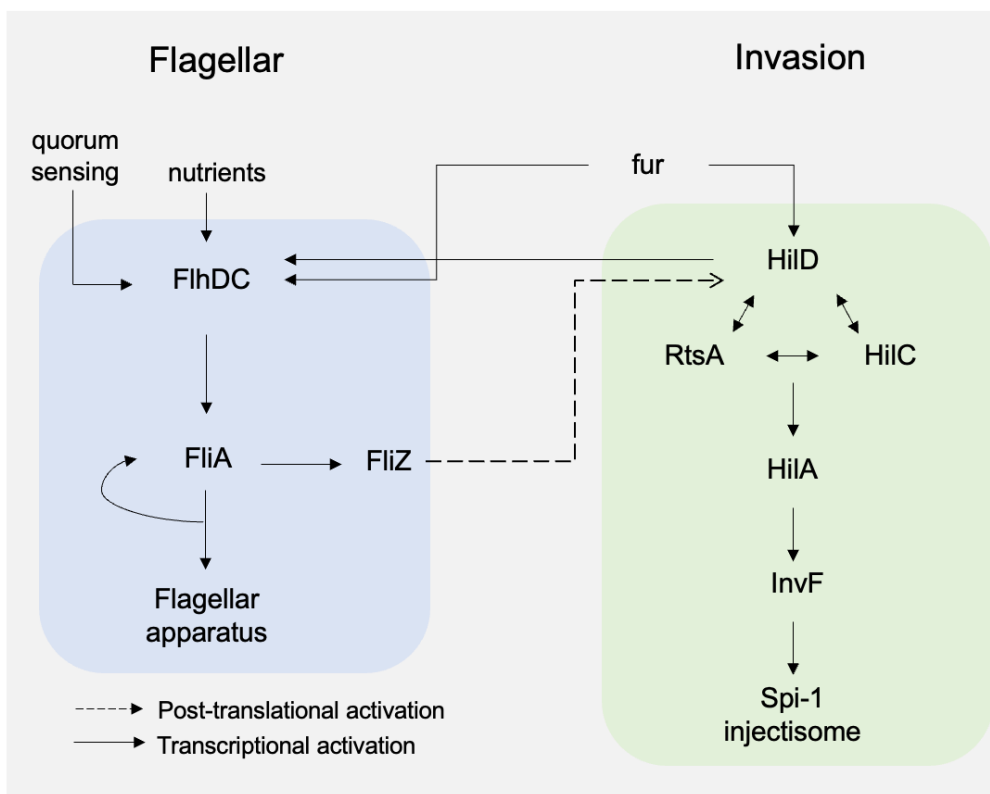


Figure 1. Overview of flagellar and invasion pathways.

Simplified regulatory flagellar and invasion cascades are depicted with FlhDC and HilD at the top of the respective cascades where they work in a feed-forward manner to regulate downstream transcription factors.

A lesser understood but important regulator involved in the flagellar pathway is LrhA, a LysR-type transcriptional regulator (LTTR). Early experimentation in *Escherichia coli* shows that in a *lrhA*⁻ strain, there is an increase in over 40 genes at the transcription level anywhere from 3- to 80-fold compared to WT levels³⁷. These 40 genes had varying roles in flagellar, chemotaxis, and motility pathways. In exploring the role of LrhA, it was discovered to bind directly to *flhD* but not its downstream flagellar genes (e.g. *fliA*, *fliC*, *trg*), and the LrhA LTTR-binding consensus sequence was found to be -105 to -117 nucleotides upstream of the *flhD* +1 transcriptional start

site. LrhA was also shown to bind to *lrhA* and positively autoregulates its transcription³⁷. Due to its inability to bind to downstream flagellar genes, LrhA is a regulator of flagellar, motility and chemotaxis pathways via repression of *flhDC*. Additionally, LrhA has been found to repress RpoS translation through Hfq-mediated sRNA pairing³⁸. *rpoS* encodes for a sigma factor 38 and is used for signalling stress and transitioning the cell into stationary phase and is part of biofilm formation^{20,39}. Being a part of the LTTR family of regulators, which is the most abundant family of proteins with over 800 members based on amino acid sequence, they contain a conserved helix-turn-helix (HTH) domain of 20-90 amino acids from the N-terminus⁴⁰⁻⁴². Traditionally, if a transcription factor is an activator or repressor, the HTH domain is found at the C- or N-terminus, respectively. In contrast, the LTTR-HTH domain is found at the N-terminus regardless of type. Additionally, LTTRs have a conserved co-inducer binding domain that can be found at the C-terminus. There is currently no known co-inducer for LrhA⁴⁰.

1.2 Sulphur assimilation and cysteine metabolism

Sulphur, which is essential for life, is predominantly stored in bacteria in the form of cysteine in cytoplasmic pools. The majority of intracellular cysteine is found in the reduced form, the oxidized form being cystine, and the reduction of cystine to cysteine is dependent on glutathione levels^{43,44}. Cysteine production is derived from sulphur assimilation pathways in which inorganic ions such as sulphate and thiosulphate donate sulphur atoms to N-acetyl serine (NAS) which then gets converted into cysteine (Figure 2A). In *E. coli* and *Salmonella*, the expression of the genes in the sulphur assimilation pathway and the cystine transporters is regulated by the transcriptional activator CysB⁴⁵⁻⁴⁸. Similar to LrhA, CysB is another LTTR and its known co-inducers is O-acetyl serine (OAS) or its isomer NAS. In contrast, sulphide and thiosulphate are anti-inducers switching CysB to its repressive state^{40,49-51}. It autoregulates by binding to its promoter and preventing transcription (Figure 2B)⁵². Under conditions of low cysteine and sulphur, CysB is activated by binding OAS in the C-terminal co-inducer binding domain, CysB can bind to the promoter region in the absence of a co-inducer but CysB-dependent transcription requires the binding of acetyl serine^{51,52}. The binding of OAS to CysB allows for *cysB* transcription^{45,47,49,52}; CysB-regulated genes are referred to collectively as components of the Cys regulon⁴⁸.

YdjN, a component of the Cys regulon, is considered the major cystine transporter in *E. coli*, transporting cystine from the periplasm into the cytoplasm where it is rapidly converted to cysteine^{43,45,46,53}. Importantly, cystine transport by YdjN is not subject to negative feedback control⁴³. When cells are grown in a poor sulphur source such as sulphate, YdjN expression is upregulated. Subsequent addition of cystine to the media causes unregulated import of cystine and thus high levels of cysteine in the cytoplasm, which are toxic to the cell. The cell limits intracellular cysteine as it can contribute to the Fenton reaction (equations 1 and 2)^{43,54}. These reactions are typically dealt with through DNA repair enzymes and lead to minimal damage, but in the presence of high cysteine levels, this can lead to the continuous production of hydroxyl radicals for the cell to deal with, thus a larger amount of DNA damage.



As levels higher than 1 mM of cysteine are toxic⁵⁵, the cell's survival depends on the rapid conversion of cysteine to other products (such as methionine, glutathione, ammonia, pyruvate, and hydrogen sulphide) and the transport of cysteine out of the cells. This phenomenon is referred to as the futile import/reduction/export cycle⁴³. In addition to YdjN, *E. coli* and *Salmonella* possess a second cystine transport system in the form of an ATP-binding cassette (ABC) transporter. This transporter consists of the periplasmic binding protein (FliY) and transport proteins YecS, the transmembrane domain (TMD), and YecC, the nucleotide-binding domain (NBD), that catalyzes the cleavage of phosphate from ATP^{45,46,56}. Unlike YdjN, which is a Na⁺/H⁺ symporter, FliY-YecSC relies on ATP for its transport. ABC-transporters are widely used among prokaryotic and eukaryotic populations to move substrates across their concentration gradients^{57,58}. ABC-transporters all share a commonality in their structure, which the basis comprises of two TMD that shift between inward and outward conformations upon the binding of the substrate and hydrolysis of ATP, and two NBD to hydrolyze the cleavage of ATP^{56,59}. The binding of the substrate is facilitated through a periplasmic-binding protein, and the periplasmic-binding protein binds to the TMD, therefore determining the transporter's specificity and affinity for specific molecules. It has been shown in *E. coli* that the FliY-YecSC

transport system has a higher affinity for cystine and is energetically favourable in conditions with low substrate levels compared to YdjN⁴³. *fliY* is directly downstream of the *fliAZ* operon with the arrangement of *fliY*, *yecS*, and *yecC* seen in Figure 2C. Although *fliY* transcription is subject to CysB regulation, *yecSC* is not, and their expression is possibly controlled through the CysB-independent *yedO* promoter that lies between *fliY* and *yecSC*^{30,43,46}. YedO encodes cysteine desulphhydrase, which takes cysteine and water and breaks it down into ammonium, pyruvate, and hydrogen sulphide⁶⁰. It is currently unknown if upregulation of one or more components of the FliY-YecSC cystine transporter also drives the futile import/reduction/export cycle. However, given the lack of feedback inhibition seen with YdjN, one could question if a similar situation would occur with the upregulation of the secondary transport system.

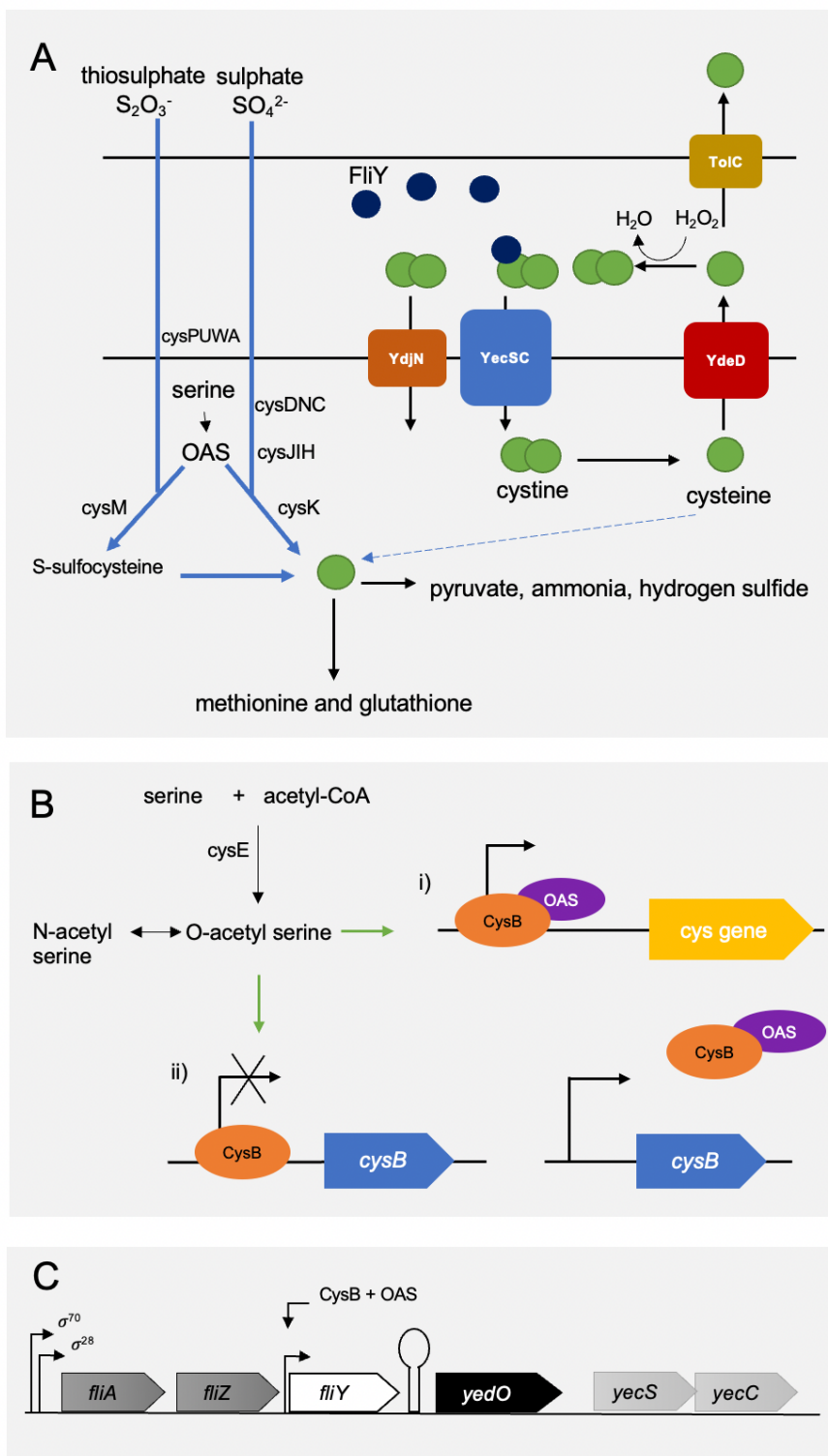


Figure 2. Sulphur assimilation pathways and cystine transport, Cys regulon regulation, and FliA operon.

(A) Cysteine biosynthesis via inorganic sulphur sources and transport of cystine into the cytoplasm. YdjN is the primary transporter of cystine whereas YecSC utilizes a periplasmic-binding protein, FliY, and ATP. (B) Genes that fall under *cysB* regulation in the cysteine regulon require acetyl serine, either N- or O- produced by serine acetyltransferase, for CysB to bind to the upstream promoter region (seen in i)). CysB also negatively autoregulates (seen in ii), where the addition of OAS binds to CysB and preventing it from binding to its promoter, allowing for transcription of *cysB*. (C) The gene structure of the FliA operon is shown along with the genes immediately downstream. FliA is under FlhDC (sigma 70) and its own (sigma 28) transcriptional regulation whereas FliY possesses its promoter under CysB control.

1.3 Small non-coding RNA and IS200

Small non-coding RNAs (sRNA) have been shown to provide important regulatory control for many pathways such as metabolism and virulence⁶¹⁻⁶³. The insertion sequence 200 (IS200) is a small dormant transposon roughly 700 base pairs in size. It is conserved throughout enterica bacteria with orthologs sharing >90% sequence identity^{64,65}. IS200 produces little to none of its transposase protein, TnpA, despite its high conservation and low transcription levels being maintained⁶⁵. At the 5' end, two products are encoded from IS200: the *tnpA* mRNA and its anti-sense RNA (*art200*), where anti-sense is in reference to the *tnpA* transcript (Figure 3A)^{64,66}. The pathogenic SL1344 strain of *Salmonella* encodes 7 copies of IS200. Given that the IS200 is highly conserved yet dormant, it points to a possible selective pressure from the bacterial host to maintain it. sRNA contributes to regulatory cascades through sRNA-mRNA base-pairing to influence translation and/or transcript stability, typically facilitated through a chaperone such as Hfq or ProQ⁶⁷⁻⁷⁰. For example, the sRNA Spot42 (encoded by *spf*) is regulated through the global regulator CRP and base-pairs to the 3' UTR of *hilD* to positively regulate *hilD* transcription and stability⁷¹. Previously, our lab has proposed that IS200 has been maintained in bacterial populations because both/either IS200-encoded RNAs (*tnpA* or *art200*) have been integrated into host gene expression regulatory networks^{64,72}. To support this hypothesis, our lab has found that overexpression of the highly structured 5' end of *tnpA* affects the expression of over 70 genes in the LT2 *Salmonella* strain⁶⁴. Of these 70 genes, one of them was *invF*, the transcription factor downstream of HilA in the SPI-1 invasion cascade (Figure 1), which was downregulated about 2-fold. In contrast, an SL1344 strain with a portion of the IS200 encoding the 5' *tnpA* end deleted in all 7 copies of IS200 (hereafter Δ IS200) seen in Figure 3B resulted in

the *invF* mRNA expression being upregulated about 15-fold. The Δ IS200 strain also exhibited a hyper-invasion phenotype in studies where HeLa cells were infected but failed to show an increase in infectivity in the GI tract tissue in a mouse infection model⁶⁴.

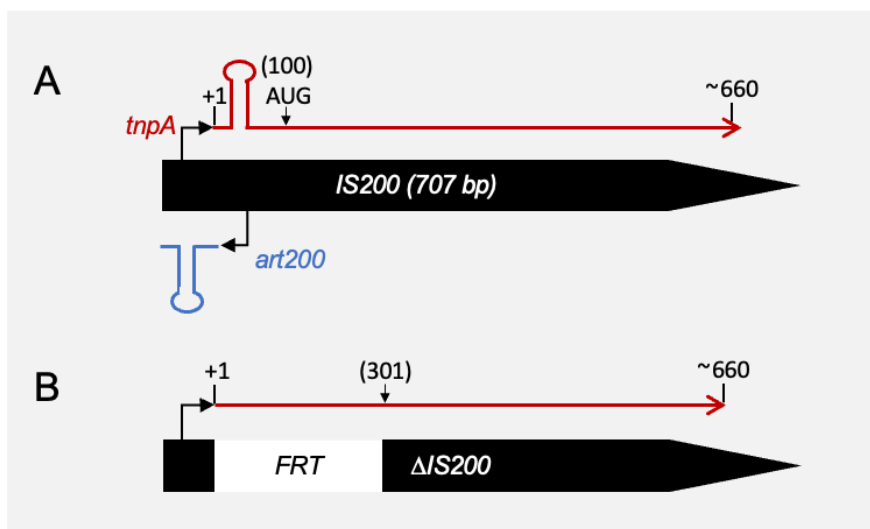


Figure 3. Simplified IS200 element.

The structure of IS200 is depicted with the two major transcription units shown. (A) The *tnpA* transcript encodes transposase and the 5' portion of this transcript (~110 nt), which includes a stable hairpin structure, forms a stable sRNA. *art200* is an anti-sense RNA that is known to pair with and destabilize the sRNA. (B) Schematic overview of the removal of the 5' portion of IS200 containing the FRT site⁷³ with IS200 transcription starting at 301 nt.

As a follow-up to our earlier study where we overexpressed the 5' portion of the IS200 *tnpA*, the Haniford lab has more recently performed differential gene expression analysis on the SL1344 Δ IS200 strain relative to WT. This included a targeted analysis of SPI-1 and flagellar genes by quantitative reverse-transcription polymerase chain reaction (qRT-PCR) and a more global RNA-seq analysis where differential gene expression was evaluated at early (EE), mid (ME), and late (LE) exponential growth phases in rich media (Supplemental Table 4-6). This work, which is still in progress, revealed that the Δ IS200 isolate of SL1344 (strain DH415) displays premature induction of flagellar and SPI-1 pathways. This is seen through an increase at the mRNA level of the master regulators of each of these pathways where there is roughly a 2-fold increase in *flhDC* transcript (regulator of the flagellar pathway) and a 5-fold increase in *hilD* transcript (regulator of the SPI-1 pathway). In addition to the increase in *flhDC* mRNA levels, a

small (30%) but statistically significant decrease in the mRNA levels of an important negative regulator of *flhDC* expression, *lrhA*, was also observed. While these changes in gene expression were revealed through qRT-PCR analysis, the RNA-seq analysis provided data supporting the possibility that deletion of the 5' portion of IS200 in all 7 copies present in the SL1344 strain causes an induction of the Cys regulon at LE growth. Typically, the Cys regulon is not induced in SL1344 grown in rich media until stationary phase. The focus of my thesis work has been to understand the mechanism through which IS200 deletion in *Salmonella* leads to premature Cys regulon induction. Gaining insight into this problem will help uncover how IS200 has integrated itself into the metabolic machinery of a medically important pathogen.

1.4 Objective and models

This project has aimed to validate and further characterize the impact of partial deletion of IS200 in *Salmonella*, focusing on understanding the mechanism behind the induction of the Cys regulon at LE phase. As described below, I developed and tested 3 different models for how IS200 might impact the control of the Cys regulon. These models are the futile import/export model, the metabolic burden model, and the LrhA deprivation model. Notably, the first two models were developed and tested first. The third model came to fruition only after a serendipitous finding that the transcription factor LrhA binds to the promoter regions of two Cys regulon genes and the recognition that LrhA mRNA levels are reduced in the Δ IS200 strain.

Model 1, the futile import/export model, centres around the observed increase in expression of *fliY* (the periplasmic-binding protein in cystine transport) in the Δ IS200 revealed by qRT-PCR analysis. I propose that this increase is explained by the increase in *flhDC* expression, which in turn drives *fliA* expression (Figure 1 and 2C). As previously noted, the upregulation of the functionally related protein YdjN led to a futile cycle of cystine import and cysteine depletion under conditions of high cystine availability⁴³. In model 1, I consider the possibility that increased production of a component of the FliY-YecSC ABC-transporter similarly drives cysteine depletion by over-importing cystine, which would lead to a rapid depletion of cysteine through the detoxification pathways (Figure 2A). Accordingly, the depletion of otherwise toxic cysteine levels is predicted to activate CysB and turn on the Cys regulon. Furthermore, I predict

that the *fliY* gene expression is upregulated due to its tight linkage to the class 2 flagellar operon *fliAZ* operon, which is overexpressed in Δ IS200 due to increases in *fliHDC* expression.

In model 2 (metabolic burden model), I speculate that the upregulation and early induction of the flagellar and invasion pathways in Δ IS200 imposes a metabolic burden on cells as they prepare to transition into stationary phase. The increased protein translation associated with the early induction of these pathways could lead to a shortage of cysteine, which could become a burden during late growth. Metabolic burden is a familiar concept in biotechnology and engineering pathways to produce compounds for industry and pharmaceutical use^{74,75}. Altering these native pathways for bioproduction requires energy and building blocks that place constraints on the host cell, thus altering both the resource and energy storage. This imbalance in the gene expression leads to under- or overexpression of different enzymes and pathways⁷⁵. This can be seen in the metabolic engineering of amorphaadiene, a precursor in the production of artemisinin the anti-malaria drug, where a multitude of genes have altered gene expression regulated through various inducible promoters in both *E. coli* and *Saccharomyces cerevisiae*⁷⁶.

The flagellar and invasion pathways consist of a large number of genes. Over 50 genes belong to about 17 operons that are part of the flagellar cascades divided among them based on the feed-forward regulation (class 1, 2 and 3) or sometimes referred to as early, mid or late depending on when they are expressed in the temporal induction^{28,29,77}. In contrast, the invasion cascade encoding SPI-1, associated chaperones and effector proteins consist of 39 genes under the feed-forward regulatory cascade under HilD regulation^{8,16}. Sulphur assimilation and cystine availability are important factors in bacterial growth as they are the primary suppliers for cysteine levels necessary for protein production. The metabolic burden model proposes that the upregulation of the complex flagellar and SPI-1 invasion cascades that start early during growth in Δ IS200 imposes a burden as the cell transitions to stationary phase. This could cause a possible shortage of cysteine and thus leads to the induction of the Cys regulon during LE phase.

In model 3, I predict that the reduced expression of LrhA drives the increased expression of Cys regulon genes observed in LE phase in the Δ IS200 strain. The development of this model is based primarily on a ChIP analysis where it was found that LrhA binds to 2 different Cys

regulon genes promoters, *cysD* and *ydjN*. Taken together with qRT-PCR data, which suggests that LrhA expression is reduced by 30% during early, mid, and late exponential phase, these observations led to considering that LrhA plays an active role in repressing the Cys regulon and this capacity for repression is reduced in an Δ IS200 strain. In considering the roles of LrhA, it is known to repress *flhDC* and *rpoS* transcription, where RpoS encodes a stress signal sigma factor for transitioning to stationary phase^{37,38}. In addition to its role as a global regulator, it is part of the LTTR family, the same family as the transcriptional activator of CysB. CysB, along with its co-activator OAS, and LrhA share an LTTR-consensus binding sequence, leading to the possibility that LrhA is involved in keeping the regulon turned off. Therefore, due to preliminary results showing that LrhA binds to *ydjN* and *cysD*, its role as a global regulator, and its similarity to the transcription factor CysB, the LrhA deprivation model proposes that LrhA actively represses the Cys regulon, and because its activity is reduced in Δ IS200 this leads to the induction of the Cys regulon come LE phase.

2 Materials and Methods

2.1 Strains, plasmids and growth conditions

A list of strains, plasmids and oligonucleotides can be seen in the Supplemental Tables (Appendix). *Salmonella* Typhimurium SL1344, or DBH347, was considered WT strain with all mutant strains derived from it. DBH415 ($\Delta tnpA_{1-7}$ referred to as $\Delta IS200$) was constructed by the lab via transducing individual IS200 knockout alleles into a single strain. Mutant strains of SL1344 were constructed via Lambda Red recombineering and checked via polymerase chain reaction (PCR)⁷³. *Escherichia coli* DH5 α cells were used for routine cloning and plasmid propagation. All strains were grown at 37°C with shaking in Lennox broth (LB; 5 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract) unless otherwise noted. Cultures grown overnight for RNA extractions were diluted to 1:100 and split into aliquots of 2 ml per culture tube. Where appropriate the following antibiotics were used: kanamycin (Kan), 25 μ g/ml; streptomycin (Str), 150 μ g/ml; chloramphenicol (Cm), 20 μ g/ml; ampicillin (Amp), 100 μ g/ml. For cystine supplementation, cells were grown in standard Lennox broth overnight and diluted 1:100 in Lennox broth with 0.5 mM L-cystine. For growth curves, cells were grown in standard Lennox broth for 7 hours and diluted 1:50 into fresh media. For ChIP analysis, 0.02% arabinose was added to the growth media during 1:100 subculture growth.

The promoter region of the green fluorescent protein (GFP) present on pDH1039 was replaced with the promoter of the gene of interest via routine cloning. The desired promoter and GFP protein lacking the added destabilizing protein was cloned into pACYC184 to form a functional, long-lasting form of GFP under the promoter of interest (Supplemental Figure 1A). For *cysD* GFP, the amplified promoter region of *cysD* was digested with SphI and cloned into digested SphI pDH1039. pDH1039 *cysD* GFP was then digested with XbaI and StuI and cloned into pACYC184 digested with XbaI and FspI. This was repeated in the construction of *ydjN* GFP except that the *ydjN* insert and pDH1039 were digested with EcoRI and SphI before cloning into pACYC184.

For FLAG-tag constructs, pDH1089 was used as a template for recombineering chromosomal kan marked C-term 3X-FLAG tag to amplify the insert (Supplemental Figure 1B). pDH740 (a

temperature-sensitive arabinose-inducible plasmid), encoding the recombinering proteins, was transformed into LT2. Cells were grown ON with 0.02% arabinose at 30°C. Following subculture, 30 minutes prior to harvesting 0.2% arabinose was added, and the cells were made electrocompetent. 300 ng of the FLAG insert was transformed. pDH740 was removed through growth on M9 minimal media (1X M9 salts, 40 µg/ml histamine, 1 µg/ml thiamine, 1mM MgCl₂, 0.2% glucose) and introduced into DBH347 and DBH415 by P22 transduction.

2.2 RNA isolation and qRT-PCR

For RNA extractions, the hot phenol method was used⁷⁸. The isolated RNA was followed up with a DNase treatment, and RNA (2 µg) was converted to cDNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). cDNA was diluted to 30 ng/µl in TE (50 mM Tris-HCl, pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA)) and stored at -20°C. Three biological replicates were analyzed with two technical replicates, and 16S rRNA (*rrsA*) was used as a reference gene for relative quantification. qRT-PCR was performed in 20 µl reactions; 10 ng cDNA, 500 nM of each primer, and PowerUp SYBR Green Master Mix (Applied Biosystems). ViiA 7 Real-Time PCR system was used with standard settings apart from the annealing/extension step adjusted to 60.5°C. The relative gene expression for each target was calculated using the efficiency corrected method⁷⁹.

2.3 Reverse-Transcription PCR

Reactions (20 µl) contained 1.5 ng/µl cDNA, 1 µM of each primer, and 10 µl GoTaq Green Master Mix (Promega). For reads greater than 2 kb, Q5 High-Fidelity 2X Master Mix (New England Biolabs) was used along with 0.5 µM of each primer. 2 µl of PCR reaction was then subjected to electrophoresis on 1% agarose gel in 1X TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) buffer. UV illumination was used to visualize the products following ethidium bromide staining.

2.4 Oxidative stress

S. Typhimurium cells were exposed to H₂O₂ to induce oxidative stress at mid- and late exponential growth phase. Briefly, cells were grown to either stage (1:100 dilution in 2 ml subcultures). Reactions (1.25 ml) contained 55 mM H₂O₂ for 15 minutes at 37°C with no shaking. Control cells were subjected to 1X phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄) in place of H₂O₂. Following incubation, cells were kept on ice. Cells were centrifuged at 4°C for 2 minutes at 20 000 x g, the supernatant was removed, and cells were washed in 1X PBS. This was repeated twice more with a final resuspension in 1 ml of 1X PBS. Control cells were subject to the same treatment. Washed cells were diluted to 10⁻⁶ for control cells and 10⁻⁴ for treated cells, and 100 µl of each dilution was plated on solid media and incubated at 37°C overnight. Colony-forming units (CFU) were counted following 16 hours of growth, and dilution was accounted for. Total count was average among treated and non-treated for each growth phase and plotted with the standard error on the mean. Two-way ANOVA was used to check for significant differences with a p<0.05.

2.5 Chromatin immunoprecipitation

Cells were grown overnight and diluted to 1:100 subcultures to various growth phases. Formaldehyde was added to a final concentration of 1% and shaken via end-over-end rotation for 15 minutes at room temperature. Glycine was added to a final concentration of 0.2M and shaken for an additional 10 minutes at room temperature. Cells were centrifuged at 4°C for 5 minutes at 6 000 x g, and the supernatant was discarded. 3 washes in ice-cold 1X Tris-buffered saline (TBS) (20 mM Tris, 150 mM NaCl) were performed with repeat centrifugation at 6 000 x g for 5 minutes. After subsequent washes, the pellets were frozen at -20°C.

Cells were lysed in 400 µl FA lysis buffer (50 mM HEPES pH 7.0, 150 mM NaCl, 1mM EDTA, 1% TritonX-100 [v/v], 0.1% sodium deoxycholate [w/v], 0.1% SDS) with 1X Protease inhibitor (0.284 µg/ml leupeptin, 1.37 µg/ml pepstatin A, 0.17 µg/ml phenylmethylsulfonyl fluoride (PMSF), 0.33 µg/ml benzamide) and transferred to a 2 ml tube. 400 µl glass beads (0.1mm in

diameter; zirconia/silica) were added, and the samples were vortexed in intervals of 30 seconds vortex and 10 seconds on ice, repeated 10 times. 800 μ l FA lysis buffer was added and vortexed for 30 seconds. The samples were then centrifuged at 20 000 x g for 10 minutes at 4°C, and the supernatant was transferred to a clean 1.5 ml microfuge tube. 1 μ l of RNaseA (20 mg/ml) was added, and the sample was incubated at 37°C for 30 minutes. Following incubation, three aliquots of 210 μ l and 100 μ l were prepared and stored at -80°C.

20 μ l of Anti-FLAG M2 MagBeads were equilibrated with two washes of 200 μ l of FA lysis buffer in 2 ml microfuge tubes. 200 μ l of the sample was added with 200 μ l of FA lysis buffer to the equilibrated beads and incubated for 1.5 hours at room temperature with end-over-end rotation. The samples were subsequently washed in 1 ml each in the following order: 3 times FA lysis buffer, 1 time FA500 buffer (FA lysis buffer with the exception of 500 mM NaCl), 1 time LiCl wash (10 mM Tris-Cl pH 8.0, 250 mM LiCl, 0.5% NP-40 [v/v], 0.1% sodium deoxycholate [w/v], 0.1% SDS), 2 times TBS. Following washes, 75 μ l of ChIP elution buffer (100 mM Tris-Cl pH 8, 200 μ M EDTA, 1% SDS [w/v]) was added to the samples and incubated at 37°C and eluate was transferred to a new 1.5 ml microfuge tube. This elution was performed twice for a total eluted sample of 150 μ l.

For reverse cross-linking, 200 mM NaCl and 1 μ l Proteinase K (20 mg/ml) was added to the immunoprecipitated samples and 50 μ l of ChIP elution buffer was added to 100 μ l input control and the samples were left to incubate at 65°C for 1 hour. Qiaquick PCR Purification Kit was used to purify both treated and control samples, and 2 μ l was used for gene-specific PCR amplification.

2.6 Growth curves

Growth was measure in a BioTek H1 Synergy microplate spectrophotometer for GFP growth curves and Multiskan Go microplate spectrophotometer was used for standard growth curves. Three biological replicates of each strain were grown in standard Lennox media for 7 hours and then diluted once (1:50). 200 μ l of each dilution was added to two wells (technical replicates) of

a 96-well clear bottom black microplate for GFP curves and a standard F-bottom clear 96-well plate for non-GFP growth curves. Cultures were grown at 37°C with continuous shaking for 9.3-12 hours and OD₆₀₀ and fluorescence emission (excitation 485nm, emission 528nm) was measured every 5 minutes for GFP curves. Non-GFP growth curves measured OD₆₀₀ every 15 minutes. GFP emission growth curves were plotted as fluorescence versus OD₆₀₀ and additional plots of growth with OD₆₀₀ versus time is included in the Appendix (Supplemental Figures 3-6).

2.7 Western blot

DBH673 (WT *yecC*::3X-FLAG-kan), DBH674 (Δ IS200 *yecC*::3X-FLAG-kan), DBH 675 (WT *fliY*::3X-FLAG-kan) and DBH676 (Δ IS200 *fliY*::3X-FLAG-kan) were grown to OD₆₀₀=0.4 and cells from 1 ml of culture was collected via centrifugation and then resuspended in 800 μ l of 1X TE. 15 μ l of sample and 15 μ l of sodium dodecyl sulphate (SDS) sample buffer (60 mM Tris-HCl, pH 6.8, 2% SDS [w/v], 0.1% bromophenol blue [w/v], 1% β -mercaptoethanol [v/v]) were mixed and boiled for 5 minutes. Samples (10 μ l) were then resolved on a 15% polyacrylamide gels and electroblotted to a polyvinylidene difluoride (PVDF) membrane. Membranes were incubated overnight with 5% milk solution in Tris Buffered Saline Tween 20 (TBST) (1X TBS, 0.5% Tween-20) at 4°C. Membranes were then incubated with primary antibody (1:5000 dilution: mouse α -FLAG M2, Sigma; mouse α -DnaK, Enzo) followed by incubation with the secondary antibody (1:5000 dilution; α -mouse, Promega). Bands were detected using ECL kit (Thermoscientific) and imaged using Bio-Rad imager. Membranes were then stripped and re-probed for loading control (DnaK). AlphaView was used to quantify bands and the amount of YecC-3X-FLAG and FliY-3X-FLAG. Samples were normalized to internal standard (DnaK) then relative to WT.

2.8 Glutathione determination

Total glutathione measurements was based on the 96-well plate adaptation⁸⁰ of the method described by Tietze⁸¹. Briefly, cells were grown to OD₆₀₀=0.4 and lysed by adding 143 mM sodium phosphate buffer (pH 7.4, 6.3 mM EDTA, 240 mM dibasic sodium phosphate, 45 mM

monobasic sodium phosphate), 0.01% Triton X-100 and 2 mg/ml lysozyme with periodic vortexing.

Glutathione reaction buffer (14 ml sodium phosphate buffer, 2.39 mM NADPH, 200 U/ml glutathione reductase) was prepared and placed on ice until ready. Standard glutathione samples were prepared in 5% sulphosalicylic acid ranging from 4 to 20 μ M. The reaction buffer minus DTNB was warmed to 30°C, 1 ml of 0.01 M 5,5'-dithiobis(2-nitrobenzoic acid) was added and let sit for 30 seconds. 125 μ l of the glutathione reaction buffer was added to 25 μ l of lysed sample or glutathione standard in a standard F-bottom clear 96-well plate and absorbances were read at 405 nm using the Multiskan Go microplate spectrophotometer.

Protein concentration was calculated via Bradford assay using Protein reagent dye (BioRad). Standards were prepared using bovine albumin serum (BSA) at 0.01, 0.1, 0.2, and 0.3 mg/ml. 200 μ l of Protein reagent dye was added to 25 μ l of sample or standard in a standard F-bottom clear 96-well plate. The samples were mixed with continuous shaking for 5 minutes and absorbance was read at 595 nm.

Standard curves were used to calculate glutathione and protein concentration and was reported as glutathione concentration per mg of protein.

3 Results

3.1 Validating the premature induction of the Cys regulon in *Salmonella* SL1344 with a disruption of IS200

Previous differential gene expression analysis using RNA-seq provided evidence that *Salmonella* strain SL1344 Δ IS200 prematurely induces genes belonging to the Cys regulon. Specifically, in LE growth Cys regulon genes including *yciW*, *cysK*, *cysC*, *cysW*, *cysN*, *cysU*, *cysD* and *cysP*, exhibit the largest increase in expression versus the WT strain of all 95 genes scored as differentially expressed at LE phase (Supplemental Table 6). The predicted increase in gene expression for these genes ranged from 10.5- to 50-fold. Therefore, I set out to validate the predicted early induction of the Cys regulon in LE phase in Δ IS200 using qRT-PCR. I chose to monitor several genes that play different roles in cysteine metabolism: sulphur assimilation (*cysD*), transcriptional control of the regulon (*cysB*), cystine import (*fliY*, *yecS*, *yecC* and *ydjN*) and cysteine detoxification (*metC* and *yciW*). YciW has been shown to be involved in converting cysteine to other products (such as glutathione or methionine) and is a novel gene under CysB control^{55,82}. MetC was chosen as a gene of interest due to its role in methionine metabolism. MetC encodes for the cystathionine β -lyase/cysteine desulphhydrase⁸³. Cystathionine β -lyase cleaves cystathionine into homocysteine in the methionine production pathway and degrades cysteine by acting as a cysteine desulphhydrase which reduces cysteine into ammonium, pyruvate and hydrogen peroxide⁸³. I show in Figure 4, aside from *metC* (which is not part of the Cys regulon), the expression of the genes noted above is upregulated in Δ IS200 anywhere from 2.5-fold to ~90-fold. Of these upregulated genes, there are two distinct patterns of overexpression. *fliY*, *yecS* and *yecC* are upregulated in all three growth phases monitored (EE, ME and LE), whereas *cysB*, *cysD*, *yciW* and *ydjN* are only overexpressed in LE phase. I also monitored the expression of two other genes, *hilD* and *fliA*. These genes were also predicted by RNA-seq to be upregulated in Δ IS200 strain and represent critical regulators of pathogenesis. It is seen in Figure 4 that *fliA* is upregulated throughout all three growth phases which is particularly important for the development of model 1 as will be discussed below. Of the genes examined in Figure 4, the expression pattern of *hilD* is unique because it is the only gene whose overexpression is apparent in earlier growth but then decreases in the later growth phases. Recall, that HilD is the master

regulator of the SPI-1 invasion cascade and this premature induction of *hilD* expression is relevant to the development of model 2.

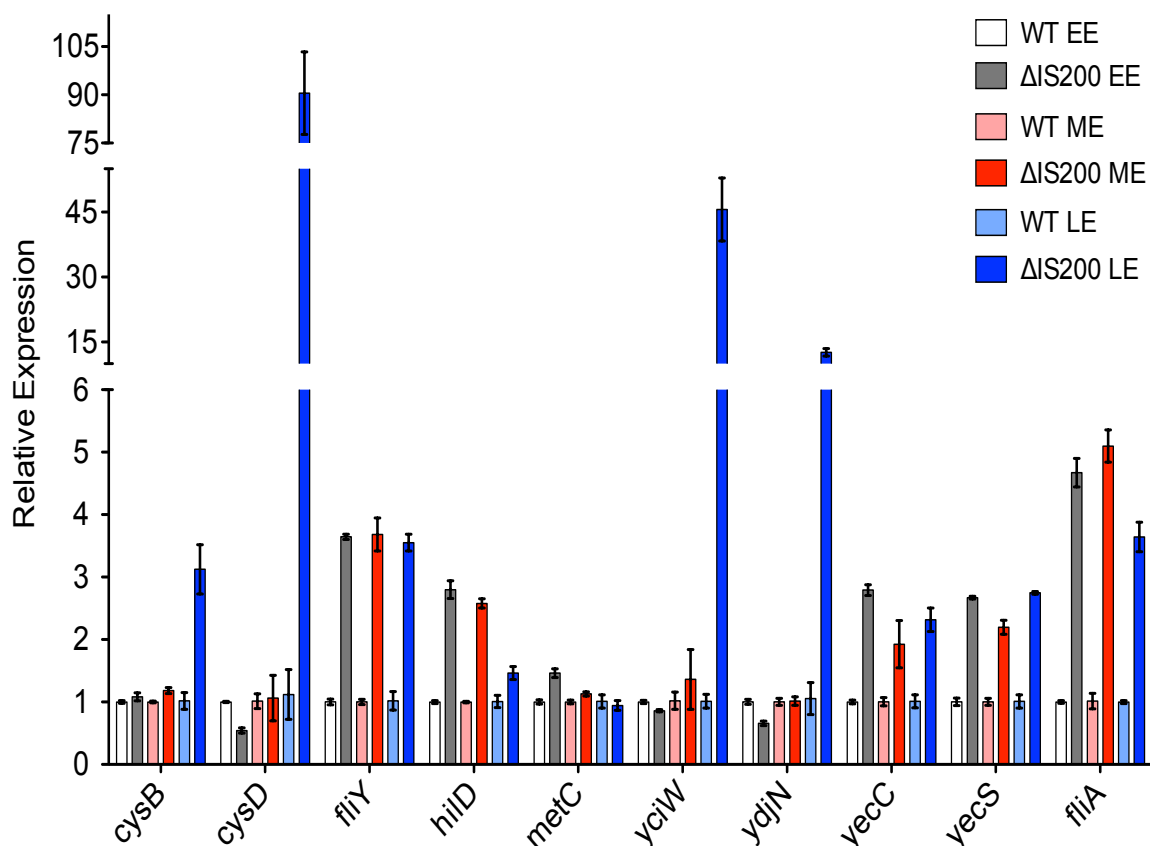


Figure 4. Removal of 5' end of IS200 leads to Cys regulon induction at LE phase.

qRT-PCR was performed on *Salmonella* SL1344 expressing full length IS200 (WT) or with the 5' end deleted in all 7 copies (Δ IS200). Strains were grown to early exponential (EE, $OD_{600}=0.25$), mid-exponential (ME, $OD_{600}=0.3$) or late exponential phase (LE, $OD_{600}=0.4$). Target genes included a variety of genes involved in cysteine metabolism, a gene in the SPI-1 pathway (*hilD*) and a gene in the flagellar pathway (*fliA*). Error bars show the standard error on the mean for three biological replicates.

To complement and expand on the findings of the qRT-PCR, I utilized GFP transcriptional fusions to gain a more extensive temporal coverage of two Cys regulon genes during growth as the qRT-PCR analysis was limited to only 3 time points. The promoter regions of either *cysD* or *ydjN* were amplified through PCR and cloned into pDH1039 containing the form of *gfp* which includes a destabilizing protein. The resulting pDH1039 plasmids with the Cys regulon gene

promoters were then digested to clone the promoter and GFP gene region lacking the destabilizing protein into pACYC184 plasmid (Supplemental Figure 1A). This transfer into the pACYC184 backbone gave rise to a functional GFP gene. These GFP transcriptional fusions plasmids were transformed into WT and Δ IS200 forms of SL1344. GFP expression and cell growth were monitored every 5 minutes through the fluorescent emission (excitation 485nm, emission 528nm) and light scattering, respectively, for 9.3 hours. Plate reader assays were performed in biological triplicate with two technical replicates per sample. Read outputs were averaged and the emission was plotted against the OD₆₀₀ which ranged from 0.2 to 1.0 OD₆₀₀ in the absence of a pathlength corrector. Growth phases were defined based on the plotted growth curve (Supplemental Figure 2) and ranged as follows: 0.2-0.25 OD₆₀₀ for EE, 0.26-0.35 OD₆₀₀ for ME, 0.36-0.45 OD₆₀₀ for LE and 0.46-0.61 for ES.

The results from the GFP growth curve supported the findings of the qRT-PCR data shown above, where there is a greater expression of GFP with the *cysD* or *ydjN* promoter in Δ IS200 compared to WT specifically in later growth stages. This approach not only validates the occurrence of the Cys regulon induction in Δ IS200 but provides a more precise timing during growth as to when the two forms of *Salmonella* SL1344 (WT and Δ IS200) differ in their expression of *cysD* and *ydjN*. Shown in Figure 5, both *cysD* and *ydjN* expression are induced earlier in Δ IS200 as seen by the offset of the emission curves. The expression of *cysD* is seen to differ between WT and Δ IS200 at OD₆₀₀~0.42 whereas *ydjN* diverges OD₆₀₀~0.475. While *cysD* expression is induced earlier, it also reaches a greater emission peak in both WT and Δ IS200 of ~260 and ~250 A.U, respectively, compared to *ydjN* emission peaks of ~195 and ~210 A.U for WT and Δ IS200, respectively. *cysD* emission peaks are greater than *ydjN* emission by about 50 units when comparing the same background [WT versus WT and Δ IS200 versus Δ IS200] (Figure 5). Additionally, growth is reduced in the Δ IS200 strain (maximum OD₆₀₀~0.8) compared to WT strain (maximum OD₆₀₀~0.925). The difference between *cysD* and *ydjN* inductions is presumed to reflect the different roles in cysteine metabolism; that is, the former is part of sulphur assimilation and the latter cystine transport. This analysis provided a more detailed understanding of when the Cys regulon induction occurs that could not be obtained solely through qRT-PCR analysis.

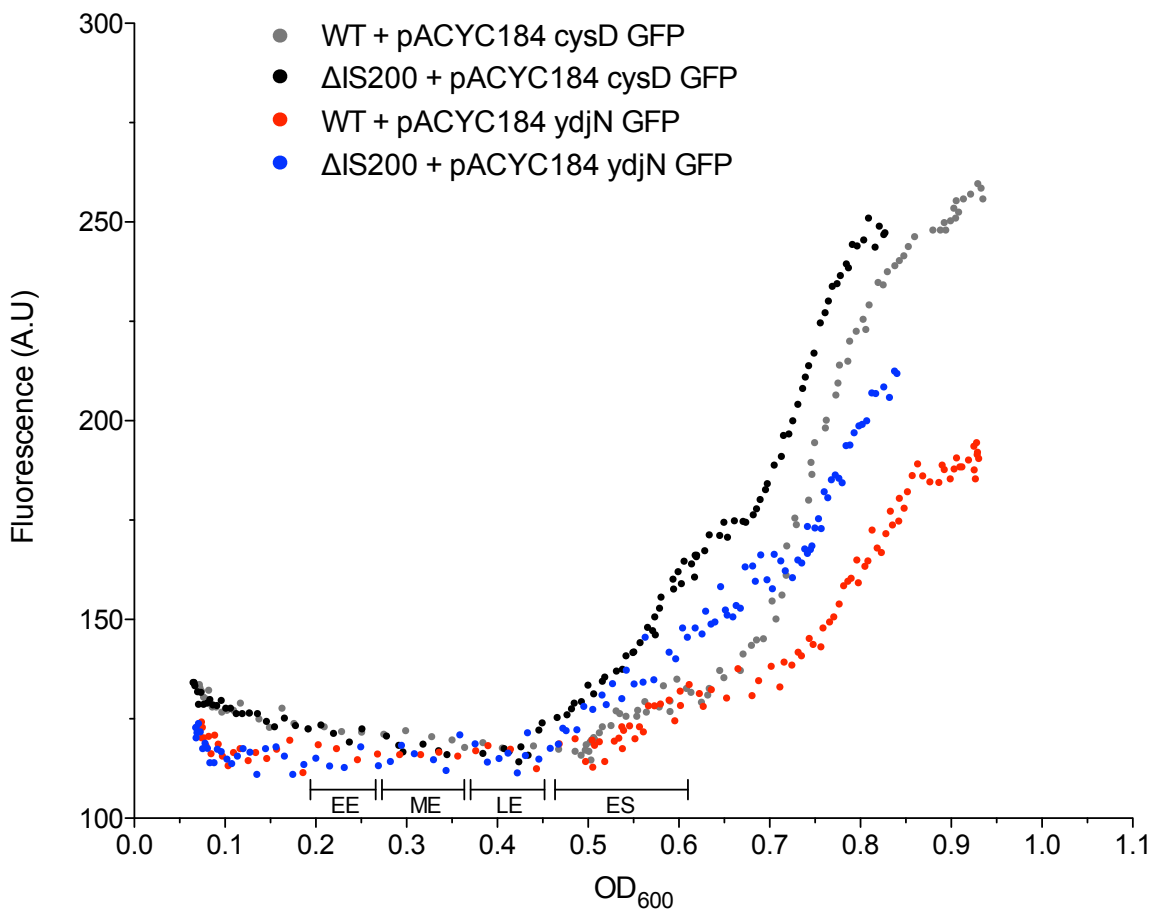


Figure 5. GFP monitoring of *cysD* and *ydjN* in *Salmonella* SL1344 WT and Δ IS200 strains grown in rich media.

pACYC184 transcriptional fusion contained the promoter region of either *cysD* [black and grey] or *ydjN* [red and blue] upstream of the GFP long-lasting form of the gene and transformed into WT [grey and red] and Δ IS200 [black and blue]. Cells were grown for 7 hours and diluted 50-fold into fresh Lennox media. 3 biological replicates and 2 technical replicates were added to a 96-well plate and fluorescence was measured every 5 minutes for roughly 9.3 hours. Average values for 3 biological replicates are shown. Growth phases (defined in the text) are shown on the x-axis. Fluorescence, measured in Arbitrary Units (A.U.), is shown on the y-axis. Note that the OD_{600} reading was not adjusted for path length.

Despite clearly depicting that the Cys regulon induction occurs in ES phase in Δ IS200 as seen by monitoring the fluorescence in the growth curve analysis, the following experiments were conducted at LE phase because the GFP transcriptional fusions were only successfully made

very late in the project. Thus, I could not observe the precise timing that the Cys regulon induction differs in the Δ IS200 from the WT strain. In retrospect, knowing that there is a more significant difference in *cysD* and *ydjN* expression between Δ IS200 and WT in early and late stationary growth, it would have been more beneficial to perform the subsequent analyses later than LE phase where the two strains diverge. Nevertheless, due to the initial findings that there is an upregulation at LE phase (Figure 4 and Supplemental Table 6) while troubleshooting the construction of these GFP fusions, I continued the project with the knowledge that it is possible to observe the premature Cys regulon induction at LE phase.

3.2 Providing exogenous cystine or disruption of CysB represses the Cys regulon induction in the Δ IS200 strain

Cysteine production is derived from sulphur assimilation pathways where sulphur is provided through ions such as sulphate and thiosulphate to donate sulphur atoms to NAS through a series of steps mediated by Cys regulon genes. The activation of the Cys regulon requires CysB binding, which is activated in the presence of OAS or NAS, and sulphur limitation. Without sulphur limitation, cysteine will inhibit the production of NAS at levels greater than 1 mM preventing the Cys regulon from being activated and preventing cysteine toxicity⁵⁵. I hypothesized that the growth specific Cys regulon induction in the Δ IS200 strain reflects a cysteine shortage. Accordingly, I expected that by adding exogenous cystine, the induction of the Cys regulon genes would be suppressed in the Δ IS200 strain. I tested this by providing 0.5 mM L-cystine to rich media and used qRT-PCR to monitor the relative gene expression of select Cys regulon genes (*cysB*, *cysD*, *cysP* and *ydjN*). As seen in Figure 6, cystine supplementation did suppress the induction of this subset of Cys regulon genes at LE phase in the WT and Δ IS200 strains. *cysB* expression is unchanged between WT and Δ IS200 when exogenous cystine was added while being reduced by ~20% than WT in standard media. All the other genes tested (e.g. *cysD*, *cysP* and *ydjN*) are down-regulated anywhere from 2- to 4-fold, or reduced by 50-75%, in the supplementation condition (both WT + *cys* and Δ IS200 + *cys*, pink and red bars) relative to WT with no L-cystine added (white bars) (Figure 6).

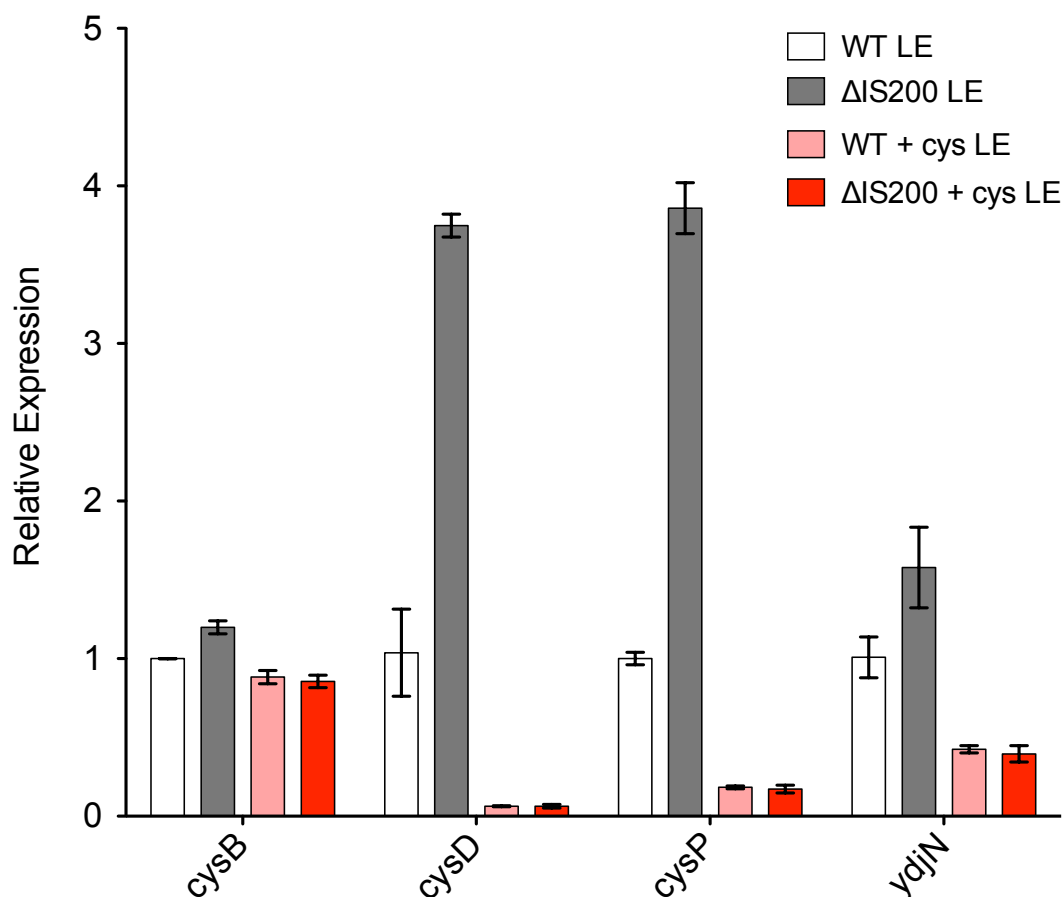


Figure 6. Impact of adding exogenous cystine on the expression of select Cys regulon genes at LE phase.

qRT-PCR was performed on *Salmonella* WT and Δ IS200 strains grown to late exponential phase (LE, $OD_{600}=0.4$) in rich media (white and grey) or rich media supplemented with 0.5 mM L-cystine (pink and red). Target genes were those involved in cysteine biosynthesis and cystine transport (*ydjN*). Error bars show the standard error on the mean for three biological replicates.

Given the role of CysB in regulating sulphur assimilation for cysteine biosynthesis, cystine transport and cysteine detoxification, I hypothesized that removal of *cysB* will prevent the induction of the Cys regulon during LE in the Δ IS200 strain. By testing the removal of the transcription factor CysB, I ask if the Cys regulon induction seen in the Δ IS200 strain reflects IS200 affecting the canonical pathway of activating cysteine metabolism. Accordingly, I

expected that if the Cys regulon induction is still present in the Δ IS200 Δ *cysB* strain, it points to IS200 effecting a non-canonical pathway. Using qRT-PCR I tested this by monitoring the relative gene expression of select genes involved in cysteine metabolism (*cysP*), cysteine detoxification (*yciW*), and cystine transport (*fliY*, *yecS*, *yecC*, *ydjN*) in Δ IS200 Δ *cysB* relative to Δ *cysB*.

Results from the qRT-PCR analysis confirm that there is no induction of *cysP*, *yciW* or *ydjN* in Δ IS200 Δ *cysB* at LE phase (Figure 7). Specifically, *cysP*, *yciW*, and *ydjN* expression in the double mutant Δ IS200 Δ *cysB* show no induction relative to Δ *cysB* in all three growth phases. This experiment is limited to Δ *cysB* mutants with no WT and Δ IS200 strains present with the expression relative to Δ *cysB* in each growth phase. If WT were present the relative expression of *cysP*, *yciW*, and *ydjN* in Δ *cysB* and Δ IS200 Δ *cysB* would be a 100% decrease compared to WT (0 expression in Δ *cysB* versions relative to 1 in WT). I have shown in Δ IS200 strain the upregulation of *cysD* (which behaves similarly to *cysP*), *yciW*, and *ydjN* range from 90-fold to 15-fold increase relative to WT in LE phase (Figure 4). The expression of *cysP*, *yciW*, and *ydjN* in Δ IS200 Δ *cysB* compared to Δ IS200 would be anywhere from 9000% to 1500% decrease [0 expression in Δ IS200 Δ *cysB* compared to 90- to -15-fold increase Δ IS200 relative to 1 in WT]. In addition, the expression of *fliA*, *fliY*, *yecS*, and *yecC* is not affected by the removal of *cysB* where their expression is seen to be induced throughout EE, ME, and LE growth phases in Δ IS200 Δ *cysB* relative to Δ *cysB* (Figure 7). *fliA* and *fliY* show similar patterns of expression where the relative expression is higher at ME than EE and LE. *yecS* and *yecC* are seen to be upregulated ~2-fold through all three growth phases monitored. This demonstrates that in the absence of the transcriptional activator CysB the Cys regulon was not induced in Δ IS200 at LE phase, and genes that do not belong to the Cys regulon (*fliA*, *yecS*, *yecC*) along with *fliY* are shown to be CysB-independent. In providing evidence that the Cys regulon induction reflects cysteine shortage and a dependency on CysB regulation, I show that Δ IS200 is not affecting any non-canonical mechanisms of activation.

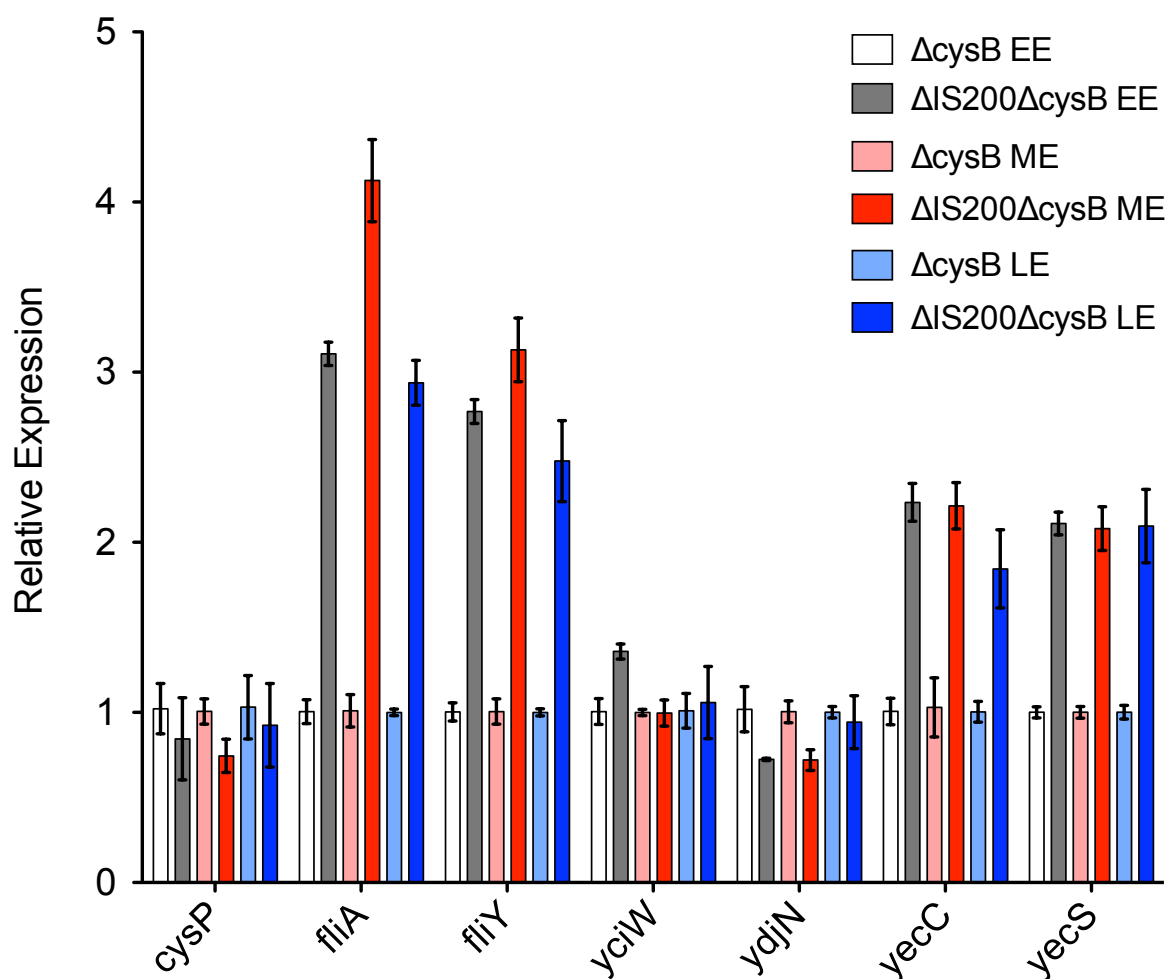


Figure 7. Impact of *cysB* deletion on the Cys regulon induction during growth.

qRT-PCR was performed on *Salmonella* $\Delta cysB$ and $\Delta IS200\Delta cysB$ grown to early exponential (EE, $OD_{600}=0.25$), mid-exponential (ME, $OD_{600}=0.3$) and late exponential (LE, $OD_{600}=0.4$) phases. Expression is relative to $\Delta cysB$ and targets include genes involved in cysteine biosynthesis, cysteine degradation (*yciW*), cysteine transport and a gene in the flagellar pathway (*fliA*). Error bars show the standard error on the mean of 3 biological replicates.

3.3 Cysteine biosynthesis and oxidative stress

Cysteine plays an important role in cellular response to oxidative stress as it is a building block for antioxidants such as glutathione and, through its conversion to cystine in the periplasm (Figure 2A), provides reducing power for the conversion of hydroxyl radicals to water^{84,85}. At

the same time, over-import of cystine (a consequence of increased expression of cystine importers or lack of feedback inhibition) puts the cells at greater risk of oxidative damage because it would drive a depletion of cysteine in the cytoplasm. Thus, in part of characterizing the impact of removing 5' end of IS200, I hypothesized that induction of Cys regulon is indicative of a cysteine depletion and will leave the cells hypersensitive to oxidative stress. I tested this by treating WT and Δ IS200 cells with hydrogen peroxide to induce oxidative stress at either EE or LE growth phases and looked for evidence of differential cell survival by plating treated and untreated cells for colony-forming units (CFUs) (Figure 8). The results show no difference in survival between treated cells and untreated across both backgrounds in EE phase, which is before the Cys regulon induction in the Δ IS200 strain (Figure 8A). In contrast, Δ IS200 cells have reduced survival (roughly 5-fold) relative to WT cells in LE phase (Figure 8B). This strongly supports the idea that the removal of IS200 causes cysteine depletion, that, in turn sensitizes the cells to oxidative stress.

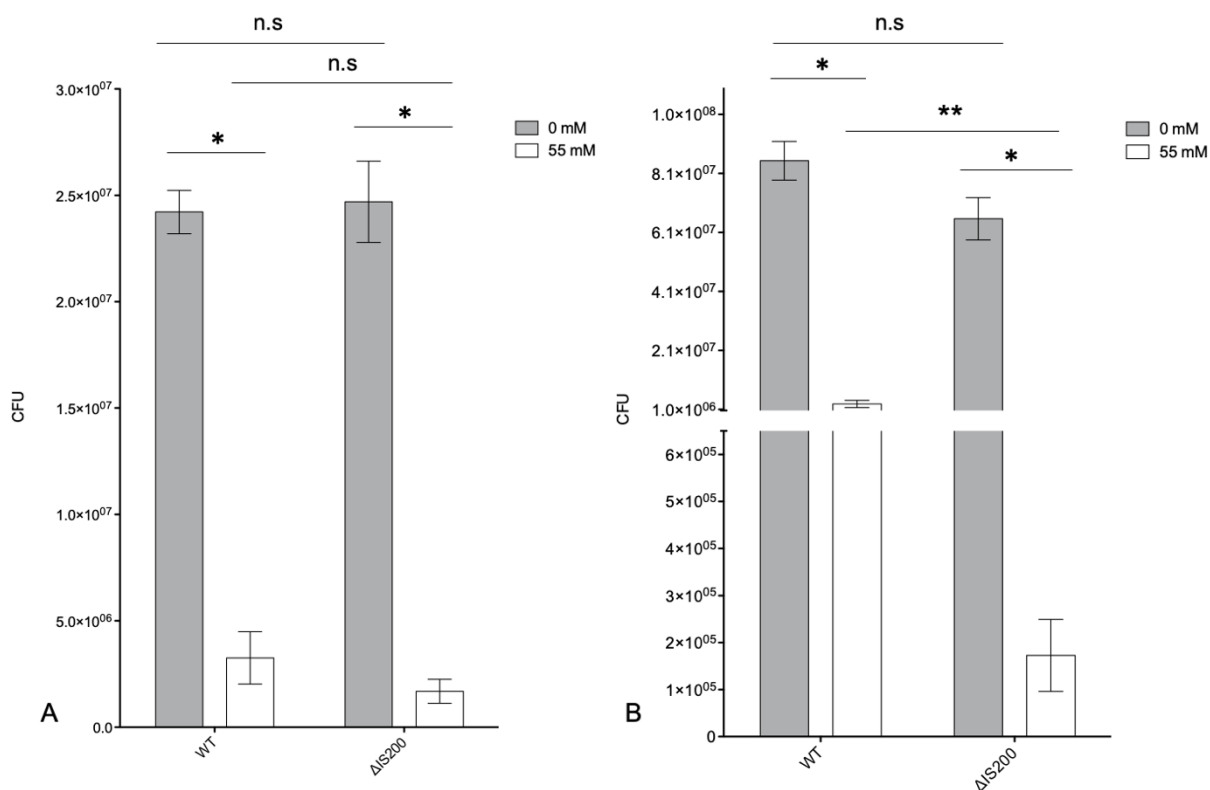


Figure 8. *Salmonella* WT and Δ IS200 exposure to oxidative stress.

Cells were grown to either (A) early exponential ($OD_{600}=0.3$) or (B) late exponential ($OD_{600}=0.4$) phase and oxidative stress was applied through exposure of the cells to hydrogen peroxide (55 mM) for 15 minutes. Cells were then diluted to 10^{-4} - 10^{-6} and plated on rich media agar plates and let grow at 37°C for 16 hours. Colony-forming units (CFUs) were counted, and error bars represent the standard error on the mean for 6 (A) and 12 (B) biological replicates. n.s for not significant; * $p < 0.0001$ and ** $p < 0.05$, two-way ANOVA.

3.4 Model 1 – the futile import/export model

To this point I have provided evidence that validates the presence of a premature induction of the Cys regulon present in the *Salmonella* SL1344 Δ IS200 strain during LE-ES phases due to a transient shortage of cysteine and/or sulphur. Previous research has shown that increased expression of YdjN, the primary cystine transporter, can drive a futile cystine import/cysteine export cycle that drives the depletion of cytoplasmic cysteine⁴³. Given this precedent, I was

intrigued with the observation that all the components of the second cystine transporter system, FliY-YecSC, are overexpressed throughout all three growth points in the Δ IS200 strain which starts early in growth (Figure 4). This raised the possibility that persistent overexpression of the FliY-YecSC cystine transporter may be driving a cysteine depletion through a similar futile cystine import/cysteine export cycle (Figure 9).

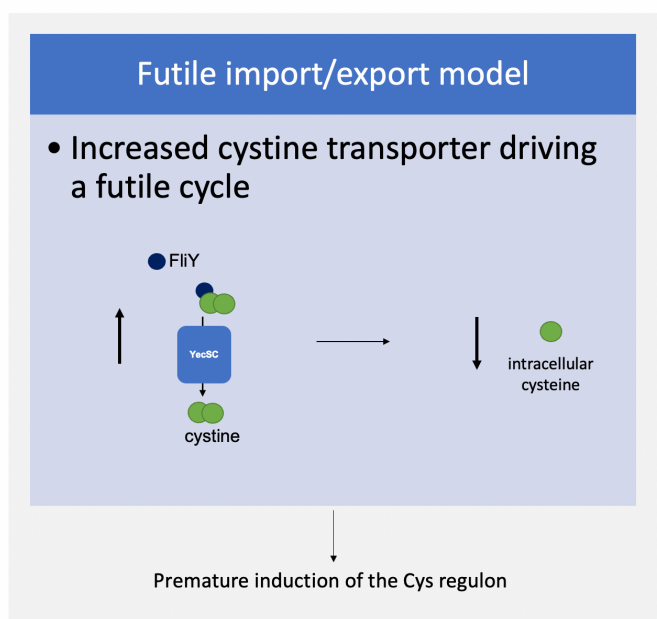


Figure 9. The futile import/export model.

A simplified diagram of the proposed futile import/export model (model 1) where an increase in expression of the cystine transporter, FliY-YecSC, drives a depletion of intracellular cysteine leading to the premature induction of the Cys regulon.

To test this, I wanted to look at the dependence of the Cys regulon induction on FliY in the Δ IS200 strain. In addition, I wanted to gain some insight into the mechanism through which FliY-YecSC overexpression occurs in the Δ IS200 strain. Regarding the latter, it seemed likely that the increased expression of FliA observed in Δ IS200 could be driving increased expression of FliY-YecSC. In *Salmonella*, the *fliY*, *yecS* and *yecC* genes are located immediately downstream of the FliA operon (Figure 2C). Accordingly, transcriptional read-through from the FliA operon could, in theory, drive coordinate regulation of these down-stream genes. To test this possibility, I performed crude mapping of transcripts through the FliA operon. This involved preparing cDNA with primer sets that hybridize to different portions of the *fliA-fliY-yecSC* locus.

For example, if there is transcriptional read-through from the *fliA* gene to the *fliY* gene, a combination of the *fliA* and *fliY* primers would give an RT-PCR product. To ensure that any PCR products detected were the result of amplification of cDNA as opposed to genomic DNA, PCR reactions were also included wherein the reverse transcriptase was left out of the cDNA synthesis reaction.

I first performed RT-PCR analysis described above using primer sets that would amplify transcripts from *fliA*, *fliA-fliY*, *fliA-yedO*, and *fliA-yecS* in WT and Δ IS200 at LE phase. I show in Figure 10 that there is more *fliA* transcript in Δ IS200 compared to WT, which is to be expected with the overexpression of flagellar synthesis in an Δ IS200 background (Figure 4). Additionally, I show that there are transcripts that span from *fliA* to *fliY* and from *fliA* to *yedO* but not from *fliA* to *yecS*, and that these read-through transcripts are generated only in the Δ IS200 strain. From this data it seems likely that upregulation of FliA in the Δ IS200 strain drives increased expression of *fliY* but not *yecS* or *yecC*.

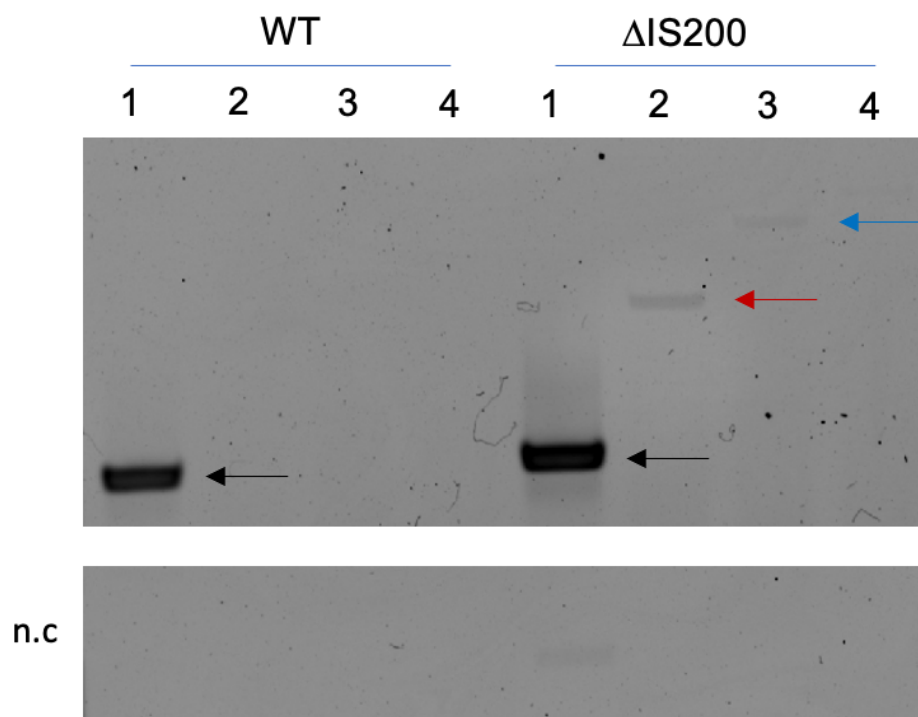


Figure 10. Transcript mapping downstream of the FliA operon.

RT-PCR was performed on *Salmonella* WT and Δ IS200 strains grown to LE ($OD_{600}=0.4$) and analyzed on 1% agarose gel. Primer sets included *fliA* 5' to (1) *fliA* 3', (2) *fliY* 3', (3) *yedO* 3' and (4) *yecS* 3'. A negative control

(n.c) wherein the reverse transcriptase was left out of the PCR reactions was included. Transcripts were present for *fliA-fliA* (black arrow), *fliA-fliY* (red arrow) and *fliA-yedO* (blue arrow).

Since FliY works in conjunction with YecSC to drive cystine import, I also wanted to test the possibility that there is coordinated regulation of these genes being driven by *fliY* transcription. To this end, I performed RT-PCR analysis with primer sets anchored in *fliY* to *yecS* and *fliY* to *yecC* genes. I show in Figure 11A that there is no transcriptional read-through from *fliY* to either *yedO*, *yecS*, or *yecC*. I included cDNA made from both Δ *cysB* (Figure 11B) and Δ *fliAZ* (Figure 11C) strains in this analysis to further investigate the dependence of *fliY* gene expression on *cysB* and *fliA*. This analysis supports the idea that *fliY* transcription is not exclusively driven by *fliA* in the Δ IS200 strain because *fliY* transcript was detected in the Δ *fliAZ* strain. In fact, *fliY* has its own promoter and this promoter is under CysB regulation^{30,31,43}. Based on my observation that *fliY* transcript is produced in a Δ *cysB* strain in both the crude transcript mapping and qRT-PCR experiments, I can conclude that *Salmonella fliY* transcription is not solely dependent on CysB. The results of this transcript mapping analysis supports that FliA overexpression in the Δ IS200 strain at least contributes to *fliY* overexpression, but curiously not to *yecS* or *yecC* overexpression, which appears to be independent from *fliA* and *fliY*.

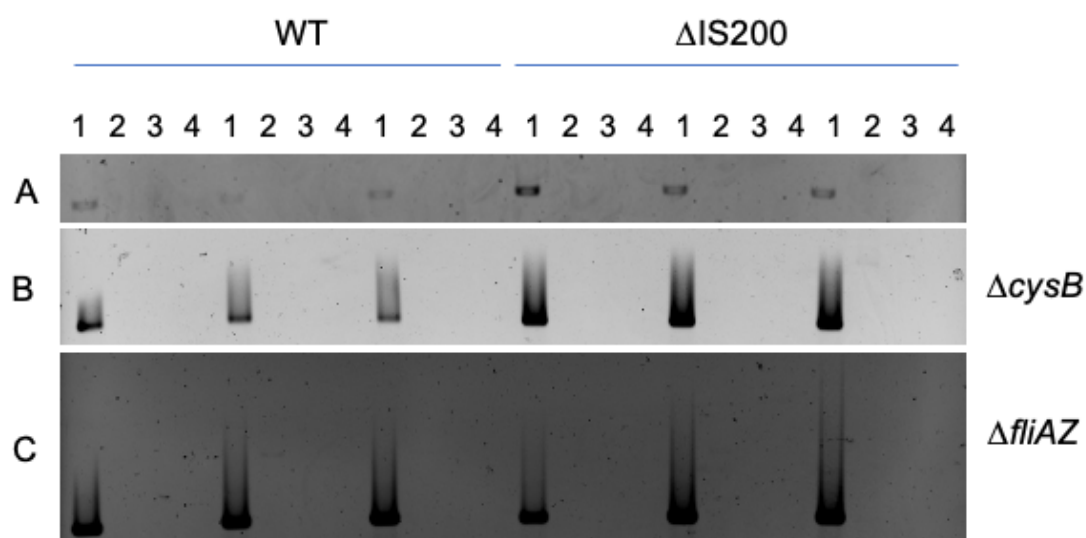


Figure 11. Transcript mapping downstream of *fliY*.

RT-PCR was performed on *Salmonella* WT and Δ IS200 in triplicate samples grown to LE ($OD_{600}=0.4$) and analyzed on 1% agarose gel. Primer sets included *fliY* 5' to (1) *fliY* 3', (2) *yedO* 3', (3) *yecS* 3' and (4) *yecC* 3'. Transcripts were present for *fliY-fliY* in both the WT and Δ IS200 strains.

Next, because of the hypothesized implication of a futile import/export cycle driving cysteine depletion and Cys regulon induction, it was important to show increased FliY and YecC protein expression in the Δ IS200 strain. To do so, I constructed 3X-FLAG tagged *fliY* and *yecC* genes in the Δ IS200 strain through recombineering. I then performed a Western blot analysis on these strains grown to LE phase. As shown in Figure 12, the YecC and FliY protein levels are increased \sim 5.5-fold and \sim 3-fold, respectively, in the Δ IS200 strain background. This was a positive outcome that further solidified the possibility that the FliY-YecSC cystine transport system could drive a futile cystine import/cysteine export cycle in the Δ IS200 strain.

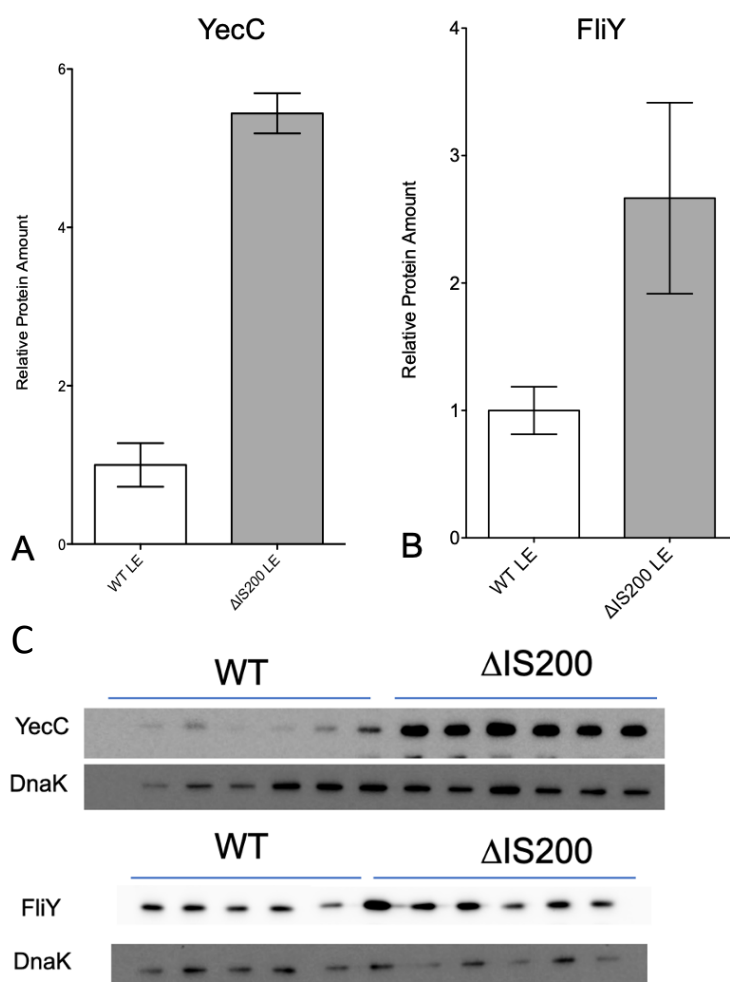


Figure 12. Western blot analysis to measure FliY and YecC levels.

Salmonella WT and Δ IS200 cells contained a 3X-FLAG tag added on to the end of *fliY* and *yecC* genes and cell extracts were prepared at LE phase ($OD_{600}=0.4$) for both the Δ IS200 and WT SL1344 strains and probed with anti-FLAG antibody. The quantified protein amount of (A) YecC and (B) FliY present in Δ IS200 relative to WT shown. Probed blots shown in (C) with DnaK as a loading control. Error is the standard error of the mean of 6 biological replicates with exception of WT FliY with 5 biological replicates.

To test whether the overexpression of FliY-YecSC transport system contributes to the early induction of the Cys regulon seen in the Δ IS200 strain at LE-ES, I tested the direct removal of *fliY* on Cys regulon induction. The removal of *fliY* in this cystine transport system is sufficient to remove all ATPase activity, and therefore its ability to transport substrates⁵⁶. It is expected that if the overexpression of FliY-YecSC cystine transporter drives a depletion of available cystine and leads to the induction of the Cys regulon at LE-ES phase, then inhibition of transporter function (through the knockout of the *fliY* gene encoding the cystine binding protein) would prevent the Cys regulon induction seen at LE-ES phase in the Δ IS200 strain. I tested this hypothesis using two different methods. I first performed a qRT-PCR experiment on the relevant Δ *fliY* strains, and second, I used the plasmid encoding Cys regulon genes (*cysD* and *ydjN*) fused to GFP in Δ *fliY* strains and measured GFP expression as a function of growth phase. The qRT-PCR results presented in Figure 13 show that knocking out the *fliY* gene partially inhibited *cysD* and *cysP* induction in the Δ IS200 strain background, thus suggesting that *cysD* and *cysP* induction in the Δ IS200 strain is at least partially dependent on *fliY*. Unexpectedly, this experiment also demonstrated that Δ *fliY* increases both *cysD* and *cysP* expression at the mRNA level in the single knockout strain where the expression of these two genes were increased 4-fold relative to WT. In this experiment, I also analyzed the expression of two genes (*fliZ* and *hilD*) whose expression is not expected to be affected by the *fliY* knockout but is affected by the disruption in IS200. As shown, the expression of either *fliZ* or *hilD* was not influenced by the removal of *fliY* (Figure 13).

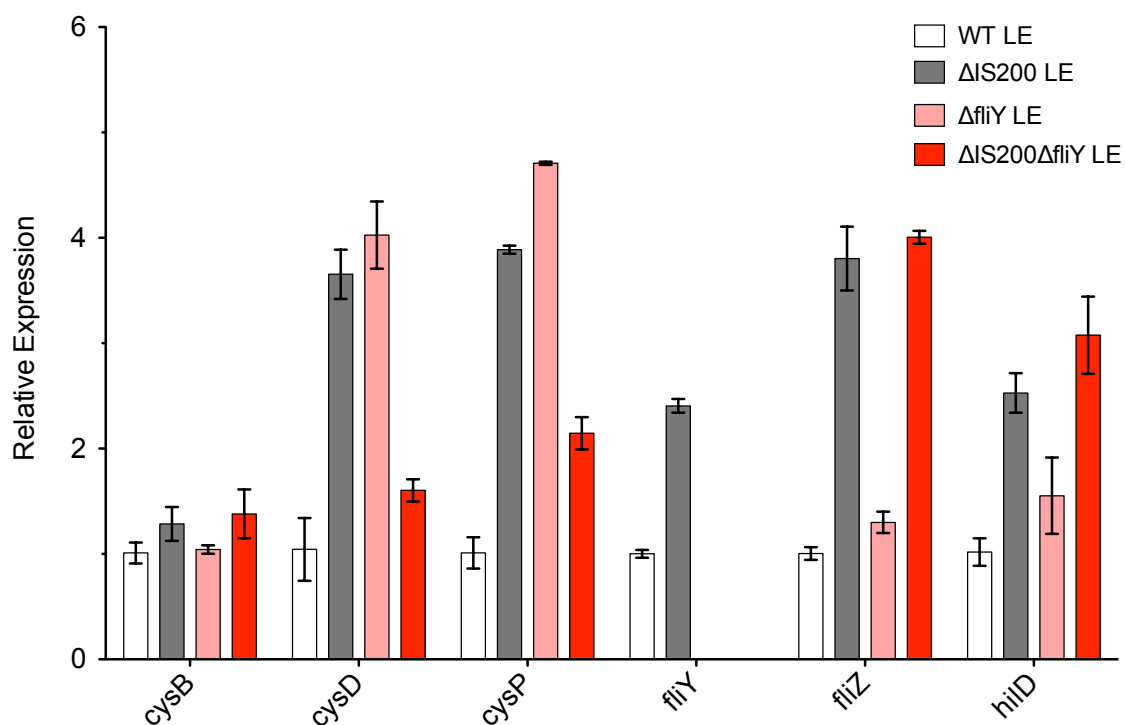


Figure 13. Impact of *fliY* deletion on Cys regulon induction.

qRT-PCR was performed on *Salmonella* SL1344 WT or Δ IS200 strains grown to late exponential ($OD_{600}=0.4$). Target genes included those belonging to the Cys regulon and markers of the SPI-1 pathway (*hilD*) and flagellar (*fliZ*). Error bars are the standard error of the mean for three biological replicates.

The second method to test if the Cys regulon induction is dependent on *fliY* used pACYC184 *cysD* GFP and pACYC184 *ydjN* GFP transformed into Δ *fliY* and Δ IS200 Δ *fliY* strains and I monitored the gene expression during growth (Figure 14). Given the results shown in Figure 13, it was expected that introducing the *fliY* deletion into the Δ IS200 strain would suppress *cysD* expression relative to WT on its own and this would manifest as a right-ward shift in the GFP expression curve. However, as shown in Figure 14A the expression of *cysD* was not significantly different between the Δ IS200 and Δ IS200 Δ *fliY* strains. Notably, both strains showed an early induction of *cysD* relative to WT strain seen in the left-ward shift of the GFP expression curve. In addition to *cysD*, I checked a second Cys regulon gene (*ydjN*) in this assay and again failed to see the expected suppression of the early Cys regulon induction in the Δ IS200 Δ *fliY* strain.

At the current time, it is unclear what is responsible for the discordance in the results presented in Figures 13 and 14 regarding the apparent dependency of early Cys regulon induction on *fliY* (see Figure 13). The conflicting results between these two approaches gave cause to consider alternative models for early Cys regulon induction in the Δ IS200 strain.

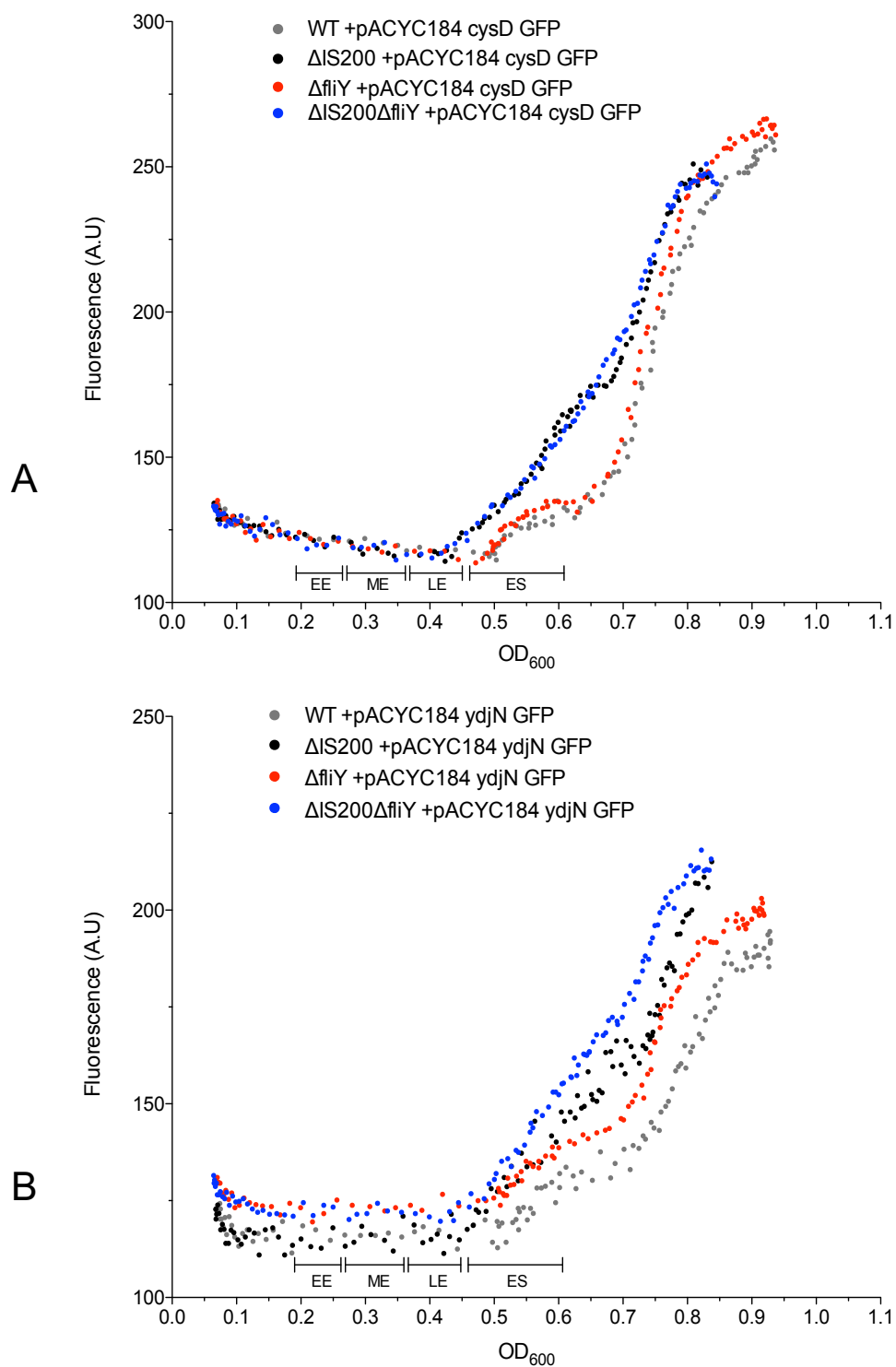


Figure 14. GFP monitoring of *cysD* and *ydjN* in WT and Δ IS200 strains containing Δ *fliY* grown in rich media.

pACYC184 transcriptional fusion containing the promoter region of *cysD* (A) or *ydjN* (B) is upstream of the long-lasting form of GFP in either WT [grey], Δ IS200 [black], Δ *fliY* [red], Δ IS200 Δ *fliY* [blue]. Cells were grown for 7 hours, diluted 50-fold into fresh Lennox media. 3 biological replicates and 2 technical replicates were added to a 96-well plate and fluorescence was measured every 5 minutes for roughly 9.3 hours. Average values for 3 biological replicates are shown. Growth phases (defined in the text) are shown on the x-axis. Fluorescence, measured in Arbitrary Units (A.U), are shown on the y-axis. Note that the OD₆₀₀ reading was not adjusted for path length.

3.5 Model 2 – the metabolic burden model

In model 2, I considered the possibility that the upregulation and premature induction of the flagellar and SPI-1 cascades due to the disruption of IS200, leads to a temporary depletion of available cysteine because of the increased demand for cysteine under conditions of increased protein synthesis (Figure 15). Additionally, because cysteine and indirectly methionine (as cysteine is a component required for its production) relies on sulphur availability, producing these amino acids could become rate limiting in conditions where levels of protein synthesis are exceptionally high.

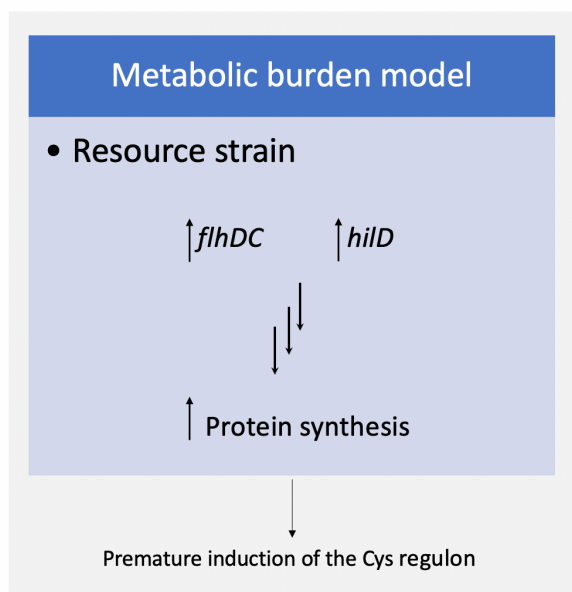


Figure 15. The metabolic burden model.

A simplified diagram depicting the metabolic burden model (model 2) where increased expression of the flagellar (*flhDC*) and SPI-1 (*hilD*) cascades lead to a resource strain on the cell, which is proposed to cause a premature induction of the Cys regulon.

This type of metabolic burden is commonly encountered in metabolic engineering research where entire pathways are being optimized and overexpressed in heterologous hosts and typically manifest as a slowing of growth. Notably, the *Salmonella* Δ IS200 strain exhibits a slow growth phenotype, which could indicate experiencing a metabolic burden. On rich solid media, the Δ IS200 strain displays smaller colony size compared to WT and in rich liquid media, the Δ IS200 cultures exhibit reduced saturation of overnight growth compared to WT cultures. The Δ IS200 strain also transitions into stationary phase earlier than WT (Figure 16); while both WT and Δ IS200 have similar growth during the first 180 minutes, the Δ IS200 cells lag in the remaining growth and transition into stationary phase at half the OD₆₀₀ compared to WT. This reduced growth is also seen in the GFP assay, where WT reaches a maximum OD₆₀₀~0.925 whereas the maximum for Δ IS200 is OD₆₀₀~0.8 (Figure 5).

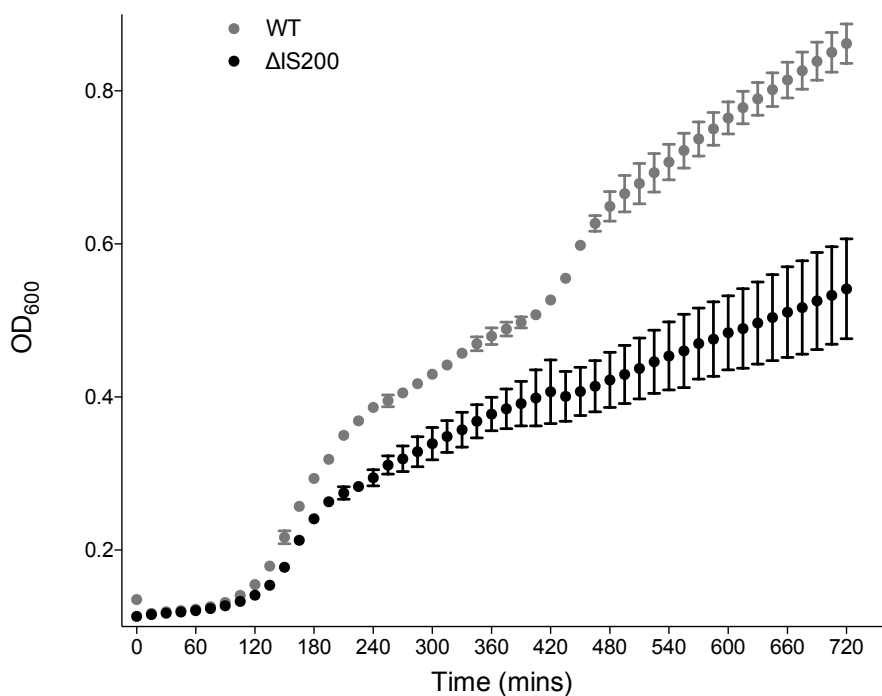


Figure 16. Growth of *Salmonella* SL1344 WT and Δ IS200 in rich media.

Growth of indicated strains was measured in a 96-well microplate spectrophotometer every 15 minutes at 37°C for 12 hours. Strains were grown in Lennox media. Error represents standard error on the mean of three biological replicates. OD₆₀₀ was measured in the absence of a path length corrector.

To test the metabolic burden model, I chose two approaches. First, I asked if cysteine levels fluctuated significantly during growth of the Δ IS200 strain relative to WT. Thus, I indirectly measured glutathione levels at the growth point (LE-ES) when the Cys regulon induction occurs in the Δ IS200 strain. Notably, I did not have access to an assay to measure the intracellular cysteine concentration directly. Instead, I measured intracellular glutathione levels, which has several links to cysteine to provide an indirect measure of cysteine levels. Second, I asked if blocking either the flagellar or the SPI-1 cascade, through removing a key transcription factor at the top of the regulatory cascade, influences the timing of Cys regulon induction in the Δ IS200 strain. Blocking the expression of either of these two pathways could reduce the predicted metabolic burden associated with expressing these pathways.

Glutathione is an important antioxidant produced in *Salmonella*, and its synthesis can be broken down into a 2-step process. First, glutamate and cysteine form γ -glutamylcysteine, which then forms glutathione through the addition of glycine. There are several important links between cysteine and glutathione. These include the availability of cysteine being the rate-limiting step for glutathione production^{44,86}, intracellular conversion of cystine to cysteine is dependent on glutathione⁴³, and glutathione is a stored source of cysteine via the γ -glutamyl cycle (Supplemental Figure 7)⁸⁷. As the formation of γ -glutamylcysteine is the rate-limiting step in glutathione biosynthesis, key determinants of glutathione production are cysteine availability and the activity of glutamate cysteine ligase. Due to cysteine being a limiting component for glutathione production, glutathione levels are directly related to cysteine levels. Utilizing Ellman's reagent for thiol determination, I measured the total glutathione levels in WT and Δ IS200 at LE-ES phases and found no significant difference in the intracellular glutathione levels; for both WT and Δ IS200 glutathione was present at roughly 10 μ M/mg protein (Figure 17). Ellman's reagent, or 5,5'-dithiobis(2-nitro-benzoic acid) (DTNB), has been adapted multiple times for enzymatic determination of glutathione^{80,81,88-90}. Ellman's reagent is an aryl disulphide compound where the addition of a thiol group, in the form of reduced glutathione, cleaves the disulphide bridge forming 5-thio-2-nitrobenzoic acid (TNB), which can be monitored at 405 nm. The rate at which TNB is formed (measured through colorimetric change) is proportional to the amount of glutathione present. This enzymatic method allows for specific determination of

glutathione thiols and has been adapted to be performed in a 96-well plate^{80,88}. The predominant form of glutathione found is oxidized glutathione (>98%) and is a large source of stored cysteine⁸⁷. As there was no difference in glutathione levels between Δ IS200 and WT strains, this measurement did not provide evidence supporting cysteine depletion in the Δ IS200 strain, at least in the single growth point examined. A limitation to this experiment is that it is reserved to a single growth point. Additional glutathione measurements at EE and ME phase would provide a complete understanding of any changes in the glutathione, and indirectly cysteine, levels in the Δ IS200 strain compared to WT.

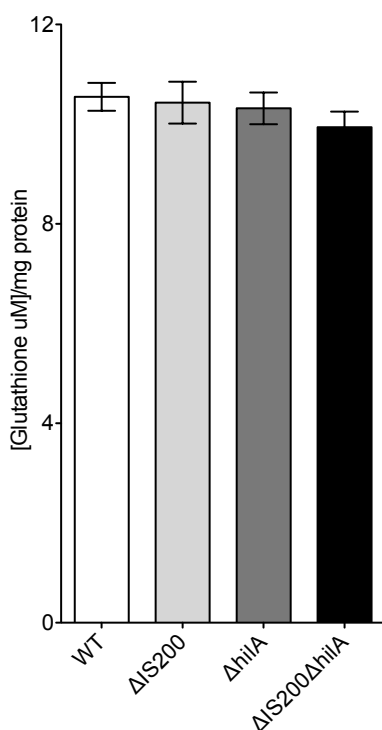


Figure 17. Glutathione levels in *Salmonella* SL1344 WT and Δ IS200 with and without Δ hilA at LE phase.

Total intracellular glutathione levels were measured using Ellman's reagent. Glutathione presented in μ M per mg of protein. Protein levels were determined via Bradford assay. Error shows the standard error on the mean of three biological replicates.

Next, I used *cysD* and *ydjN* GFP transcriptional fusions to monitor the Cys regulon induction in Δ IS200 strains that contained a second gene disruption that blocked the SPI-1 turn on (Δ hilA) or

flagellar cascade turn on ($\Delta flhDC$). As shown in Figure 18A, the early induction of *cysD* is seen in the $\Delta IS200$ strain (black dots) compared to the WT strain (grey dots). This leftward shift in the GFP expression profile was also observed in the $\Delta IS200\Delta hilA$ strain (blue dots) but not in the $\Delta hilA$ strain (red dots). Similar results were obtained when the reporter gene was *ydjN* seen in Figure 18B. Notably, blocking the SPI-1 induction through the removal of *hilA* in either the WT or $\Delta IS200$ strains also did not impact intracellular glutathione levels (Figure 17). These results taken together indicate that blocking the SPI-1 induction does not suppress the early Cys regulon induction occurring in the $\Delta IS200$ strain and accordingly does not support the metabolic burden model.

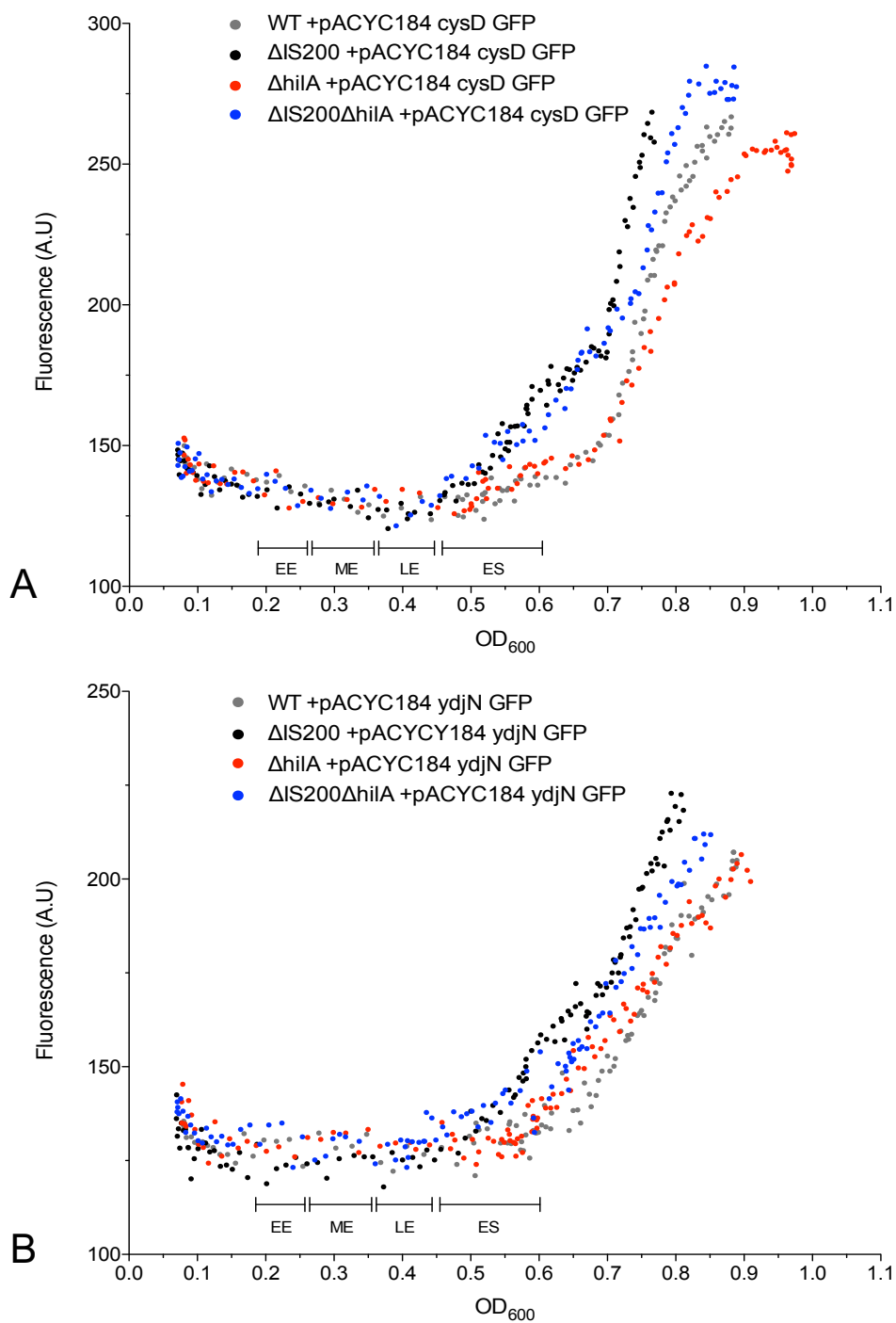


Figure 18. Removal of *hilA* on *cysD* and *ydjN* expression via GFP monitoring in *Salmonella* SL1344 WT and Δ IS200 strains grown in rich media.

pACYC184 transcriptional fusion to either (A) *cysD* or (B) *ydjN* promoter region upstream of the long-lasting form of GFP transformed into WT [grey], Δ IS200 [black], Δ hilA [red] or Δ IS200 Δ hilA [blue]. Cells were grown for 7

hours, diluted 50-fold into fresh media. Three biological replicates and two technical replicates were added to a 96-well plate and fluorescence was measured every 5 minutes for ~9.3 hours. Average values for 3 biological replicates are shown. Growth phases (defined in the text) are shown on the x-axis. Fluorescence (A.U) are shown on the y-axis. Note that the OD₆₀₀ reading was not adjusted for path length.

Finally, testing to see if a disruption in the flagellar cascade turn-on ($\Delta flhDC$) influences the timing of the Cys regulon induction in the $\Delta IS200$ strain using *cysD* and *ydjN* GFP transcriptional fusions showed that the dominant phenotype of $\Delta flhDC$ is removing the Cys regulon induction entirely. As seen in Figure 19, the early induction of *cysD* and *ydjN* observed in the $\Delta IS200$ strain (black dots) is seen by the leftward shift in GFP expression profile compared to the WT strain (grey dots). Interestingly, the disruption in *flhDC* resulted in no induction of either of the monitored *cysD* or *ydjN* genes in both the $\Delta flhDC$ and $\Delta IS200\Delta flhDC$ strains. This dominant phenotype where removal of *flhDC* removes the Cys regulon induction points to an unexpected connection between the two, which will be considered further below (see discussion). Thus, blocking the flagellar pathway via $\Delta flhDC$ cannot be used to measure any effect on reducing a possible metabolic burden associated with overexpressing this pathway in the $\Delta IS200$ strain.

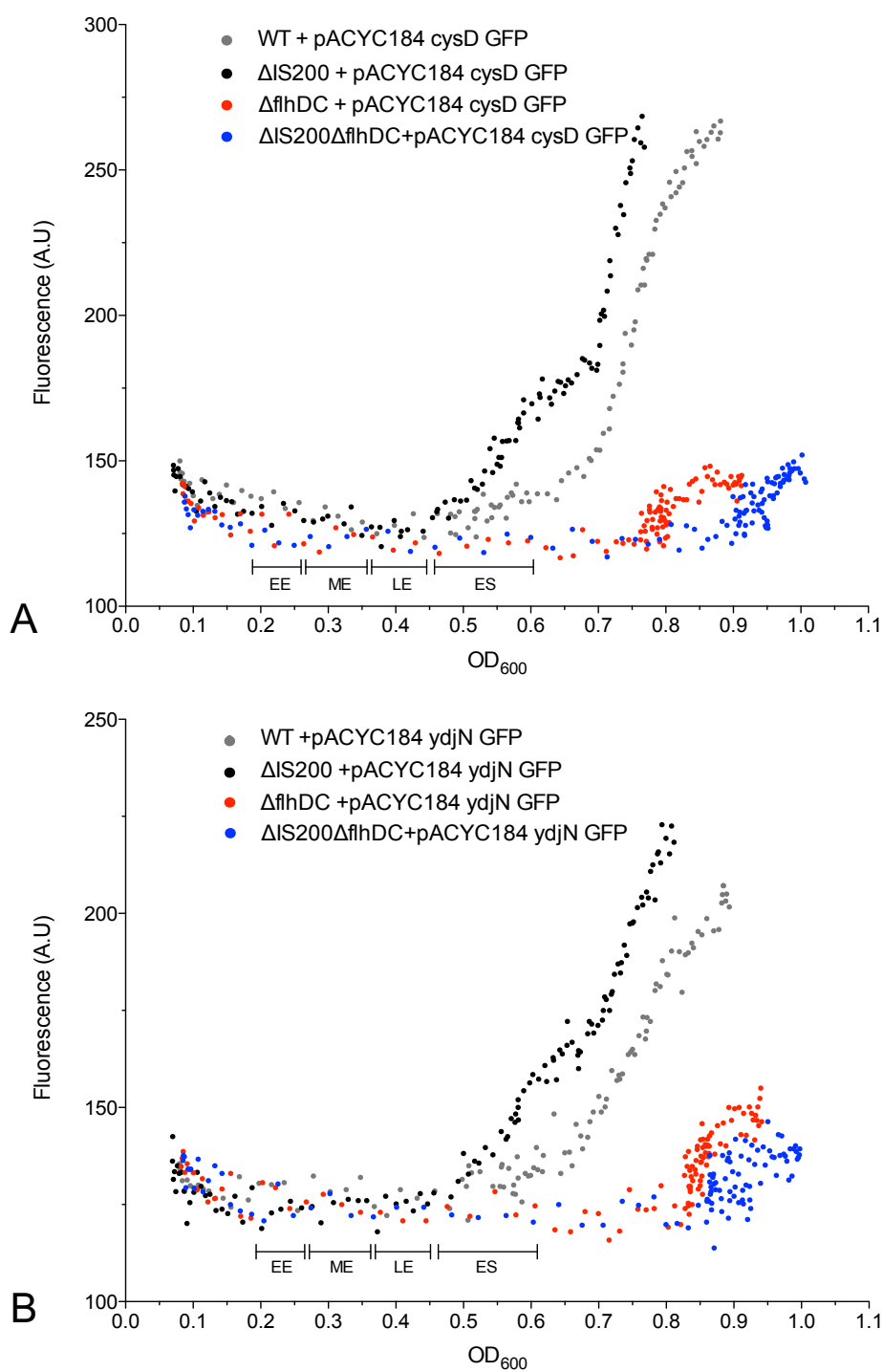


Figure 19. Removal of *flhDC* on *cysD* and *ydjN* expression via GFP monitoring in *Salmonella* SL1344 WT and Δ IS200 strains in rich media.

pACYC184 transcriptional fusion to either (A) *cysD* or (B) *ydjN* promoter region upstream of the long-lasting form of GFP transformed into WT [grey], Δ IS200 [black], Δ *flhDC* [red] or Δ IS200 Δ *flhDC* [blue]. Cells were grown for 7

hours, diluted 50-fold into fresh media. 3 biological replicates and 2 technical replicates were added to a 96-well plate and fluorescence was measured every 5 minutes for ~9.3 hours. Average values for 3 biological replicates are shown. Growth phases (defined in the text) are shown on the x-axis. Fluorescence (A.U) are on the y-axis. Note that the OD₆₀₀ reading was not adjusted for path length.

In conclusion, the *Salmonella* ΔIS200 strain undoubtedly has an added metabolic pressure with SPI-1 and flagellar cascades being overexpressed. The effects of exceptionally high protein synthesis, seen by the upregulation of downstream genes such as *invF* (SPI-1 cascade) and *motAB* (flagellar cascade), can be seen phenotypically through reduced colony size and overall reduced growth (Figure 16). An effort to test the pressures of the SPI-1 pathway by blocking its turn-on (*ΔhilA*) did not support the metabolic burden model. In addition, blocking the flagellar pathway resulted in the complete removal of *cysD* and *ydjN* expression in the *ΔflhDC* and ΔIS200Δ*flhDC* strains. This dominant effect does not allow for measuring any metabolic burden that overexpression of the flagellar pathway may cause. Therefore, in an effort to understand the mechanism behind Cys regulon induction at LE-ES phase in the ΔIS200 strain, I did not continue this line of inquiry. Instead, I proceeded to test and develop the third model presented below.

3.6 Model 3 – the LrhA deprivation model

The final model developed and tested during this project was the ‘LrhA deprivation’ model. LrhA is a less characterized global regulator belonging to the family of LTTRs and is known to repress *flhDC* transcription³⁷. Previous unpublished findings from our lab using qRT-PCR to look at differential gene expression and followed up with Western blotting showed that LrhA RNA and protein levels are reduced by about 30% in *Salmonella* SL1344 ΔIS200 compared to WT (data not shown). One possibility the Haniford lab is currently exploring is that the reduced levels of LrhA could be responsible for the early SPI-1 induction and increased expression of the flagellar pathway in the ΔIS200 strain. Notably, increased expression of FlhDC is expected to increase expression of both the flagellar and SPI-1 cascade genes because FlhDC is the master regulator of the former and indirectly responsible for producing the activator (FlhZ) of HilD³⁵, which is the master regulator of the SPI-1 cascade⁸. In this third and final model, I explore the possibility that LrhA is a negative regulator of Cys regulon genes. If this were the case, then the

reduced expression of LrhA in the Δ IS200 strain could account for the early induction of the Cys regulon genes (Figure 20). This possible connection between LrhA and the regulation of Cys regulon genes came serendipitously by including a Cys regulon gene (*cysD*) as a control in an LrhA ChIP experiment. It was found that there was a high occupancy of LrhA on the ‘front’ end of the *cysD* gene. In addition to validating the aforementioned ChIP result, I also asked if the expression of Cys regulon genes, namely *cysD* and *ydjN*, is repressed by overexpression of LrhA. As described below, my results support the possibility that LrhA binds to the front end of different Cys regulon genes and, in doing so, inhibits the expression of these genes.

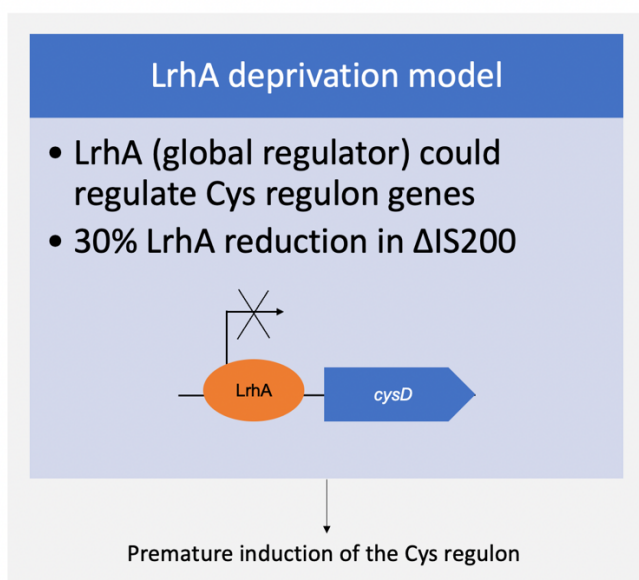


Figure 20. The LrhA deprivation model.

Depiction of the LrhA deprivation model (model 3) which proposes that because preliminary findings that LrhA can bind to Cys regulon genes, the reduction in LrhA levels could be responsible for the premature induction of the Cys regulon.

First, I set out to validate the previous LrhA-Cys regulon gene interaction via ChIP analysis by repeating the ChIP experiment, although in this case I looked for possible LrhA-Cys regulon interactions at multiple growth phases in WT *Salmonella* SL1344. In addition to looking for LrhA interactions with *cysD*, *cysP*, and *ydjN*, I also included *flhDC* as a positive control as it is known that LrhA represses *flhDC* transcription. In parallel experimentation, the Haniford lab looked at a strain with LrhA lacking the FLAG epitope tag in a ChIP analysis. This untagged

LrhA strain produced little signal. In conjunction with the untagged control strain, the 16S rRNA gene was included as a negative control for LrhA binding, and the 16S rRNA gene also showed a weak signal. In this experiment, the eluted samples were normalized to their input controls and then normalized to the non-FLAG tagged LrhA strain.

For the ChIP analysis presented here, the LrhA overexpression plasmid pBAD-LrhA-FLAG, which is induced in the presence of arabinose, was transformed into the WT strain. Cultures from EE, ME and ES phases were cross-linked with formaldehyde, the cells were lysed, and the DNA fragments were sheared by continuous agitation with glass beads. Anti-FLAG M2 MagBeads were equilibrated, and the lysed samples were incubated followed up with a series of washes (see Materials and Methods for further details). Protein-DNA complexes were eluted, and reverse cross-linking was performed by incubating with Proteinase K. Reverse cross-linking was also performed on the control pre-immunoprecipitated samples. Samples were purified using Qiaquick PCR Purification Kit (Qiagen), and 2 μ l was used for gene-specific PCR. Note that the amount of purified DNA was not quantified as this method has low yield. Quantifying would require concentrating the sample and reducing the number of gene-specific targets analyzed per sample. Primers were designed to amplify the promoter region of the intended target; that is, the primers are complementary to sequences of a couple hundred nucleotides upstream and downstream of the transcription start site.

I predicted that I would see LrhA bind to the same Cys regulon genes with a greater affinity in earlier growth when the Cys regulon is actively being repressed by CysB and potentially LrhA. I show in Figure 21 that LrhA binds to all Cys regulon targets examined (*cysP*, *cysD* and *ydjN*). Table 1 details the amount of PCR product found in the eluted (E) samples normalized to the input pre-immunoprecipitated (PI) control. When comparing LrhA binding to the positive control target (*flhDC*), in EE the normalized occupancy amounts are roughly the same (0.95 [*cysD*]) or greater than (1.97 [*cysP*] and 2.66 [*ydjN*]) than 1.04 for *flhDC*. However, the higher normalized ChIP signal for *flhDC* at ME phase (5.7 versus 2.68 [*cysP*], 2.4 [*cysD*] and 3.12 [*ydjN*]) is consistent with LrhA having a higher affinity for the *flhDC* promoter than for any of the other promoter regions tested. Notably, at ES phase, the occupancy level for the Cys regulon

genes is much lower than *flhDC* (less than 0.84 compared to an occupancy of 2.05 for *flhDC*). Thus, in general my results both support the idea that LrhA binds to Cys regulon promoters.

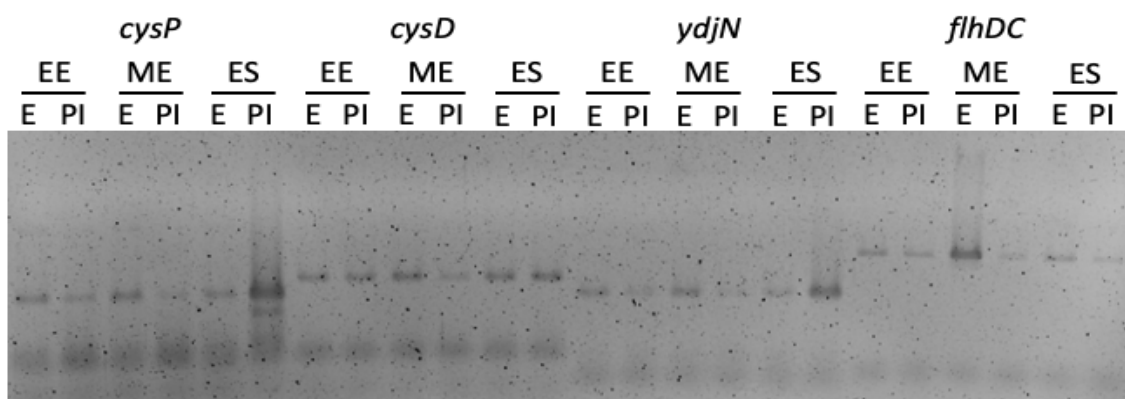


Figure 21. ChIP analysis of *Salmonella* SL1344 WT showing LrhA binding to select Cys regulon genes at EE, ME, and ES phase.

Gene specific PCR was performed for *cysP*, *cysD*, *ydjN* and *flhDC* on eluted protein-DNA (E) and pre-immunoprecipitated inputs (PI) samples.

Table 1. Normalized signal of LrhA-DNA targets.

Eluted signal normalized to the pre-immunoprecipitated input rounded to the nearest hundredth. Normalized signals can be compared across targets within a phase but not across phases.

Growth phase	Gene target	Normalized signal (E/PI)
EE	<i>cysP</i>	1.97
	<i>cysD</i>	0.95
	<i>ydjN</i>	2.66
	<i>flhDC</i>	1.04
ME	<i>cysP</i>	2.68
	<i>cysD</i>	2.40
	<i>ydjN</i>	3.12
	<i>flhDC</i>	5.70
ES	<i>cysP</i>	0.43
	<i>cysD</i>	0.84
	<i>ydjN</i>	0.32
	<i>flhDC</i>	2.05

Next, I asked if LrhA can influence the expression of Cys regulon genes. An LrhA overexpression plasmid (pDH1092 LrhA) was constructed on a low-copy plasmid containing the native LrhA promoter and was transformed into *Salmonella* SL1344. The level of LrhA expression at the RNA level was determined for WT/pDH1092 (LrhA) versus WT/pDH1093 (empty vector) by qRT-PCR and found to be 30-fold higher in the former strain (data not shown). A second compatible plasmid encoding a Cys regulon gene, either *cysD* or *ydjN*, fused to GFP was introduced into the above strains. The level of GFP expression was monitored against the growth phase using the method previously described. These assays were done with three biological and two technical replicates. The read outputs were averaged, and emission was plotted versus OD₆₀₀, which ranged from 0.1 to 1.0 without a path length corrector.

I show in Figure 22 that both *cysD* and *ydjN* expression is repressed in the presence of overexpressed LrhA. *cysD* has an overall greater expression compared to *ydjN*, where the emission peaks are greater by about 50 units when comparing pDH1092 LrhA *cysD* (grey) to pDH1092 LrhA *ydjN* (red), and pDH1093 EV *cysD* (black) to pDH1093 EV *ydjN* (blue) (Figure 22). Furthermore, while in both cases with the overexpression of LrhA present (pDH1092 LrhA), the induction of the Cys regulon gene occurs at OD₆₀₀~0.6, there is an earlier diverging point of *ydjN* expression in the cells containing the empty vector (blue dots). That is, for *ydjN* the expression diverges early on around OD₆₀₀~0.25 compared to *cysD* at OD₆₀₀~0.5. In conclusion, I show that the overexpression of LrhA results in the repression of *cysD* and *ydjN* expression monitored via GFP. This along with the ChIP analysis supports model 3, which proposes that LrhA binds to Cys regulon genes and inhibits their expression and that the reduced levels of LrhA expression in the Δ IS200 strain could account for the premature induction of the Cys regulon.

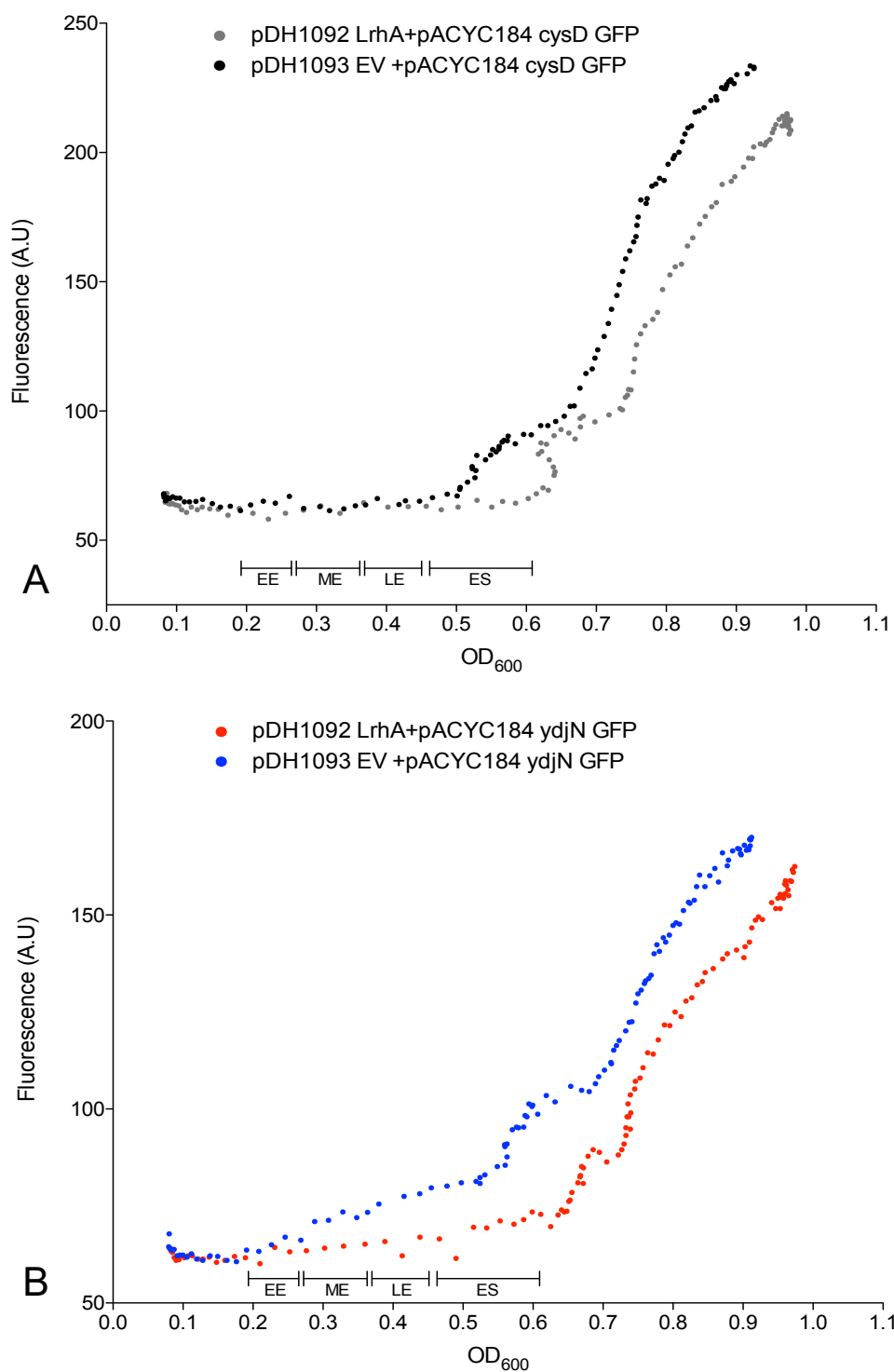


Figure 22. GFP monitoring of *cysD* and *ydjN* expression in *Salmonella* SL1344 WT in the presence of LrhA overexpression.

cysD and *ydjN* show repressed expression in the presence of LrhA overexpression in WT strain grown in rich media. pACYC184 transcriptional fusion contained the promoter region of either (A) *cysD* or (B) *ydjN* upstream of the

long-lasting form of GFP. pDH1092 LrhA overexpression plasmid [grey and red] or pDH1093 empty vector [black and blue] were transformed into WT. Cells were grown for 7 hours and diluted 50-fold into fresh Lennox media. 3 biological and 2 technical replicates were added to a 96-well plate and fluorescence was measured every 5 minutes for 9.3 hours. Average values for 3 biological replicates are shown. Growth phases (define in text) are shown on the x-axis and fluorescence (A.U) is shown on the y-axis. Note that the OD₆₀₀ reading was not adjusted for pathlength.

4 Discussion

In the current work, I have shown that deletion of the 5' portion of all 7 copies of IS200 in *Salmonella* SL1344 results in the premature induction of the Cys regulon. I have characterized this induction, and I show that it occurs at LE-ES phase, it is dependent on the presence of CysB, and can be suppressed with additional cystine supplemented. Its dependence on CysB provides evidence that IS200 is not affecting any non-canonical mechanisms of activation. In addition, the Cys regulon induction depends on cytoplasmic cysteine levels being low, and I show that cystine supplementation represses the premature Cys regulon induction. The apparent depletion of the intracellular cysteine pool at LE-ES phase in Δ IS200 cells results in increased cell sensitivity to oxidative stress. I have developed and tested three possible models explain how the removal of IS200 may impact the regulation of the Cys regulon. The three models I have presented are the 1) futile import/export model, 2) the metabolic burden model, and 3) the LrhA deprivation model.

First, I provide evidence that the upregulation in the cystine transporter *fliY-yecSC* could stem from increased *fliA* levels through RT-PCR transcript mapping. In addition, I show that the upregulation of *fliY* and *yecC* at the RNA level is carried to the protein level. Conflicting Δ *fliY* results were presented between qRT-PCR and GFP growth analysis and will be discussed further. Second, the metabolic burden model provided little insight as there were no changes in the glutathione levels between the WT and Δ IS200 strains at LE phase. The removal of *hilA* did not alter the premature Cys regulon induction nor the glutathione levels. Curiously, the removal of *flhDC* eliminated a Cys regulon response from both WT and Δ IS200 strains when monitoring *cysD* and *ydjN* expression through GFP growth analysis. This points to a deeper connection between the flagellar pathway and cysteine metabolism and is discussed below. Finally, the

LrhA deprivation model provides promising initial results. I provide evidence that the global regulator LrhA binds to multiple Cys regulon genes, and overexpression of LrhA leads to repression of *cysD* and *ydjN*. This discussion will further expand on the limitations and findings provided in each model and possible steps for future research where needed.

4.1 Δ IS200 leads to a premature induction of the Cys regulon

In a WT strain, the induction of the Cys regulon occurs in late stationary phase. Previous findings by the lab have seen that in the Δ IS200 strain, there is a premature induction of the Cys regulon at LE phase (Supplemental Table 6). I first validated the occurrence of Cys regulon induction by monitoring gene expression through two different methods: qRT-PCR and GFP transcriptional fusion. Figure 4 shows a wide range of gene targets of interest, and in the Δ IS200 strain, there is an upregulation of Cys regulon genes ranging from 2.5- to 90-fold at LE phase. To complement this finding and learn the precise timing in which the Cys regulon induction differs in Δ IS200 compared to WT, I performed growth curve analysis monitoring the expression of *cysD* and *ydjN* through GFP transcriptional fusions (Figure 5). *cysD* and *ydjN* were chosen because they represent different parts of the cysteine metabolism, with *cysD* involved in sulphur assimilation and part of an operon, and *ydjN* involved in cystine transport and represents a standalone gene. I show that the expression of these genes is induced earlier, as seen by the leftward shift in the expression curves. The GFP growth curve analysis supports the qRT-PCR analysis and validates the occurrence of a premature Cys regulon induction. The GFP growth curve analysis was limited to *cysD* and *ydjN*, creating an assumption that these genes are accurate reporters of the entire Cys regulon. This assumption is based on the regulation of the Cys regulon (Figure 2B), the different roles of *cysD* and *ydjN*, and their increase in expression seen in the qRT-PCR analysis (Figure 4). Finally, I check the expected dependencies of the Cys regulon induction. I expect that Cys regulon induction reflects cysteine shortage, and I tested this by adding exogenous cystine. In addition, I expect that the Cys regulon induction is dependent on CysB regulatory control, and I tested this through Δ *cysB*. These dependencies were confirmed and show what was expected; providing exogenous cystine suppressed the Cys regulon induction, and removal of CysB removes any expression of the monitored Cys regulon genes except for *fliY*.

4.2 IS200 is predicted to be a critical component of *Salmonella* fitness during macrophage invasion

IS200 is a highly conserved transposon throughout enteric bacteria and is found in various copy numbers amongst the *Salmonella* species. Despite its high conservation, it rarely transposes, which points to it being important for maintaining it in bacterial populations. During salmonella infection process when invading phagocytotic cells, salmonella must deal with reactive oxygen species (ROS) as one of the methods of attack used by macrophages to cause cell damage⁹¹.

Here I show that when IS200 is disrupted, the cells become sensitive to oxidative stress (Figure 8) at LE phase when there is a depletion of available cysteine. Cysteine is a critical amino acid in several antioxidants (e.g. glutaredoxins and thioredoxins) and is important for maintaining thiol pools and redox potential⁸⁴. In addition, it has been shown in *E. coli* that the cystine transporter FliY-YecSC has been implemented in the cyclic shuttle system that converts cystine to cysteine in the periplasm, which provides reducing power as a defence against oxidative stress⁸⁵. They show that causing oxidative stress through the addition of hydrogen peroxide induces the use of this cystine/cysteine shuttle system. Cystine is imported from the periplasm via FliY-YecSC, reduced intracellularly, and exported back into the periplasm as cysteine that can then be oxidized back to cystine, thus providing reducing power to combat oxidative stress in the periplasm. This cyclic shuttle of cystine/cysteine between the inner membrane is proposed to be a preventative measure against lipid peroxidation and defence against cell damage⁸⁵. Therefore, I asked if the Δ IS200 strain is at a disadvantage in the presence of oxidative stress.

Given that there is a premature induction of the Cys regulon, which has been shown to be due to reduced cysteine levels (Figure 6), I expected that Δ IS200 would not have enough cysteine/cystine to utilize this cyclic shuttle system as a defence and therefore would be at a disadvantage compared to WT. This was confirmed in Figure 8, where at LE phase, when the Cys regulon induction occurs in Δ IS200 strain but not WT, the number of CFU is significantly lower ($p < 0.05$) for Δ IS200 strain treated compared to WT treated. In contrast, there is no significant difference in the surviving CFU between Δ IS200 and WT strains at the ME phase (where cysteine levels have not yet been depleted) when comparing treatments. This suggests that IS200 would be a critical component of salmonella fitness during invasion when they are

dealing with ROS from the host defences. It would be interesting to observe the survival of Δ IS200 compared to WT in macrophage invasion assays, and I would expect, based on these preliminary findings, that Δ IS200 have a lower rate of survival. This continued line of inquiry would further support IS200 element as a necessary component for *Salmonella* invasion.

4.3 Upregulation in the flagellar pathway contributes to *fliY* expression in Δ IS200

The disruption of IS200 in *Salmonella* SL1344 leads to overexpression of flagellar (FlhDC and FliA) and SPI-1 invasions (HilD) pathways. A possible indirect consequence of FlhDC upregulation is the constitutive overexpression of the FliY-YecSC cystine transporter. The first model (futile import/export model) proposes that the induction of the Cys regulon is driven by the FliY-YecSC overexpression that promotes a futile cycle of cystine import and cysteine export (or conversion to other products). This futile cycle is proposed to lead to a depletion of cytoplasmic cysteine levels. In support of this model, I provide evidence that increased *fliY* expression in the Δ IS200 strain can be driven by increased levels of FliA through examining its expression in Δ *cysB* (Figure 7 and Figure 11B) despite the *fliY* gene possessing a CysB-regulated promoter. This finding is supported by work in the literature where an early study of the *fliA* operon in *Salmonella* shows products in a Northern blot of 2300 and 1400 nucleotides in size, representing *fliA-fliZ-fliY* and *fliA-fliZ* being co-transcribed, respectively³⁰. The *fliA-fliZ-fliY* product was much weaker in signal than *fliA-fliZ*. Noted by the authors, this indicated that the majority of transcription from the *fliA* promoter terminates between *fliZ* and *fliY*. This work was also the first to provide evidence that *fliY* was transcribed from a non-flagellar promoter³⁰. A more recent study conducted in *E. coli* that studied the regulation of the downstream genes (i.e. *fliY*, *yedO*, *yecS*, *yecC*) noted that the predominate transcriptional start site for *fliY* is its non-flagellar promoter, which CysB activates⁴³. These results, taken together with previous findings in the literature, support model 1, where increased *fliY* transcription could be from readthrough from *fliA* in Δ IS200 when there is upregulation of the flagellar pathway.

Interestingly, the units of an ABC-transporter are separated by a gene (*yedO*; encoding cysteine desulphhydrase with overlapping functions of MetC) (Figure 2C), and research in *E. coli* has concluded that the regulation of *yecSC* comes from the *yedO* promoter with little read-through

from the *fliY* promoter⁴³. To my knowledge, this ‘break’ in gene organization is an exception within ABC-transporters and provides a puzzling regulatory organization in Δ IS200. That is, upregulation of *fliA* leads to read-through to *fliY*, but *fliY*’s transporter *yecSC* regulation comes from *yedO*, which is not under CysB regulatory control nor FlhDC control. Notably, another ABC transporter involved in cysteine metabolism is transcribed in the *cysPUWA* operon⁹². *cysPUWA* imports sulphate and thiosulphate, where *cysP* is the periplasmic-binding protein, *cysUW* are components of the TMD, and *cysA* is the NBD which binds and hydrolyzes ATP⁹³. Unlike FliY-YecSC, its gene organization is not fragmented with multiple transcriptional start sites and is co-transcribed under CysB activation. This uniqueness in gene organization where *fliY* is seemingly attached to the FliA operon but also has a CysB-regulated promoter and is separated from its ABC-transporter is puzzling. In examining the role of YedO, researchers have proposed that this gene organization points to a mechanism of defence against cysteine toxicity while allowing cysteine to be used as a sulphur source⁶⁰. YedO, whose transcription is induced under sulphate limiting conditions, is bordered by FliY and YecSC, whose expression is also induced under sulphate limiting conditions. Thus, the turn-on of FliY-YecSC (which has a higher affinity for cystine than YdjN) to import cystine paired with YedO allows for the use of cysteine (once cystine is imported and reduced) as a sulphur source in limiting sulphate conditions and transcribing YedO and prevents cysteine toxicity within the cell⁶⁰.

In continuing to test model 1, I measured the protein amounts of FliY (the periplasmic-binding protein) and YecC (NBD) via Western blotting. I chose to look not only at FliY but YecC as well, where upregulation of the periplasmic-binding protein would not alone support model 1. For ABC importers, the periplasmic-binding protein (FliY) is utilized to deliver substrates to the transporters. It is typically specific to one transporter, and therefore it also dictates its specificity. FliY-YecSC is known to transport not only L-cystine but D-cystine, L-djenkolate, lanthionine, homocystine, and diaminopimelate, which are all structurally similar to the cystine dimer^{43,85,94,95}. I found that consistent with the upregulation of *fliY* and *yecC* at the RNA level, there was also an upregulation at the protein level (Figure 12). This increase in FliY and YecC provided further validation that the upregulation of FliY-YecSC could drive a futile import/export cycle and lead to a premature induction of the Cys regulon in the Δ IS200 strain.

Finally, in testing the effects of removing *fliY* in the Δ IS200 strain, I found conflicting results between the qRT-PCR analysis and monitoring the expression of *cysD* and *ydjN* through GFP transcriptional fusions. qRT-PCR analysis at LE phase showed a partial loss of the Cys regulon induction in the Δ IS200 Δ *fliY* (Figure 13). A loss of Cys regulon induction would indicate that *fliY*, and therefore cystine transport via FliY-YecSC, is essential in driving the depletion of cysteine and inducing premature expression of the Cys regulon. In contrast, the GFP growth curve analysis showed no difference in *cysD* and *ydjN* expression when comparing Δ *fliY* versus WT and Δ IS200 versus Δ IS200 Δ *fliY* strains (Figure 14). These conflicting results did not give confidence in model 1. qRT-PCR analysis has been used for gene expression analysis throughout this project as the Haniford lab has expertise in this method. I complement qRT-PCR by monitoring *cysD* and *ydjN* GFP transcriptional fusions via a growth curve, which provides additional validation in an alternative method. Accordingly, there are discrete differences and select advantages to either method, like qRT-PCR allowing for multiple gene targets in its analysis, whereas GFP fusions are limited to monitoring a single target. A significant limitation to the qRT-PCR method is that it is confined to examining single time points. Comparatively, GFP fluorescence allows for temporal monitoring of gene expression, which allows for a more precise indication of when the gene expressions differ in WT versus Δ IS200. But, as the GFP construct was only functional as the long-lasting form of GFP, these assays are limited to providing the timing in which the Δ IS200 strain differs, but there is no information on the turn-off of these genes. Due to the conflicting results presented by the qRT-PCR and GFP analysis, one could repeat such analyses and push the qRT-PCR to a later growth point (ES phase) when a greater difference in expression is observed. Throughout this project, I have chosen a select few genes belonging to the Cys regulons to act as ‘reporters’ of the regulon: *cysP*, *cysD*, *ydjN*, *yciW*, *cysB*. As GFP-transcriptional fusions are limited in a single target, one could note that monitoring *cysD* and *ydjN* are not fully representative of the regulon as a whole. In addition to repeating Δ *fliY* qRT-PCR, constructing GFP-transcriptional fusions to alternative Cys regulon genes would provide a broader test of the Cys regulon response.

4.4 Preliminary evidence points to a connection between the flagellar pathway and cysteine metabolism

In model 2, I explored the possibility of a metabolic burden because of the exceptionally high expression of the flagellar and SPI-1 pathways. High expression of these pathways is seen through a 2- and 5-fold increase of *flhDC* and *hilD* (the master regulators of these pathways), respectively, in addition to seeing downstream genes being upregulated. As a result, the Δ IS200 strain has reduced growth (Figure 16) and shows a reduction in colony size compared to WT *Salmonella* SL1344. I tested the possibility that induction of these pathways would lower cytoplasmic cysteine pools by indirectly measuring cysteine levels. As flagellar and SPI-1 pathways are fully induced in the Δ IS200 strain at LE phase, I focused on this growth phase to measure cysteine levels. As I did not have the capability to measure cysteine directly, I chose to measure glutathione for reasons previously explained. I show that I did not see evidence of differences in intracellular glutathione levels between Δ IS200 and WT strains (Figure 17). In retrospect, it may have been worthwhile to extend this analysis to other growth phases, but time constraints did not allow for this. In addition, alternative methods of measuring cysteine levels, such as mass spectrometry, should be considered.

While I was unable to support model 2 by showing that flagellar and invasion pathway induction coincides with cytoplasmic pool depletion, I provided a strong test of this model by asking if early Cys regulon induction occurred in genetic backgrounds where SPI-1 induction cannot occur; that is, a double mutant where the *hilA* gene is knocked out in the Δ IS200 strain. I show that blocking the SPI-1 invasion pathway, thus relieving a large amount of protein synthesis associated with this pathway, does not change the premature induction of the *cysD* and *ydjN* (Figure 18). In addition, blocking the SPI-1 pathway does not alter the glutathione levels (Figure 17). These results did not support the metabolic burden model. Next, I could not further assess the validity of model 2 by evaluating the impact of preventing the flagellar pathway induction by introducing an Δ *flhDC* mutation into the Δ IS200 strain. Interestingly, this mutation has a dominant-negative impact on Cys regulon induction (Figure 19). The lack of Cys regulon induction in Δ *flhDC* in WT and Δ IS200 backgrounds poses an exciting and surprising connection between the flagellar pathway and cysteine biosynthesis. Note, as this was only performed by GFP transcriptional fusions, additional method of validation should follow. A

known connection between these two pathways comes in the form of *fliY*, the cystine periplasmic-binding protein located directly downstream of *fliAZ* (Figure 2C). The elimination of *cysD* and *ydjN* expression in Δ *flhDC* points to either 1) a flagellar transcription factor or product is needed to induce the Cys regulon; 2) motility is required to induce the Cys regulon, or 3) a functioning flagellar pathway is required for functional GFP. While option 3 is unlikely, I consider options 1 and 2 as follows.

The requirement of a flagellar transcription factor to induce the Cys regulon could be possible. The canonical mechanism of Cys regulon activation via CysB is well-studied^{49,50,52,96}. CysB negatively autoregulates itself when not activating the Cys regulon. This negative autoregulation occurs through CysB binding to the *cysB* promoter, thus preventing its transcription. This transcriptional repression is removed in the presence of OAS, and CysB-OAS will positively regulate the genes belonging to the Cys regulon (Figure 2B). Interestingly, a non-canonical mechanism of activation has been reported under nitrosative stress conditions, the overproduction of nitric oxide radicals (\bullet NO). The regulatory protein of RNA polymerase, DksA, induces anti-nitrosative defences and can induce the activation of the Cys regulon⁹⁷. Research has shown that cysteine biosynthesis is upregulated when exposed to reactive nitrogen species as a mechanism of supplying additional cysteine for cysteine-containing proteins and replenishing glutathione pools to maintain redox balance. It was found that cysteine biosynthesis was not upregulated in a Δ *dksA* mutant in *Salmonella*, and *cysD* could be activated by DksA when exposed to reactive nitrogen species. This activation by DksA is reliant on CysB activation⁹⁷. Despite this activation mechanism occurring due to a stressor, it sets a precedent that alternative activation mechanisms for the Cys regulon exist. To follow up on the Δ *flhDC* finding in looking for a connection between flagellar pathway and cysteine metabolism, I would ask if removing FliA would also eliminate a Cys regulon induction. Recall, FliA is a class 2 flagellar gene, but it also produces the flagellar sigma factor. Monitoring the Cys regulon genes (*cysD* and *ydjN*) in a Δ *fliA* knockout, perhaps under different stress conditions, would test the potential importance of FliA expression in controlling the Cys regulon.

Next, the possibility that motility is required to induce the Cys regulon seems counterintuitive. That is, what does motility provide in signalling to the cell to induce biosynthesis pathways. One

would think that being non-motile, therefore unable to move to the more desirable environments for nutrient resources, would induce rather than repress a biosynthetic pathway. To test this, one could monitor *cysD* and *ydjN* expression and perform qRT-PCR in a Δ *motAB* strain. MotAB encodes the components that generate the force in the flagella⁹⁸; therefore, flagella are produced in a Δ *motAB* strain, but the cells are paralyzed and non-motile. Therefore, I would expect that if motility is required for Cys regulon induction, a Δ *motAB* strain will produce the same outcome as an Δ *flhDC* such that when monitoring select Cys regulon genes, there would be no expression.

4.5 LrhA binds to multiple Cys regulon genes

The final model, the ‘LrhA deprivation’ model, where LrhA mRNA and protein levels are reduced by 30% in the Δ IS200 strain, leads to an increased expression of the Cys regulon genes at LE-ES phase. I utilize ChIP to analyze *in vivo* DNA-protein interactions for detecting any interactions between LrhA and Cys regulon genes. ChIP assays are widely used as an *in vivo* method to study protein-DNA interactions^{99–101}. Traditional *in vitro* methods include DNA footprinting, electrophoretic mobility shift assay and yeast 1-hybrid systems. While these methods are still used today and provide their advantages, they are limited to *in vitro* conditions. For ChIP assays, DNA-protein interactions are fixed via crosslinking. Formaldehyde as the crosslinking agent is commonly used for its simplicity and effectiveness as it does not require any specific conditions and can be added directly into the culture medium^{99,102}. For native ChIP analyses, crosslinking can be omitted and provides advantages such as keeping the protein-DNA interactions in their native state. However, native ChIP is limited to strong DNA-protein interactions and is rarely used outside of histone interactions; thus, most ChIP analyses use the crosslinking method¹⁰⁰. Crosslinking has its limitations, including trapping any non-specific binding that is occurring, thus creating the possibilities of false positives. ChIP analysis also has a low precipitation yield, leading to an additional limitation of requiring a large number of cells.

The serendipitous finding the Haniford lab discovered by including *cysD* as an intended negative control for LrhA binding but produced otherwise gave rise to model 3. This finding led me to predict that validation of this analysis would show that in addition to *cysD*, LrhA would bind to other Cys regulon genes. I validated this initial finding by showing through repeat ChIP analysis

that LrhA, a global repressor in *Salmonella* and known to repress *flhDC*, binds to *cysP*, *cysD* and *ydjN* in WT *Salmonella* SL1344 over three different growth points (Figure 21).

Briefly discussed above (see results), the Haniford lab has tested and validated a non-FLAG tagged LrhA strain to measure any background signal. This test produced very weak signalling where the normalized occupancy (E/PI) would be ~ 0 ; that is, virtually no signal was present in the eluted sample in a non-FLAG tagged LrhA strain. In addition to performing a strain control, the Haniford lab looked at a negative control for LrhA binding (16S) and found a similar result where occupancy reported as E/PI would be ~ 0 at LE phase. Contrary, I show that *flhDC* (positive control) reaches a peak occupancy of 5.70. In the same precipitated sample at ME phase, the occupancy of LrhA binding to *cysD*, *ydjN* and *cysP* are 2.40, 3.12 and 2.68, respectively (Table 1). Except for *cysD* at 2.40, *ydjN* and *cysP* signal occupancies are roughly half of what is seen in the positive control, indicating that signals obtained are likely to be weak specific binding instead of a non-specific binding as the latter has a signal of roughly 0 when checking 16S binding. Given that *cysP* and *ydjN* have signal occupancies about 50% of *flhDC* in that sample, it is likely that, despite its slightly lower occupancy at 2.40, the *cysD* signal also represents weak specific binding. In this analysis, the FLAG-tagged LrhA is overexpressed; this begs the question of whether LrhA binding to Cys gene promoters would still be detected above the non-specific background at native levels of LrhA expression. It would be beneficial for additional ChIP analyses performed with LrhA under its native promoter, and instead of testing a non-FLAG tagged strain, allow for a preincubation with beads without the antibody to reduce non-specific binding. In addition to showing that LrhA can bind to select Cys regulon genes, I also demonstrate using transcriptional Cys regulon gene GFP reporters that overexpression of LrhA can effectively repress *cysD* and *ydjN* expression (Figure 22). This supports the idea that LrhA can be a repressor of the Cys regulon.

4.6 LrhA could extend its control over the Cys regulon through multiple mechanisms

LrhA's potential role in the premature induction of the Cys regulon could extend further than being a repressor of the Cys regulon, where reduced levels of LrhA provides an increased chance of activation. As previously discussed (see introduction), because LrhA is a repressor of *flhDC*,

the Haniford lab is exploring the possibility that reduced LrhA levels are responsible for the early induction of the flagellar pathway, and consequently the SPI-1 invasion pathway in the Δ IS200 strain. Given that the flagellar product FliZ, post-translationally stabilizes HilD^{12,16,35}, it is expected that an increased expression of FlhDC increases the flagellar pathway and the SPI-1 invasion cascade. This leads to the possibility that LrhA potentially repressing the Cys regulon via both directly (through LrhA binding) and indirectly (as a consequence of influencing the Δ IS200 phenotype). That is, increased flagellar and SPI-1 pathways cause a multitude of transcription factors to be turned on, which could provide alternative forms of activation as seen by inducing nitrosative stress. The activation of the Cys regulon in standard conditions (i.e., no environmental stressors) is induced in low sulphate conditions. In standard conditions, there is a baseline expression of the Cys regulon when induced. Given the evidence that LrhA can bind to Cys regulon genes and is a part of the LTTR family, suppose that LrhA potentially competes with CysB-OAS to bind to the promoter regions of Cys regulon genes. If there is a reduction in LrhA, as in Δ IS200, this could favour CysB-OAS occupancy of Cys regulon promoters and raise the overall Cys regulon expression. In addition to competing with CysB-OAS to interfere with Cys regulon gene expression, it is also possible that LrhA could play a role in preventing other possible positive regulators from binding to Cys regulon genes. On the other hand, given the proposed impact that reduced LrhA levels have on flagellar and SPI-1 pathways, it could be possible that a large part of LrhA's role in Cys regulon induction seen in Δ IS200 is due to an increase in a non-canonical activator that is being overexpressed in these conditions. Therefore, implementing LrhA as an important factor in this premature induction of the Cys regulon could stem from multiple mechanisms of action including direct LrhA-Cys gene repression proposed here.

5 Conclusion

I have shown that a disruption in the 5' end of IS200 leads to a premature induction of the Cys regulon in *Salmonella* SL1344. This phenomenon was dependent on the Cys regulon activator, CysB, through testing a Δ cysB strain. In addition, I showed that the premature induction of the Cys regulon indicated a cysteine shortage in the Δ IS200 strain. In characterizing the impact that

Δ IS200 has on the cell, I tested and showed that at LE phase Δ IS200 has greater sensitivity to oxidative stress than WT due to the lack of readily available cysteine. This sensitivity was not seen in earlier growth phases when there is no difference in Cys regulon expression between Δ IS200 and WT. When put into a biological context, these findings show the importance of IS200 in *Salmonella* host fitness when encountering oxidative stress such as ROS, which happens during the invasion of phagocytotic cells (e.g. macrophages and dendritic cells). I presented three possible models developed and tested: the futile import/export model, the metabolic burden model, and the LrhA deprivation model. This first model provides evidence that upregulation of the cystine transporter FliY-YecSC is carried to the protein level, and the upregulation in *fliA* transcript could extend downstream to contribute to *fliY* upregulation. I also present evidence that *fliY* transcription can be CysB-independent. I also provided evidence that the upregulation of the components of the FliY-YecSC cystine transporter at the mRNA level translated to the protein level by monitoring FliY and YecC levels. These findings, along with the precedent of unregulated cystine import through the second transporter (YdjN) in *E. coli*, supported the proposed futile import/export model. Although, the lack of consistency in the Δ *fliY* testing led to the pursuit of other models. In testing the second model, the metabolic burden model, I presented evidence that Δ IS200 leads to a decreased growth but no change in glutathione levels at LE phase compared to WT. In asking if blocking either the flagellar or SPI-1 pathways affects the outcome of when the Cys regulon is induced, I found the latter does not alter the premature induction of the Cys regulon and that the former eliminated the Cys regulon induction in both WT and Δ IS200 backgrounds. Finally, of these three models, the LrhA deprivation model in which the reduced levels of LrhA are proposed to lead to the premature induction of the Cys regulon seen in Δ IS200, provided the most promising outcome. LrhA was found to bind to multiple Cys regulon genes, and overexpression of LrhA was seen to repress *cysD* and *ydjN* expression. LrhA's potential role in causing the upregulation of flagellar and SPI-1 pathways characteristic in an Δ IS200 background leads to possibly implementing a broader role for LrhA, more than acting as a repressor of the Cys regulon, that contributes to the premature induction of the Cys regulon in the Δ IS200 strain. Furthermore, understanding the effect of IS200's role in regulating vital pathogenicity pathways contributes novel findings for how it has integrated itself into regulatory mechanisms of a medically relevant pathogen.

Reference

1. Stanaway, J. D. *et al.* The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect. Dis.* **19**, 1312–1324 (2019).
2. M, R. *et al.* Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nat. Immunol.* **2**, 361–7 (2001).
3. Hurley, D., McCusker, M. P., Fanning, S. & Martins, M. Salmonella-Host Interactions - Modulation of the Host Innate Immune System. *Front. Immunol.* **5**, 1–11 (2014).
4. Velge, P. *et al.* Multiplicity of Salmonella entry mechanisms, a new paradigm for Salmonella pathogenesis. *Microbiologyopen* **1**, 243–258 (2012).
5. Boumart, Z., Velge, P. & Wiedemann, A. Multiple invasion mechanisms and different intracellular Behaviors: A new vision of Salmonella-host cell interaction. *FEMS Microbiol. Lett.* **361**, 1–7 (2014).
6. Hensel, M. Evolution of pathogenicity islands of Salmonella enterica. *Int. J. Med. Microbiol.* **294**, 95–102 (2004).
7. Saroj, S. D., Shashidhar, R., Karani, M. & Bandekar, J. R. Distribution of Salmonella pathogenicity island (SPI)-8 and SPI-10 among different serotypes of Salmonella. *J. Med. Microbiol.* **57**, 424–427 (2008).
8. Lou, L., Zhang, P., Piao, R. & Wang, Y. Salmonella Pathogenicity Island 1 (SPI-1) and Its Complex Regulatory Network. *Front. Cell. Infect. Microbiol.* **9**, 1–12 (2019).
9. Hensel, M. Salmonella pathogenicity island 2. *Mol. Microbiol.* **36**, 1015–1023 (2000).
10. Gerlach, R. G. *et al.* Salmonella Pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. *Cell. Microbiol.* **9**, 1834–1850 (2007).
11. Wood, M. W. *et al.* Identification of a pathogenicity island required for Salmonella enteropathogenicity. *Mol. Microbiol.* **29**, 883–891 (1998).
12. Fàbrega, A. & Vila, J. Salmonella enterica serovar Typhimurium skills to succeed in the host: Virulence and regulation. *Clin. Microbiol. Rev.* **26**, 308–341 (2013).
13. Gorvel, J. P. & Méresse, S. Maturation steps of the Salmonella-containing vacuole. *Microbes Infect.* **3**, 1299–1303 (2001).
14. Coombes, B. K. *et al.* Analysis of the contribution of Salmonella pathogenicity islands 1

- and 2 to enteric disease progression using a novel bovine ileal loop model and a murine model of infectious enterocolitis. *Infect. Immun.* **73**, 7161–7169 (2005).
15. Stecher, B. *et al.* Motility allows *S. Typhimurium* to benefit from the mucosal defence. *Cell. Microbiol.* **10**, 1166–1180 (2008).
 16. Erhardt, M. & Dersch, P. Regulatory principles governing *Salmonella* and *Yersinia* virulence. *Front. Microbiol.* **6**, 1–20 (2015).
 17. Young, G. M., Badger, J. L. & Miller, V. L. Motility is required to initiate host cell invasion by *Yersinia enterocolitica*. *Infect. Immun.* **68**, 4323–4326 (2000).
 18. Josenhans, C. & Suerbaum, S. The role of motility as a virulence factor in bacteria. *Int. J. Med. Microbiol.* **291**, 605–614 (2002).
 19. Achouri, S. *et al.* The frequency and duration of *Salmonella*–macrophage adhesion events determines infection efficiency. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).
 20. Kim, T. J., Young, B. M. & Young, G. M. Effect of flagellar mutations on *Yersinia enterocolitica* biofilm formation. *Appl. Environ. Microbiol.* **74**, 5466–5474 (2008).
 21. Duan, Q., Zhou, M., Zhu, L. & Zhu, G. Flagella and bacterial pathogenicity. *J. Basic Microbiol.* **53**, 1–8 (2013).
 22. Wood, T. K., González Barrios, A. F., Herzberg, M. & Lee, J. Motility influences biofilm architecture in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **72**, 361–367 (2006).
 23. Schechter, L. M., Damrauer, S. M. & Lee, C. A. Two AraC/XylS family members can independently counteract the effect of repressing sequences upstream of the *hila* promoter. *Mol. Microbiol.* **32**, 629–642 (1999).
 24. Darwin, K. H. & Miller, V. L. InvF is required for expression of genes encoding proteins secreted by the SPI1 type III secretion apparatus in *Salmonella typhimurium*. *J. Bacteriol.* **181**, 4949–4954 (1999).
 25. Eichelberg, K. & Galán, J. E. Differential regulation of *Salmonella typhimurium* type III secreted proteins by pathogenicity island 1 (SPI-1)-encoded transcriptional activators InvF and Hila. *Infect. Immun.* **67**, 4099–4105 (1999).
 26. Stafford, G. P., Ogi, T. & Hughes, C. Binding and transcriptional activation of non-flagellar genes by the *Escherichia coli* flagellar master regulator FlhD2C2. *Microbiology* **151**, 1779–1788 (2005).
 27. Prüß, B. M., Liu, X., Hendrickson, W. & Matsumura, P. FlhD/FlhC-regulated promoters

- analyzed by gene array and lacZ gene fusions. *FEMS Microbiol. Lett.* **197**, 91–97 (2001).
28. Chilcott, G. S. & Hughes, K. T. Coupling of Flagellar Gene Expression to Flagellar Assembly in *Salmonella enterica* Serovar Typhimurium and *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **64**, 694–708 (2000).
 29. Fitzgerald, D. M., Bonocora, R. P. & Wade, J. T. Comprehensive Mapping of the *Escherichia coli* Flagellar Regulatory Network. *PLoS Genet.* **10**, (2014).
 30. Ikebe, T., Iyoda, S. & Kutsukake, K. Structure and expression of the fliA operon of *Salmonella typhimurium*. *Microbiology* **145**, 1389–1396 (1999).
 31. Mytelka, D. S. & Chamberlin, M. J. *Escherichia coli* fliAZY operon. *J. Bacteriol.* **178**, 24–34 (1996).
 32. Ohnishi, K., Kutsukake, K., Suzuki, H. & Iino, T. Gene fliA encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. *MGG Mol. Gen. Genet.* **221**, 139–147 (1990).
 33. Iyoda, S., Kamidoi, T., Hirose, K., Kutsukake, K. & Watanabe, H. A flagellar gene fliZ regulates the expression of invasion genes and virulence phenotype in *Salmonella enterica* serovar Typhimurium. *Microb. Pathog.* **30**, 81–90 (2001).
 34. Mouslim, C. & Hughes, K. T. The Effect of Cell Growth Phase on the Regulatory Cross-Talk between Flagellar and Spi1 Virulence Gene Expression. *PLoS Pathog.* **10**, (2014).
 35. Cott Chubiz, J. E., Golubeva, Y. A., Lin, D., Miller, L. D. & Slauch, J. M. FliZ regulates expression of the *Salmonella* pathogenicity island 1 invasion locus by controlling HilD protein activity in *Salmonella enterica* serovar typhimurium. *J. Bacteriol.* **192**, 6261–6270 (2010).
 36. Singer, H. M., Kühne, C., Deditius, J. A., Hughes, K. T. & Erhardt, M. The *Salmonella* Spi1 virulence regulatory protein HilD directly activates transcription of the flagellar master operon flhDC. *J. Bacteriol.* **196**, 1448–1457 (2014).
 37. Lehnen, D. *et al.* LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. *Mol. Microbiol.* **45**, 521–532 (2002).
 38. Peterson, C. N., Carabetta, V. J., Chowdhury, T. & Silhavy, T. J. LrhA regulates rpoS translation in response to the Rcs phosphorelay system in *Escherichia coli*. *J. Bacteriol.* **188**, 3175–3181 (2006).
 39. Gottesman, S. Trouble is coming: Signaling pathways that regulate general stress

- responses in bacteria. *J. Biol. Chem.* **294**, 11685–11700 (2019).
40. Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609–3623 (2008).
 41. Schell, M. A. Molecular Biology of the LysR Family of Transcriptional Regulators. *Annu. Rev. Microbiol.* **47**, 597–626 (1993).
 42. Zaim, J. & Kierzek, A. M. The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. *Nucleic Acids Res.* **31**, 1444–1454 (2003).
 43. Chonoles Imlay, K. R., Korshunov, S. & Imlay, J. A. Physiological Roles and Adverse Effects of the Two Cystine Importers of Escherichia coli. *J. Bacteriol.* **197**, 3629–3644 (2015).
 44. Banjac, A. *et al.* The cystine/cysteine cycle: a redox cycle regulating susceptibility versus resistance to cell death. *Oncogene* **27**, 1618–1628 (2008).
 45. Baptist, E. W. & Kredich, N. M. Regulation of L cystine transport in Salmonella typhimurium. *J. Bacteriol.* **131**, 111–118 (1977).
 46. Kredich, N. M. Biosynthesis of Cysteine. *EcoSal Plus* **3**, ecosalplus.3.6.1.11 (2008).
 47. Malo, M. S. & Loughlin, R. E. Promoter elements and regulation of expression of the *cysD* gene of Escherichia coli K-12. *Gene* **87**, 127–131 (1990).
 48. Kredich, N. M. Regulation of L-cysteine biosynthesis in Salmonella typhimurium. I. Effects of growth of varying sulfur sources and O-acetyl-L-serine on gene expression. *J. Biol. Chem.* **246**, 3474–84 (1971).
 49. Kredich, N. M. The molecular basis for positive regulation of *cys* promoters in Salmonella typhimurium and Escherichia coli. *Mol. Microbiol.* **6**, 2747–2753 (1992).
 50. Mittal, M., Singh, A. K. & Kumaran, S. Structural and biochemical characterization of ligand recognition by CysB, the master regulator of sulfate metabolism. *Biochimie* **142**, 112–124 (2017).
 51. Lochowska, A. *et al.* Identification of activating region (AR) of Escherichia coli LysR-type transcription factor CysB and CysB contact site on RNA polymerase alpha subunit at the *cysP* promoter. *Mol. Microbiol.* **53**, 791–806 (2004).
 52. Ostrowski, J. & Kredich, N. M. Negative autoregulation of *cysB* in Salmonella typhimurium: In vitro interactions of CysB protein with the *cysB* promoter. *J. Bacteriol.*

- 173**, 2212–2218 (1991).
53. Yamazaki, S., Takei, K. & Nonaka, G. ydjN encodes an S-sulfocysteine transporter required by *Escherichia coli* for growth on S-sulfocysteine as a sulfur source. *FEMS Microbiol. Lett.* **363**, fnw185 (2016).
 54. Park, S. & Imlay, J. A. High levels of intracellular cysteine promote oxidative DNA damage by driving the Fenton reaction. *J. Bacteriol.* **185**, 1942–1950 (2003).
 55. Kawano, Y. *et al.* Enhancement of l-cysteine production by disruption of yciW in *Escherichia coli*. *J. Biosci. Bioeng.* **119**, 176–179 (2015).
 56. Sabrialabed, S., Yang, J. G., Yariv, E., Ben-Tal, N. & Lewinson, O. Substrate recognition and ATPase activity of the *E. Coli* cysteine/cystine ABC transporter YecSC-FliY. *J. Biol. Chem.* **295**, 5245–5256 (2020).
 57. Moussatova, A., Kandt, C., O'Mara, M. L. & Tieleman, D. P. ATP-binding cassette transporters in *Escherichia coli*. *Biochim. Biophys. Acta - Biomembr.* **1778**, 1757–1771 (2008).
 58. Linton, K. J. & Higgins, C. F. The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Molecular Microbiology* vol. 28 5–13 (1998).
 59. Wilkens, S. Structure and mechanism of ABC transporters. *F1000Prime Rep.* **7**, 1–9 (2015).
 60. Soutourina, J., Blanquet, S. & Plateau, P. Role of D-Cysteine Desulfhydrase in the Adaptation of *Escherichia coli* to D-Cysteine. *J. Biol. Chem.* **276**, 40864–40872 (2001).
 61. Dutta, T. & Srivastava, S. Small RNA-mediated regulation in bacteria: A growing palette of diverse mechanisms. *Gene* **656**, 60–72 (2018).
 62. Babski, J. *et al.* Small regulatory RNAs in archaea. *RNA Biol.* **11**, 484–493 (2014).
 63. Papenfort, K. & Vogel, J. Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res. Microbiol.* **160**, 278–287 (2009).
 64. Ellis, M. J., Trussler, R. S., Charles, O. & Haniford, D. B. A transposon-derived small RNA regulates gene expression in *Salmonella Typhimurium*. *Nucleic Acids Res.* **45**, 5470–5486 (2017).
 65. Beuzón, C. R., Chessa, D. & Casadesús, J. IS200: An old and still bacterial transposon. *Int. Microbiol.* **7**, 3–12 (2004).
 66. Beuzón, C. R., Marqués, S. & Casadesús, J. Repression of IS200 transposase synthesis by

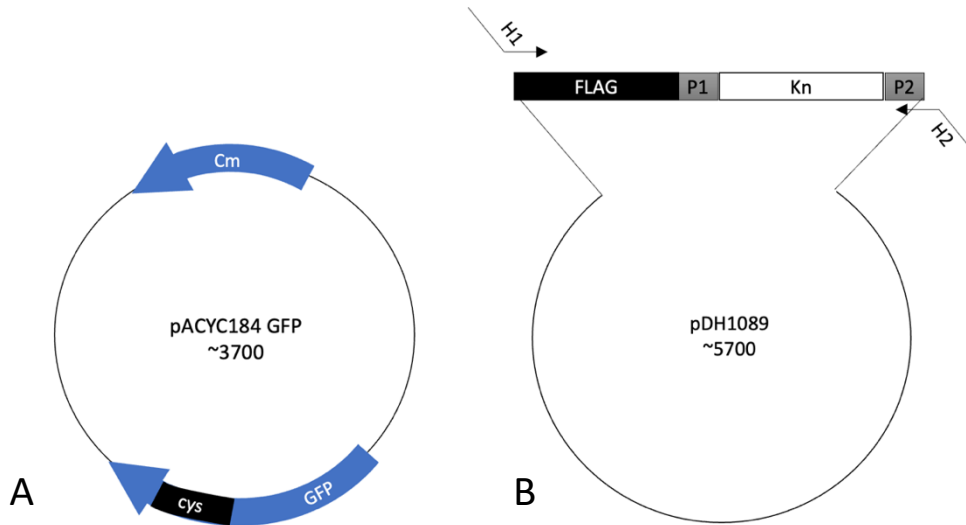
- RNA secondary structures. *Nucleic Acids Res.* **27**, 3690–3695 (1999).
67. Gottesman, S. & Storz, G. Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* **3**, 1–16 (2011).
 68. Westermann, A. J. *et al.* The major RNA-binding protein ProQ impacts virulence gene expression in salmonella enterica serovar typhimurium. *MBio* **10**, 1–21 (2019).
 69. Holmqvist, E., Li, L., Bischler, T., Barquist, L. & Vogel, J. Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3' Ends. *Mol. Cell* **70**, 971-982.e6 (2018).
 70. Holmqvist, E. & Vogel, J. RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* **16**, 601–615 (2018).
 71. El Mouali, Y. *et al.* CRP-cAMP mediates silencing of Salmonella virulence at the post-transcriptional level. *PLoS Genet.* **14**, 1–26 (2018).
 72. Ellis, M. J. *et al.* Silent but deadly: IS200 promotes pathogenicity in Salmonella Typhimurium. *RNA Biol.* **15**, 176–181 (2018).
 73. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645 (2000).
 74. Glick, B. R. Metabolic load and heterologous gene expression. *Biotechnol. Adv.* **13**, 247–261 (1995).
 75. Wu, G. *et al.* Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **34**, 652–664 (2016).
 76. Anthony, J. R. *et al.* Optimization of the mevalonate-based isoprenoid biosynthetic pathway in Escherichia coli for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab. Eng.* **11**, 13–19 (2009).
 77. Kutsukake, K., Ohya, Y. & Iino, T. Transcriptional analysis of the flagellar regulon of Salmonella typhimurium. *J. Bacteriol.* **172**, 741–747 (1990).
 78. Aiba, H., Adhya, S. & de Crombrughe, B. Evidence for two functional gal promoters in intact Escherichia coli cells. *J. Biol. Chem.* **256**, 11905–11910 (1981).
 79. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, 45e – 45 (2001).
 80. Allen, S., Shea, J. M., Felmet, T., Gadra, J. & Dehn, P. F. A kinetic microassay for

- glutathione in cells plated on 96-well microtiter plates. *Methods Cell Sci.* **22**, 305–312 (2000).
81. Tietze, F. Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: Applications to mammalian blood and other tissues. *Anal. Biochem.* **27**, 502–522 (1969).
 82. Kawano, Y. *et al.* Involvement of the *yciW* gene in l-cysteine and l-methionine metabolism in *Escherichia coli*. *J. Biosci. Bioeng.* **119**, 310–313 (2015).
 83. Rowbury, R. J. & Woods, D. D. Repression by Methionine of Cystathionase Formation in *Escherichia coli*. *J. Gen. Microbiol.* **35**, 145–158 (1964).
 84. Turnbull, A. L. & Surette, M. G. Cysteine biosynthesis, oxidative stress and antibiotic resistance in *Salmonella typhimurium*. *Res. Microbiol.* **161**, 643–650 (2010).
 85. Ohtsu, I. *et al.* Uptake of L-cystine via an ABC transporter contributes defense of oxidative stress in the L-cystine export-dependent manner in *Escherichia coli*. *PLoS One* **10**, 1–14 (2015).
 86. Ishii, T., Sugita, Y. & Bannai, S. Regulation of glutathione levels in mouse spleen lymphocytes by transport of cysteine. *J. Cell. Physiol.* **133**, 330–336 (1987).
 87. Lu, S. C. Glutathione synthesis. *Biochim. Biophys. Acta - Gen. Subj.* **1830**, 3143–3153 (2013).
 88. Anderson, M. E. Determination of glutathione and glutathione disulfide in biological samples. in *Methods* vol. 113 548–555 (1985).
 89. Griffith, O. W. Determination of glutathione and glutathione disulfide using glutathione reductase and 2-vinylpyridine. *Anal. Biochem.* **106**, 207–212 (1980).
 90. Ellman, G. L. A colorimetric method for determining low concentrations of mercaptans. *Arch. Biochem. Biophys.* **74**, 443–450 (1958).
 91. Herb, M. & Schramm, M. Functions of ROS in Macrophages and Antimicrobial Immunity. *Antioxidants* **10**, 313 (2021).
 92. Hryniewicz, M., Sirko, A. & Hulanicka, D. Identification and mapping of the sulphate permease promoter region in *Escherichia coli*. *Acta Biochim. Pol.* **36**, 353–363 (1989).
 93. Sirko, A., Zatyka, M., Sadowy, E. & Hulanicka, D. Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *J. Bacteriol.* **177**, 4134–4136 (1995).

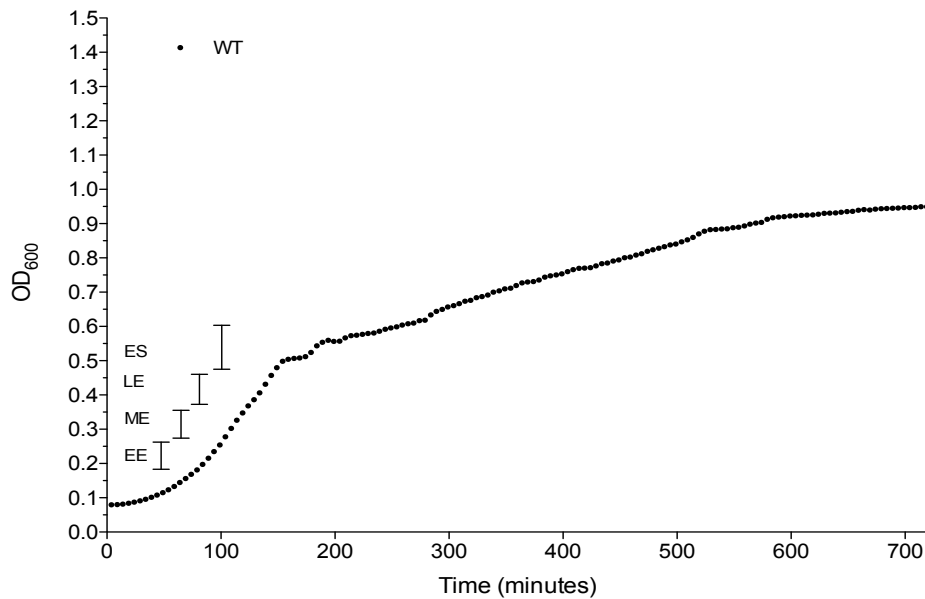
94. Berger, E. A. & Heppel, L. A. A binding protein involved in the transport of cystine and diaminopimelic acid in *Escherichia coli*. *J. Biol. Chem.* **247**, 7684–7694 (1972).
95. Kertesz, M. A. Bacterial transporters for sulfate and organosulfur compounds. *Res. Microbiol.* **152**, 279–290 (2001).
96. Hryniewicz, M. M. & Kredich, N. M. The *cysP* promoter of *Salmonella typhimurium*: Characterization of two binding sites for CysB protein, studies of in vivo transcription initiation, and demonstration of the anti-inducer effects of thiosulfate. *J. Bacteriol.* **173**, 5876–5886 (1991).
97. Crawford, M. A. *et al.* DksA-dependent transcriptional regulation in *Salmonella* experiencing nitrosative stress. *Front. Microbiol.* **7**, 1–8 (2016).
98. Berg, H. C. The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* **72**, 19–54 (2003).
99. Dey, B. *et al.* DNA-protein interactions: Methods for detection and analysis. *Mol. Cell. Biochem.* **365**, 279–299 (2012).
100. Wiehle, L. & Breiling, A. Chromatin Immunoprecipitation. in *Polycomb Group Proteins* (eds. Lanzaolo, C. & Bodega, B.) vol. 9 7–21 (Springer New York, 2016).
101. Rodríguez-Ubreva, J. & Ballestar, E. Chromatin Immunoprecipitation. in (eds. Stockert, J. C., Espada, J. & Blázquez-Castro, A.) vol. 1094 309–318 (Humana Press, 2014).
102. Das, P. M., Ramachandran, K., VanWert, J. & Singal, R. Chromatin immunoprecipitation assay. *Biotechniques* **37**, 961–969 (2004).
103. Antonio Ibarra, J. *et al.* Induction of salmonella pathogenicity island 1 under different growth conditions can affect *Salmonella*-host cell interactions in vitro. *Microbiology* **156**, 1120–1133 (2010).

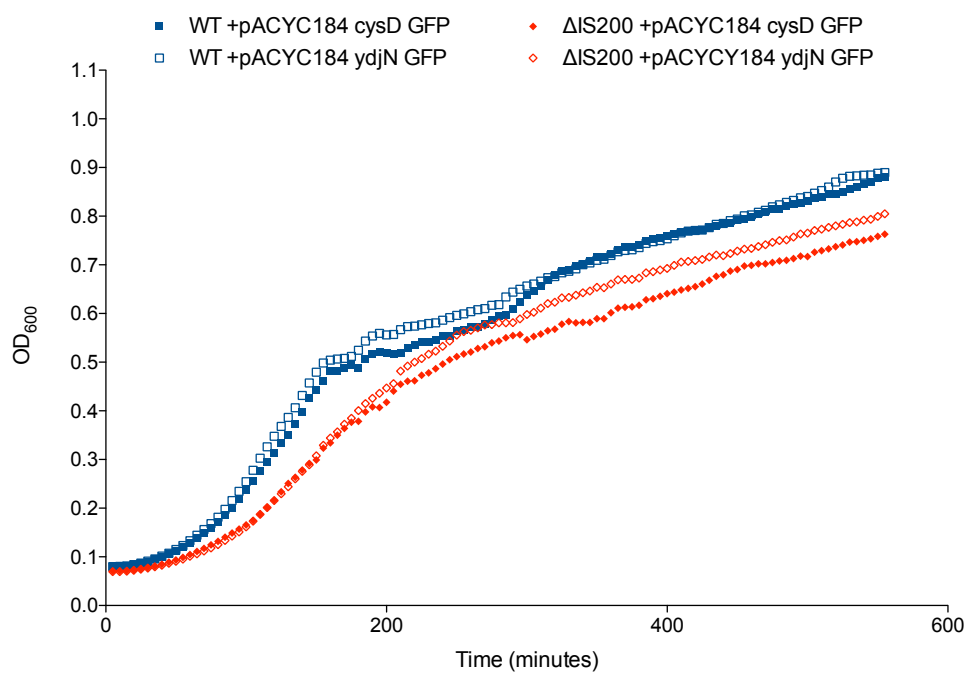
Appendix

Supplemental Figures

**Supplemental Figure 1. Schematic overview of pACYC184 and pDH1089.**

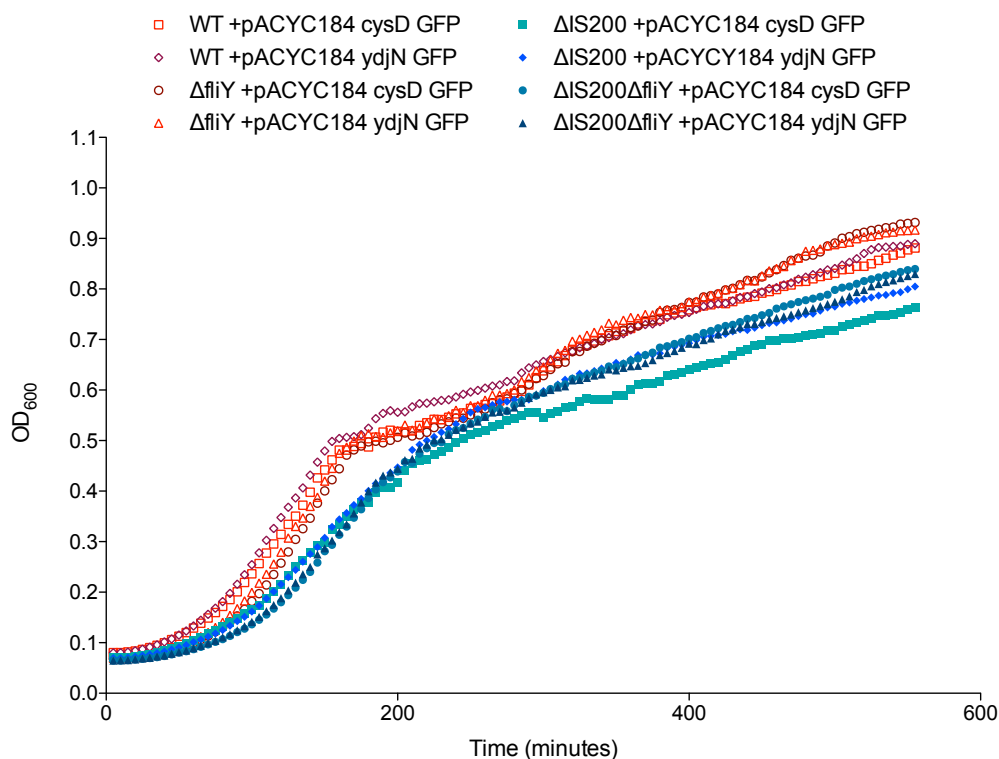
(A) Simplified plasmid map of GFP transcriptional fusions where *cys* denotes the promoter region of the Cys gene of interest, either *cysD* or *ydjN*, cloned in front of the GFP gene on pACYC184. (B) FLAG template for PCR amplification of *fliY*::3X-FLAG-kan and *yecC*::3X-FLAG-kan with H1 and H2 noted for regions of homology.

**Supplemental Figure 2. Growth curve analysis of *Salmonella* SL1344 WT.**



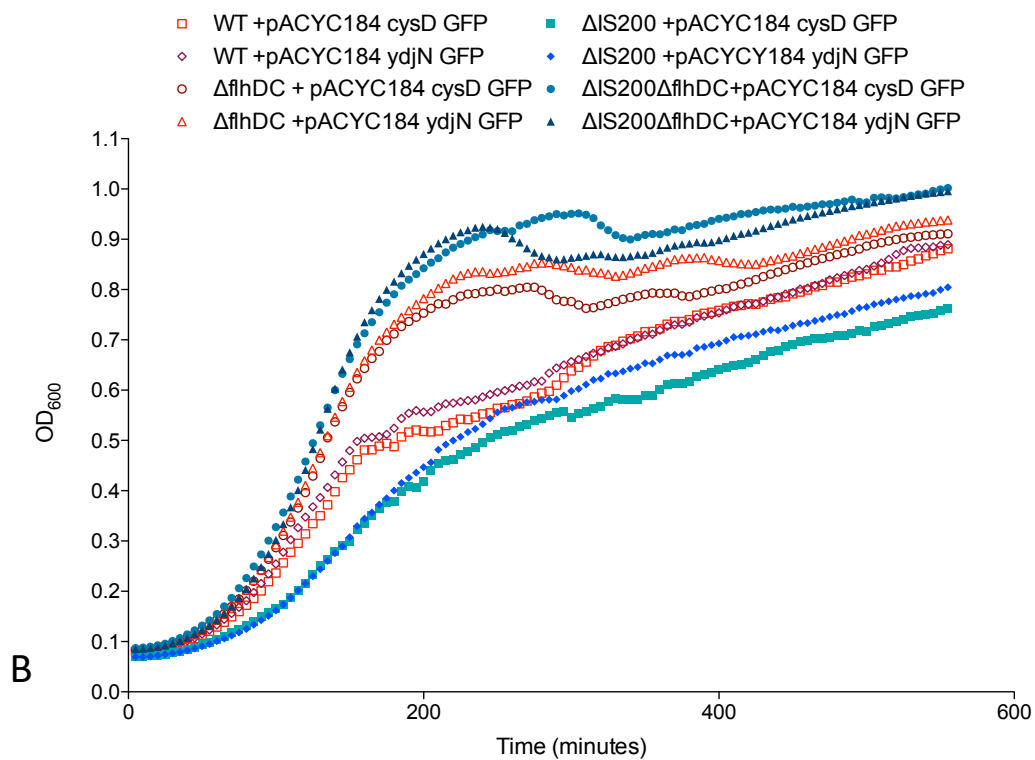
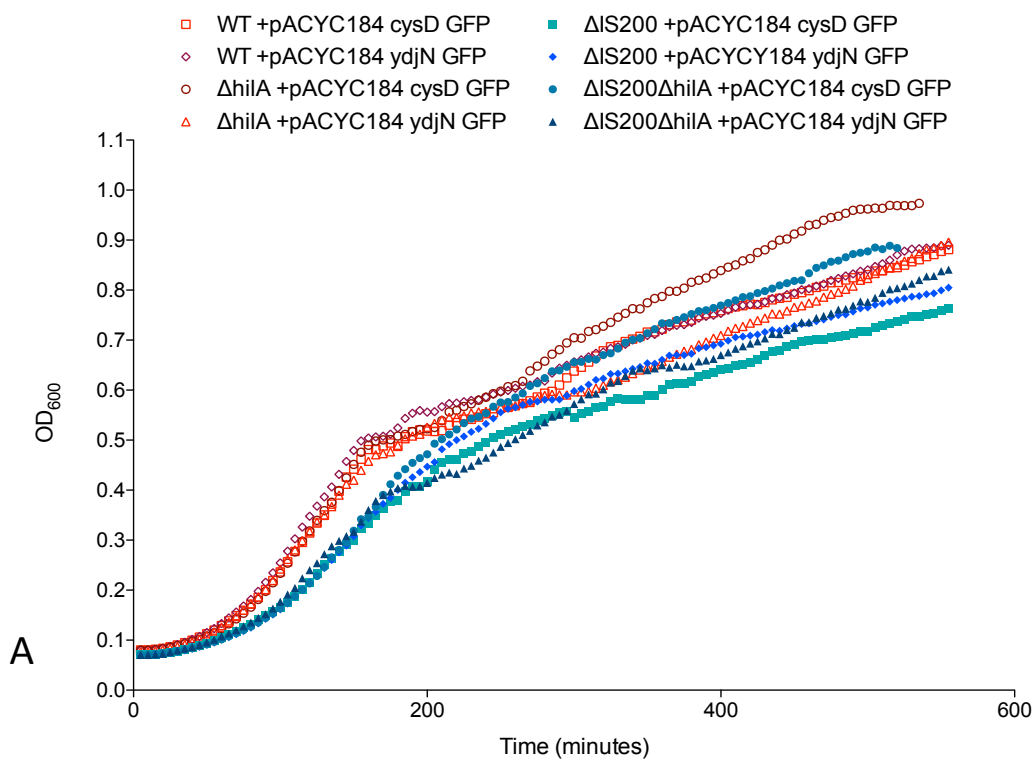
Supplemental Figure 3. Growth curves of *cysD* and *ydjN* GFP transcriptional fusions in *Salmonella* WT and Δ IS200.

Plotted OD₆₀₀ versus time (minutes) showing the growth throughout monitoring the GFP expression under *cysD* and *ydjN* promoters in *Salmonella* WT and Δ IS200. Growth was monitored every 5 minutes for 9.3 hours and the average of 3 biological and 2 technical replicates are plotted.



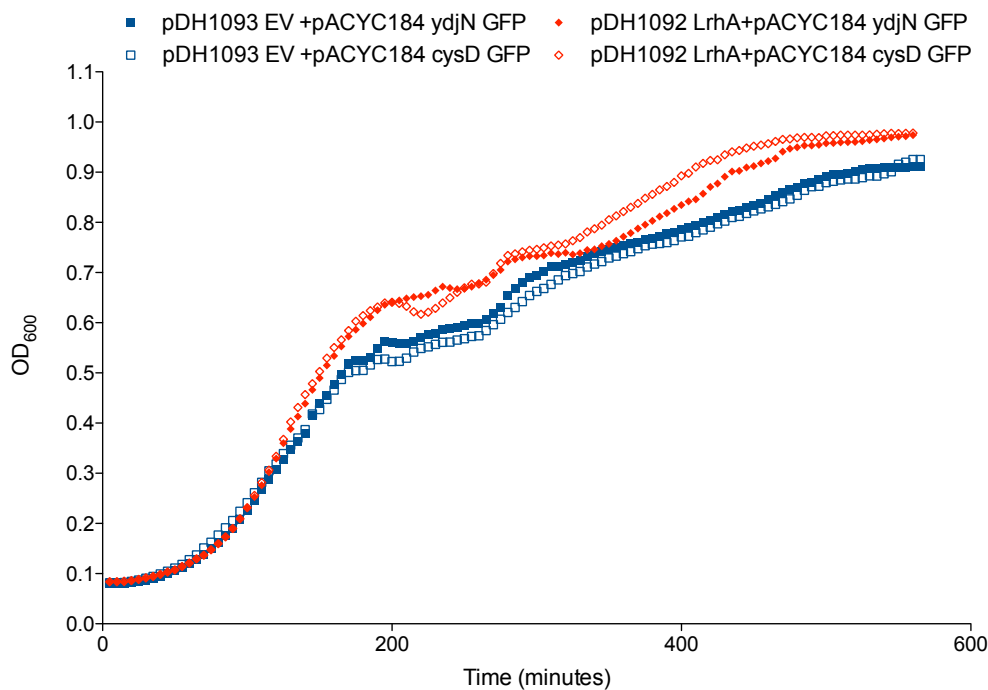
Supplemental Figure 4. Growth curve for *cysD* and *ydjN* GFP transcriptional fusions in Δ *fliY* backgrounds.

Plotted OD₆₀₀ versus time for the growth throughout monitoring GFP expression under *cysD* and *ydjN* promoters in *Salmonella* WT and Δ IS200 backgrounds with and without Δ *fliY* mutation. Growth was monitored every 5 minutes for 9.3 hours and the average of 3 biological and 2 technical replicates are plotted.



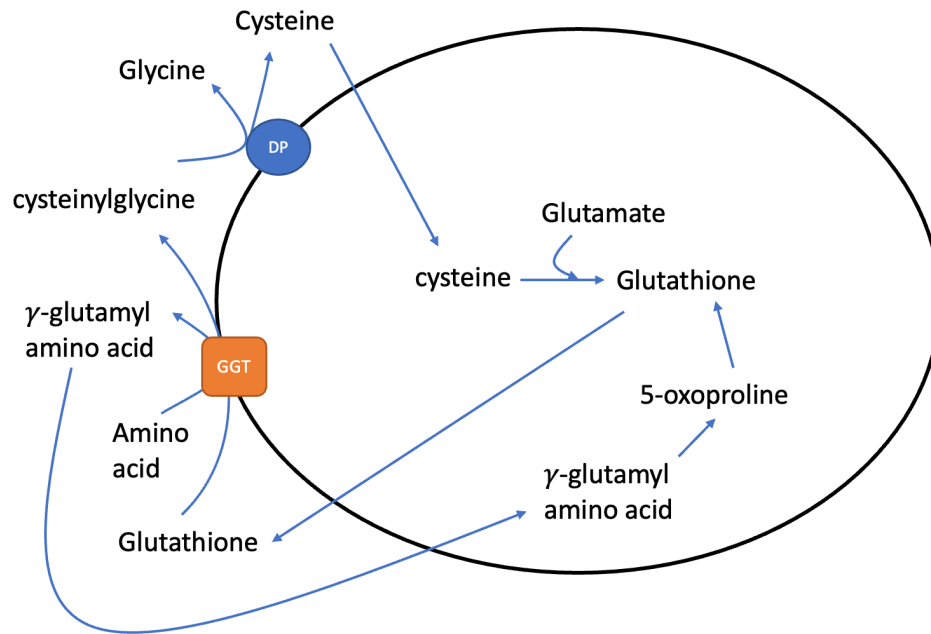
Supplemental Figure 5. Growth curves for *cysD* and *ydjN* GFP transcriptional fusions in Δ *hilA* and Δ *flhDC* backgrounds.

Plotted OD₆₀₀ versus time for growth curves monitoring GFP expression in *Salmonella* WT and Δ IS200 with and without (A) Δ *hilA* and (B) Δ *flhDC* backgrounds. Growth was monitored every 5 minutes for 9.3 hours and the average of 3 biological and 2 technical replicates are plotted.



Supplemental Figure 6. Growth curve for overexpression of LrhA with *cysD* and *ydjN* GFP transcriptional fusions in *Salmonella* SL1344 WT.

Plotted OD₆₀₀ versus time (minutes) for *Salmonella* SL1344 WT containing overexpression of LrhA or empty vector with *cysD* or *ydjN* GFP transcriptional fusions. Growth was monitored every 5 minutes for 9.3 hours and the average of 3 biological and 2 technical replicates are plotted.



Supplemental Figure 7. γ -glutamyl cycle as a source of cysteine.

DP = dipeptidase enzyme

GGT= γ -glutamylpeptidase

Supplemental Tables

Supplemental Table 1. *Salmonella* strains used in this study.

Strain	Description	Reference/source
DBH199	Wild-type LT2, Str ^S	Miguel Valvano
DBH347	Wild-type SL1344, Str ^R	Salmonella Genetic Stock Centre (SGSC438)
DBH367	DBH347 $\Delta hilA::kan$, Str ^R Kan ^R	
DBH415	DBH347 $\Delta tnpA_{1-7}$, Str ^R	64
DBH432	DBH415 $\Delta hilA::kan$, Str ^R Kan ^R	
DBH467	DBH347 $\Delta flhDC::cm$, Str ^R Cm ^R	
DBH468	DBH415 $\Delta flhDC::cm$, Str ^R Cm ^R	
DBH546	DBH347 $\Delta fliY::kan$, Str ^R Kan ^R	
DBH547	DBH415 $\Delta fliY::kan$, Str ^R Kan ^R	
DBH616	DBH415 $\Delta fliAZ::kan$, kan flipped, Str ^R Kan ^S	
DBH633	DBH347 $\Delta fliAZ::kan$, kan flipped, Str ^R Kan ^S	
DBH657	DBH347 $\Delta cysB::kan$, Str ^R Kan ^R	
DBH658	DBH415 $\Delta cysB::kan$, Str ^R Kan ^R	
DBH671	LT2 <i>yecC::3X-FLAG-kan</i> , Str ^S Kan ^R	
DBH672	LT2 <i>fliY::3X-FLAG-kan</i> , Str ^S Kan ^R	
DBH673	DBH347 <i>yecC::3X-FLAG-kan</i> , Str ^R Kan ^R	
DBH674	DBH415 <i>yecC::3X-FLAG-kan</i> , Str ^R Kan ^R	
DBH675	DBH347 <i>fliY::3X-FLAG-kan</i> , Str ^R Kan ^R	
DBH676	DBH415 <i>fliY::3X-FLAG-kan</i> , Str ^R Kan ^R	

Supplemental Table 2. Plasmids used in this study.

Plasmid	Description	Reference/source
pACYC184	Cm ^R	
pDH740	pKD46, temperature sensitive Lambda Red expression plasmid, Amp ^R	73

pDH1039	<i>prgH</i> transcriptional GFP (pMPMA3dPlac-PprgH-GFPLVA) from Steele-Mortimer Lab via addgene, Amp ^R	103
pDH1089	pBR322-derived template for recombineering chromosomal kan marked C-term 3X-FLAG tag	
pDH1092	<i>lrhA</i> expressed from own promoter cloned into pDH337, Amp ^R	
pDH1093	pDH337 with <i>LacZα</i> promoter removed, Amp ^R	

Supplemental Table 3. Oligonucleotides used in this study.

Name	Sequence (5' to 3')	Description/Use
3X FLAG		
oDH770	AGCGCTCTCTGAAAAATGGTTTGGCGCTGACGTGACCCAAGACTACAAAGACCATGACGG	Forward primer for <i>fliY</i> ::3X-FLAG-kan, contains FLAG+40nt overlap
oDH771	GCCCGGTGGCGCACTGCCTGCCGGGCAACGTAGAACATTACATATGAATATCCTCCTTA	Reverse primer for <i>fliY</i> ::3X-FLAG-kan, contains P2+40nt overlap
oDH772	TCCGCGTACCCGCCAGTTCCTTGAGAAATTTTAGTGAAAGACTACAAAGACCATGACGG	Forward primer for <i>yecC</i> ::3X-FLAG-kan, contains FLAG+40nt overlap
oDH773	CAAGGCGCGGCTAATTCGGGCGAGCGTTGAAAGGCATTACATATGAATATCCTCCTTA	Reverse primer for <i>yecC</i> ::3X-FLAG-kan, contains P2+40nt overlap
GFP		
oDH774	TACGCATGCGCGCCAGAATGACGGCAAAG	<i>cysD</i> , was also used for
oDH775	TACGCATGCTTAATTTCTCCTCTTTAATCTAGGTACCCGCTGCCACTTCACGGATAATG	ChIP gene specific PCR
oDH776	TTACGAATTCTGGTCTTCGCAGATGATTAC	

oDH777	TACGCATGCTTAATTTCTCCTCTTTAATTCT AGGTACCTTTATGACGGGCCTGCGCCAG	<i>ydjN</i> , was also used for ChIP gene specific PCR
qRT-PCR		
oDH661	TCTGGGAAACTGCCTGATGGAGG	<i>rrsA</i> (16S rRNA)
oDH662	CGTAGGAGTCTGGACCGTGTCTC	
oDH778	GAGCTTTGCCGATGAGGGTT	<i>fliY</i> , oDH778 also used in
oDH779	CAGCATTCCGTCCCATTTGGT	RT PCR
oDH780	CTCGCCTGGAGTTTATCGGC	<i>yedO</i>
oDH781	GAATCGCCCCTGCGGTTATC	
oDH782	GACGATCTGCTACAAGCGGG	<i>fliA</i>
oDH783	CTGCTCCAGTTGTCCCATCG	
oDH784	TCTGCTGCTGGGATTTGTGC	<i>yecS</i>
oDH785	GCCGCGGTATTGAGCGATAA	
oDH786	GTGCTGCACGGTATTGACCT	<i>yecC</i>
oDH787	CGTAGCTGGCGAATCAACCC	
oDH794	TGCCGGTGACCATTACGAAG	<i>hilA</i>
oDH795	AGAGAGAAGCGGGTTGGTGT	
oDH788	CGTCCTGCCGCAGATAACTT	<i>hilD</i>
oDH789	TGAGCACCAACATCCCAGGT	
oDH790	CACAAACCCCGCTTTCGCTA	<i>yciW</i>
oDH791	CGCGGTGCTCATA CGTTCAT	
oDH792	ACTGGTGTTTGACACCGTCG	<i>metC</i>
oDH793	CCTGCTCAACGAATGCCAGA	
oDH796	GCTGGAAGATGAGCTTGGCA	<i>cysB</i>
oDH797	AACGCGCCTGAGTATGGGTA	
oDH798	GCAACTGGAGGCGGAAAGTA	<i>cysD</i>
oDH799	CGCGAAAGGCGTACATCTCA	
oDH800	TGAGCAACAATGGGCGAAGG	<i>cysP</i>
oDH801	GTTGTTCGGCAGACGGCTTT	
oDH802	GCAGGCCCGTCATAAACAGT	<i>ydjN</i>
oDH803	CGAATACCAGCGGCATCACA	

oDH804	AGGTAATTGCCAAGGCGTCG	<i>fliZ</i> , oDH805 also used in
oDH805	CTACTTGACGACGCTGCCTG	RT-PCR
RT PCR		
oDH806	GTGAATTCACTGTATAACCGC	Forward primer <i>fliA</i>
oDH807	CATTATTGGGTCACGTCAGCG	Reverse primer <i>fliY</i> 3' end
oDH808	GAGGTTATACATGAGGATGGTAGGC	Reverse primer <i>yedO</i> 3' end
oDH809	CGCACTCATTTTGGCTCTCTC	Reverse primer <i>yecS</i> 3' end
oDH810	CGAGCGTTGAAAAGGCATTA	Reverse primer <i>yecC</i> 3' end
ChIP		
oDH811	ACTGAGATTCGCCTTACACG	<i>flhDC</i>
oDH812	TATGTCTTATAAATGCCACATTAATGTGAAG GAC	
oDH813	TACGCATGCGTCATTGATGGCGGCAGTAC	<i>cysP</i>
oDH814	TACGCATGCTTAATTTCTCCTCTTTAATCTA GGTACCTGCCTGCGCCTGCCCTGCCAG	

Supplemental Table 4. 165 annotated hits from unpublished RNA-seq data at early exponential phase.

Summary of gene hits in which the differential gene expression was statistically significant between *Salmonella* SL1344 WT and Δ IS200. Gene, mean expression, WT expression, Δ IS200 expression listed along with the differential expression between strain and within strains. Effect is the differential expression between divided by within, with effect sizes greater than 2 as true hits ($p < 0.05$).

Gene name	Mean Exp	WT Exp	Δ IS200 Exp	Diff between	Diff within	Effect	$p < 0.05$
STnc205	4.7757776	3.5356261	5.9570027	2.4206470	0.0611659	39.5900253	0.0034384
CheW	6.1088293	4.8860348	7.3832453	2.4904607	0.0800043	31.1788747	0.0058901
sipC	2.3273924	0.5386877	3.9271222	3.3849098	0.1230108	27.6260328	0.0055785
tcp	3.2157096	2.0801095	4.2521691	2.1714175	0.0826119	26.5037345	0.0078343
flgL	5.2222277	4.5239082	5.9260072	1.4007941	0.0538332	26.1979511	0.0079092
fliZ	4.3740241	3.2759298	5.3659090	2.1021825	0.0824716	25.5088340	0.0046347
tsr	6.4953583	5.7841581	7.2360242	1.4410585	0.0584437	24.5644508	0.0043663
STM3216	5.2029066	4.3830489	6.1337400	1.7451465	0.0720538	23.2333514	0.0199314
flgK	5.4824912	4.6961010	6.2371719	1.5204144	0.0695727	21.9542189	0.0045339

cheM	5.0645981	3.7916604	6.3249747	2.5479987	0.1183171	21.5809858	0.0120044
mobA	4.0956429	2.8032868	5.2700927	2.4927333	0.1198041	20.8474492	0.0055946
fliY	3.5198045	2.6054925	4.4236259	1.8120605	0.0917148	19.8438431	0.0072999
motB	5.1606932	3.9152795	6.3863416	2.4342024	0.1331523	18.3246325	0.0044539
sigE	0.8004894	-1.0183094	2.3376591	3.3663853	0.1861940	18.1452532	0.0117404
prgH	1.0971275	-1.0137919	2.6929273	3.7128937	0.2082554	17.9035623	0.0172723
cheY	3.8759793	2.7777725	4.8978209	2.1428833	0.1239223	17.2064077	0.0066348
invG	0.4813559	-1.6252048	2.2545057	3.8749184	0.2344899	16.4774541	0.0120414
flhB	4.0116152	3.0747647	5.0136907	1.9172859	0.1156113	16.3967994	0.0109778
flgE	4.8854161	4.4084075	5.4700733	1.0537905	0.0646682	15.7950392	0.0275624
cheA	6.2513748	5.0874508	7.5219630	2.4363579	0.1560464	15.7148929	0.0042625
cheR	4.2769690	3.1020082	5.2955987	2.2188144	0.1416435	15.7023533	0.0055714
prgJ	0.8225713	-1.0051744	2.2956267	3.2999131	0.2133580	15.3849433	0.0210707
modA	2.8524577	2.0482438	3.6364860	1.5874699	0.1048693	15.0435368	0.0101469
yebR	3.7185581	3.1996045	4.2346530	1.0411279	0.0692819	15.0292684	0.0109190
flgF	4.5950832	4.1056002	5.1696954	1.0531069	0.0679035	14.9979340	0.0220820
invA	0.6824716	-0.9365568	2.0808265	3.0185794	0.2041824	14.8589798	0.0100220
sipB	1.4888323	-0.5576007	3.1452426	3.7126497	0.2520366	14.7229198	0.0189046
cheM-							
cheR_IGR	2.2728967	1.0148571	3.3147603	2.3038858	0.1614212	14.3508566	0.0239036
flgG	5.3226724	4.7957277	5.9269625	1.1309265	0.0775538	14.2919620	0.0261373
trg	1.7245706	0.6469350	2.6944116	2.0351915	0.1428687	14.2561625	0.0109600
flgC	3.5377100	3.0017424	4.0519921	1.0419104	0.0742612	13.9846413	0.0108654
prgK	2.0099881	0.6956417	3.1710926	2.4849697	0.1792592	13.9544823	0.0118656
cheV	4.1253890	3.3627158	4.8376538	1.5063347	0.1098139	13.9026286	0.0131153
motA	5.0077558	3.6342147	6.3939339	2.7004836	0.1982806	13.6894654	0.0052798
sopE	0.4344359	-2.4285137	2.4745871	4.9335709	0.3619971	13.6819003	0.0238734
sopB	0.7519400	-1.3404171	2.4197163	3.7976564	0.2807554	13.5530593	0.0105701
ycgR	3.0777982	2.4194979	3.6667275	1.2565699	0.0954132	13.1656167	0.0107118
sicA	0.1043077	-2.0888412	1.7103109	3.8008949	0.2997774	12.7602513	0.0254082
modC	3.1620527	2.4588345	3.8722887	1.4045218	0.1108744	12.6833956	0.0143751
STM3152	4.4913442	3.8000612	5.2438141	1.4257877	0.1137345	12.4942316	0.0092881
STM1934	2.2293957	0.9147252	3.4148963	2.5358390	0.2053386	12.4804696	0.0134953
sdiA	2.3164619	1.1452219	3.3880315	2.2179539	0.1776355	12.4406632	0.0116191
STM3138	2.1654868	1.1004486	3.0005769	1.9010044	0.1547510	12.3642151	0.0314948
flhC	3.9842537	3.1108024	4.7119478	1.6272669	0.1325920	12.2076509	0.0117736
cheZ	5.3440957	4.3195995	6.2886285	2.0276747	0.1730741	11.9007724	0.0128607

STM3156	1.9911298	1.3040178	2.6221522	1.3260398	0.1136560	11.7180152	0.0127456
flhA	4.3962416	3.5239091	5.2989353	1.7507647	0.1480695	11.6517631	0.0154319
invE	0.1324770	-1.9204686	1.5691759	3.5027163	0.3045748	11.4522176	0.0210525
sopE2	-0.7156983	-3.6493218	1.2998618	4.9593730	0.4322479	11.4321065	0.0253184
srfC	3.5470167	2.8423799	4.2751879	1.4399710	0.1266487	11.3923385	0.0231496
ycgO	3.1858361	2.4917794	3.9350233	1.4247646	0.1242399	11.3635076	0.0123603
fliA	5.5458366	4.6316716	6.4825170	1.9108076	0.1705562	11.3442535	0.0143606
cheB	3.7856402	2.7342515	4.8905602	2.1554773	0.1940432	11.1388633	0.0138105
yoaA	2.4620914	1.9336168	2.9789317	1.0411811	0.0943143	11.0519772	0.0144517
STM1874	2.5913907	1.9688372	3.1232328	1.1616553	0.1049411	10.9592290	0.0140794
modB	2.6615573	1.9391828	3.4587109	1.4935058	0.1350458	10.9038722	0.0173747
sipA	1.1345437	-0.0028381	2.1179258	2.1465302	0.2019717	10.8480287	0.0136369
SL1344_R							
S24700	3.3911464	2.8480512	3.9703621	1.1214844	0.1059396	10.6273839	0.0168593
mobC	1.8660799	0.3184976	2.9304811	2.6027451	0.2372773	10.6001246	0.0370935
STM1911	1.7069918	0.9089729	2.3972692	1.4912815	0.1431016	10.4704914	0.0143736
SL1344_R							
S24580	1.5244034	0.6957991	2.1281813	1.4288003	0.1379053	10.2818030	0.0218935
invF	0.5260547	-1.7586531	2.2306560	4.0285252	0.3946802	10.2607006	0.0417155
yhjG	2.5159719	1.9531376	3.0807786	1.1177551	0.1096430	10.2233907	0.0153067
STM3154	3.3234436	2.6543249	3.8614098	1.2095911	0.1222640	10.0508275	0.0388209
flgM	4.0620443	3.3772150	4.7647638	1.3968796	0.1409372	10.0259751	0.0263350
tdcC	1.1308603	0.2733829	1.8868779	1.6160803	0.1611613	10.0193635	0.0192572
gltA	3.7289762	3.1479255	4.1907274	1.0457230	0.1037207	10.0172934	0.0190692
srfA	3.5519814	2.8448490	4.2688877	1.4326417	0.1466684	9.8118599	0.0301250
invC	-0.1262159	-1.4362472	1.0256498	2.4428090	0.2496465	9.7915958	0.0170763
sprB	-0.2127587	-1.5464077	0.8282194	2.3871812	0.2442066	9.6931092	0.0193945
flgN	4.5882687	3.9101108	5.2608291	1.3414691	0.1389715	9.6182118	0.0127527
orgA	2.0529631	1.0047560	3.0208355	2.0736379	0.2212785	9.4572683	0.0135024
srfB	3.9412437	3.2223394	4.6457872	1.4126163	0.1490436	9.3802265	0.0162231
flgD	4.1883475	3.6279376	4.7315156	1.1192499	0.1218735	9.3117733	0.0262101
hilA	0.4621654	-1.0688782	1.8447732	2.9463360	0.3202533	9.2556293	0.0206261
flgB	3.9735240	3.4968759	4.5235861	1.0165813	0.1095916	9.2075154	0.0195116
SL1344_R							
S24725	0.7917721	-0.4188791	1.6945779	2.1423732	0.2390664	9.0261822	0.0164235
SL1344_R							
S24695	3.1543140	2.5720065	3.7183787	1.1450790	0.1275045	8.9635968	0.0157829

yebF	1.3124562	0.6195132	1.9140438	1.2914445	0.1454873	8.9503644	0.0190474
lpxR	-0.3841771	-1.9844393	0.6859573	2.6805261	0.3003685	8.9407068	0.0304792
STM3155	1.9916651	1.2147302	2.5728243	1.3463530	0.1491941	8.8778807	0.0347529
cutC	1.2421122	0.6318708	1.8866443	1.2435448	0.1394558	8.8410927	0.0187379
tdcE	1.1284972	0.4372357	1.7167215	1.2741629	0.1450582	8.7579478	0.0202637
ptrB	1.4288542	0.8772209	2.0164991	1.1312815	0.1315443	8.5910197	0.0191981
prgI	0.0661180	-2.1819169	1.5562416	3.7511988	0.4322065	8.5579580	0.0433409
flgA	4.0298018	3.5769621	4.5836391	1.0086077	0.1157927	8.4941630	0.0411108
spaO	-0.4787503	-2.1574089	0.6424313	2.7999564	0.3327392	8.4563981	0.0293302
RyeF	0.4535661	-0.3073202	1.1331386	1.4392016	0.1714745	8.3813201	0.0208544
STM1873	2.7463603	2.1365517	3.1939459	1.0648758	0.1303243	8.2110227	0.0229555
yhjH	4.0060146	3.2004222	4.7953187	1.5171882	0.1895570	8.0875349	0.0121576
flhE	2.3015573	1.4533444	3.1031851	1.5823567	0.1961095	8.0134264	0.0182923
InvR	2.4912440	1.5578077	3.2541729	1.7071869	0.2209925	7.9707317	0.0481478
spaN	-0.2200070	-1.6189362	0.8441939	2.5251432	0.3182831	7.9117931	0.0199452
STnc1405	0.1171241	-0.9184134	1.0599544	2.0175302	0.2640688	7.7527067	0.0216228
yghA	1.8822966	1.2636066	2.4891585	1.2306752	0.1601527	7.7053660	0.0199407
aer	2.1658886	1.3380863	2.9009652	1.5381517	0.2002642	7.6442097	0.0277691
spaP	-0.6379001	-2.4196062	0.5076240	2.9288580	0.3861261	7.6169369	0.0430455
STnc840	1.9592394	1.2722192	2.5587336	1.2776084	0.1693403	7.6118899	0.0198280
yoaB	3.1440173	2.6393736	3.6664912	1.0062632	0.1327488	7.5641530	0.0233259
STnc1285	2.5313253	1.4140313	3.6106902	2.1862259	0.2884833	7.5634952	0.0460210
spaQ	-1.1338766	-2.7711707	0.0684317	2.8108459	0.3750173	7.5080396	0.0270676
STM1300	2.3380655	1.6075373	3.0293399	1.4225343	0.1928213	7.4930691	0.0297233
flgH	3.4459898	2.8748192	3.9065423	1.0230768	0.1366312	7.4602427	0.0203308
ydcI	0.9662324	0.2858477	1.4525007	1.1651564	0.1574050	7.3268053	0.0300173
nrfB	-0.0063627	-0.8978083	0.7705197	1.6891447	0.2297147	7.2999452	0.0224670
hilC	2.0012940	1.2296198	2.7585336	1.5344404	0.2138410	7.2660827	0.0385486
rtsB	-1.2442888	-3.4990573	0.0899424	3.5898803	0.4899678	7.1793690	0.0377749
STM4534	1.3387095	0.7785424	1.8614345	1.0895220	0.1503376	7.1790040	0.0252718
yecM	0.9051977	0.2409277	1.4440459	1.2074134	0.1690323	7.1506955	0.0322473
yecC	1.4813625	0.7985153	2.0782023	1.2748418	0.1803944	7.0825694	0.0235686
STM3032	0.8443589	-0.0484974	1.6258236	1.6729554	0.2388439	7.0811648	0.0332513
rtsA	0.4529027	-1.0623211	1.7926108	2.9415578	0.4375190	6.8999318	0.0319971
STM1554	1.4757455	0.8321066	1.8977745	1.0569580	0.1496928	6.8507286	0.0411958
SraC	3.4127844	2.5116255	4.0176451	1.5175672	0.2193454	6.8324446	0.0256237
STM1832	2.8073540	2.2656118	3.3689204	1.0964960	0.1614735	6.8220952	0.0226770

yncL	3.5105089	2.9091353	4.1606768	1.2072476	0.1745228	6.8178335	0.0286186
fliB	2.9392257	2.2458287	3.5098616	1.2550463	0.1849384	6.7920650	0.0236704
iagB	0.4905936	-0.7956904	1.3588903	2.2184003	0.3327006	6.7379208	0.0243305
rflM	2.8699841	2.3194156	3.4452876	1.1646749	0.1791262	6.5575793	0.0279910
yebE	0.9695126	0.3665128	1.4915352	1.1310280	0.1762945	6.5413222	0.0393890
STM0567	2.1913082	1.5652654	2.7285496	1.1546706	0.1753872	6.5333857	0.0231840
adhP	1.5030382	0.9007348	1.9878654	1.1018967	0.1691211	6.4636202	0.0291026
sipD	-0.0852663	-1.7697751	1.0662349	2.8427018	0.4308167	6.4422570	0.0301770
hilC-							
sprB_IGR	-0.4379160	-1.7373760	0.6456326	2.3950926	0.3730300	6.3464027	0.0261170
sopA	-0.7180003	-2.1913972	0.3088780	2.5377722	0.4054575	6.2789632	0.0357091
nrfC	0.1748668	-0.7535176	0.8246654	1.5767266	0.2482677	6.2691393	0.0295743
SLnc0014	0.7898209	0.0841698	1.3070617	1.2208943	0.1945696	6.2598116	0.0468837
yghW	1.7608983	1.1320063	2.5441480	1.3620992	0.2157232	6.2057655	0.0283460
yedO	1.5287144	0.8738124	2.1325669	1.2799058	0.2119831	6.1452005	0.0339181
spaR	-1.3181860	-2.6529560	0.0279090	2.6260616	0.4280773	6.1153465	0.0303995
STM1857	-0.1039577	-1.0914350	0.7441643	1.8254372	0.3051265	6.0105398	0.0422039
siiB	-1.0069599	-2.2736694	-0.1737211	2.0998733	0.3520770	5.9870935	0.0443912
STnc600	0.9957770	-0.1517166	1.6514758	1.8160179	0.2969421	5.9340405	0.0435985
gcvH	1.9418783	1.3271211	2.4328233	1.0696048	0.1806257	5.9033809	0.0281449
hilD	2.3771875	1.7656222	2.9173048	1.2216177	0.2100019	5.9026447	0.0237825
rpsV	1.8058653	1.2781989	2.4589763	1.1452225	0.1975876	5.7249778	0.0340936
StyR.305	-1.0372925	-1.8872031	-0.2922617	1.5942570	0.2841906	5.5782462	0.0376614
siiD	-0.9834694	-2.0542244	-0.0566884	1.9778188	0.3582441	5.5216309	0.0339068
csgC	0.9769965	0.1785619	1.6081320	1.5110129	0.3001819	5.1497080	0.0396210
siiA	-2.0307179	-3.4177984	-1.0728298	2.3526652	0.4616879	5.0996932	0.0396498
ymdA	1.0404399	0.2492807	1.6601241	1.3404175	0.2674354	4.9716839	0.0365023
purT	0.4832728	-0.0364555	1.1024243	1.1228013	0.2229424	4.9093638	0.0450136
STM1841	-1.0942727	-2.1196515	-0.1436845	1.9593262	0.4282072	4.5894358	0.0423950
iroB	-0.6835411	-1.5487431	0.3253881	1.8075733	0.3898684	4.5673432	0.0454461
STnc3480	-1.2487816	-2.1987947	-0.4805451	1.7480026	0.3853570	4.4806411	0.0438599
STM1859	-1.0041193	-1.6652726	-0.4261300	1.2345657	0.2804787	4.3041550	0.0493504
cspH	0.4536230	0.9742813	-0.3264256	-1.3168060	0.3194814	-4.1387208	0.0480528
yhjR	-0.1756356	0.4202552	-0.9369607	-1.3718760	0.2841055	-4.8055582	0.0448743
cysA	2.3748489	2.8967591	1.8544424	-1.1301275	0.2335025	-4.9266566	0.0385621
isrE	-0.1737740	0.2856047	-0.7519336	-1.0399823	0.1961072	-5.2014157	0.0380492
yjiY	2.0678575	2.5956338	1.5220751	-1.1189254	0.1988555	-5.6059481	0.0276897

cysJ	3.9614208	4.5181894	3.2920396	-1.2763657	0.1891843	-6.8369870	0.0348737
cysI	4.0561641	4.6862540	3.3470551	-1.3958327	0.2040928	-6.9060438	0.0464928
wzzB	2.6565642	3.1898280	2.1238768	-1.0753606	0.1565660	-6.9308169	0.0289536
cspA	6.4190627	7.0972250	5.6422876	-1.4881110	0.2082095	-7.0676788	0.0140465
yjcO	0.7339369	1.3027223	0.0710263	-1.2278676	0.1694571	-7.1802617	0.0261499
yegD	1.1056701	1.7327087	0.4484253	-1.2868003	0.1609639	-7.9673815	0.0215276
ugd	1.0030890	1.5379909	0.3804140	-1.1561096	0.1420090	-8.1626870	0.0244947
ydgI	2.4527405	2.9073630	1.8943307	-1.0247756	0.1214495	-8.4530498	0.0204373
bcsA	2.0048765	2.4911519	1.4628315	-1.0422871	0.1228616	-8.6085641	0.0232635
yhjQ	1.4100545	1.9238030	0.7873987	-1.1376897	0.1192808	-9.4896724	0.0194487
cysW	2.3953118	2.9673923	1.7429384	-1.2313342	0.1244141	-9.7669740	0.0261149
eptA	3.1749098	3.7221176	2.4492119	-1.2671583	0.1198872	-10.4330608	0.0195211
ptpS	1.9141325	2.5221760	1.2672145	-1.2641315	0.1164826	-10.8257666	0.0145281
malT	2.6702353	3.2014417	2.0745074	-1.1350736	0.1015697	-11.2654133	0.0260452

Supplemental Table 5. 63 annotated hits from unpublished RNA-Seq data at mid-exponential phase.

Summary of gene hits in which the differential gene expression was statistically significant between *Salmonella* SL1344 WT and Δ IS200. Gene, mean expression, WT expression, Δ IS200 expression listed along with the differential expression between strain and within strains. Effect is the differential expression between divided by within, with effect sizes greater than 2 as true hits ($p < 0.05$).

Gene name	Mean Exp	WT Exp	Δ IS200 Exp	Diff between	Diff within	Effect	p<0.05
cheZ	5.2670019	4.3121173	6.2143573	1.9157751	0.0890853	21.5458641	0.0161020
sipC	2.5283679	0.5845236	4.3633573	3.7517492	0.1870156	19.9032220	0.0155176
	1.7117418	0.3972633	2.9063959	2.5026222	0.1385867	18.0398831	0.0227129
STnc205	4.9847497	3.9501277	5.9999659	2.0489146	0.1177398	17.3153408	0.0155151
spaN	1.5525019	0.1118872	2.6836172	2.5726751	0.1530546	16.6289153	0.0357327
invE	1.2027010	-0.2309793	2.3472115	2.6016375	0.1607038	16.1445174	0.0203799
sipB	1.9252995	0.2305561	3.4717673	3.2661592	0.2218119	14.8457792	0.0258741
CheW	6.2590604	5.3626458	7.3974034	2.0298439	0.1360031	14.4971471	0.0439135
invF	0.9857652	-0.5281452	2.0670169	2.5970865	0.1828588	14.0486436	0.0428099
cheB	3.8103431	2.9430749	4.7578861	1.8172686	0.1352600	13.4705451	0.0494895
flgB	3.8707521	3.3162397	4.3765613	1.0726052	0.0852186	12.6620472	0.0272627
prgK	2.2349710	1.0765687	3.3944858	2.2971320	0.1874070	12.2310380	0.0258071

orgA	2.0825794	0.9797887	3.0966075	2.1222664	0.1724229	12.2174440	0.0229205
spaO	1.2824126	-0.1288843	2.4988525	2.6274832	0.2165746	12.1251160	0.0246835
srfC	3.2566932	2.5650546	4.0228705	1.4276747	0.1210628	11.7266047	0.0314885
sopB	1.1366038	-0.7684114	2.8838864	3.6565262	0.3134396	11.6679989	0.0250199
tsaB	2.0348473	1.4393317	2.5182765	1.0808444	0.0946893	11.2765558	0.0310723
prgH	1.3102809	-0.1137509	2.6756765	2.7700318	0.2498156	11.1603976	0.0268573
sicA	0.3187946	-1.5442163	1.9281060	3.5219559	0.3304092	10.7244559	0.0289045
STM3032	0.3323171	-0.6094708	1.1606066	1.7808483	0.1714355	10.4627518	0.0314544
STnc1405	0.5583179	-0.6020443	1.4786849	2.0943019	0.2014278	10.4184175	0.0347791
invC	0.1697927	-0.9559375	1.1882728	2.1556847	0.2099553	10.2285752	0.0303258
yebF	1.2895243	0.6826461	1.8721326	1.1886106	0.1183828	10.0759054	0.0338613
sipA	0.1148032	-1.3637579	1.1690743	2.5504571	0.2552558	9.9094383	0.0382634
cheM-							
cheR_IGR	2.3880383	1.3522051	3.3395845	1.9733992	0.1984499	9.9056832	0.0269042
SL1344_R							
S24695	3.0257135	2.3520525	3.5674258	1.2190104	0.1238792	9.8238372	0.0271217
yoaA	2.0064442	1.4770936	2.4976937	1.0312869	0.1044538	9.7825679	0.0376903
fliZ	4.6622763	3.6655537	5.5432374	1.9539044	0.2105002	9.4054721	0.0205739
sopE	0.9259932	-1.2051904	2.7707295	3.9761638	0.4170998	9.3931968	0.0320211
motB	5.1758840	4.0082676	6.2259831	2.1137086	0.2348249	9.2347059	0.0201769
invG	1.7024642	0.2558528	2.9424346	2.7597127	0.3023781	9.2159402	0.0256691
STnc600	1.6615944	0.8399498	2.4950693	1.6567388	0.1827632	9.0293119	0.0453357
cheR	4.2763391	3.2233023	5.2422038	1.9110275	0.2161839	8.9681867	0.0216582
motA	1.1853464	-0.1571050	2.3341056	2.5433440	0.2875715	8.9611695	0.0348074
hilA	1.2626883	0.1936048	2.3596938	2.1089420	0.2372747	8.9113635	0.0386448
yeaY	1.5514322	1.0078975	2.1190966	1.0986680	0.1227963	8.9075092	0.0401893
invA	-0.2140046	-1.6405199	1.0191008	2.6234945	0.3002199	8.7156309	0.0405136
yecC	1.1887180	0.5342533	1.8060047	1.2881785	0.1495304	8.5995053	0.0388848
fliA	5.3756673	4.4654694	6.3801675	1.9705161	0.2395104	8.2271383	0.0379433
sigE	1.0308208	-0.8106257	2.8485206	3.5050047	0.4342693	8.1882222	0.0280262
mobA	4.1497968	2.8070888	5.1384735	2.2828191	0.2798755	8.1845179	0.0387510
nrfB	-0.0464889	-1.0730077	0.8513451	1.9154491	0.2363536	8.0664051	0.0407692
cheM	5.3130962	3.9450699	6.3274825	2.2938317	0.2825461	8.0582113	0.0441054
spaM	-0.0601211	-1.7531390	1.2515520	3.0939170	0.4001106	7.9422341	0.0373859
STM1911	1.0500081	0.2395983	1.6453329	1.4272587	0.1898613	7.6449352	0.0415497
iagB	0.9637832	-0.1294711	1.9260383	2.0055832	0.2642201	7.5544340	0.0375148
srfB	3.6559524	2.8938272	4.3911991	1.4692223	0.1932123	7.4810622	0.0366453

srfA	3.2259442	2.4743557	4.1080317	1.5657005	0.2069099	7.3895636	0.0441768
prgJ	0.3853627	-1.0677381	1.8765238	2.9122820	0.3920864	7.3782518	0.0377085
prgI	-0.5978556	-2.3642917	0.7398841	3.0131906	0.4123082	7.2706159	0.0499912
STM1934	5.0675745	3.6104295	6.1483309	2.5299788	0.3539379	7.1910655	0.0332325
sopA	-0.9357802	-2.1129245	0.0534973	2.1999115	0.3072009	7.0986236	0.0466303
flgD	4.0291606	3.5431103	4.5415678	1.0694287	0.1551345	6.9377125	0.0334321
flhB	3.4036424	2.4000153	4.3359530	2.0023898	0.3228136	6.5489412	0.0434157
flgK	5.4999887	4.7482932	6.1304928	1.3810927	0.2666397	5.3729771	0.0472118
ycgR	2.6775715	2.0915194	3.3451785	1.1551713	0.2257491	5.3018974	0.0478393
yoaB	3.3533363	2.8877431	3.9228468	1.0290591	0.2084278	5.0911161	0.0492557
STM3152	3.8667116	3.1599284	4.6994148	1.3883192	0.3093656	4.9056746	0.0485691
STM3216	5.5467347	4.7412036	6.1143002	1.4257907	0.3261224	4.6787566	0.0467669
malE	4.9350079	5.4125400	4.3690517	-1.0059472	0.1424976	-7.1423634	0.0377151
eco	0.6581829	1.4233763	-0.1467115	-1.5957753	0.2248611	-7.1595898	0.0449839
speF	0.5770017	1.3863438	-0.2763232	-1.6552329	0.1711223	-9.7015768	0.0351704
wzzB	2.5055086	3.2763162	1.8143979	-1.4525133	0.1328001	-10.8808066	0.0338611

Supplemental Table 6. 95 annotated hits from unpublished RNA-Seq data at late exponential phase.

Summary of gene hits in which the differential gene expression was statistically significant between *Salmonella* SL1344 WT and Δ IS200. Gene, mean expression, WT expression, Δ IS200 expression listed along with the differential expression between strain and within strains. Effect is the differential expression between divided by within, with effect sizes greater than 2 as true hits ($p < 0.05$).

Gene name	Mean Exp	WT Exp	Δ IS200 Exp	Diff between	Diff within	Effect	$p < 0.05$
cysP	2.3155845	-0.9359389	5.3972659	5.6403185	0.9194802	6.7039856	0.0257233
cysD	2.1101545	-1.0590731	4.7951924	5.4666440	0.7428563	7.7701802	0.0275228
cysU	1.7722971	-1.1410261	4.0103285	4.8645842	0.7463191	6.6471001	0.0369511
cysN	3.4399150	0.7357263	5.6314723	4.6937667	0.5719007	7.2394517	0.0400621
cysW	2.6592884	0.1776177	4.6634168	4.3579725	0.4900917	8.2429326	0.0406424
cysC	2.6498392	0.0526279	4.4916441	4.3397593	0.4242982	9.7940841	0.0492531
cysK	5.3279250	3.2260045	7.1581382	3.5699330	0.4915255	6.9537349	0.0288580
yciW	0.9128701	-0.8152505	2.6724662	3.3861075	0.4671808	7.2613957	0.0297162
ygbE	0.8370044	-0.8692086	2.2379229	3.1226720	0.2355450	13.4474373	0.0349831

mobA	3.6919197	2.5597304	4.7563515	2.2379701	0.1380811	16.4319972	0.0228297
mobC	2.2425674	0.9999724	3.2168783	2.1935240	0.2881294	7.5448982	0.0497247
STM1528	-1.4556047	-2.5077866	-0.6404267	1.8863875	0.3709098	5.0715141	0.0475753
STM1934	2.4319892	1.4653419	3.3477868	1.8748752	0.1030252	18.1303877	0.0174152
motA	5.1448308	4.3234108	6.1583134	1.8397516	0.2348076	7.9955307	0.0250483
motB	5.2332585	4.3506778	6.1857710	1.8149179	0.1552929	11.8898515	0.0172047
cysM	5.3125662	4.4598588	6.1346580	1.7728585	0.1857044	9.4546732	0.0337927
sdiA	1.2748912	0.3691678	2.1717852	1.7479471	0.2153740	8.1104096	0.0282792
cheW	3.0768841	2.1999900	3.9146557	1.7222033	0.1716252	10.0950441	0.0307883
STM3032	6.6230551	5.8013311	7.4427385	1.7133689	0.1603441	10.6825159	0.0187796
sopE2	0.6439700	-0.2800047	1.4225431	1.7053158	0.1313254	12.8616840	0.0242229
yghW	2.4060718	1.6105985	3.3637725	1.6964753	0.2112233	7.6840803	0.0422959
cheA	0.7647197	-0.0275155	1.7145798	1.6837593	0.2699920	6.1729686	0.0420302
cheR	6.5203163	5.7699917	7.4314330	1.6665138	0.1810071	9.3899176	0.0299874
STM3216	4.3430665	3.5237985	5.1883839	1.6641430	0.1114878	14.9259737	0.0182036
STnc1065	1.8998780	1.1057946	2.7257047	1.6292498	0.2657205	6.1479460	0.0367844
cheM-							
cheR_IGR	2.9536240	2.1573612	3.7077951	1.5847389	0.1483718	10.8598540	0.0244309
cheM	5.4068203	4.6527513	6.1721872	1.5831314	0.1359961	11.6939940	0.0188761
flhA	5.8915970	5.1730701	6.7256274	1.5535003	0.1282872	11.8035955	0.0316727
flhB	4.3886655	3.5755990	5.1236130	1.5497970	0.0913596	16.9336315	0.0159920
rflM	3.7524625	2.9669413	4.4797079	1.5108900	0.0852850	17.6954201	0.0169402
cheB	0.7412276	0.0062020	1.4777078	1.4626509	0.2022707	7.2480910	0.0353432
fliA	3.9964986	3.2515922	4.6769042	1.4428725	0.1689368	8.6552213	0.0264824
sraC	5.6909897	5.0539393	6.4976689	1.4356048	0.2066823	6.9911519	0.0255575
fliY	2.3309265	1.6526131	3.0718475	1.4151018	0.1742961	7.9287863	0.0375244
cheZ	4.7129908	4.0087039	5.3724202	1.4094380	0.1584818	9.0420028	0.0407084
flhE	5.3284273	4.6504211	6.0444290	1.3956457	0.157961	8.6706809	0.044167
trg	2.5354327	1.8290929	3.2114876	1.3810431	0.110628	12.492507	0.022374
SL1344_R							
S08105	5.1569133	4.5174913	5.8453254	1.3588947	0.205448	6.8220327	0.031999
SL1344_R							
S09745	-0.4647523	-1.2136022	0.1334835	1.3399452	0.226194	5.8685795	0.045897
dadA	1.7470585	1.0222795	2.3826747	1.3295496	0.176820	7.4949616	0.034906
yecF	1.1359434	0.5223875	1.8391850	1.3081373	0.173310	7.6071323	0.037755
cheY	4.3519225	3.7158951	4.9358867	1.2993379	0.175684	7.4026825	0.048208
srfA	3.6451490	2.9769127	4.2434850	1.2817205	0.1022373	12.6288029	0.0243785

dadX	1.6885196	0.9475546	2.2360477	1.2796615	0.1522757	8.3136530	0.0453722
yecC	1.8538459	1.1641737	2.4323624	1.2675642	0.0979965	12.9474004	0.0249912
ycgR	2.4194027	1.8180514	3.0098263	1.2666471	0.2024748	6.4468464	0.0387543
SL1344_R							
S24580	3.0397270	2.4233852	3.6892337	1.2534181	0.1126770	11.0891271	0.0282886
cyoE	4.2766004	3.5192510	4.9045285	1.2352566	0.2389550	5.0525969	0.0417814
yecS	0.8731913	0.2487449	1.4728596	1.2330196	0.1905244	6.5080305	0.0385061
modC	3.5668505	2.9610365	4.1662204	1.2156758	0.1272873	9.6931677	0.0326314
modB	3.5184141	2.9709939	4.2039775	1.2093746	0.1341703	8.7519114	0.0332831
cheV	5.1962179	4.5900910	5.8633508	1.1982593	0.1393653	8.5640999	0.0276390
ruvA	1.0150549	0.3824299	1.5738142	1.1856122	0.1619919	7.1722712	0.0335417
yedO	1.8853228	1.3129774	2.4592844	1.1589774	0.1267968	9.1664685	0.0312011
ymdA	2.4203490	1.8255769	2.9260792	1.1361191	0.1454677	7.9384235	0.0388843
flgB	4.2244483	3.6706107	4.7739711	1.1227746	0.1091039	10.2709671	0.0225542
STM3156	1.8757032	1.2936355	2.3806924	1.1058655	0.1363366	8.0688101	0.0384339
fadD	1.3990174	0.8542058	1.9294532	1.0762227	0.1418585	7.6857157	0.0317107
STM3155	1.4532883	0.8455731	1.8986577	1.0720911	0.1787459	6.0897769	0.0374441
gltJ	0.4145809	-0.1715366	0.9079210	1.0690198	0.1865976	5.7178868	0.0446116
SL1344_R							
S24700	2.9536809	2.4210622	3.4537091	1.0530980	0.1036103	10.2072755	0.0253298
STM1300	3.3901930	2.7863871	3.8412754	1.0519175	0.1853752	5.6782405	0.0402865
SL1344_R							
S24710	2.4907041	2.0090167	3.0801235	1.0491572	0.1782883	5.8403124	0.0436003
SL1344_R							
S24585	1.0441763	0.5637544	1.5973448	1.0275252	0.1886718	5.3941194	0.0430912
sigE	4.3230930	3.7482791	4.7690504	1.0231870	0.2224176	4.8286712	0.0412230
invR	5.4049283	4.8051605	5.9458745	1.0193200	0.2685342	3.8553272	0.0460660
SL1344_R							
S24695	2.7463550	2.2277893	3.2130261	1.0020818	0.1159607	8.6818887	0.0347271
silD	1.9177444	2.3630907	1.3350980	-1.0054853	0.2213579	-4.4830222	0.0451794
STM4305	2.4193406	2.8932525	1.8848949	-1.0255293	0.1306306	-7.9472893	0.0302761
wzzB	2.4627605	2.9365307	1.8864048	-1.0506635	0.0864093	-12.1289065	0.0287432
cbiK	1.2148254	1.7260343	0.6703794	-1.0551445	0.1463132	-7.2262164	0.0364411
dmsB	0.5148806	1.0004016	-0.0870861	-1.0693500	0.2164391	-4.8469825	0.0497281
pfkB	2.8877805	3.4812457	2.3168378	-1.0861034	0.1943194	-5.6496737	0.0350891
cbiC	0.9286161	1.4763484	0.4029920	-1.0863754	0.1410878	-7.7145486	0.0343108
cbiL	1.2487743	1.7281998	0.6464601	-1.0919335	0.2153083	-5.0793276	0.0428609

cbiH	1.7968579	2.3157166	1.2132046	-1.1062426	0.2191265	-5.0838213	0.0434302
yeaG	2.5667759	3.1694969	2.0368926	-1.1422445	0.1211890	-9.5897287	0.0335337
orf48	-0.0856519	0.4800918	-0.6791170	-1.1598656	0.1660595	-6.9853497	0.0385204
ybeL	1.9615265	2.4786319	1.3250957	-1.1623301	0.1983494	-5.9087801	0.0438298
cbiG	1.5645192	2.1689023	0.9937580	-1.1762567	0.1963862	-6.1085850	0.0427806
glgS	1.9481645	2.4820847	1.3634620	-1.1877652	0.1935764	-6.1414356	0.0337450
tdcA	0.3521106	0.9078799	-0.2875850	-1.1903755	0.1672982	-7.1069002	0.0360673
STM4519	1.9371317	2.5526571	1.3217040	-1.2355460	0.1595662	-7.7432767	0.0306003
cbiF	1.1673571	1.6548504	0.3246438	-1.3635959	0.3416956	-4.0825486	0.0481515
csiE	0.9447392	1.6412130	0.2429036	-1.3857471	0.1673324	-8.1918828	0.0323464
STM4535	0.8108688	1.4093354	-0.0018534	-1.3859302	0.2075030	-6.5874914	0.0421310
STM4539	1.6312991	2.2501074	0.8271177	-1.4326526	0.1699094	-8.4820922	0.0283496
STM4538	1.3121802	2.0696799	0.5667292	-1.4998370	0.1342463	-11.1783590	0.0256544
STM4536	0.4919620	1.1913678	-0.3209243	-1.5450814	0.2241517	-6.9748015	0.0421148
STM4537	1.0172247	1.6744196	0.1158541	-1.5568634	0.1672952	-9.3135168	0.0416487
yjfN	1.1231564	1.8393380	0.2292343	-1.6307221	0.1970984	-8.3633555	0.0287552
yjfO	2.7595929	3.7080779	1.8834247	-1.8032311	0.2928956	-6.1383756	0.0358562
phsC	0.5855642	2.5389868	-1.2202502	-3.6902305	0.4161986	-8.9577719	0.0324752
phsB	-0.4306555	1.5965706	-2.3306998	-3.7943069	0.5265284	-7.2209721	0.0352893
phsA	1.2909279	3.6932779	-0.7906747	-4.4098646	0.3228692	-13.3119044	0.0370904

Curriculum Vitae

Name: Naomi-Jean Qiu-Yi Scherba

Post-secondary Education and Degrees: Carleton University
Ottawa, Ontario, Canada
2015-2019 B.Sc. Honours

The University of Western Ontario
London, Ontario, Canada
2019-2021 M.Sc.

Honours and Awards: Gordon Robertson Scholarship
2018

Related Work Experience Teaching Assistant
Biochemistry 2280
The University of Western Ontario
2019-2020

Committees Biochemistry Graduate Student Association
Biochemistry Department
2019-2021

Graduate Peer Support Chair
Society of Graduate Students
2020-2021

Orientation and Social Committee
Society of Graduate Students
2020-2021