



University of  
**Nottingham**  
UK | CHINA | MALAYSIA

# Utilizing the Zinc Homeostasis System of *Escherichia coli* as a Novel Inducible Promoter System

by  
Joseph Ingram *BSc, MSc*

Thesis submitted to The University of Nottingham for the degree of  
Doctor of Philosophy

September 2020

*For my Grandad  
George Ingram*



## Acknowledgements

I would first like to thank my three supervisors; Jon Hobman, Phil Hill, and Dov Stekel for their support and guidance throughout my PhD. Dov, for his support with the Bayesian analysis of flow cytometry data. Phil, for his knowledge and help with all things fluorescent and luminescent. And Jon, for his constant wisdom, and for giving me this opportunity to pursue my PhD. Thanks to all the technical and administrative staff at Food Science.

I would like to thank all my friends I made during my PhD, “The Food Sci Crew”, for making my time during my PhD fun, and for making me smile on the difficult days. Alex, for being my lab buddy; helping me during the crazy lab experiments; bouncing ideas off; and for listening to me rant when my experiments fail. Holly, for being a great house mate; musical theatre buddy; and most importantly introducing me to Half Birthdays.

A special thanks to Cellan for his constant support, encouragement and most importantly, for believing in me even when I didn’t believe in myself.

Thanks to my parents, brother and all my family for their support and always pushing me to achieve my goals. To my Mom, for believing in me from a child that I could achieve anything I wanted.

## Abstract

$\text{Zn}^{2+}$  is an essential post-transition metal found in all life, however, at high concentrations  $\text{Zn}^{2+}$  can become toxic, causing oxidative stress and inactivation of essential enzymes by replacing other metals in catalytic centres of proteins. *Escherichia coli* cells can control their internal zinc to a femtomolar concentration, equivalent to one to two free  $\text{Zn}^{2+}$  ions per cell. This tightly controlled zinc homeostasis system is regulated by two transcription factors, ZntR and Zur, which regulated the expression of the major zinc export and zinc acquisition genes. This zinc homeostasis system of *E. coli* has the potential to be developed as a novel inducible promoter system.

The ZntR regulated promoter ( $P_{zntA}$ ) demonstrated a strong correlation between increasing zinc concentrations and promoter induction, as well as showing a comparable induction level to the IPTG inducible promoter,  $P_{trc}$ . Six Zur regulated promoters showed a variation of induction level when induced with the zinc chelator TPEN ( $P_{c1265} > P_{ykgM} > P_{znuA} > P_{znuCB} > P_{plIG} > P_{zinT}$ ). Promoters  $P_{c1265}$  and  $P_{ykgM}$  demonstrated more desirable characteristics than the IPTG inducible  $P_{trc}$ ; showing lower basal expression, higher induced expression and higher fold induction. Both ZntR and Zur regulated promoters show strong potential to be utilized as either a zinc or TPEN inducible promoter for use in both research and biotechnology.

Flow cytometry data of *E. coli zntA::rfp* in combination with Bayesian analysis demonstrated that *E. coli* mounts a heterogenous gene expression of *zntA*. This analysis further showed that with increasing zinc concentration, *E. coli* shifts the heterogenous *zntA* expression, reducing low level gene expression and increasing high level gene expression.

*In silico* analysis of the uncharacterised Zur regulated C1265-7 suggested that C1265 is a TonB-dependent receptor which translocates zinc, C1266 may be involved in bacteria-host adhesion, and C1276 is likely a COG0523 protein. *In vitro* analysis did not suggest a phenotype and it is likely that the true phenotype of C1265-7 can only be observed in an infection model.

# Contents

<b>Acknowledgements .....</b>	<b>ii</b>
<b>Abstract .....</b>	<b>iii</b>
<b>Contents .....</b>	<b>v</b>
<b>Figures .....</b>	<b>ix</b>
<b>Tables .....</b>	<b>xii</b>
<b>Abbreviations .....</b>	<b>xiii</b>
<b>1. Introduction .....</b>	<b>1</b>
<b>1.1. Gene Regulation .....</b>	<b>2</b>
1.1.1. Transcription Cycle .....	3
1.1.2. Bacterial Promoters .....	5
1.1.3. RNA Polymerase .....	6
1.1.4. RNA Polymerase Holoenzyme ( $\sigma^{70}$ ) .....	8
1.1.5. Transcription Factors .....	10
1.1.6. Alternative RNAP Holoenzyme Sigma Factors .....	16
<b>1.2. Plasmid Biology .....</b>	<b>18</b>
1.2.1. Reporter Plasmids .....	18
<b>1.3. Biotechnology .....</b>	<b>21</b>
1.3.1. Inducible Promoter Systems .....	22
<b>1.4. Zinc Homeostasis in <i>E. coli</i> .....</b>	<b>25</b>
1.4.1. Low Affinity Zinc Import and Export systems .....	26
<b>1.5. Zur (Zinc Uptake Regulator) .....</b>	<b>28</b>
1.5.1. Zur Structure .....	31
1.5.2. Zur Regulated Genes .....	33
<b>1.6. ZntR .....</b>	<b>37</b>
1.6.1. ZntR- $P_{zntA}$ .....	38
1.6.2. ZntA .....	39
1.6.3. Stochastics of ZntA Regulation .....	40
<b>1.7. Aims and Objectives .....</b>	<b>42</b>
<b>2. Materials and Methods .....</b>	<b>43</b>
<b>2.1. Materials .....</b>	<b>44</b>
2.1.1. Bacterial Strains .....	44
2.1.2. Plasmids .....	45
2.1.3. Primers .....	47
2.1.4. Media .....	58
2.1.5. Antibiotic Stock .....	61
2.1.6. Metal Solutions .....	62
2.1.7. Solutions .....	64
2.1.8. Buffers .....	65
2.1.9. Enzymes .....	67
<b>2.2. Methods .....</b>	<b>68</b>

2.2.1. <i>Escherichia coli</i> Propagation and Maintenance .....	68
2.2.2. <i>Cupriavidus necator</i> Propagation and Maintenance .....	70
2.2.3. DNA Manipulation and Purification .....	72
2.2.4. RNA Purification and Quantification.....	76
2.2.5. Polymerase Chain Reaction (PCR).....	76
2.2.6. Nitric Acid Soaking of PMP Erlenmeyer Flasks.....	78
2.2.7. Nitric Acid Soaking of Magnetic Followers .....	78
2.2.8. Zinc Depleted Neidhardt's MOPS Minimal Media (NH) Preparation .....	78
2.2.9. pJI100 Plasmid Series Construction .....	82
2.2.10. pJI300 Reporter Plasmid Construction .....	85
2.2.11. pJI300 Plasmid Series Promoter Inserts.....	86
2.2.12. pJI300 Promoter Insert Confirmation via Colony PCR .....	87
2.2.13. $P_{zntA}$ (pJI301) Site Directed Mutagenesis.....	87
2.2.14. Random Mutagenesis of $P_{zntA}$ .....	88
2.2.15. Gene Coupling of <i>mFRP1</i> to <i>zntA</i> by Gene Doctoring.....	91
2.2.16. Temporal Luminescence Assays .....	93
2.2.17. Luminescence Assay Measurement.....	94
2.2.18. End Point Fluorescence Assay.....	95
2.2.19. BacLight LIVE:DEAD Assay.....	96
2.2.20. Flow Cytometry.....	98
2.2.21. C1265-7 Phenotype Growth Experiments .....	101
2.2.22. ICP-MS Metal Analysis of Media .....	103
2.2.23. RT-qPCR to Determine Relative Abundance of <i>zntA</i> .....	104
<b>3. Fundamental Understanding of the ZntR Transcriptional Regulation of the <i>zntA</i> Promoter <math>P_{zntA}</math>.....</b>	<b>107</b>
<b>3.1. Introduction.....</b>	<b>108</b>
3.1.1. Aims and Objectives .....	109
<b>3.2. Results .....</b>	<b>110</b>
3.2.1. Nomenclature .....	110
3.2.2. Validation of Zinc Depleted Neidhardt's MOPS Minimal Media (NH) .....	110
3.2.3. ICP-MS Analysis of Growth Media .....	111
3.2.4. BacLight Toxicity Tests .....	114
3.2.5. pJI300: A Dual Reporter Plasmid.....	116
3.2.6. $P_{zntA}$ Mutants: Construction and Validation .....	119
3.2.7. End Point Fluorescence Assays in Lysogeny Broth (LB) .....	121
3.2.8. End Point Assays in NH Broth .....	124
3.2.9. Comparison of End Point Assays in LB and NH .....	127
3.2.10. End Point Assay of $P_{zntA}$ Mutants in LB Broth.....	129
3.2.11. RT-qPCR .....	132
3.2.12. RT-qPCR: mRNA Variation of <i>zntA</i> .....	134
3.2.13. Temporal Luminescence Assay in NH Broth .....	139
3.2.14. Temporal Luminescence in LB .....	143
3.2.15. Temporal Luminescence: Comparison of LB and NH .....	147
3.2.16. Alternative Bacteria .....	148
<b>3.3. Discussion and Conclusion.....</b>	<b>150</b>
3.3.1. Growth Media Validation .....	150
3.3.2. End Point Fluorescence Assays .....	151
3.3.3. End Point Fluorescence Assay of Mutant $P_{zntA}$ .....	152
3.3.4. mRNA Analysis of <i>zntA</i> Through RT-qPCR.....	154
3.3.5. Temporal Luminescence Assays .....	156

3.3.6. Alternative Bacteria .....	157
3.3.7. Concluding Remarks .....	158
3.3.8. Future work .....	158
<b>4. Fundamental Understanding of Zur Regulated Promoters .....</b>	<b>160</b>
<b>4.1. Introduction.....</b>	<b>161</b>
4.1.1. Aims and Objectives .....	162
<b>4.2. Results .....</b>	<b>163</b>
4.2.1. Zur Regulated Promoters.....	163
4.2.2. End Point Fluorescence Assay In NH.....	167
4.2.3. End Point Fluorescence Assays in LB .....	171
4.2.4. Temporal Luminescence Assay in NH .....	177
4.2.5. Temporal Luminescence Assays in LB.....	179
4.2.6. Further understanding of $P_{c1265}$ .....	184
<b>4.3. Discussion and Conclusion.....</b>	<b>186</b>
4.3.1. End Point Fluorescence Assay.....	186
4.3.2. Temporal Luminescence Assays .....	188
4.3.3. Concluding remarks .....	189
4.3.4. Future work .....	190
<b>5. Understanding the Dynamic Response of the Transcriptional Regulation of <i>zntA</i> .....</b>	<b>192</b>
<b>5.1. Introduction.....</b>	<b>193</b>
5.1.1. Aim and Objectives .....	193
<b>5.2. Results .....</b>	<b>194</b>
5.2.1. Construction and Validation of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> .....	194
5.2.2. <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> Test of Validation.....	197
5.2.3. Flow Cytometry.....	198
5.2.4. Flow Cytometry of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> in LB.....	204
5.2.5. Flow Cytometry of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> in NH .....	208
5.2.6. A Bayesian Background Subtraction Approach for Flow Cytometry Data .....	213
5.2.7. Bayesian Background Subtraction: True Foreground Fluorescence of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> in LB .....	218
5.2.8. Bayesian Background Subtraction: True Foreground Fluorescence of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> in NH .....	220
<b>5.3. Discussion and Conclusion.....</b>	<b>223</b>
5.3.1. Construction of <i>E. coli</i> MG1655 <i>zntA::rfp::kan<sup>R</sup></i> .....	223
5.3.2. Raw Flow Cytometry Fluorescence Data .....	223
5.3.3. Bayesian Background Subtraction of Flow Cytometry Fluorescence Data .....	225
5.3.4. Concluding Remarks .....	227
5.3.5. Future Work.....	227
<b>6. Determining the Phenotype of the Novel Zur Regulated Operon, C1265-7 .....</b>	<b>230</b>
<b>6.1. Introduction.....</b>	<b>231</b>
6.1.1. Aim and Objectives .....	232
<b>6.2. Results .....</b>	<b>233</b>
6.2.1. Promoter Activity.....	233
6.2.2. Bioinformatics.....	234

6.2.3. Microtiter Plate Growth Assay.....	239
6.2.4. Shake Flask Growth Curve Assay .....	240
6.2.5. Construction and validation of pJI100A and pJI102: C1265-7 Expression Plasmid.....	243
6.2.6. Growth Assays with the <i>c1265-7</i> Expression Vector (pJI102).....	245
6.2.7. BacTiter-Glo Cell Viability Assay .....	249
<b>6.3. Discussion and Conclusion.....</b>	<b>253</b>
6.3.1. Bioinformatics.....	253
6.3.2. Growth Curve Experiments.....	255
6.3.3. Concluding Remarks .....	256
6.3.4. Future Work.....	257
<b>7. General Discussion, Conclusion, and Future Work .....</b>	<b>259</b>
<b>7.1. Discussion .....</b>	<b>260</b>
7.1.1. Zinc Depleted Neidhardt's MOPS Minimal Media (NH).....	261
7.1.2. A Novel $P_{zntA}$ Zinc Inducible Promoter.....	262
7.1.3. Zur Regulated Promoters for Novel Inducible Promoters .....	265
7.1.4. Heterogeneous <i>zntA</i> Gene Expression.....	267
7.1.5. Bayesian Background Subtraction Approach to Flow Cytometry Data.....	268
7.1.6. <i>In silico</i> Approach to Determine the C1265-7 Phenotype .....	270
7.1.7. <i>In vitro</i> Approach to Determine the C1265-7 Phenotype .....	272
<b>7.2. Conclusion .....</b>	<b>275</b>
<b>7.3. Future Direction of Work .....</b>	<b>277</b>
<b>8. References .....</b>	<b>279</b>
<b>9. Appendix .....</b>	<b>294</b>
<b>9.1. Supplementary Sequences .....</b>	<b>295</b>
9.1.1. pJI100 Sequence .....	295
9.1.2. pJI100A Sequence.....	296
9.1.3. pJI102 Sequence .....	297
9.1.4. pJI300 Sequence .....	299
9.1.5. pDOC- <i>zntA::rfp</i> Sequence .....	302
9.1.6. Sequence of Promoter Inserts for of pJI300 Plasmid Series .....	304
<b>9.2. Supplementary Figures.....</b>	<b>305</b>
9.2.1. Alternative Media .....	305
9.2.2. Alternative Metal Inducers .....	308
9.2.3. pH of NH Media .....	309
9.2.4. Error Prone PCR .....	310
9.2.5. Flow Cytometry.....	312
9.2.6. BacLight Control.....	314
9.2.7. Alternative Temporal Luminescence Graphs.....	315
9.2.8. Additional End Point Fluorescence Assays.....	316
<b>9.3. Manuscript in Production for Submission .....</b>	<b>317</b>

## Figures

Figure 1.1 Diagram of The Central Dogma.....	2
Figure 1.2 Diagram of the Transcription Cycle.....	4
Figure 1.3 Diagram of the Consensus Promoter Sequence .....	6
Figure 1.4 Diagram of RNA Polymerase Holoenzyme.....	7
Figure 1.5 Diagram of Group 1-4 Sigma Factors .....	9
Figure 1.6 Diagram of Class-I and Class-II Transcription Factor Activation.....	13
Figure 1.7 Diagram of MerR Underwinding of DNA .....	14
Figure 1.8 Diagram of the Mechanisms of TF Repression .....	16
Figure 1.9 Plasmid Map of a Generic Reporter Plasmid .....	19
Figure 1.10 Diagram of The Lac Operon .....	23
Figure 1.11 Diagram of Zinc Homeostasis .....	26
Figure 1.12 Diagram of Zur Repression .....	29
Figure 1.13 Promoter Alignment of Zur Regulated Promoters.....	30
Figure 1.14 Model of (Zur <sub>2</sub> ) <sub>2</sub> -DNA Complex.....	32
Figure 1.15 Zn <sup>2+</sup> Interaction Sites with Zur .....	33
Figure 1.16 Diagram of <i>znuABC</i> Operon Structure .....	34
Figure 1.17 Diagram of ZntR Regulation.....	37
Figure 1.18 Amino Acid Sequence ZntR.....	38
Figure 1.19 Promoter Sequence of <i>zntA</i> .....	39
Figure 1.20 Structure of ZntA <i>Shigella sonnei</i> .....	40
Figure 2.1 Diagram of NEBuilder HiFi DNA Assembly .....	75
Figure 2.2 Plasmid Map of pJI100A.....	83
Figure 2.3 Plasmid Map of pJI102.....	84
Figure 2.4 Plasmid Map of pJI300.....	85
Figure 2.5 Plasmid Map of pDOC- <i>zntA:rfp</i> .....	91
Figure 2.6 Ingram Function.....	101
Figure 3.1 Promoter Sequence of <i>zntA</i> .....	108
Figure 3.2 Molar Concentration of Zinc During <i>E. coli</i> Growth .....	113
Figure 3.3 Zinc Toxicity on <i>E. coli</i> During Mid Log Growth Phase .....	115
Figure 3.4 TPEN Toxicity on <i>E. coli</i> During Mid Log Growth Phase .....	116
Figure 3.5 Plasmid Map of pJI300.....	117
Figure 3.6 PCR Product Gel Image for Promoter Inserts .....	119
Figure 3.7 Sequence Alignment of P <sub><i>zntA</i></sub> Mutations.....	121
Figure 3.8 End Point Fluorescence Assay of ZntR Regulated P <sub><i>zntA</i></sub> in LB with Increasing ZnSO <sub>4</sub> .....	122
Figure 3.9 Regression Analysis; End Point Fluorescence of P <sub><i>zntA</i></sub> in LB.....	123
Figure 3.10 End Point Fluorescence Assay of P <sub><i>zntA</i></sub> and P <sub><i>trc</i></sub> in LB Induced or Uninduced .....	124
Figure 3.11 End Point Fluorescence Assay of ZntR Regulated P <sub><i>zntA</i></sub> in NH with Increasing ZnSO <sub>4</sub> .....	125
Figure 3.12 Regression Analysis: End Point Fluorescence of P <sub><i>zntA</i></sub> in NH.....	126
Figure 3.13 End Point Fluorescence Assay of P <sub><i>zntA</i></sub> and P <sub><i>trc</i></sub> in NH .....	127
Figure 3.14 End Point Fluorescence Assay Comparison of P <sub><i>zntA</i></sub> in Both LB and NH .....	128
Figure 3.15 End Point Fluorescence Assay of P <sub><i>zntA</i></sub> Mutants: Extended Spacer .....	130
Figure 3.16 End Point Fluorescence Assay of P <sub><i>zntA</i></sub> Mutants: Consensus Promoter Elements .....	132
Figure 3.17 RT-qPCR Calibration Curve of <i>rrsA</i> .....	134
Figure 3.18 RT-qPCR Plot of <i>rrsA</i> .....	135
Figure 3.19 qPCR Plot for of <i>zntA</i> and <i>rrsA</i> .....	136
Figure 3.20 $\Delta C_q$ Plot for mRNA of <i>zntA</i> .....	137
Figure 3.21 Comparative C <sub>q</sub> (2 <sup>-<math>\Delta\Delta C_q</math>) Plot for mRNA of <i>zntA</i> .....</sup>	138
Figure 3.22 Comparison of Fold Change; RT-qPCR vs End Point Fluorescence.....	139



Figure 3.23 Temporal Luminescence Assay of ZntR Regulated $P_{zntA}$ in NH Media .....	140
Figure 3.24 Temporal Luminescence Assay of $P_{zntA}$ In NH + 400 $\mu$ M ZnSO <sub>4</sub> : Absorbance and Luminescence .....	140
Figure 3.25 Temporal and Accumulative Assay of $P_{zntA}$ in NH + 400 $\mu$ M ZnSO <sub>4</sub> .....	141
Figure 3.26 Temporal Luminescence Assay of $P_{zntA}$ and $P_{trc}$ in NH .....	142
Figure 3.27 Luminescence Expression Rate of $P_{zntA}$ and $P_{trc}$ in NH.....	143
Figure 3.28 Temporal Luminescence Assay of ZntR regulated $P_{zntA}$ in LB .....	144
Figure 3.29 Temporal Luminescence Assay of $P_{zntA}$ In LB + 400 $\mu$ M ZnSO <sub>4</sub> : Absorbance and Luminescence .....	145
Figure 3.30 Temporal Luminescence of $P_{zntA}$ and $P_{trc}$ in LB .....	145
Figure 3.31 Luminescence Expression Rate of $P_{zntA}$ and $P_{trc}$ in LB.....	146
Figure 3.32 Temporal Luminescence of Induced $P_{zntA}$ in LB and NH.....	147
Figure 3.33 End Point Assay of ZntR Regulated $P_{zntA}$ : Expressed in <i>C. necator</i> .....	149
Figure 4.1 Promoter Alignment of Zur Regulated Promoters.....	166
Figure 4.2 End Point Fluorescence Assay of Zur Regulated Promoters in NH .....	168
Figure 4.3 End Point Fluorescence Assay of Zur Regulated Promoters in NH with Increasing ZnSO <sub>4</sub> Concentrations.....	170
Figure 4.4 End Point Fluorescence Assay of Zur Regulated Promoters in LB.....	172
Figure 4.5 End Point Fluorescence Assay of Zur Regulated Promoters in LB with Increasing ZnSO <sub>4</sub> Concentrations.....	176
Figure 4.6 Temporal Luminescence Assay of Zur Regulated Promoters in NH.....	178
Figure 4.7 Temporal Assay of $P_{c1265}$ in NH; Luminescence and Absorbance.....	179
Figure 4.8 Temporal Luminescence Assay of Zur Promoters in LB + 20 $\mu$ M TPEN .....	181
Figure 4.9 Temporal Luminescence Assay of Zur Regulated Promoters in LB: Induced; Uninduced; and Supressed .....	183
Figure 4.10 Luminescence Expression Rate of Zur Promoters in LB + 20 $\mu$ M TPEN .....	184
Figure 4.11 End Point Fluorescence Assay of $P_{c1265}$ in NH with Increasing ZnSO <sub>4</sub> .....	185
Figure 5.1 Plasmid Map of pDOC-zntA:rfp.....	195
Figure 5.2 PCR Product Gel Image for Conformation of $zntA:rfp::kan^R$ Fusion .....	197
Figure 5.3 End Point Fluorescence Assay of Chromosomally Fused $zntA:rfp::kan^R$ .....	198
Figure 5.4 Flow Cytometry Gate for Injection Rate .....	200
Figure 5.5 Flow Cytometry Gate for Whole Cells.....	201
Figure 5.6 Flow Cytometry Gate for Single Cells.....	202
Figure 5.7 Flow Cytometry Gate for Fluorescence .....	203
Figure 5.8 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> MG1655 (wt and $zntA:rfp::kan^R$ ) in LB Only.....	205
Figure 5.9 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> MG1655 (wt and $zntA:rfp::kan^R$ ) in LB with Added ZnSO <sub>4</sub> .....	206
Figure 5.10 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> MG1655 $zntA:rfp::kan^R$ in LB + 600 $\mu$ M ZnSO <sub>4</sub> .....	207
Figure 5.11 Q-Q plot <i>E. coli</i> MG1655 wt and $zntA:rfp::kan^R$ in LB .....	208
Figure 5.12 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> MG1655 (wt and $zntA:rfp::kan^R$ ) in NH Only .....	209
Figure 5.13 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> MG1655 wt in NH with or without 400 $\mu$ M ZnSO <sub>4</sub> .....	210
Figure 5.14 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> (wt and $zntA:rfp$ ) grown in NH with increasing ZnSO <sub>4</sub> .....	211
Figure 5.15 Histogram (Flow Cytometry Fluorescence Data) of <i>E. coli</i> $zntA:rfp$ in NH + 400 $\mu$ M ZnSO <sub>4</sub> .....	212
Figure 5.16 Q-Q plot <i>E. coli</i> MG1655 wt and $zntA:rfp::kan^R$ in NH .....	213
Figure 5.17 Illustration of Simple Foreground Minus Background Mean.....	214
Figure 5.18 Bayesian Background Subtraction .....	216
Figure 5.19 Bayesian Background Subtraction: Proof of Principle .....	217

Figure 5.20 Bayesian Background Subtraction: Fluorescence of <i>E. coli zntA::rfp</i> Grown in LB Media.....	220
Figure 5.21 Bayesian Background Subtraction: Fluorescence of <i>E. coli zntA::rfp</i> Grown in NH Media.....	222
Figure 6.1 End Point Fluorescent Assay of $P_{c1265}$ in LB and NH (Summary) .....	233
Figure 6.2 Swiss-Model Structure of C1265.....	236
Figure 6.3 Amino Acid Sequence Alignment of C1265 and ZnuD .....	237
Figure 6.4 Amino Acid Sequence Alignment of C1267 and YjiA .....	239
Figure 6.5 Microtiter Growth Curve Assay of <i>E. coli</i> CFT073 (wt and KO strains) in NH Media .....	240
Figure 6.6 Shake Flask Growth Curve Assay of <i>E. coli</i> CFT073 in NH .....	241
Figure 6.7 Shake Flask Growth Rate of <i>E. coli</i> CFT073 (wt and KO) in NH .....	242
Figure 6.8 Shake Flask Growth Curve Assay of <i>E. coli</i> CFT073 (wt and KO) in NH + 50 $\mu$ M ZnSO <sub>4</sub> .....	243
Figure 6.9 Plasmid Map of pJI100A.....	244
Figure 6.10 Plasmid Map of pJI102 .....	245
Figure 6.11 Microtiter Growth Assay of <i>E. coli</i> MG1655 wt (pJI100A or pJI102) in NH .....	246
Figure 6.12 Microtiter Growth Assays of <i>E. coli</i> MG1655 $\Delta$ znuA & $\Delta$ znuCB (pJI100A or pJI102) in NH .....	247
Figure 6.13 Microtiter Growth Assay of <i>E. coli</i> MG1655 $\Delta$ znuA & $\Delta$ znuCB (pJI100A, pJI102) in NH, Five Biological Repeats .....	248
Figure 6.14 BacTiter-Glo Assay of <i>E. coli</i> MG1655 wt, $\Delta$ znuA and $\Delta$ znuCB (pJI100A, pJI102) in NH .....	251
Figure 6.15 Growth Rate of <i>E. coli</i> MG1655 wt, $\Delta$ znuA and $\Delta$ znuCB (pJI100A or pJI102).....	252
Supplementary Figure 9.1 End Point Fluorescence Assay of ZntR Regulated $P_{zntA}$ in Various Standard Growth Mediums .....	305
Supplementary Figure 9.2 End Point Fluorescence Assay of ZntR Regulated $P_{zntA}$ in LB and M17 .....	306
Supplementary Figure 9.3 Temporal Luminescence of $P_{zntA}$ in M17.....	307
Supplementary Figure 9.4 End Point Fluorescence Assay of ZntR Regulated $P_{zntA}$ in NH with Various Metal Inducers.....	308
Supplementary Figure 9.5 pH of NH with Increasing ZnSO <sub>4</sub> concentrations .....	309
Supplementary Figure 9.6 Random Mutagenesis of $P_{zntA}$ , Bioluminescence Image .....	310
Supplementary Figure 9.7 End Point Fluorescence Assay of Randomly Mutated $P_{zntA}$ .....	311
Supplementary Figure 9.8 Histogram (Flow Cytometry Fluorescence Data) of $P_{zntA}$ reporter Plasmid (pJI301) Expressed in LB .....	312
Supplementary Figure 9.9 Regression Analysis: Flow Cytometry data.....	313
Supplementary Figure 9.10 BacLight Standard Curve .....	314
Supplementary Figure 9.11 Temporal Luminescence Assay of $P_{c1265}$ in NH .....	315
Supplementary Figure 9.12 End Point Fluorescence of Zur Regulated Promoter in NH, Expressed in <i>E. coli</i> wt and $\Delta$ zur .....	316

## Tables

Table 2.1 Table of Bacterial Strains .....	44
Table 2.2 Table of Plasmids .....	46
Table 2.3 Table of Primers .....	57
Table 2.4 Neidhardt's MOPS Minimal Media: Molar Composition.....	61
Table 2.5 Q5 Thermocycle Conditions .....	77
Table 2.6 DreamTaq Thermocycle Conditions .....	78
Table 2.7 10x MOPS Composition.....	79
Table 2.8 Neidhardt's Media Composition .....	81
Table 2.9 Q5 Colony PCR Thermocycle Conditions .....	87
Table 2.10 Q5 SDM Thermocycle Conditions .....	88
Table 2.11 Error Prone PCR Composition .....	89
Table 2.12 Error Prone PCR Thermocycle .....	89
Table 2.13 qPCR Thermocycle Conditions .....	105
Table 2.14 RT-qPCR Samples & Controls .....	106
Table 3.1 Molar Concentration of Different Growth Media .....	112
Table 3.2 Fold Increase of Promoter Activity.....	129
Table 4.1 Transcriptomics data of <i>E. coli</i> CFT073 .....	164
Table 4.2 Fold Increase of Zur Promoter Activity in LB.....	174

## Abbreviations

Absorbance <sub>600</sub>	Absorbance measured at 600 nm
Amp	Ampicillin
ATP	Adenosine Triphosphate
bp	Base pair
<i>C. necator</i>	<i>Cupriavidus necator</i>
CDF	Cation Diffusion Family
Chlor	Chloramphenicol
CTD	C-Terminal Domain
dH <sub>2</sub> O	Distilled water
DNA	Deoxyribonucleic Acid
dsDNA	double stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
em	Emission nm
EDTA	Ethylenediaminetetraacetic Acid
<i>et al</i>	And others
ex	Excitation nm
fMet-tRNA <sup>Met</sup>	N-formyl-methionine tRNA
FMN	Oxidised flavin mononucleotide
FMNH <sub>2</sub>	Reduced flavin mononucleotide
FRT	Flp-FRT recombinase site
Fur	Ferric uptake regulator
g	Gram(s)
gDNA	genomic DNA
GFP	Green Fluorescent Protein
GMQE	Global Model Quality Estimation
GTP	Guanosine Triphosphate
H <sub>2</sub> O	Water
HEPES	4-(2-Hydroxyethyl)-1- piperazineethanesulfonic acid
HPLC	High Performance Liquid Chromatography
hr	Hour(s)
HTH	Helix-Turn-Helix
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
Kan	Kanamycin
KO	Knout Out gene
L	Litre(s)
LB	Lysogeny Broth (a.k.a Luria Broth)
M	Molar
m	Milli
MCS	Multiple Cloning Site
Mfd	Mutation frequency decline
min	Minute(s)

mol	Mole
MOPS	3-(N-morpholino)propanesulfonic acid
mRFP1	Monomeric Red Fluorescent Protein
mRNA	messenger RNA
n	Nano
<i>N. meningitidis</i>	<i>Neisseria meningitidis</i>
NH	Zinc depleted Neidhardt's MOPS minimal media
NRAMP	Natural Resistance Associated Macrophage Protein (metal ion transporters)
NTD	N-Terminal Domain
OD <sub>600</sub>	Optical Density measured at 600 nm
ori	Origin of replication
PCR	Polymerase Chain Reaction
P <sub>i</sub>	Inorganic phosphate
PlIG	Periplasmic lysozyme inhibitor of g-type lysozyme
PMP	Polymethylpentene
Pol III	DNA Polymerase 3
(p)ppGpp	Guanosine tetraphosphate and guanosine pentaphosphate
pRNA	Primer RNA
QMEAN	Qualitative Model Energy Analysis
R.E	Restriction Enzyme
RCHO	Long chain aldehyde
RCP	Rolling Circle Replication
RFP	Red Fluorescent Protein
RNA	Ribonucleic Acid
RNAP	RNA Polymerase
RNAPc	RNA Polymerase closed complex
RNAPo	RNA Polymerase open complex
RPM	Revolutions Per Minute
RT-qPCR	Reverse Transcriptase - quantitative Polymerase Chain Reaction
s	Second(s)
<i>S. dysenteriae</i>	<i>Shigella dysenteriae</i>
ssDNA	single stranded DNA
TDP	Adenosine Diphosphate
TF	Transcription Factors
TPEN	N,N,N',N'-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine
tRNA	transfer RNA
UPEC	Uropathogenic <i>Escherichia coli</i>
V/V %	Volume to Volume, percentage
W/V %	Weight to Volume, percentage
Zn <sup>2+</sup>	Zinc (II) ion
ZnSO <sub>4</sub>	Zinc Sulphate

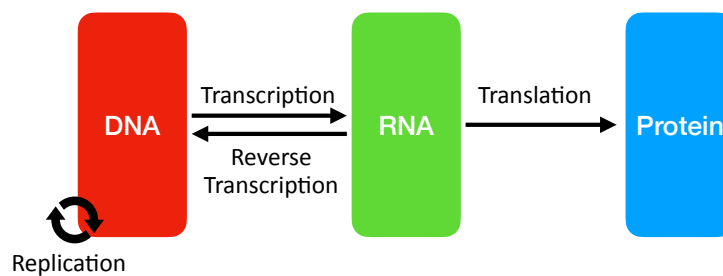
Zur	Zinc uptake regulator
$\alpha$	Alpha
$\beta$	Beta
$\Delta$	Delta (deletion)
$\Theta$	Theta
$\lambda$	Lambda
$\mu$	Mu (micro)
$\sigma$	Sigma
$\omega$	Omega

# Chapter 1

## Introduction

### 1.1. Gene Regulation

Bacteria tightly control gene expression, allowing them to conserve energy and resources. Gene regulation is controlled via two essential steps; transcription and translation. This thesis will focus on gene regulation via transcription. It is essential for bacteria to control gene expression, it improves efficiency of energy and resources, reduces production of potential toxic intermediate compounds, and in pathogenic bacteria it helps maintain virulence (Stoebel, Dean and Dykhuizen, 2008; Bervoets and Charlier, 2019). Gene regulation via transcription is controlled predominantly via RNA Polymerase (RNAP), which transcribes DNA to RNA. At its most basic, the central dogma explains the “*residue by residue transfer of sequential information*”, the transfer of information from DNA to RNA to protein (Figure 1.1) (Crick, 1970).



**Figure 1.1 Diagram of The Central Dogma**

Diagram of the Central Dogma, showing the “*residue by residue transfer of sequential information*”, (Crick, 1970).

Many external factors control gene regulation, such as: temperature; nutrient availability; metal ion availability; oxygen concentration; antibiotics; and many more. Transcriptional gene regulation is controlled by many factors, such as: RNA polymerase; sigma factors; up and down stream DNA; transcription factors; and DNA methylation, to name a few.



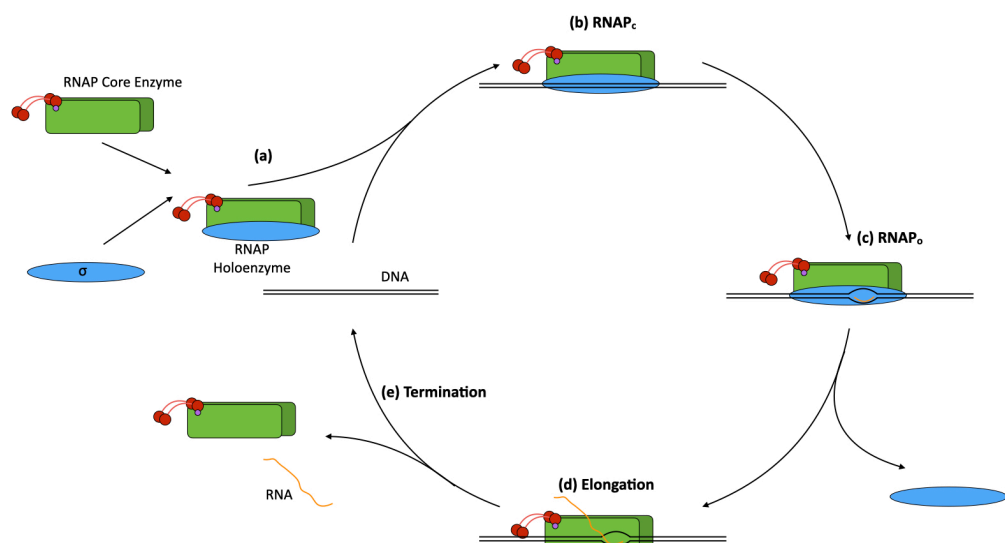
### 1.1.1. Transcription Cycle

The transcription cycle describes the process of transcription in bacteria. Simply, RNA Polymerase (RNAP) binds to a DNA promoter which leads to RNA transcription, upon transcription termination RNAP is released from the DNA and thus free to bind to a new promoter (Harden *et al.*, 2020).

The cycle is initiated with the formation of the RNAP holoenzyme, binding the RNAP core enzyme and a sigma factor (discussed further in 1.1.3) (Figure 1.2a). RNAP holoenzyme must first bind to promoter DNA for transcription to occur; there are a plethora of mechanisms to aid in RNAP-DNA binding, which will be discussed in more detail throughout this chapter. The mechanisms of transcription initiation are a key driver in gene regulation. This binding leads to the RNAP closed complex (RNAP<sub>c</sub>) (Figure 1.2b). A conformational change occurs in the RNAP, which causes DNA to wrap around the RNAP<sub>c</sub>. RNAP<sub>c</sub> unwinds DNA at the +1 site, leading to the RNAP open complex (RNAP<sub>o</sub>) (Figure 1.2c). Approximately 13 bp of DNA are unwound, referred to as 'the transcription bubble'. Residing in the RNAP transcription bubble is the active site, where RNAP catalyses the formation of phosphodiester bonds between ribonucleotides (DNA to RNA). A DNA complementary nucleoside triphosphate binds to the DNA in the active site of RNAP, at this point the sigma factor is released from the RNAP complex. Polymerisation occurs at the 3'-OH of the nucleoside triphosphate, in which the adjacent complementary nucleoside triphosphate binds, and thus elongation. To help with RNA stability during elongation, approximately 9 bp of the polymerised RNA chain binds to the corresponding DNA within the RNAP<sub>o</sub> transcription bubble. As well as this DNA-RNA interaction, a further 5 bp upstream the RNA binds to RNAP in the exit channel (Figure 1.2d). Transcription termination occurs to ensure that only the specific gene, or genes within an operon are transcribed, and not unwanted

genes downstream. There are three main mechanisms of termination: intrinsic termination; Rho dependent termination; and Mfd termination. At termination, both RNA and RNAP core enzyme are released; RNAP is now free to form a new holoenzyme and initiate transcription (Figure 1.2e). (Roberts, Shankar and Filter, 2008; Browning and Busby, 2016).

The transcription cycle is coupled with translation, the process in which the transcribed RNA is translated into polypeptides. Translation occurs in the bacterial 70s ribosome, where complementary amino acids (carried by tRNA) binds to RNA codon, leading to the formation of a polypeptide (Madigan *et al.*, 2012).



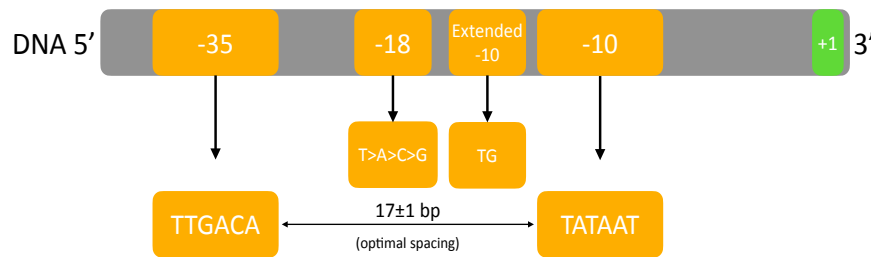
**Figure 1.2 Diagram of the Transcription Cycle**

Diagram of the basic steps of the transcription cycle. **(a)** RNAP core enzyme and  $\sigma^{70}$  bind together to form RNAP holoenzyme. **(b)** RNAP binds to promoter DNA. **(c)** initiation of elongation, with release of  $\sigma^{70}$ . **(d)** RNA polymerisation occurs. **(e)** Termination occurs, and RNAP is released from the promoter. RNA is free and is translated into amino acids by ribosomes.

### 1.1.2. Bacterial Promoters

A bacterial promoter is a region of DNA found upstream of the start site (+1), which RNAP binds. The promoter can be found upstream of a single gene, or a gene operon. Therefore, promoters are the key driving factor for gene regulation. Promoters are essential for RNAP recognition of DNA, and this is essential for initiation of transcription. Promoters allow bacteria to turn on and off, as well as up-regulate or down-regulate a specific gene, or a gene operon. This ability allows bacteria to rapidly adapt to new environmental conditions, as well as conserving energy by switching off gene regulation of undesired/non-essential genes (Browning and Busby, 2016).

For transcription to occur, RNAP needs to bind upstream of a gene. Often RNAP recognises and binds to three key motifs: the -35 the -10, and the extended -10 element (Figure 1.3). These three elements are recognised, more specifically, by the sigma factor ( $\sigma$ ) subunit of RNAP holoenzyme (discussed further in 1.1.4). The -35 and -10 elements are sometimes referred to as hexamers (six-mer, six nucleotide sequence). It is worth noting that all sigma factors, apart from  $\sigma^N$ , recognise the -35 and -10 regions of the promoter.  $\sigma^N$  recognise different promoter elements, the -24 and -12 regions. Only  $\sigma^{70}$  RNAP holoenzyme will be discussed in detail, as  $\sigma^{70}$  regulates most of the promoters to be studied as part of this thesis. RNAP binding to promoter DNA can be aided by transcription factors, proteins which can both increase and decrease RNAP affinity to specific DNA (Browning and Busby, 2004).



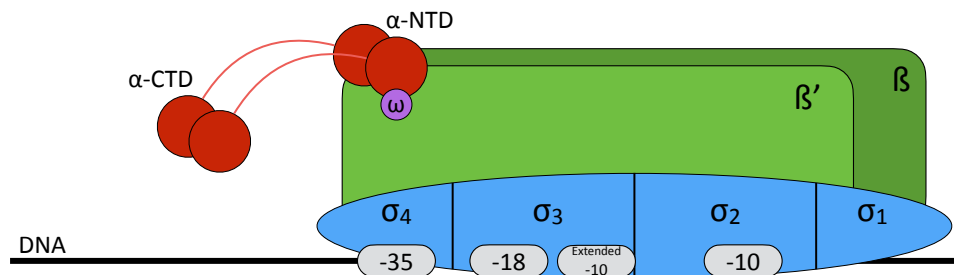
**Figure 1.3 Diagram of the Consensus Promoter Sequence**

Diagram of the key promoter elements (and the DNA consensus sequence) found upstream of the +1 start site; which  $\sigma^{70}$  of the RNAP holoenzyme complex binds to, leading to transcription initiation.

### 1.1.3. RNA Polymerase

RNA Polymerase (RNAP) is an essential enzyme found in all bacteria, which transcribes DNA to RNA, including mRNA, tRNA, and rRNA. RNAP is present in cells in two forms, the core enzyme, and the holoenzyme. RNAP core enzyme consists of five subunits, Alpha, Alpha, Beta, Beta prime, and Omega ( $\alpha_2\beta\beta'\omega$ ). Whereas the RNAP holoenzyme consists of an addition subunit, Sigma ( $\sigma$ ), with a total of six subunits  $\alpha_2\beta\beta'\omega\sigma$  (Figure 1.4). This additional subunit,  $\sigma$ , is essential in RNAP recognition of promoter DNA, and allowing correct and successful binding. There are many variations of  $\sigma$ , known as sigma factors, the majority of these in *Escherichia coli* (*E. coli*) recognise the -35 and -10 element (discussed in further detail later, 1.1.4) (Burgess, 1969; Vassylyev *et al.*, 2002; Browning and Busby, 2016).

The affinity of RNAP to promoter DNA is the major factor involved in transcriptional gene regulation, which heavily influences the rate of transcription. Many of the mechanisms which bacteria employ to increase or decrease gene expression involve means to alter the affinity of RNAP to the promoter DNA.



**Figure 1.4 Diagram of RNA Polymerase Holoenzyme**

Schematic diagram of  $\alpha_2\beta\beta'\omega\sigma$  subunits of RNA Polymerase Holoenzyme, showing Group 1 sigma factor  $\sigma^{70}$  and common upstream promoter regions. Adapted from (Lee and Busby, 2012).

#### 1.1.3.1. Alpha ( $\alpha$ ) Subunit of RNAP

RNAP consists of two identical alpha subunits ( $\alpha_2$ ) and each  $\alpha$  is comprised of 329 amino acids, the second smallest subunit after  $\omega$ . The  $\alpha$  subunits are essential in recognition of DNA-protein complex upstream of the -35 element, which includes both class-I and class-II transcription factor activation (discussed in further detail later, 1.1.5.1).  $\alpha$  is composed of two domains, the C-terminal domain ( $\alpha$ -CTD) and the N-terminal domain ( $\alpha$ -NTD), which are linked via a flexible protein linker domain. The  $\alpha$ -NTD forms a strong dimerization bond within the RNAP complex. The  $\alpha$ -CTD also forms a dimerization bond, weaker than the  $\alpha$ -NTD. The  $\alpha$ -CTD interacts directly with DNA-Protein complexes upstream of the -35 element, essentially helping to guide RNAP to gene promoters. Further,  $\alpha$  plays a vital role in RNAP assembly (Ebright and Busby, 1995; Browning and Busby, 2004)

### 1.1.3.2. Omega ( $\omega$ ) Subunit of RNAP

The subunit  $\omega$  was discovered in 1969, however its role in transcription and RNAP was ambiguous. For many years,  $\omega$  was considered a unnecessary subunit, ironically by its discoverer Burgess, (1969). It has now been shown that  $\omega$  plays a vital role in  $\beta'$  folding, RNAP assembly, RNAP stability, and prevents aggregation of  $\beta'$  subunits (Gentry and Burgess, 1993; Zhang *et al.*, 1999; Ghosh, Ishihama and Chatterji, 2001; Ghosh, Ramakrishnan and Chatterji, 2003).

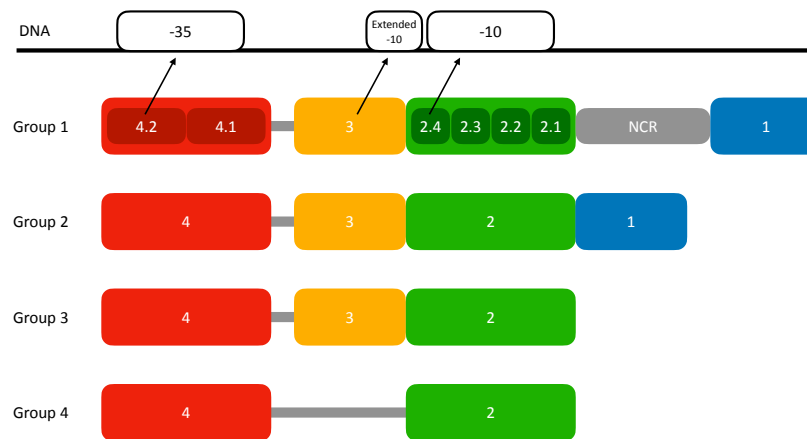
### 1.1.3.3. RNA Polymerase Assembly

$\omega$  binds to  $\beta'$ , both to the CTD and NTD terminal ( $\omega\beta'$ ). Interestingly  $\omega$  only binds to both  $\beta'$  CTD and NTD in tandem, and not individually. Simultaneously,  $\alpha$  subunits dimerise ( $\alpha_2$ ), followed by the addition of  $\beta$  ( $\alpha_2\beta$ ). These two complexes then come together to form the RNAP core enzyme ( $\alpha_2\beta + \omega\beta' \rightarrow \alpha_2\beta\beta'\omega$ ). The addition of a sigma factor ( $\sigma$ ) with the RNAP core enzyme leads to the RNAP holoenzyme complex formation ( $\alpha_2\beta\beta'\omega\sigma$ ) (Ito, Iwakura and Tshihama, 1975; Ghosh, Ishihama and Chatterji, 2001; Murakami and Darst, 2003; Mathew and Chatterji, 2006).

### 1.1.4. RNA Polymerase Holoenzyme ( $\sigma^{70}$ )

Sigma factors are a subunit of the RNAP holoenzyme, which are responsible for recognition and binding to DNA promoter sequences. This recognition of promoter DNA is essential for RNAP to initiate transcription of DNA to RNA. *E. coli* has seven different sigma factors;  $\sigma^{70}$  is the most common and regulates the largest proportion of genes. There are four groups of sigma factors, 1 to 4 (Figure 1.5). Group 1 sigma factors, such as  $\sigma^{70}$ , are mainly involved in the regulation of housekeeping genes, such as growth and structural genes. Groups 2-4 sigma factors generally regulate stress

response associated genes; stationary phase associated genes; and motility associated genes (Maeda, H; Fujita, N; Ishihama, 2005; Bervoets and Charlier, 2019).



**Figure 1.5 Diagram of Group 1-4 Sigma Factors**

Diagram of groups 1 to 4 sigma factors, showing sub-regions. NCR non-conserved region. Adapted from (Paget, 2015; Bervoets and Charlier, 2019)

$\sigma^{70}$  is the principle sigma factor in *E. coli*, often referred to as the housekeeping sigma factor, as it controls transcription of most genes required for bacterial growth.  $\sigma^{70}$  is a member of the group 1 sigma factors.  $\sigma^{70}$  can be further divided into four subregions,  $\sigma_{1-4}$ , which are linked via a flexible linker; these subregions recognise specific promoter elements (Figure 1.5).

The -35 element, recognised by  $\sigma_{4.2}$  has a consensus sequence of 5'-TTGACA-3', and the -10 element, recognised by  $\sigma_{2.4}$  has a consensus sequence of 5'-TATAAT-3' (Figure 1.3). These two elements are generally well conserved throughout the *E. coli* chromosome, with the -10 element seemingly more conserved than the -35 element (Barne *et al.*, 1997).

Another element known as the -10 extended element, also known as the TG motif, which is recognised by  $\sigma_3$ , and as the name suggests, has a consensus sequence of 5'-TG-3'. This element is located upstream of the -10 element, at -14 to -15, and can

also have a second TG motif at -16 to -17. The -10 extended element is not essential for RNAP binding and is only found in ~10% of *E. coli* promoters (Burr *et al.*, 2000).

A further promoter element was established recently. It has been shown that the side chain of  $\sigma^{70}$  recognises another DNA element between the -35 and -10 element, at -18. The -18 element, sometimes referred to as the wobble base, is not essential for RNAP recognition, but shows a higher affinity to DNA in the following order: T>A>C>G (Singh *et al.*, 2011).

### 1.1.5. Transcription Factors

Transcription factors (TFs) are proteins which bind to DNA, usually upstream of the +1 start site of a promoter, which can alter the affinity or binding of RNAP to DNA, and thus can up or down regulate gene transcription. There are two main types of transcription factors, activators and repressors, which up-regulate and down-regulate expression, respectively (Browning and Busby, 2004). The regulation of *E. coli* gene expression by transcription factors is one of the most well understood transcriptional regulation systems in the bacterial kingdom. *E. coli* K-12 has been shown to have 314 different transcription factors, of which 35% are activators, 43% repressors, and 22% dual regulators (Perez-Rueda, 2000). Of these 314 TFs, there are seven classed as global transcription factors: CRP, FNR, IHF, FIS, ArcA, NarL, and Lrp, which account for the regulation of ~50% of all *E. coli* genes. CRP (cAMP receptor protein), also known as CAP (catabolite activator protein), has been shown to regulate over 250 *E. coli* genes, the most of all *E. coli*'s TFs. CRP works in tandem with sigma factors  $\sigma^{70}$ ,  $\sigma^{38}$ ,  $\sigma^{32}$ , and  $\sigma^{24}$  (Martínez-Antonio and Collado-Vides, 2003).

It is essential that TFs can be activated and deactivated, which will affect the promoter activity of a gene. There are three main mechanisms which control transcriptional



activation by TFs. The first is ligand activation, where a ligand molecule binds reversibly to a TF. Ligand activation generally leads to a conformational change in the TFs, which leads to either increased binding or disassociation of the TF to the promoter DNA (Browning and Busby, 2004).

Secondly, the transcription factors itself can be activated via a secondary system called a two-component regulatory system (TCS). A TCS, as its name suggests is composed of two components: a sensor histidine kinase, and a response regulator (often a transcription factor). The histidine kinase is embedded in the inner membrane, and its role is to sense external stimuli. In the presence of a specific stimulus, the histidine kinase will auto-phosphorylate specifically at a histidine residue. This in turn phosphorylates the response regulator, leading to a conformation change altering the DNA binding affinity, and thus altering gene regulation (Stock, Robinson and Goudreau, 2000).

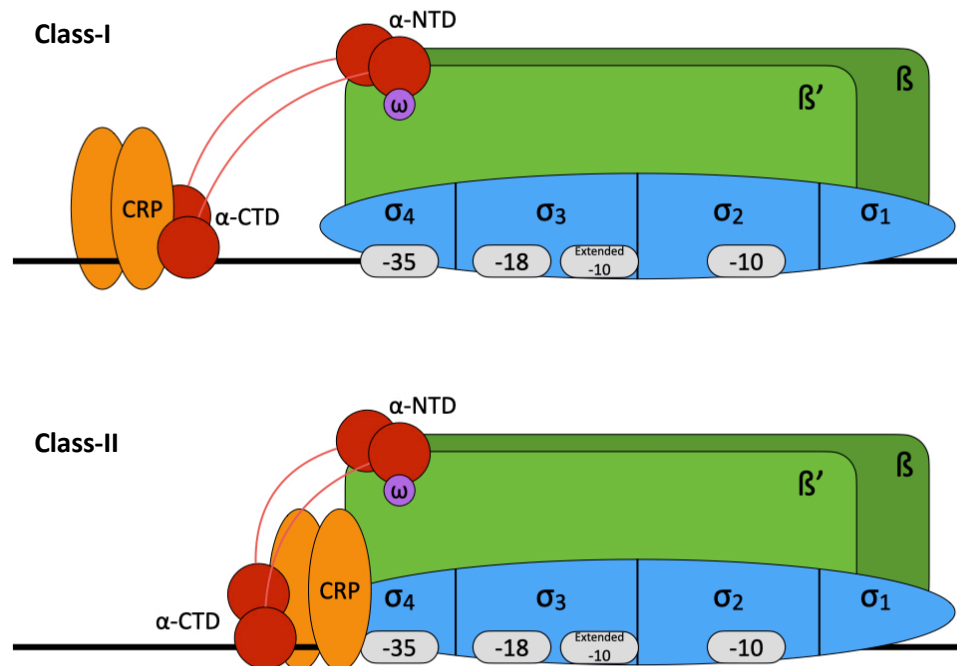
Finally, the internal concentration of a transcription factor can be altered through the regulation of said transcription factor, or through enzymatic degradation of the transcription factor (Dempsey, 1996; Browning and Busby, 2004).

#### 1.1.5.1. Activators

Activator transcription factors upregulate gene expression. Generally, they do this by increasing RNAP affinity to the DNA promoter sequence, which in turn initiates transcriptional gene expression. There are three main mechanisms by which activator TFs can cause an increase in gene expression. This is done by either increasing RNAP affinity to a promoter by causing TF-RNAP interactions (Class-I and Class-II activation) (Figure 1.6), or TFs cause a physical change in the DNA, leading to a more favourable interaction between RNAP and the promoter DNA.

Class-I activation occurs when a TF binds upstream of the -35 hexamer on a promoter and interacts directly with the  $\alpha$ -CTD of RNAP. CRP is the most abundant TF of *E. coli*, and the most well understood. It has been shown that CRP has an optimal binding position (-71.5 to -82.5), with a 22 bp (11-11) palindromic binding region with a 5'-TGTGA-3' motif (5'-AAATGTGATCTAGATCACAATTT-3') (Busby and Ebright, 1999). This activation aids the recruitment of RNAP to DNA. Often Class-I activators are located on the same helix face as the potential binding site for the RNAP-DNA complex; this allows easier interactions between the  $\alpha$ -CTD and the TF, and thus aids in RNAP recruitment (Kolb, 1993; Zheng *et al.*, 2004; Mitchell *et al.*, 2007; Zhou *et al.*, 2014).

Class-II activation is like Class-I. However, Class-II activation occurs when TFs bind to DNA overlapping the -35 element. This leads to more interaction with the TF and RNAP. Three different points of interaction occur between the TF and RNAP, specifically with the  $\alpha$ -CTD,  $\alpha$ -NTD and  $\sigma_4$ . The  $\alpha$ -CTD preferentially binds to the upstream (5') end of the TFs, allowing further interaction between the TF and the RNAP. This can additionally allow dual activation with both Class-I and Class-II activation (Niu *et al.*, 1996; Busby and Ebright, 1999).

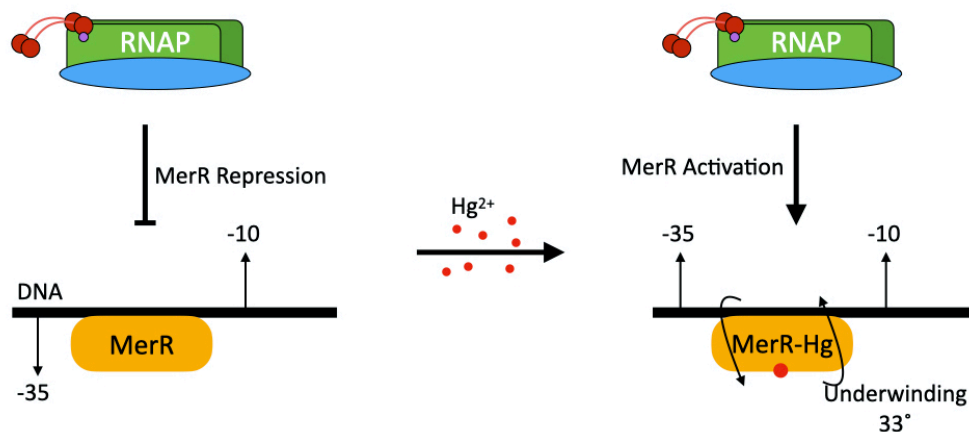


**Figure 1.6 Diagram of Class-I and Class-II Transcription Factor Activation**

Diagram of Class-I and Class-II transcription factor activation of RNA polymerase on a DNA promoter. Class-I activation; TF bind to DNA upstream of the -35 element and interacts with the  $\alpha$ -CTD. Class-II activation; TF bind near the -35 element, leading to multiple interaction with the  $\alpha$ -CTD,  $\alpha$ -NTD and the sigma factor ( $\sigma_4$ ). Adapted from (Browning and Busby, 2004).

The third mechanism of activation causes a conformational change in the DNA, which in turn increases the RNAP affinity to the DNA. A well-studied example of this form of activation is MerR, as well the ZntR which forms a large study area of this thesis (Brocklehurst *et al.*, 1999; Brown *et al.*, 2003). It has been shown that the optimal spacing between the -35 and -10 hexamer in well expressed promoters is  $17 \pm 1$  bp (Hawley and McClure, 1983). Often TFs which cause DNA distortion bind between the -35 and -10 element of promoters with non-optimal spacing. This non optimal spacer hinders correct RNAP-DNA binding, such that the  $\sigma_{4.2}$  and  $\sigma_{2.4}$  cannot successfully interact with the -35 and -10, respectively. TFs that cause DNA distortion effectively alter the non-optimal spacing, such that under the correct conditions the -35 and -10 elements in the promoter are realigned. Using the example of MerR

(which can act both as an activator and repressor depending on the presence of  $\text{Hg}^{2+}$ ), the MerR dimer constitutively binds to the DNA in the non-optimal spacer (19 bp) between the -35 and -10 element. In the absence of  $\text{Hg}^{2+}$ , the MerR dimer simply binds to the DNA, causing mild repression of gene expression. The presence of  $\text{Hg}^{2+}$  causes a conformation change in the structure of MerR, which underwinds the DNA by  $\sim 33^\circ$ . This underwinding brings the -35 and -10 element closer, and aligns them on the same helical face, thus allowing correct  $\sigma_{2.4}$  and  $\sigma_{4.2}$  interaction (Figure 1.7) (Lund and Brown, 1989; Ansari, Cheal and O'Halloran, 1992; Brown *et al.*, 2003).



**Figure 1.7 Diagram of MerR Underwinding of DNA**

Diagram of MerR activation via DNA underwinding. Adapted from. MerR binds constitutively to the sub optimal 19 bp spacer. RNAP cannot bind to the promoter due to sub-optimal spacing and mild steric hindrance caused by MerR.  $\text{Hg}^{2+}$  binds reversibly to MerR, causing a conformation change, which underwinds the DNA by  $\sim 33^\circ$ , bringing the -10 and -35 elements in perfect alignment for interaction the sigma factor of the RNAP. Adapted from (Ansari, Cheal and O'Halloran, 1992).

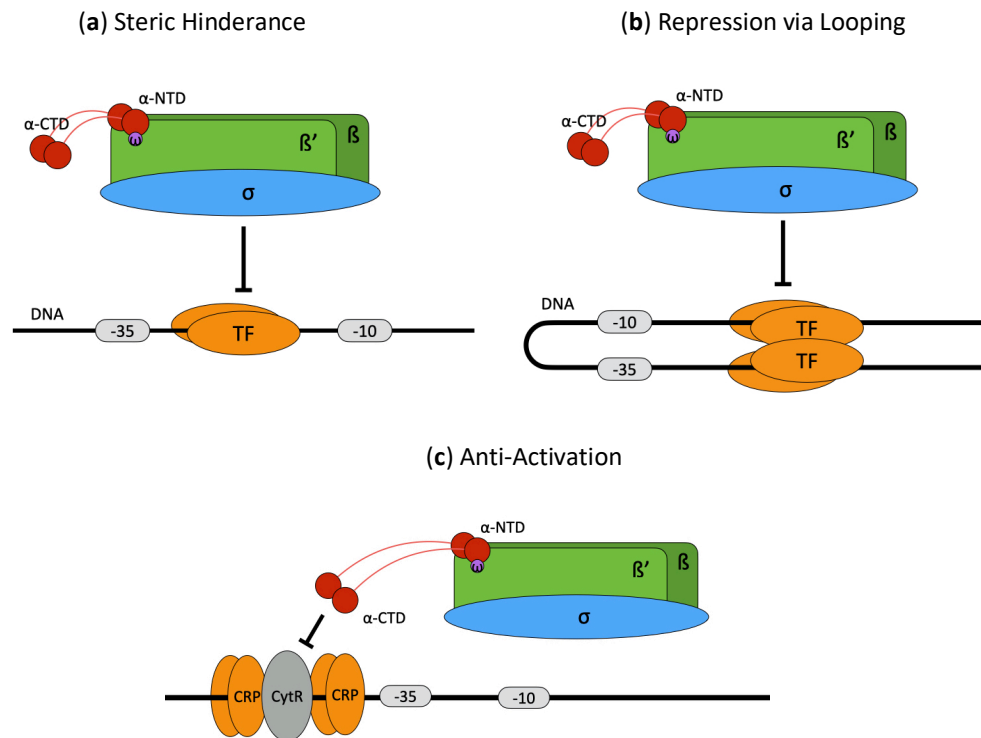
### 1.1.5.2. Repressors

As well as activating gene expression, TFs can also down regulate gene expression; these are known as repressors. There are three main mechanisms by which TFs can cause repression (Figure 1.8). The first is the simplest, steric hindrance (Figure 1.8a). A transcription factor binds to the promoter DNA, often between the -35 and -10 hexamer, which physically blocks RNAP interacting with the DNA. Zur, a transcription

factor used in this study, represses gene regulation via steric hinderance in the presence of  $Zn^{2+}$  (Patzer and Hantke, 1998; Browning and Busby, 2016).

Repression looping is another example of transcription repression, where TFs bind to multiple locations in DNA upstream and downstream of the +1 start site (Figure 1.8b). These TFs form a complex by binding to each other, causing looping of the DNA. This loop then impedes correct RNAP binding to the DNA, by physically obstructing the -35 and -10 interactions with the sigma factor (Cournac and Plumbridge, 2013).

The final mechanism of repression is anti-activation (Figure 1.8c). Anti-activation occurs when an anti-activator TF binds to DNA in close proximity to an activator TF and blocks the effect of the activator. A well-studied example of this is CytR-CRP anti-activation. CytR binds between two CRP molecules, which are approximately 50 bp apart, and upstream of the -35 element. The CytR protein causes two interactions, a protein-protein interaction (CytR-CRP), and a DNA-protein interaction (CytR-DNA). CytR occupies the site which  $\alpha$ -CTD usually binds, thus sterically hindering  $\alpha$ -CTD-CRP binding (Valentin-Hansen, S gaard-Andersen and Pedersen, 1996; Busby and Ebright, 1999).



**Figure 1.8 Diagram of the Mechanisms of TF Repression**

Diagram of three main type of repression via transcription factors. **(a)** Repression via steric hinderance; the TF physically blocks TF-DNA interactions. **(b)** Repression via looping of DNA; which alters the optimal interactions between the promoter DNA and the RNAP. **(c)** Repression via anti-activation; where a TF binds near an activator TF and physically blocks the interaction with RNAP. Adapted from Browning and Busby, (2004)

### 1.1.6. Alternative RNAP Holoenzyme Sigma Factors

*E. coli* has seven different sigma factors:  $\sigma^{70}$  ( $\sigma^A$ );  $\sigma^{54}$  ( $\sigma^N$ );  $\sigma^{38}$  ( $\sigma^S$ );  $\sigma^{32}$  ( $\sigma^H$ );  $\sigma^{28}$  ( $\sigma^F$ );  $\sigma^{24}$  ( $\sigma^E$ ); and  $\sigma^{19}$  ( $\sigma^{FecI}$ ). The superscript-number associated with the sigma factors (e.g. 70 of  $\sigma^{70}$ ) refers to the mass (kDa) of the sigma factor.  $\sigma^{70}$  generally regulates housekeeping genes, whereas the other six sigma factors tend to be involved with regulating transcription of a more specific subset of genes.

$\sigma^{54}$  ( $\sigma^N$ ) is associated with the regulation of genes expressed during nitrogen stress and alternative carbon sources. Interestingly, this is the only *E. coli* sigma factor which does not recognise the -35 and -10 element, but rather the -24 and -12, and is not a member of groups 1-4 sigma factors (Merrick, 1993; Wigneshweraraj *et al.*, 2008; Bervoets and

Charlier, 2019).  $\sigma^{38}$  ( $\sigma^S$ ), a group 2 sigma factor, regulates genes associated with stress response, and genes expressed during stationary phase (Landini *et al.*, 2014).  $\sigma^{32}$  ( $\sigma^H$ ), a group 3 sigma factor, is associated with genes expressed during heat shock (Morita *et al.*, 1999).  $\sigma^{28}$  ( $\sigma^F$ ), a group 3 sigma factor, regulates genes associated with motility (flagella) and chemotaxis receptors (Helmann, 1991).  $\sigma^{24}$  ( $\sigma^E$ ), a group 4 sigma factor, regulates genes associated with intracytoplasmic stress, and heat shock (Erickson and Gross, 1989; De Las Penas, Connolly and Gross, 1997).  $\sigma^{19}$  ( $\sigma^{FecI}$ ), a group 4 sigma factor, is associated with intracytoplasmic function.  $\sigma^{19}$  ( $\sigma^{FecI}$ ) is only known to regulate the *fecABCDE* operon (ferric acid transport system). Interestingly, the regulation of the sigma factor *fecI* ( $\sigma^{19}$ ) itself is regulated by Fur, a  $\text{Fe}^{2+}$  repressed transaction factor, which is similar to Zur (Tanaka *et al.*, 1995; Angerer and Braun, 1998).

## 1.2. Plasmid Biology

Bacteria can carry additional nucleic acid outside of the chromosome, known as plasmids. These plasmids replicate independently of the chromosome, and generally carry additional genes that are not essential for basic cellular function. Plasmids are most commonly found as circular dsDNA, however they have also been found as linear dsDNA in *Streptomyces spp*, and as circular ssDNA in *Myxococcus spp* (Casali and Preston, 2003). Plasmids play an important role in evolution; they allow horizontal gene transfer between bacteria. Though most genes found on plasmids are not essential, they often aid survival in specific environments, this could include: resistant genes; virulence genes; and alternative synthetic pathways (Brinkmann *et al.*, 2018; Rodríguez-Beltrán *et al.*, 2021).

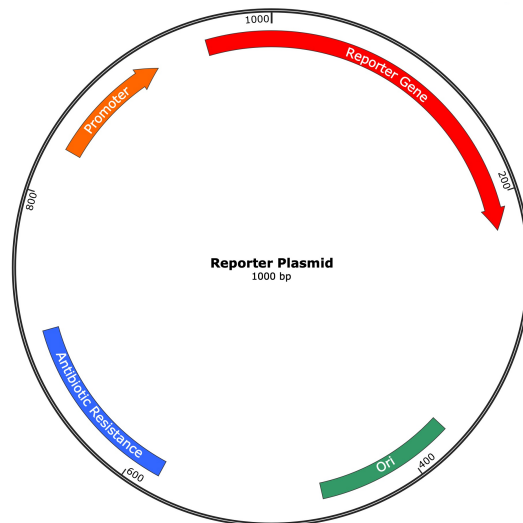
Plasmids have been adapted, by humans, to be used both in biotechnology and for research. Reporter plasmids are commonly used in research, as seen extensively throughout this thesis. Plasmids can also be designed to used introduce foreign DNA into a host, as well as used to delete chromosomal DNA.

### 1.2.1. Reporter Plasmids

Reporter plasmids are an essential tool used to understand gene regulation. A reporter plasmid consists of four main elements: a reporter gene/element; a promoter sequence; an origin of replication; and an antibiotic resistance gene (Figure 1.9). A reporter gene is located downstream of a gene promoter, which has been cloned into the plasmid. The regulation of the gene promoter will directly impact the expression of the reporter element, which can then be quantified. An antibiotic resistance gene is used to select for bacterial cells containing the reporter plasmid, and an origin of



replication is needed for plasmid replication; the *ori* must be compatible with the bacterial strain used and any other plasmids carried by the host.



**Figure 1.9 Plasmid Map of a Generic Reporter Plasmid**

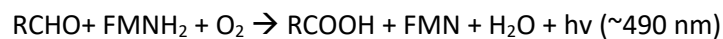
Plasmid map of a generic reporter plasmid showing essential elements.

There are three main reporter elements used in reporter plasmids: fluorescence; bioluminescence (i.e. Lux); and  $\beta$ -galactosidase. All the elements have their strengths and weaknesses. Generally, fluorescent proteins have strong stability and a long half-life, whereas bioluminescence elements have a shorter half-life and faster maturation. Both fluorescence and luminescence assays are commonly used due to their ease of use.  $\beta$ -galactosidase assays are a robust system, with a long half-life, however, it is currently used less as a reporter due to the time consuming and laborious protocol required in a  $\beta$ -galactosidase assay. Reporter plasmids in this work use *mRFP1* and *luxCDABE*; both for their ease of use, and because they provide accumulative and temporal assay data, respectively.

#### 1.2.1.1. Bioluminescence

The *luxCDABE* operon is a commonly used reporter element, which contains all genes required for bioluminescence. *luxAB* encodes the  $\alpha$  and  $\beta$  subunits of the enzyme

luciferase; *luxC* encodes fatty acid reductase; *luxD* encodes an acyl transferase; and *luxE* encodes an acyl protein synthetase (Iqbal *et al.*, 2017). LuxCDABE has a short half-life and a quick maturation, making it ideal for temporal gene expression analysis. LuxCDE together form the fatty acid reductase complex, which synthesises the long chain aldehyde (RCHO). LuxAB (luciferase) catalyses the reduced flavin mononucleotide (FMNH<sub>2</sub>) and RCHO to an oxidized flavin mononucleotide (FMN) and RCOOH. This reaction produces light in the process (~490 nm) (Meighen, 1991; Brodl, Winkler and Macheroux, 2018).



#### 1.2.1.2. Fluorescence

The discovery of the green fluorescent protein (GFP), isolated from the jelly fish *Aequorea victoria*, was first published in 1962 (Johnson *et al.*, 1962). Fluorescent proteins have many uses in biology, which include their use in promoter activity assays. Fluorescent proteins have the unique ability to absorb light of a shorter wavelength and emit light at a longer wavelength; this unique property of fluorescence is known as Stokes shift. Many alterations from the original GFP have been created and discovered, including mRFP1 (monomeric red fluorescent protein) which is used in this study. mRFP1 matures 10x faster than its predecessor (DsRed) and shows similar brightness in cells. The excitation and emission wavelength of mRFP1 has improved cells and tissue penetration; further, the excitation wavelength is longer than the excitation wavelength of GFP which often causes auto fluorescence of both cells and media. This leads to lower background fluorescence compared to GFP. mRFP1 has a maximum excitation of 584 nm and a maximum emission at 607 nm (Campbell *et al.*, 2002).

### 1.3. Biotechnology

Merriam-Webster defines Biotechnology as;

“The manipulation (as through genetic engineering) of living organisms or their components to produce useful usually commercial products (such as pest resistant crops, new bacterial strains, or novel pharmaceuticals)” (Merriam-Webster, 2020).

This technology has a major impact on many areas, including health care, biofuels, pharmaceuticals, chemical production, and the food industry. Biotechnology allows useful recombinant genes to be expressed in easy to manipulate biotechnology relevant organisms, such as *E. coli*, *Bacillus subtilis*, and *Saccharomyces cerevisiae*. This advancement in recombinant DNA technology has allowed the expression of genes from higher eukaryotes to be expressed in prokaryotes and lower eukaryotes (Walker and Rapley, 2000). An example of biotechnology is the production of Human insulin through recombinant DNA expression in *E. coli* and *S. cerevisiae* (Baeshen *et al.*, 2014). Recombinant DNA expression can be achieved via two mechanisms: transformation of a plasmid expressing a gene of interest; or chromosomal insertion of a gene into a host organism.

The use of plasmids in biotechnology allows an easy and simple mechanism to insert recombinant DNA into a host strain. It is essential that these expression vectors contain the following: an *ori* site; a selectable marker gene (frequently an antibiotic resistance gene); a controllable promoter; a site to insert cloned sequences (frequently a multiple cloning site); and terminator sequence. It can also be essential to have affinity tag sequences, such as poly-His tags for downstream purification (Sørensen and Mortensen, 2005; Rosano and Ceccarelli, 2014). There are a variety of plasmid vectors which can be used for recombinant DNA expression, including pBAD, pET, and pUC

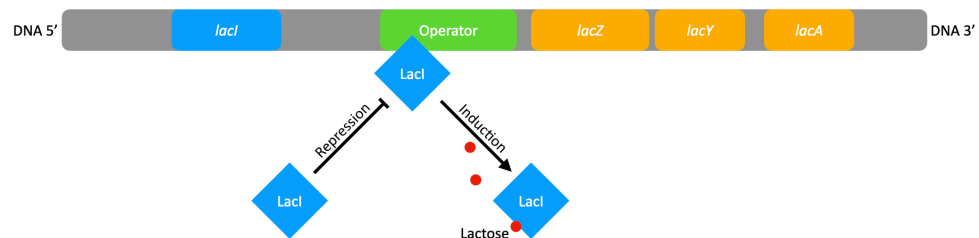
series plasmids. The plasmid vector copy number is essential for consideration, as too high a level of expression of a recombinant gene may lead to toxic side effects on the host. pBAD vectors (p15A replicon) maintain 10-12 copies of the plasmid in *E. coli* K12 per cell, pET (pBR322 replicon) maintains 16-60 copies, and pUC vectors (pMB1 derivative replication) maintain the highest at 500-700 copies per cell. Occasionally, low copy number plasmids with *ori* pSC101 are used, at <5 per cell (Rosano et al., 2014). The expression of recombinant Human mini-proinsulin is expressed, on mass in batch cultures, in *E. coli* using a pET expression vector (Shin *et al.*, 1997). As well as considering which expression plasmid to use, one must also consider the promoter strength. Inducible promoters are an ingenious mechanism to control recombinant gene expression.

### 1.3.1. Inducible Promoter Systems

For both biotechnology and laboratory research, it is vital to be able to control recombinant gene expression. Inducible promoter systems are used to control gene expression, turning selected genes on and off when desired. Most of these inducible promoter systems are based on gene expression systems found in bacteria, with some exceptions, such as T7 promoter system based on T7 bacteriophage RNA polymerase which recognises the specific T7 phage promoter sequence (Madigan *et al.*, 2012). It is essential that inducible promoters produce high levels of expression and ideally a low level of uninduced basal expression. A major disadvantage with most inducible promoter systems is that they are not controllable and offer an 'on/off' level of expression. However, the 'off' inducible promoters rarely offer zero expression, as most promoters are leaky or show low basal expression (Hannig and Makrides, 1998).

### 1.3.1.1. The Lac Operon

The *lac* operon, found in *E. coli*, is a classic example of controlled gene expression, which has now been manipulated to be used as an inducible promoter system. Simply, a TF (LacI) binds to an operator region downstream of the promoter, blocking RNAP transcription, in the absence of lactose. In the presence of lactose, LacI disassociates from the operator, allowing transcription of the *lac* operon (*lacZYA*) (Figure 1.10) (Ullmann, 2009). The *lac* operon has been manipulated so that it can be used as an inducible promoter system and is still frequently used in plasmids and biotechnology. The *lac* promoter ( $P_{lac}$ ) including the operator region, is cloned upstream of an inducible gene of interest. The allosteric inducer, IPTG, is used rather than lactose as an inducer for  $P_{lac}$ , due to higher cell permeability, and IPTG is not cleaved by  $\beta$ -galactosidase. This allows IPTG to continuously act as an inducer, unlike lactose which is metabolized (Rosano and Ceccarelli, 2014).



**Figure 1.10 Diagram of The Lac Operon**

Diagram of the Lac operon *LacZYA* of *E. coli*. In the absence of lactose, the transcription factor LacI will bind to the Lac operator, blocking transcription. In the presence of lactose, lactose binds to LacI, leading to LacI dissociating from the promoter and thus allowing transcription.

### 1.3.1.2. The *lacUV5* Promoter ( $P_{lacUV5}$ )

To further improve on  $P_{lac}$  for an inducible expression system, a mutated promoter named *lacUV5* was created. *lacUV5* contains a 2 bp mutation in the -10 element, which gives a higher similarity to the -10 consensus sequence. This has led to improved

induction levels of expression of cloned genes, when induced with either lactose or IPTG (Ullmann, 2009).

#### 1.3.1.3. The Hybrid trp-lac Promoter ( $P_{trc}$ )

The promoter  $P_{trc}$  (or  $P_{tac}$ ) is a hybrid promoter, combining the -35 element of the  $P_{trp}$ , and the -10 element of the  $P_{lacUV5}$ .  $P_{trc}$  also contains the LacO (lac operator region) meaning that the promoter is IPTG or lactose inducible.  $P_{trc}$  shows further improvement and increased induction expression compared to  $P_{lacUV5}$  or  $P_{trp}$  solely (de Boer, Comstock and Vasser, 1983). The  $P_{trc}$  promoter is used as a control, a comparison of a standard inducible system, throughout the work in this thesis.

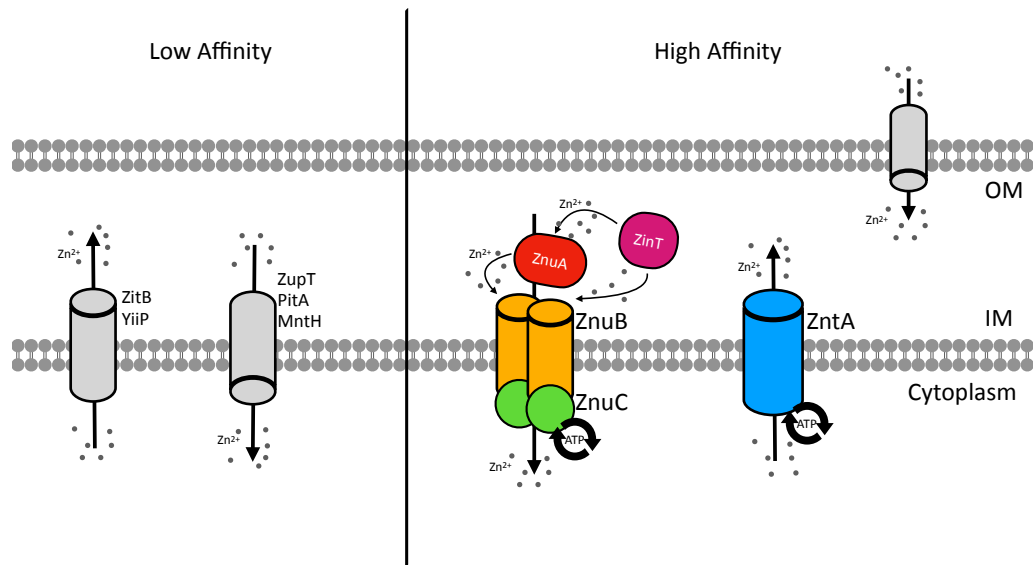
#### 1.4. Zinc Homeostasis in *E. coli*

Zinc ( $\text{Zn}^{2+}$ ) is a post-transition metal, often referred to simply as a transition metal, found in group 12 of the Periodic table, with an atomic number of 30, and an atomic mass of 65.38. Zinc is an essential metal for *E. coli*, and for all known life. Zinc plays an important role in enzyme activity and is a co-factor in all six classes of enzyme: oxidoreductases; transferases; hydrolases; lyases; isomerases; and ligases. However, zinc can become toxic to bacteria if internal cellular concentrations become too high. Toxic levels of zinc lead to oxidative stress in the cells. Further, zinc has been shown to compete with and replace other transition metals interactions in essential protein and enzyme, which leads to inactivation (Webb, 1992; Coleman, 1998; David, 2011; Lemire, Harrison and Turner, 2013).

Due to this, it is essential for bacteria to tightly control their internal concentration of zinc, to a level which does not cause toxicity, but also at a concentration high enough to allow for optimum cellular function. *E. coli* possess a suite of genes which tightly controls the internal zinc concentration (zinc homeostasis). *E. coli* tightly control their internal free zinc concentration to the femtomolar level, which equates to one or two free  $\text{Zn}^{2+}$  ions per cell. It has been shown that the total internal zinc in log phase *E. coli* grown in minimal media is in the range of 10-40  $\mu\text{M}$ , or  $2 \times 10^5 \text{ Zn}^{2+}$  per cell, most of which is found in cellular components (Outten and O'Halloran, 2001; Wang *et al.*, 2011).

This tight regulation of internal zinc concentrations is controlled by two major transcription factors, Zur and ZntR (Patzner and Hantke, 1998; Brocklehurst *et al.*, 1999). Zur upregulated the zinc import system *znuABC*, as well as other genes zinc acquisition genes when zinc is depleted, whereas ZntR upregulates the zinc export gene, *zntA*, when internal zinc levels are too high. ZnuABC and ZntA work in tandem as the major

zinc import and export system for *E. coli*. However, there are several low affinity zinc transport systems, some regulated and some constitutively expressed, as well as other mechanisms to scavenge zinc. Together these form the tightly controlled zinc homeostasis system (Figure 1.11).



**Figure 1.11 Diagram of Zinc Homeostasis**

Diagram of the *E. coli* zinc homeostasis system. ZnuABC, Zur regulated. ZntA, ZntR regulated. OM, outer membrane. IM, inner membrane.

#### 1.4.1. Low Affinity Zinc Import and Export systems

ZntA and ZnuABC are the major zinc import and export systems of *E. coli*. However, *E. coli* has some other systems, which are not the dominant transport system, but still play a role in *E. coli* zinc homeostasis. These are often constitutively expressed, and not regulated by either Zur or ZntR.

##### 1.4.1.1. Low Affinity Import Systems

ZupT, previously known as YqiE, is a member of the ZIP (ZRT, IRT-like protein) transport family, a cytoplasmic membrane protein involved in the import of divalent metal ions. ZupT imports a broad range of divalent metal ions, including;  $Zn^{2+}$ ,  $Fe^{2+}$ , and  $Co^{2+}$ , but with a higher affinity to  $Zn^{2+}$  (Grass *et al.*, 2002; Grass, Franke, *et al.*, 2005).



PitA, a member of the Pit system, is a low affinity, high velocity inorganic phosphate ( $P_i$ ) transport system. Divalent cations are essential for the transport of  $P_i$ , allowing a soluble neutral metal phosphate complex for import. It has been shown that PitA also constitutently imports  $Zn^{2+}$  into the cytoplasm (Beard *et al.*, 2000; Harris *et al.*, 2001; Jackson *et al.*, 2008).

MntH, previously YfeP, is a member of the NRAMP (natural resistance associated macrophage protein) metal ion transporters. MntH predominantly imports  $Mn^{2+}$ , but has been shown to import other divalent metal cations with low specificity (Makui *et al.*, 2000; Haemig and Brooker, 2004).

### 1.4.1.2. Low Affinity Export Systems

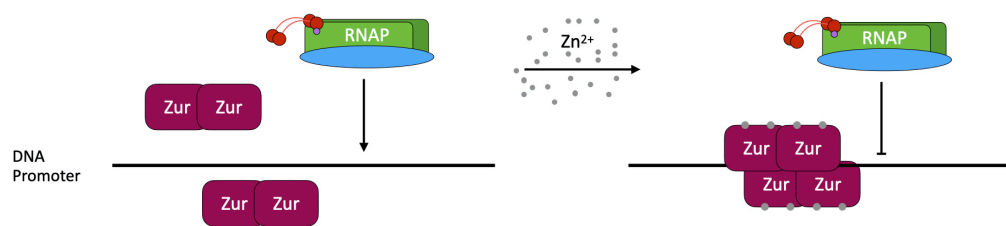
ZitB, previously YbgR, is a member of the CDF (cation diffusion family) family. ZitB has been shown to export  $Zn^{2+}$ . As well as  $Zn^{2+}$ , ZitB also demonstrates low affinity efflux of  $Cd^{2+}$ ,  $Cu^{2+}$  and  $Ni^{2+}$  (Grass *et al.*, 2001; Rahman *et al.*, 2008).

YiiP, also known as FieF, is also a member of the CDF family. YiiP has been shown to facilitate translocation of  $Zn^{2+}$ , and  $Cd^{2+}$  out the cell (Grass, Otto, *et al.*, 2005; Wei and Fu, 2005).

### 1.5. Zur (Zinc Uptake Regulator)

Zur (Zinc Uptake Regulator), a zinc binding metalloregulator protein, is a transcription factor that regulates zinc uptake systems. Zur is found in many different organisms, including: *E. coli*; *Mycobacterium smegmatis*; *Enterococcus faecalis*; *Streptomyces coelicolor*; *Bacillus subtilis*; and *Listeria monocytogenes*. Zur has been shown to regulate a variety of genes involved in zinc management; including zinc import systems, and alternative zinc-free proteins (Gaballa and Helmann, 1998; Dalet *et al.*, 1999; Shin *et al.*, 2007; Latorre *et al.*, 2015; Goethe *et al.*, 2020). Zur is a member of the FUR (ferric uptake regulator) superfamily of TFs, which are involved in metal ion homeostasis. FUR family members include: Fur ( $\text{Fe}^{2+}$ ); Zur ( $\text{Zn}^{2+}$ ); Nur ( $\text{Ni}^{2+}$ ); Mur ( $\text{Mn}^{2+}$ ); PerR (oxidative stress); and Irr (heme-detection). Generally, all FUR members use metal as the ligand for ligand activation, which stabilises the TF interaction with DNA (Fillat, 2014).

Zur works as an active repressor, in the presence of  $\text{Zn}^{2+}$  Zur will bind to a gene promoter, causing steric hinderance of RNAP binding, and thus reduces gene expression (Figure 1.12). In the absence of  $\text{Zn}^{2+}$ , Zur is naturally found as a dimer  $\text{Zur}_2$ . When in the presence of  $\text{Zn}^{2+}$ , each Zur monomer binds with two  $\text{Zn}^{2+}$  ions, which leads to the formation of a dimer of dimers  $(\text{Zur}_2)_2$  which binds to DNA, forming the  $(\text{Zur}_2)_2$ -DNA complex. This sterically hinders RNAP binding to DNA (Patzner and Hantke, 1998; Outten and O'Halloran, 2001).



**Figure 1.12 Diagram of Zur Repression**

Diagram to represent Zur repression on a bacterial gene promoter. In the absence of  $\text{Zn}^{2+}$ , Zur does not bind to DNA, allowing RNAP binding and transcription. In the presence of  $\text{Zn}^{2+}$ , each Zur monomer binds two  $\text{Zn}^{2+}$  ions, leading to the  $(\text{Zur}_2)_2$ -DNA complex, sterically hindering RNAP binding, and thus no transcription.

The Zur binding location, 'Zur box', was originally discovered through a DNA footprint assay in *E. coli* K-12, which showed a 30 bp protected region between *znuA* and *znuCB* [5'-GAAATGTTATA-N-TATAACATTTC-3'] (Patzer and Hantke, 2000). However, more recently, a more accurate motif based *E. coli* Zur consensus was identified. Each Zur monomer binds to a Purine-N-N-N-Pyrimidine (RNNNY) motif. This motif is separated by 3 bp, giving a repeated motif of RNNNYxxxRNNNY, which is recognised by the Zur dimer. A combination of crystallography of Zur-DNA and biochemical data was used to define this Zur binding motif (Gilston *et al.*, 2014).

A combination of *in vitro* and *in silico* analysis has shown Zur to regulate six promoters in *E. coli*: *znuA*; *znuCB*; *zinT*; *pliG*; *ykgMO*; and *c1265-7* (Patzer and Hantke, 1998; Clayton, 2012; Gilston *et al.*, 2014). The specific  $(\text{Zur}_2)_2$ -DNA binding position varies between the six genes; some bind up stream, and some downstream of a +1 start site (Figure 1.13).

	-35	-10	+1
<i>znuA</i>	<b>CCGAGAC</b>	<b>ATTTTCCAGG</b>	<b>GAAACCAGACT</b>
<i>znuCB</i>	<b>ATCGTA</b>	<b>ATGAATATGAGAAGTGTGAT</b>	<b>TATTATAACATT</b>
<i>zinT</i>	<b>TTGCAT</b>	<b>TTTGCTATATGTTACAAT</b>	<b>TATAACATTACACATC</b>
<i>pliG</i>	<b>TTAGGGATTGTTAT</b>	<b>ATTATAACAGTTCATCGTACTCATTCTGA</b>	<b>ACAGGAG</b>
<i>ykgMO</i>	<b>TTGAAATGATCCGGATGAGCATGTATCTT</b>	<b>TATGGTTATGTTATAACATAACAGGTAA</b>	
<i>c1265</i>	<b>AAAATGTTATGATATAACATCACAATCACAAC</b>	<b>ATTTGTTTTTCGAA</b>	

**Figure 1.13 Promoter Alignment of Zur Regulated Promoters**

Sequence alignment of six Zur regulated promoters. **Bold** = -35 and -10 element. Underline = Zur binding site. **Green** = +1 start site. A combination *in silico* and *in vitro* analysis was used to locate the Zur binding site, and the -35 and -10 elements.

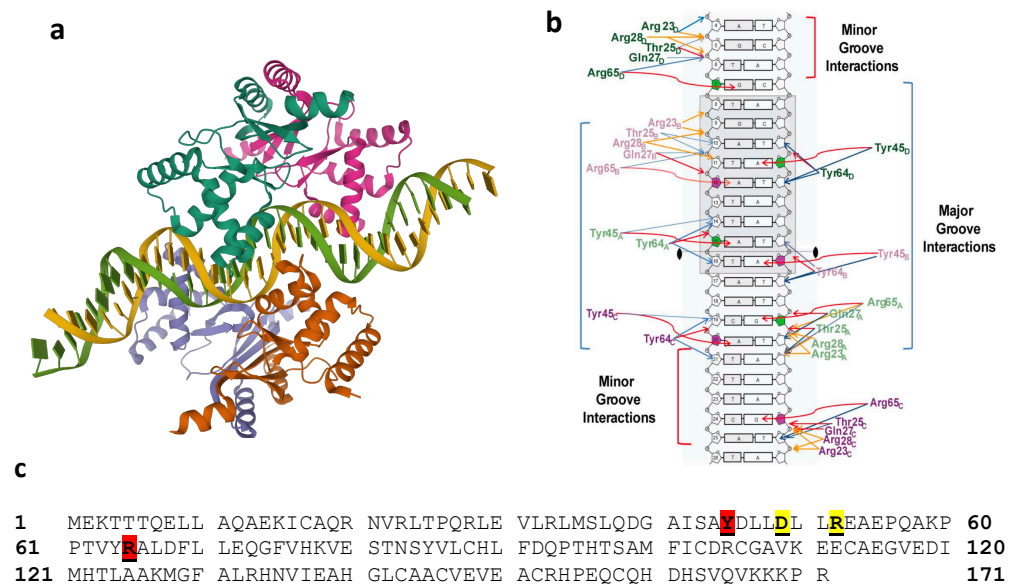
### 1.5.1. Zur Structure

Zur is naturally found as a homodimer; two identical Zur monomers dimerize (Zur<sub>2</sub>). Zur homodimerization has been shown to occur independently of zinc concentrations (Patzner and Hantke, 2000). Each Zur monomer contains two domains: the N-terminal domain, which interacts with DNA; and the C-terminal domain, which causes dimerization. The DNA binding domain contains a helix-turn-helix (HTH), which is often seen in the DNA binding domain of TFs. Each HTH of the Zur<sub>2</sub> binds to the major grooves of DNA. Two Zur<sub>2</sub> dimers bind on opposite sides of the DNA, essentially encircling the DNA, creating the (Zur<sub>2</sub>)<sub>2</sub>-DNA complex, with over 100 interaction (Figure 1.14a). Zur dimers in the (Zur<sub>2</sub>)<sub>2</sub> complex forms a salt bridge when bound to DNA, leading to the increased stability of the (Zur<sub>2</sub>)<sub>2</sub>-DNA complex. This salt bridge occurs at amino acid residues D49 and R52 (Figure 1.14c). Each Zur monomer makes two hydrogen bonds with the DNA, at amino acid residues R64 and Y45, this interaction defines the RNNNY motif (Figure 1.14b) (Gilston *et al.*, 2014).

(Zur<sub>2</sub>)<sub>2</sub>-P<sub>znuABC</sub> forms over 100 interactions, mainly through Zur amino acid (aa) side chains in the DNA phosphate backbone. Each Zur monomer forms two hydrogen bonds with conserved purine bases at Y45 and R65, totalling eight hydrogen bonds within (Zur<sub>2</sub>)<sub>2</sub>-P<sub>znuABC</sub> (Figure 1.14b). A combination of *in vitro* and *in silico* analysis suggests that R65 is highly conserved throughout the FUR family, suggesting its essential role in DNA binding, whereas Y45 is only seen in *E. coli* Zur (Gilston *et al.*, 2014).

Each Zur monomer can bind up to two Zn<sup>2+</sup> ions, with two distinct binding sites. In the first zinc binding site (A site) Zn<sup>2+</sup> is bound to four sulphur atoms at C103, C106, C143, and C146 (Figure 1.15a). In the second zinc binding site (B site) Zn<sup>2+</sup> is bound by residues in a nitrogen/oxygen rich site at H77, C888, H96, and G111 (Figure 1.15b). The sulphur rich binding site (A site) has shown to have stronger binding affinity to Zn<sup>2+</sup>

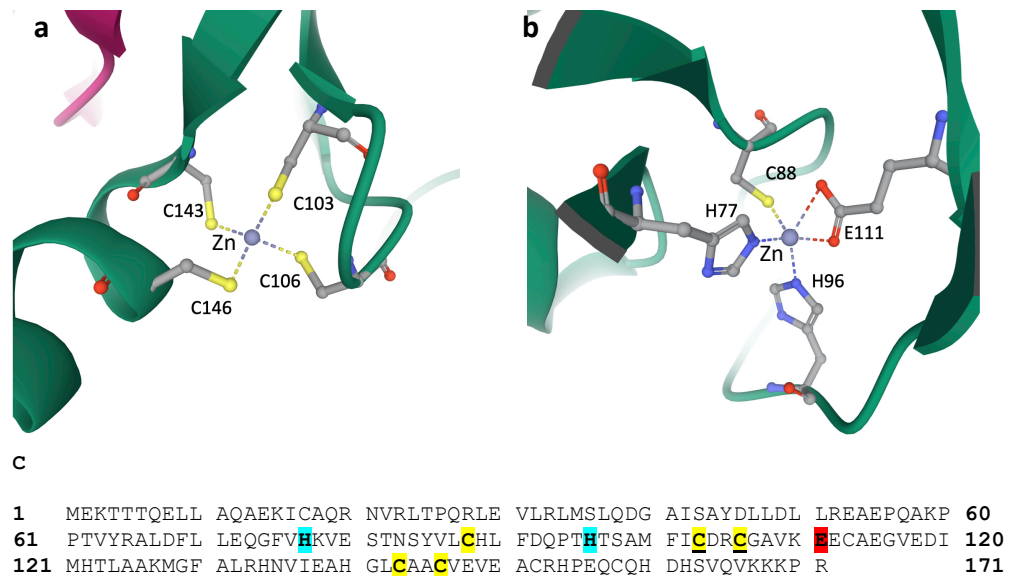
than at the B site. It has been shown that it is essential for both A and B sites to be occupied with  $\text{Zn}^{2+}$  for correct  $(\text{Zur}_2)_2$ -DNA binding (Gilston *et al.*, 2014). Outten *et al.*, (2001) suggested that the B site can bind both  $\text{Co}^{2+}$  and  $\text{Zn}^{2+}$ , but with a higher affinity to  $\text{Zn}^{2+}$ .



**Figure 1.14 Model of  $(\text{Zur}_2)_2$ -DNA Complex**

(a) Computer generated model of  $(\text{Zur}_2)_2$ -DNA complex, based on  $P_{\text{ZnuACB}}$  promoter, from X-ray diffraction [accession, 4MTD]. DNA shown in the centre, surround by four Zur monomers, each monomer represented by a different colour. (b)  $(\text{Zur}_2)_2$ -DNA interactions. Red arrows = Hydrogen bonds, blue arrow = hydrophobic interactions, orange arrows = electrostatic interactions (Gilston *et al.*, 2014). (c) Amino acid sequence of Zur. **Red** = amino acid involved in Hydrogen-DNA bond (Y45, R65). **Yellow** = amino acid involved in  $(\text{Zur}_2)_2$  salt bridge (D49, R52). Adapted from (Gilston *et al.*, 2014).

Gilston, B. A., Wang, S., Marcus, M. D., Canalizo-Hernández, M. A., Swindell, E. P., Xue, Y., Mondragón, A. and O'Halloran, T. V. (2014) 'Structural and Mechanistic Basis of Zinc Regulation Across the E. coli Zur Regulon', PLoS Biology, 12(11). Reprinted with permission from PLoS



**Figure 1.15 Zn<sup>2+</sup> Interaction Sites with Zur**

Diagram to show Zn<sup>2+</sup> interaction sites with *E. coli* Zur. (a) Zur, zinc A site. Sulphur binding interactions with Zn<sup>2+</sup>-Zur. (b) Zur, zinc B site. Nitrogen/oxygen binding interactions with Zn<sup>2+</sup>-Zur. (c) Amino Acid sequence of Zur monomer, showing Zn<sup>2+</sup> binding sites. **Underline Bold** indicates Zn<sup>2+</sup>-Zur in A site. **Bold only** indicates Zn<sup>2+</sup>-Zur in B site. Adapted from (Gilston *et al.*, 2014)

### 1.5.2. Zur Regulated Genes

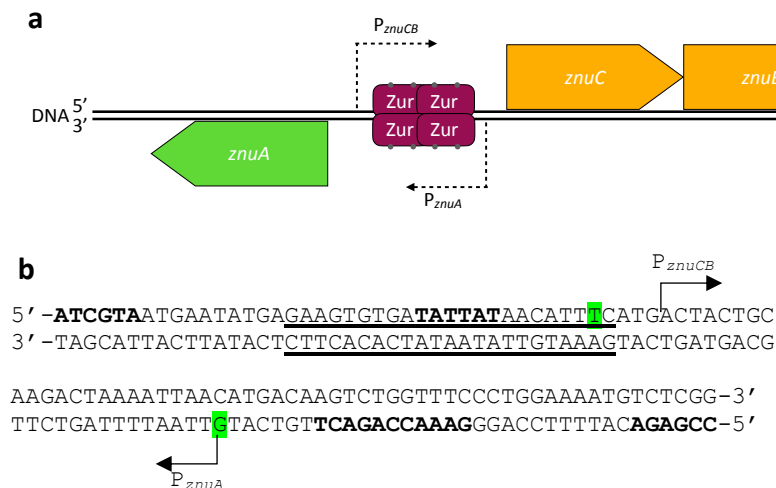
For the purpose of this thesis, six *E. coli* Zur regulated promoters will be analysed and discussed; P<sub>znuA</sub>, P<sub>znuCB</sub>, P<sub>zinT</sub>, P<sub>pliG</sub>, P<sub>ykgM</sub>, and P<sub>c1265</sub>.

#### 1.5.2.1. ZnuABC

It is important to note, that when talking about the promoter region of *znuABC*, two different promoters are considered, P<sub>znuA</sub>, and P<sub>znuCB</sub>. *znuABC* is a divergently transcribed operon, under the regulation of Zur (Figure 1.16). As *znuA* and *znuCB* are transcribed divergently, both divergent promoters are considered for this thesis.

Zur binds to a central region between *znuA* and *znuCB*. Interestingly, only one Zur binding site regulates both divergently transcribed genes. Zur binds to the DNA sequence 5'-GAAGTGTGATATTATAACATTTC-3'. The Zur binding site is not central

between the divergent genes; Zur overlaps the -10 and +1 start of  $P_{znuCB}$ , whereas Zur binds +26 bp upstream of the +1 start site of  $P_{znuA}$  (Figure 1.16).



**Figure 1.16 Diagram of *znuABC* Operon Structure**

(a) Diagram of the gene structure of the *znuABC* operon of *E. coli* MG1655. Diagram shows  $(Zur_2)_2$ - $P_{znuABC}$  binding between the divergent genes. (b) DNA Sequence of divergently transcribed  $P_{znuA}$  and  $P_{znuCB}$ . **Bold** = -35 and -10 element. Underline = Zur binding site. Green = +1 start site. Adapted from (Keseler *et al.*, 2017)

ZnuABC (Zinc Uptake), previously known as YebLMI, was first associated with Zur regulated by Patzer and Hantke, (1998). ZnuABC is a member of the ATP-Binding Cassette transporters system (ABC transporters) with a structure of  $AB_2C_2$  (Patzer and Hantke, 1998; Locher, 2009).

ZnuA is a periplasmic binding protein with a high affinity to  $Zn^{2+}$ . It has been shown that ZnuA binds at least two  $Zn^{2+}$ . This  $Zn^{2+}$ -ZnuA complex causes a conformational change which stabilises the complex and may facilitate recognition of ZnuB. ZnuA does not exclusively bind  $Zn^{2+}$ , it also binds  $Co^{2+}$ ,  $Ni^{2+}$ ,  $Cu^{2+}$  and  $Cd^{2+}$ , though it shows a higher affinity to  $Zn^{2+}$  (Chandra, Yogavel and Sharma, 2007; Yatsunyk *et al.*, 2008). ZnuB is an integral membrane protein and ZnuC is an ATPase. ZnuB receives  $Zn^{2+}$  from ZnuA and is responsible for its transport across the inner membrane to the cytoplasm. This  $Zn^{2+}$



translocation is driven by ATP hydrolysis of the ZnuC ATPase (Figure 1.11) (Patzner and Hantke, 1998).

#### 1.5.2.2. YkgMO

*ykgM* is part of the *ykgMO* operon, which is regulated by Zur. YkgM is a paralogue of the ribosomal protein L31. A single  $\text{Zn}^{2+}$  ion binds to a 'zinc ribbon' within the L31 protein. The YkgM paralogue does not require  $\text{Zn}^{2+}$  to function, but carries out the same cellular role as L31, which allows  $\text{Zn}^{2+}$  to be released from L31 (Graham *et al.*, 2009; Hemm *et al.*, 2010; Hensley *et al.*, 2012). It has also been shown that *ykgO* is a zinc free paralogue of the ribosomal protein L36. Like YkgM, YkgO replaces L36, and allows L36 bound  $\text{Zn}^{2+}$  to be released (Ueta, Wada and Wada, 2020)

#### 1.5.2.3. PlIG

pliG (periplasmic lysozyme inhibitor of g-type lysozyme) is upregulated by Zur in zinc depleted conditions. Lysozymes are part of the innate immune system and are antimicrobial enzymes which hydrolyse peptidoglycan; specifically, the 1,4- $\beta$  links between MurNAc and GlcNAc, leading to cell lysis. As the name suggests, PlIG inhibits g-type lysozyme, increasing protection from host cells (Vanderkelen, Van Herreweghe, Vanoirbeek, *et al.*, 2011; Gilston *et al.*, 2014).

#### 1.5.2.4. ZinT

ZinT, previously known as YodA, is repressed by Zur, but has also been shown to be regulated by Fur and SoxS. ZinT is a zinc chaperone for the major zinc import system, ZnuABC. ZinT is not essential for ZnuABC function but can aid in metal acquisition during zinc depleted states. Further, ZinT has been shown to bind not just to zinc, but

also nickel, cadmium and mercury (Kershaw, Brown and Hobman, 2007; Gabbianelli *et al.*, 2011; Colaço *et al.*, 2016).

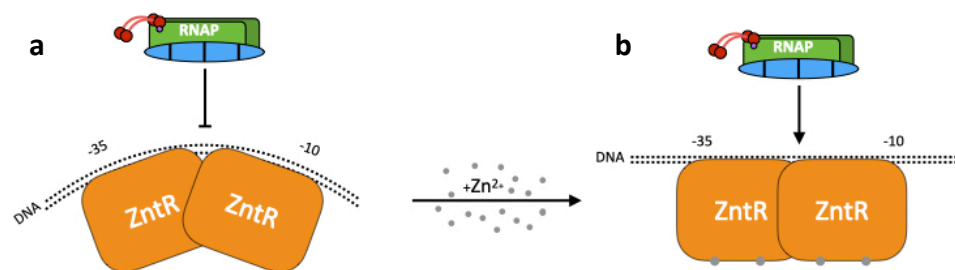
#### 1.5.2.5. C1265-7

*c1265-7* is a novel Zur regulated operon found chromosomally only in uropathogenic *E. coli* (UPEC) strains. A combination of *in vitro* and *in silico* analysis suggests that *c1265-7* is regulated by Zur. Transcriptomics analysis showed that in *E. coli* CFT073, when expressed in  $\Delta zur$  compared to *wt*, *c1265-7* showed up to 60-fold increase in gene transcripts (Clayton, 2012).

Further analysis has shown that homologues of C1265-7 are only found in exPEC (extra-intestinal *E. coli*) and not in K-12 lab strains. *In silico* amino acid analysis suggested C1265-7 may be involved in zinc acquisition. C1265 shows homology to the TonB-dependent transporter ShuA from *Shigella dysenteriae*, a heme transporter. C1266 shows no homology to any genes. C1267 has significant homology to COG0523 proteins, which are involved in zinc homeostasis (Haas *et al.*, 2009; Clayton, 2012). The combination of transcriptomics, *in silico* homology analysis, and predictive Zur binding upstream of *c1265* suggests Zur regulation, with the potential for involvement in zinc acquisition.

## 1.6. ZntR

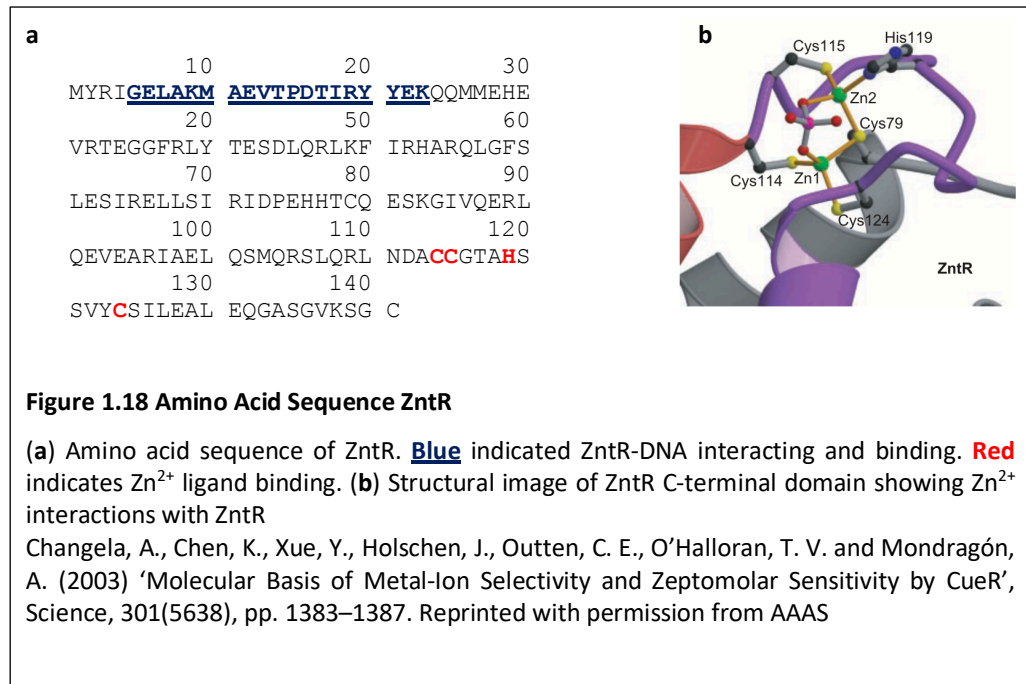
ZntR, previously known as YhdM, is a member of the MerR family of transcription factor regulators, which regulates the expression of the P-type ATPase ZntA. Like MerR, ZntR binds permanently upstream of *zntA*. The promoter region of *zntA* has sub-optimal spacing between the -35 and -10 hexamer at 20 bp, optimal is  $17 \pm 1$ . Once the ZntR dimer is activated with  $\text{Zn}^{2+}$ , a conformational change occurs, underwinding the DNA in the promoter which brings the spacing between the -35 and -10 hexamer to optimal spacing, allowing RNA Polymerase (RNAP) to bind. In turn, this upregulates gene expression (Figure 1.17) (Rensing, Mitra and Rosen, 1997; Brocklehurst *et al.*, 1999; Outten, Outten and O'Halloran, 1999).



**Figure 1.17 Diagram of ZntR Regulation**

(a) ZntR binds to a 22 bp 11-11 bp palindromic repeats, causing slight repression of RNAP. (b) Excess  $\text{Zn}^{2+}$  binds to ZntR at the C-terminal end, causing a conformational change, which in turn underwinds the DNA, allowing for increase DNA-RNAP affinity, upregulating gene expression.

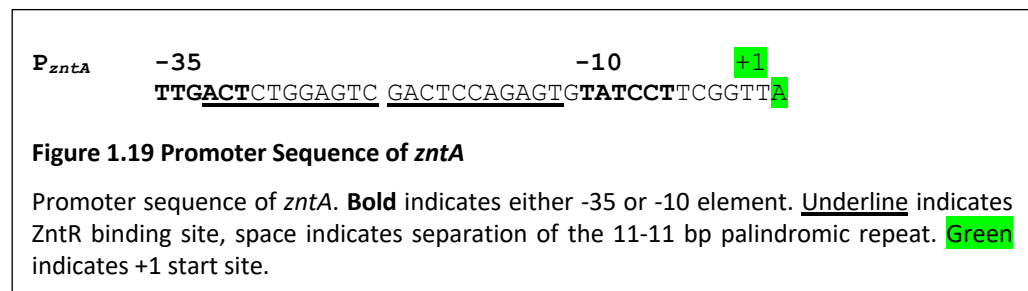
ZntR is composed of two monomers, each monomer is a 141-amino acid polypeptide which has an N-terminal domain which binds to DNA, and a C-terminal domain, which interacts with  $\text{Zn}^{2+}$ . ZntR requires two  $\text{Zn}^{2+}$  ions per monomer to activate transcription. ZntR-DNA interaction occurs at the N-terminal domain, which has a helix-turn-helix structure; DNA interaction occurs at amino acids 4-23. ZntR- $\text{Zn}^{2+}$  interaction occurs at amino acids C114, C115, H119 and C124 (Figure 1.18) (Khan *et al.*, 2002; Changela *et al.*, 2003).



### 1.6.1. ZntR- $P_{zntA}$

ZntR binds upstream of *zntA*, in the *zntA* promoter region ( $P_{zntA}$ ). ZntR binds to a 22 bp sequence, which consists of an 11-11 bp palindromic repeat, specifically between -37 to -12 bp. Each monomer of the ZntR dimer binds to an 11 bp sequence of the palindrome. The 5' 11 bp repeat overlaps the -35 hexamer of  $P_{zntA}$  (Figure 1.19). It is important to note that  $P_{zntA}$ -RNAP requires  $\sigma^{70}$  for promoter recognition. ZntR is a member of the MerR family regulators and the mechanism of induction is similar. The spacing between the -35 and -10 element of  $P_{zntA}$  is suboptimal at 20 bp, meaning that RNAP binds, but cannot form the open complex. As the spacing is not optimal the  $\sigma^{70}$  subunits  $\sigma_{4.2}$  and  $\sigma_{2.4}$  cannot successfully make interactions with the -35 and -10 elements, respectively. The ZntR- $P_{zntA}$  complex weakly suppresses the expression of *zntA*. Once  $\text{Zn}^{2+}$  bind to the C-terminal of ZntR, it causes a conformational change in the structure of ZntR. The  $\text{Zn}^{2+}$ -ZntR- $P_{zntA}$  complex underwinds the DNA ( $\sim 33^\circ$ ), straightening the helix backbone, allowing correct  $\sigma^{70}$  interaction with the promoter DNA, and thus forms the RNAP open complex. Once the RNAP open complex is

achieved, transcription of *zntA* can start. (Brocklehurst *et al.*, 1999; Outten, Outten and O'Halloran, 1999; Brown *et al.*, 2003; Wang *et al.*, 2012).



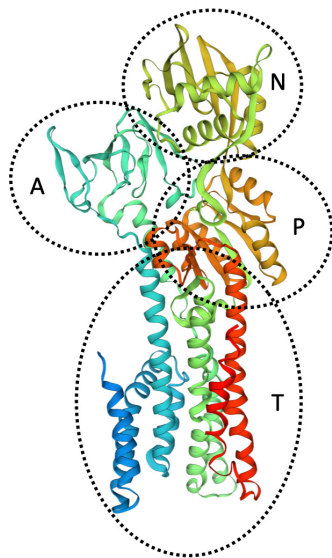
It has been proposed that ligand controlled proteolysis of ZntR may play a role in zinc homeostasis. Pruteanu, Neher and Baker, (2007) proposed a novel feedback loop between proteolysis of ZntR and zinc. Essentially, results indicate that ZntR-DNA has a higher potential to be degraded by Lon protease than  $Zn^{2+}$ -ZntR-DNA. When zinc levels are high, the  $Zn^{2+}$ -ZntR-DNA complex is more stable ( $t_{1/2} > 60$  min), increasing ZntR stability, which further increases expression of ZntA and thus increases zinc export. When zinc levels are low, the ZntR-DNA complex is more unstable ( $t_{1/2} \sim 30$  min), increasing proteolysis by Lon, which further decreases expression of ZntA. Interestingly, it is not the  $Zn^{2+}$  ligand binding C-terminal domain of ZntR which alters the potential for proteolysis, but rather the DNA binding N-terminal domain. It is thought that the proteolysis by Lon occurs via recognition of the N-terminal domain, which changes conformation with  $Zn^{2+}$  ligand binding at the C-terminal domain (Pruteanu, Neher and Baker, 2007).

### 1.6.2. ZntA

ZntA, previously known as YhhO, is a P-type ATPase associated with the translocation of zinc from the cytoplasm to the periplasm. ZntA is a major component of the zinc homeostasis system, and is involved in the export of zinc when internal concentrations are too high (Brocklehurst *et al.*, 1999; Outten and O'Halloran, 2001). ZntA is a P-type

ATPase, a family of integral membrane proteins, specifically a P1B-ATPase, associated with translocation of monovalent and divalent metals (Palmgren and Nissen, 2011).

There is currently not crystal structure of ZntA, however, *E. coli* ZntA shows significant similarity to ZntA from *Shigella sonnei*, (99 % aa identity) (Figure 1.20) (Wang *et al.*, 2014). P1B-ATPase show a primary structure; a nucleotide binding domain (N), a phosphorylation domain (P), an actuator domain (A), and a transmembrane domain (T). P-type ATPases can self-phosphorylate; the N domain interacts with ATP and phosphorylates the P domain. This phosphorylation leads to a conformation change, and translocation of the metal ion (e.g.  $\text{Zn}^{2+}$ ) (Rensing, Mitra and Rosen, 1997; Palmgren and Nissen, 2011).



**Figure 1.20 Structure of ZntA *Shigella sonnei***

ZntA structural model of ZntA from *S. sonnei* in a phosphoenzyme ground state (E2P) (accession, 4umv.1.A). *S. sonnei* ZntA show 99.04% amino acid identity to *E. coli* ZntA. (N) nucleotide binding domain. (P) phosphorylation domain. (A) actuator domain. (T) transmembrane domain. Adapted from (Wang *et al.*, 2014)

### 1.6.3. Stochastics of ZntA Regulation

A study by Takahashi *et al.*, (2015) hypothesised that in the presence of zinc, *E. coli* mount a heterogenous regulation response of *zntA*, rather than a homogenous

response. During increasing zinc concentrations, ZntA gene expression is upregulated, through transcriptional regulation aided by the active activator, ZntR (Brocklehurst *et al.*, 1999). A homogenous gene regulation response within a homogenous population would suggest that all *E. coli* within a population respond to zinc in the same way, equally regulating the transcriptional expression of ZntA. Takahashi *et al.*, (2015) proposed that *E. coli* mounts a heterogenous response; sub-population of *E. coli* responds to zinc with a variety of expression levels of ZntA, some express at high levels, others low. Takahashi *et al.*, (2015) further suggested that some sub-population may even die or stop growing. This allows the population to 'bet hedge' and more readily prepare for a response, and potential toxic zinc concentrations (Takahashi *et al.*, 2015).

### 1.7. Aims and Objectives

There are four main aims for this thesis and each aim has a results chapter dedicated to its findings.

- i. To further understand the adaptive transcriptional regulation of the ZntR regulated  $P_{zntA}$ , with the intention to utilise  $P_{zntA}$  as a novel zinc inducible system.
- ii. To further understand the regulation of Zur regulated promoters, with the intention to develop a novel inducible or repressible promoter system.
- iii. To further understand the stochastic response to the ZntR regulated  $P_{zntA}$ , and determine if *zntA* shows a homogenous or heterogenous transcriptional regulation.
- iv. To determine if the phenotypic function of the novel Zur regulated *c1265-7* operon of UPEC strains is, as hypothesised, is associated with zinc acquisition.



# Chapter 2

## Materials and Methods

## 2.1. Materials

## 2.1.1. Bacterial Strains

Bacterial strain	Characteristics	Source
<i>E. coli</i> K-12, MG1655	Laboratory Strain F-, lambda-, rph-1	(Blattner <i>et al.</i> , 1997) Laboratory stock
<i>E. coli</i> CFT073	UPEC (Uropathogenic <i>Escherichia coli</i> )	(Welch <i>et al.</i> , 2002)
<i>E. coli</i> CFT073 $\Delta c1265-7$	UPEC $\Delta c1265-7::kan^R$	(Clayton, 2012)
<i>E. coli</i> CFT073 $\Delta c1265-7 \Delta znuABC$	UPEC, $\Delta c1265-7 \Delta znuABC::kan^R$	(Clayton, 2012)
<i>E. coli</i> MG1655 <i>zntA::rfp::kanR</i>	<i>mRFP1</i> chromosomally fused to the C-terminal end of <i>zntA</i> . Kanamycin resistant	This Work (2.2.15)
<i>E. coli</i> MG1655 $\Delta zur$	Laboratory strain <i>E. coli</i> . $\Delta zur Kan^S$	(Clayton, 2012)
<i>E. coli</i> MG1655 $\Delta znuA::kan^R$	Laboratory strain <i>E. coli</i> . $\Delta znuA::kan^R$	(Clayton, 2012)
<i>E. coli</i> MG1655 $\Delta znuCB::kan^R$	Laboratory <i>E. coli</i> . $\Delta znuCB::kan^R$	(Clayton, 2012)
<i>E. coli</i> K-12, NEB 5-alpha	Highly efficient chemically competent cells. DH5-alpha derivative	NEB #C2987
<i>Cupriavidus necator</i> H16	<i>Previously known as; Ralstonia eutropha, Alcaligenes eutrophus, Wautersia eutropha, and Hydrogenomonas eutropha.</i>	(Pohlmann <i>et al.</i> , 2006)

Table 2.1 Table of Bacterial Strains

## 2.1.2. Plasmids

Plasmid	Characteristics	Source
pLUX	Reporter plasmid, with <i>luxCDABE</i> downstream of a MCS. Kan <sup>R</sup> . <i>ori</i> pSC101	(Burton <i>et al.</i> , 2010)
pJI100	Low copy number expression plasmid. Kan <sup>R</sup> . <i>ori</i> pSC101.	This work (2.2.9.1)
pJI100A	Low copy number expression plasmid. Amp <sup>R</sup> . <i>ori</i> pSC101.	This work (2.2.9.2)
pJI102	<i>c1265-7</i> expression plasmid, based on pJI100A. Amp <sup>R</sup> . <i>ori</i> pSC101	This work (2.2.9.3)
pJI300	Reporter plasmid, with <i>luxCDABE</i> and <i>mRFP1</i> downstream of a MCS. Kan <sup>R</sup> . <i>ori</i> pSC101	This work (2.2.10)
pJI301	P <sub>zntA</sub> cloned into pJI300	This work (2.2.11)
pJI302	P <sub>trc</sub> cloned into pJI300	This work (2.2.11)
pJI303	P <sub>znuA</sub> cloned into pJI300	This work (2.2.11)
pJI304	P <sub>znuCB</sub> cloned into pJI300	This work (2.2.11)
pJI305	P <sub>zinT</sub> cloned into pJI300	This work (2.2.11)
pJI306	P <sub>pliG</sub> cloned into pJI300	This work (2.2.11)
pJI307	P <sub>ykgM</sub> cloned into pJI300	This work (2.2.11)
pJI308	P <sub>c1265</sub> cloned into pJI300	This work (2.2.11)

Plasmid	Characteristics	Source
pUltraRFP-KM	Constitutively expressed <i>mRFP1</i> expression plasmid. Kan <sup>R</sup> . <i>ori</i> p15A. <i>mRFP1</i> .	(Mavridou <i>et al.</i> , 2016)
pDOC-C	Gene doctoring cloning vector Amp <sup>R</sup> . <i>sacB</i> .	(Lee <i>et al.</i> , 2009)
pDOC-G	Gene doctoring, amplification of <i>gfp</i> for gene fusion. Amp <sup>R</sup> . Kan <sup>R</sup> . <i>sacB</i> . <i>eGFP</i> .	(Lee <i>et al.</i> , 2009)
pDOC-K	Gene doctoring, amplification of Kan <sup>R</sup> for gene deletion. Kan <sup>R</sup> . Amp <sup>R</sup> . <i>sacB</i> .	(Lee <i>et al.</i> , 2009)
pACBSR	Gene doctoring helper plasmid. Lambda ( $\lambda$ ) red induction for chromosomal insertion. Chlor <sup>R</sup> . <i>araBAD</i> promoter.	(Herring, Glasner and Blattner, 2003)
pDOC- <i>zntA::rfp</i>	Amplified <i>mRFP1</i> from pUltraRFP-KM, with flanking regions of <i>zntA</i> , cloned into pDOC-C.	This work (2.2.15.1)

Table 2.2 Table of Plasmids

### 2.1.3. Primers

All DNA oligonucleotides for PCR and sequencing were supplied by Sigma-Aldrich (Merck, Darmstadt, Germany), with 0.025  $\mu$ mole synthesis, desalt purification, and provided at a concentration of 100  $\mu$ M in nuclease free H<sub>2</sub>O. Working stocks of 10  $\mu$ M were prepared using sterile HPLC grade H<sub>2</sub>O, and stored at -20°C.

All qPCR fluorescently labelled probes were supplied by Eurofins Genomics (Ebersberg, Germany), with 0.01  $\mu$ mole synthesis, HPLC purification, and made up to 100  $\mu$ M using supplied qPCR probe dilution buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). For the 5' fluorophore FAM (ex 495 nm | em 520 nm), 3' quencher BHQ-1 (Black Hole Quencher one), working stocks of 10  $\mu$ M were prepared in HPLC grade H<sub>2</sub>O, aliquoted to reduce freeze/thaw degradation, and stored at -20°C in a black out storage bag.

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
seq_204_F	CGATCCTCATCCTGTCTC	Plasmid sequencing	62
rfp_seq_R	TACCTTCGTACGGACGAC		
rrnB_seq_R	CGGATTTGTCCTACTCAGGA		
pDOC-C_seq_F	TATGCTTCCGGCTCG	Plasmid sequencing	N/A

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
pDOC-C_seq_R	GGATGTGCTGCAAG		
pDOC-C_seq_F	TATGCTCCGGCTCG	Plasmid sequencing	N/A
pDOC-C_seq_R	GGATGTGCTGCAAGG		
pJI100_F	GGATCCGCGGCCGCAACTAGAGGC	Cloning	72
pJI100_R	GGATCCCCGGGCTCGAG		
pJI100A_lin_F	GCGAAACGATCCTCATCC	Cloning	69
pJI100A_lin_R	GCGGGACTCTGGGGTTCG		
pJI100A_amp_F	CTCGAACCCAGAGTCCCGCTTACCAATGCTTAATCAGTGAGG	Cloning	66
pJI100A_amp_R	CAGGATGAGGATCGTTTCGCCGCGGAACCCCTATTTGTTTATTTTC		
pJI102_lin_F	ATTTGTGTATAAGAGACAGTACTTATC	Cloning	72
pJI102_lin_R	TGCAATATCGCGTTAATC		
pJI102_c12657_F	GTCTTCACCTCGAGCCCGGGATCACAACATTTGTTTTTCG	Cloning	59

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
pJI102_c12657_R	CTAGTTGCGGCCGCGGATCCTGCAATATCGCGTTAATC		
pJI300_Lin_F	TCTAGTTAGTTAGTAAGGAGTTTACCATGGCAAATATGACTAAAAAATTC	Cloning	72
pJI300_Lin_R	ATTTGTTTTTTTAAACTCCTTACTAACTAACTAGAGGATCCCCGGGCTCGAG		
pJI300_mRFP_F	CGGGGGATCCTCTAGTTAGTTAGTAAGGAGTTTAAAAAACAAATATGGCAAGTAGTGAAGACG	Cloning	63
pJI300_mRFP_R	CTCCTTACTAACTAACTAGATTAAGCACCGGTGGAGTG		
pJI301(zntA)_F	GAGGCCCTTTCGTCTTCACCTTTGCCGGTCACTTCCTG	Cloning	64
pJI301(zntA)_R	TCCTTACTAACTAACTAGAGTTCTTGCCGTGATTGTCAG		
pJI302(trc)_F	GAGGCCCTTTCGTCTTCACCGCGCGAAGGCGAAGCGGC	Cloning	69
pJI302(trc)_R	TCCTTACTAACTAACTAGAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCC		
pJI303(znuA)_F	GAGGCCCTTTCGTCTTCACCGGCGTTGGCCAAAAGAAAC	Cloning	62

Primer	Sequence 5'-3'	Use	Tm x°C
pJI303(znuA)_R	TCCTTACTAACTAACTAGAGCGGATAATGCTGCGAAAAG		
pJI304(zunCB)_F	GAGGCCCTTTCGTCTTCACCGAGCGGCGGATAATGCTG	Cloning	64
pJI304(zunCB)_R	TCCTTACTAACTAACTAGAGCCAGGGAAACCAGACTTG		
pJI305(zinT)_F	GAGGCCCTTTCGTCTTCACCATAAGATAGATAAGTAGAACTGAG	Cloning	56
pJI305(zinT)_R	TCCTTACTAACTAACTAGAGAACAGCCAGTTTGTAAG		
pJI306(pliG)_F	GAGGCCCTTTCGTCTTCACCAATGCCGCAACGTGATTTTAC	Cloning	63
pJI306(pliG)_F	TCCTTACTAACTAACTAGAGATACAGCCTTCCTGATGC		
pJI307(ykgM)_F	GAGGCCCTTTCGTCTTCACCGCTAACAATGCCAGAGTTC	Cloning	59
pJI307(ykgM)_R	TCCTTACTAACTAACTAGAGATTGGGCTTCATCATTTTTAC		
pJI308(c1265)_F	GAGGCCCTTTCGTCTTCACCCGACAGAACATTATTAACAGAG	Cloning	59



Primer	Sequence 5'-3'	Use	Tm x°C
pJI308(c1265)_R	TCCTTACTAACTAACTAGAGGTACAAATGACAGTTCTGATTAC		
PzntA(-25 <sup>^</sup> -24insT)_F	GACTCCAGAGTCAAGTTTTATC	Site directed mutagenesis	61
PzntA(-25 <sup>^</sup> -24insT)_R	TGGACTCCAGAGTGTATCCTTC		
PzntA(-25 <sup>^</sup> -24insTG)_F	TGACTCCAGAGTGTATCCTTC	Site directed mutagenesis	61
PzntA(-25 <sup>^</sup> -24insTG)_R	TGGACTCCAGAGTGTATCCTTC		
PzntA(-12_-7>TGACA)_F	TCCAGAGTGTTGACATCGGTTAATGAGAAAAAAC	Site directed mutagenesis	60
PzntA(-12_-7>TGACA)_F	GTCGACTCCAGAGTCAAG		
PzntA(-33T>A)_F	AAAACCTTGACACTGGAGTCGAC	Site directed mutagenesis	60
PzntA(-33T>A)_R	ATCAGAGATACAGCGAGC		
PzntA(-14T>G)_F	GACTCCAGAGGGTATCCTTCG	Site directed mutagenesis	61
PzntA(-14T>G)_R	GACTCCAGAGTCAAGTTTTATC		
PzntA(-18A>G)_F	AGTCGACTCCTGAGTGTATCC	Site directed mutagenesis	58
PzntA(-18A>G)_R	CCAGAGTCAAGTTTTATCAG		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
PzntA(-11_-7>GTCAA)_F	CTCCAGAGTGGTCAAATCGGTTAATGAGAAAAAACTTAAC	Site directed mutagenesis	59
PzntA(-11_-7>GTCAA)_R	TCGACTCCAGAGTCAAGT		
PzntA(-9_-8>AA)_F	CAGAGTGTATAATTCGGTTAATGAGAAAAAAC	Site directed mutagenesis	62
PzntA(-9_-8>AA)_R	GAGTCGACTCCAGAGTCA		
PzntA(-9_-8>AA, -33T>A)_F	CCAGAGTGTATAATTCGGTTAATGAGAAAAAAC	Site directed mutagenesis	55
PzntA(-9_-8>AA, -33T>A)_R	AGTCGACTCCAGTGTCAAGTTTTATCAGAGATAC		
PzntA(-36_-33>TACA)_F	GATAAACTTTTACACTGGAGTCGACTCCAG	Site directed mutagenesis	63
PzntA(-36_-33>TACA)_R	AGAGATACAGCGAGCGGA		
PzntA(-36_-33>TACA, -9_-8>AA)_F	TCCAGAGTGTATAATTCGGTTAATGAGAAAAAAC	Site directed mutagenesis	55
PzntA(-36_-33>TACA, -9_-8>AA)_R	GTCGACTCCAGTGTAAGTTTTATCAGAGATACAG		
PzntA(-35A<C)_F	ATAAACTTGCCTCTGGAGTCGACTC	Site directed mutagenesis	63
PzntA(-35A<C)_R	CAGAGATACAGCGAGCGG		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
PzntA(-34C<A)_F	TAAAACTTGAATCTGGAGTCGAC	Site directed mutagenesis	62
PzntA(-34C<A)_R	TCAGAGATACAGCGAGCG		
PzntA(-33T<G)_F	AAAACCTTGACGCTGGAGTCGAC	Site directed mutagenesis	62
PzntA(-33T<G)_R	ATCAGAGATACAGCGAGC		
PzntA(-32C<A)_F	AAACTTGACTATGGAGTCGAC	Site directed mutagenesis	59
PzntA(-32C<A)_R	TATCAGAGATACAGCGAG		
PzntA(-31T<G)_F	AACCTTGACTCGGGAGTCGACTC	Site directed mutagenesis	59
PzntA(-31T<G)_R	TTATCAGAGATACAGCGAG		
PzntA(-30G<T)_F	ACTTGACTCTTGAGTCGACTC	Site directed mutagenesis	57
PzntA(-30G<T)_R	TTTATCAGAGATACAGCG		
PzntA(-29G<T)_F	CTTGACTCTGTAGTCGACTCC	Site directed mutagenesis	57
PzntA(-29G<T)_R	TTTTATCAGAGATACAGCG		
PzntA(-28A<C)_F	TTGACTCTGGCGTCGACTCCA	Site directed mutagenesis	59
PzntA(-28A<C)_R	GTTTTATCAGAGATACAGCG		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
PzntA(-27G<T)_F	TGACTCTGGATTCTGACTCCAG	Site directed mutagenesis	57
PzntA(-27G<T)_R	AGTTTTATCAGAGATACAGC		
PzntA(-26T<G)_F	GACTCTGGAGGCGACTCCAGA	Site directed mutagenesis	63
PzntA(-26T<G)_R	AAGTTTTATCAGAGATACAGCGAG		
PzntA(-25C<A)_F	ACTCTGGAGTAGACTCCAGAG	Site directed mutagenesis	60
PzntA(-25C<A)_R	CAAGTTTTATCAGAGATACAGC		
PzntA(-24G<T)_F	CTCTGGAGTCTACTCCAGAGTG	Site directed mutagenesis	58
PzntA(-24G<T)_R	TCAAGTTTTATCAGAGATACAG		
PzntA(-23A<C)_F	TCTGGAGTCGCCTCCAGAGTG	Site directed mutagenesis	59
PzntA(-23A<C)_R	GTCAAGTTTTATCAGAGATACAG		
PzntA(-22C<A)_F	CTGGAGTCGAATCCAGAGTGT	Site directed mutagenesis	61
PzntA(-22C<A)_R	AGTCAAGTTTTATCAGAGATACAG		
PzntA(-21T<G)_F	TGGAGTCGACGCCAGAGTGTA	Site directed mutagenesis	62
PzntA(-21T<G)_R	GAGTCAAGTTTTATCAGAGATACAG		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
PzntA(-20C<A)_F	GGAGTCGACTACAGAGTGTATC	Site directed mutagenesis	60
PzntA(-20C<A)_R	AGAGTCAAGTTTTATCAGAGATAC		
PzntA(-19C<A)_F	GAGTCGACTCAAGAGTGTATC	Site directed mutagenesis	58
PzntA(-19C<A)_R	CAGAGTCAAGTTTTATCAGAG		
PzntA(-18A<C)_F	AGTCGACTCCCGAGTGTATCC	Site directed mutagenesis	58
PzntA(-18A<C)_R	CCAGAGTCAAGTTTTATCAG		
PzntA(-17G<T)_F	GTCGACTCCATAGTGTATCCTTC	Site directed mutagenesis	57
PzntA(-17G<T)_R	TCCAGAGTCAAGTTTTATC		
PzntA(-16A<C)_F	TCGACTCCAGCGTGTATCCTTC	Site directed mutagenesis	58
PzntA(-16A<C)_R	CTCCAGAGTCAAGTTTTATC		
PzntA(-15G<T)_F	CGACTCCAGATTGTATCCTTC	Site directed mutagenesis	59
PzntA(-15G<T)_R	ACTCCAGAGTCAAGTTTTATC		
PzntA(-14T<G)_F	GACTCCAGAGGGTATCCTTCG	Site directed mutagenesis	61
PzntA(-14T<G)_R	GACTCCAGAGTCAAGTTTTATC		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
pDOC_ZntA:rfp_Lin_F	GGCAAACCGATCGCAACATTGAGCGCGATCGGTCCCCTCGGATCCCGGGTACCCACAG	Linearization of pDOC-C for fusion plasmid pDOC-ZntA:rfp	65
pDOC_ZntA:rfp_Lin_R	TCTCTGCGCAACAATCTTAACGCATTCGCTGTCACCAGCCAAGCTTGAATTCATCGATTACCC		
pDOC_ZntA:rfp_KanR_F	CGGTGCTTAAACCGGTCAATTGGCTGG	Amplification of <i>kan<sup>R</sup></i> for fusion plasmid pDOC-ZntA:rfp	72
pDOC_ZntA:rfp_KanR_F	CGAGGGGACCGATCGCGCTCAATGTTGCGATCGGTTTGCCACTAGTCGACGCTAGCATATG		
pDOC_ZntA:rfp_RFP_F	GCTGGTGACAGCGAATGCGTTAAGATTGTTGCGCAGGAGAATGGCAAGTAGTGAAGACG	Amplification of <i>mRFP1</i> for fusion plasmid pDOC-ZntA:rfp	63
pDOC_ZntA:rfp_RFP_R	TTGACCGGTTTTAAGCACCGGTGGAGTG		
ZntA:RFP_Scan_F	AATATCCGCCAGAACATCACTATTG	Confirmation of chromosomal fusion of <i>zntA:rfp:kan<sup>R</sup></i>	64
ZntA:RFP_Scan_R	GACTGCCTTATCGTTATTTGATTCGTAAAG		
zntA_qPCR_F	ACCGCCGACGCAGCATTAAAC	RT-qPCR analysis of <i>zntA</i>	N/A
zntA_qPCR_R	GAGGAAGATCCCTTTCAGCCCCAG		

Primer	Sequence 5'-3'	Use	T <sub>m</sub> x°C
ZntA_qPCR_Probe	[FAM]-AACTGGCACGCGCCACTCACGCCAATA-[BHQ1]		
mRFP1_qPCR_F	ACCGCTAAACTGAAAGTTACC	RT-qPCR analysis of <i>mRFP1</i>	N/A
mRFP1_qPCR_R	ACACGTTCCCATTTGAAACC		
mRFP1_qPCR_Probe	[FAM]-CCGCTGCCGTTGCTTGGGACAT-[BHQ1]		
rrsA_qPCR_F	TTACTGGGCGTAAAGCGCAC	RT-qPCR reference gene <i>rrsA</i> (control)	N/A
rrsA_qPCR_R	TCTACGCATTTACCGCTACAC		
rrsA_qPCR_Probe	[FAM]-CCGGGCTCAACCTGGGAACTGCAT-[BHQ1]		

**Table 2.3 Table of Primers**

#### 2.1.4. Media

Distilled H<sub>2</sub>O (dH<sub>2</sub>O) was used, unless stated otherwise. Media was dispensed into suitable sized containers (DURAN® GL 45) and sterilized by autoclaving at 121°C, 15 psi for 20 min. All chemicals were ultrapure grade and supplied by Sigma Aldrich (Merck), unless otherwise stated. Media components were supplied by Oxoid, unless otherwise stated.

##### 2.1.4.1. Lysogeny Broth (LB), According to Lennox

10 g tryptone, 5 g yeast extract, and 5 g NaCl were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving (Bertani, 1951).

##### 2.1.4.2. Lysogeny Broth with Agar (LB Agar), According to Lennox

15 g bacteriological agar was added to 1 L LB broth, prior to autoclaving.

##### 2.1.4.3. Lysogeny Broth Double Strength (LB 2x), According to Lennox

20 g tryptone, 10 g yeast extract, and 10 g NaCl were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

##### 2.1.4.4. Nutrient Broth (NB)

1 g Lab Lemco powder, 2 g yeast extract, 5 g peptone, and 5 g NaCl were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

##### 2.1.4.5. Nutrient Broth with Agar (NA)

15 g bacteriological agar was added to 1 L NB broth, prior to autoclaving.



#### 2.1.4.6. Nutrient Broth Number 2 (NB N°2)

10 g Lab Lemco powder, 10 g peptone, and 5 g NaCl were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

#### 2.1.4.7. Buffered Peptone Water (BPW)

10 g peptone, 5 g NaCl, 3.5 g Na<sub>2</sub>HPO<sub>4</sub>, and 1.5 g KH<sub>2</sub>PO<sub>4</sub> were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

#### 2.1.4.8. Mueller-Hinton (MH)

2 g beef extract, 17.5 g casein hydrolysate, and 1.5 g starch were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

#### 2.1.4.9. M17 Broth

5 g tryptone, 5 g soya peptone, 5 g Lab Lemco powder, 2.5 g yeast extract, 0.5 g ascorbic acid, 0.35 g MgSO<sub>4</sub>, 19 g di-sodium-glycerophosphate were dissolved in dH<sub>2</sub>O, and made up to final volume of 950 mL, prior to autoclaving. 50 mL lactose 10% (w/v) was added post autoclave, once cooled to ≤ 50°C.

#### 2.1.4.10. Tryptone Soya Broth (TSB)

17 g pancreatic digest of casein, 3 g enzymatic digest of soya bean, 5 g NaCl, 2.5 g K<sub>2</sub>HPO<sub>4</sub>, and 2.5 g glucose were dissolved in dH<sub>2</sub>O, and made up to final volume of 1 L, prior to autoclaving.

#### 2.1.4.11. Terrific Broth (TB)

24 g yeast extract, 20 g tryptone, and 4 mL glycerol were dissolved in dH<sub>2</sub>O, and made up to final volume of 900 mL, prior to autoclaving. 100 mL TB phosphate buffer was added post autoclave, once cooled to  $\leq 50^{\circ}\text{C}$ .

#### 2.1.4.12. Super Optimal Broth (SOB)

20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 20 mL KCl (250 mM) was dissolved in 950 mL dH<sub>2</sub>O, and the pH was adjusted to 7.0 with NaOH (5 M). The volume was adjusted to 1 L prior to autoclaved. 5 mL of sterile MgCl<sub>2</sub> (2 M) was added post autoclave, before use.

#### 2.1.4.13. Super Optimal Broth with Catabolite Repressor (SOC)

Prior to use, 20 mL of sterile glucose (1 M) was added to each litre of SOB.

#### 2.1.4.14. Zinc Depleted Neidhardt's MOPS Minimal Media (NH)

Zinc depleted Neidhardt's MOPS minimal media (NH) was used as a zinc depleted grown medium. NH media made up using ultra-pure HPLC grade H<sub>2</sub>O (Fisher, 10449380), Table 2.4 shows the final concentration of NH media.

Component	Concentration (mM)
MOPS (3-(N-Morpholino)propanesulfonic acid)	40
Tricine (N-[Tris(hydroxymethyl)methyl]glycine)	4.00
Iron sulphate ( $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ )	0.01
Ammonium chloride ( $\text{NH}_4\text{Cl}$ )	9.5
Potassium sulphate ( $\text{K}_2\text{SO}_4$ )	0.276
Calcium chloride ( $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ )	$5 \times 10^{-4}$
Magnesium chloride ( $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ )	0.525
Sodium chloride ( $\text{NaCl}$ )	50.0
Potassium phosphate ( $\text{K}_2\text{HPO}_4$ )	1.32
D-(+)-Glucose ( $\text{C}_6\text{H}_{12}\text{O}_6$ )	22.21
Ammonium molybdate ( $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ )	$7.3 \times 10^{-9}$
Cobalt chloride ( $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ )	$7.6 \times 10^{-8}$
Cupric sulphate ( $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ )	$2.4 \times 10^{-8}$
Manganese chloride ( $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ )	$2.0 \times 10^{-7}$
MEM amino acids solution (50x) liquid [without L-Glutamine, without L-glutamine] (Invitrogen Gibco, 11570386)	1x

**Table 2.4 Neidhardt's MOPS Minimal Media: Molar Composition**

### 2.1.5. Antibiotic Stock

#### 2.1.5.1. Kanamycin $50 \text{ mg mL}^{-1}$ (1000x)

500 mg kanamycin sulphate (Kan) (Gibco, 11578676) was dissolved in 10 mL Milli-Q  $\text{H}_2\text{O}$ , filter sterilized through a  $0.22 \mu\text{m}$  filter (Sartorius) and stored in 1 mL aliquots at  $-20^\circ\text{C}$  in sterile 1.5 mL microcentrifuge tube.

#### 2.1.5.2. Ampicillin 100 mg mL<sup>-1</sup> (1000x)

1 g ampicillin sodium salt (Amp) (Sigma-Aldrich, A0166) was dissolved in 10 mL Milli-Q H<sub>2</sub>O, filter sterilized through a 0.22 µm filter, and stored in 1 mL aliquots at -20 °C in sterile 1.5 mL microcentrifuge tube.

#### 2.1.5.3. Chloramphenicol 35 mg mL<sup>-1</sup> (1000x)

350 mg chloramphenicol (Chlor) (Sigma-Aldrich, C0378) was dissolved in 10 mL 95% ethanol, and stored in 1 mL aliquots at -20 °C in sterile 1.5 mL microcentrifuge tube.

#### 2.1.6. Metal Solutions

Filter sterilization was conducted using a 0.22 µm PES filter (Sartorius), unless otherwise stated. All chemicals were ultrapure grade and supplied by Sigma-Aldrich (Merck), unless stated otherwise. Metal solutions were stored in sterile 50 mL virgin Polypropylene (PP) centrifuge tubes (Sarstedt, Nümbrecht, Germany) at room temperature.

##### 2.1.6.1. 100 mM ZnSO<sub>4</sub>·7H<sub>2</sub>O

2.875 g ZnSO<sub>4</sub>·7H<sub>2</sub>O (99.995% trace metals basis) (Merck, 204986) was dissolved in HPLC grade H<sub>2</sub>O, made up to a final volume of 100 mL, and filter sterilized.

##### 2.1.6.2. 50 mM AgNO<sub>3</sub>

849.35 mg AgNO<sub>3</sub> was dissolved in HPLC grade H<sub>2</sub>O, made up to a final volume of 100 mL, and filter sterilized. Stored in the dark.

2.1.6.3. 500 mM  $\text{CuSO}_4 \cdot 7\text{H}_2\text{O}$

7.98 g  $\text{CuSO}_4 \cdot 7\text{H}_2\text{O}$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

2.1.6.4. 400 mM  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$

6.81 g  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

2.1.6.5. 10 mM  $\text{HgCl}_2$

271.5 mg  $\text{HgCl}_2$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

2.1.6.6. 10 mM  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$

237.9 mg  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

2.1.6.7. 10 mM  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$

197.9 mg  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

2.1.6.8. 100 mM  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$

2.78 g  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  was dissolved in HPLC grade  $\text{H}_2\text{O}$ , made up to a final volume of 100 mL, and filter sterilized.

### 2.1.7. Solutions

#### 2.1.7.1. 10 M NaOH

40 g NaOH pellets were dissolved in dH<sub>2</sub>O and made up to a final volume of 100 mL.

#### 2.1.7.2. 30 % D-Glucose (C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) (% W/V)

150 g glucose (Sigma-Aldrich, G7528) was dissolved in dH<sub>2</sub>O, made up to a final volume of 500 mL, and filter sterilized through a 0.22 µm filtration unit (Nalgene).

#### 2.1.7.3. 30 % Sucrose (C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>) (% W/V)

150 g sucrose (Sigma Aldrich, C0389) was dissolved in dH<sub>2</sub>O, made up to a final volume of 500 mL, and filter sterilized through a 0.22 µm filtration unit (Nalgene).

#### 2.1.7.4. 10 % Lactose (C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>) (% W/V)

10 g lactose was dissolved in dH<sub>2</sub>O and made up to a final volume of 100 mL, prior to autoclaving.

#### 2.1.7.5. 10% Arabinose (C<sub>5</sub>H<sub>10</sub>O<sub>5</sub>) (% W/V)

1 g L-Arabinose (Sigma Aldrich, 10839) was dissolved in dH<sub>2</sub>O, made up to a final volume of 10 mL in dH<sub>2</sub>O, and filter sterilized. Stored at room temperature.

#### 2.1.7.6. 1 M IPTG (Isopropyl β-D-1-thiogalactopyranoside)

2.38 g IPTG (Sigma Aldrich, I6758) was dissolved in 10 mL HPLC grade H<sub>2</sub>O, filter sterilized, and stored at -20°C for long term storage.

**2.1.7.7. 10% Glycerol (% V/V)**

12.6 g (10 mL) glycerol (Sigma Alrich, G5516) was added to 90 mL dH<sub>2</sub>O, and autoclaved.

**2.1.7.8. 50% Glycerol (% V/V)**

63 g (50 mL) glycerol was added to 50 mL dH<sub>2</sub>O, and autoclaved.

**2.1.7.9. 20 mM TEPN (1000x)**

426 mg TPEN [N,N,N',N' tetrakis(2 pyridinylmethyl) 1,2 ethanediamine] (Alfa Aesar, 15460347) was dissolved in Ethanol and made up to a final volume of 50 mL. Stored at 4°C for long term storage.

**2.1.7.10. 4% Formaldehyde in PBS pH 6.9 (% W/V)**

4% formaldehyde was purchased, premade, for safety (Merk, 1004968350) and stored at 4°C.

**2.1.7.11. 2.22% Nitric acid**

32.6 mL 68% nitric acid (Fisher, 10050270) was combined with dH<sub>2</sub>O, and made up to a final volume of 1 L.

**2.1.8. Buffers**

**2.1.8.1. 50x TAE Buffer**

242 g tris base, 57.1 mL acetic acid (glacial), and 37.2 g Na<sub>2</sub>EDTA·2H<sub>2</sub>O was dissolved in dH<sub>2</sub>O, adjust to pH 8.5, and made up to a final volume of 1 L (Ausubel *et al.*, 2003).

#### 2.1.8.2. 1x TAE Buffer

100 mL 50x TAE buffer was combined with 4900 mL dH<sub>2</sub>O.

#### 2.1.8.3. 1 M Tris-HCl pH 8.0

121.14 g tris was dissolved in 800 mL dH<sub>2</sub>O, adjusted to pH 8.0 with HCl, and made up to a final volume of 1 L, prior to autoclaving.

#### 2.1.8.4. 0.5M EDTA pH 8.0

186.1 g EDTA was dissolved in 800 mL dH<sub>2</sub>O, adjusted to pH 8.0 with NaOH pellets, and made up to a final volume of 1 L, prior to autoclaving.

#### 2.1.8.5. Tris-EDTA (TE) Buffer pH 8.0

10 mL 1 M Tris-HCl pH 8.0, 2 mL 0.5 M EDTA pH 8.0, 982 mL dH<sub>2</sub>O was combined, and autoclaved.

#### 2.1.8.6. Phosphate Buffered Saline (PBS)

8 g NaCl, 0.2 g KCl, 1.44 g Na<sub>2</sub>HPO<sub>4</sub>, and 0.24 g KH<sub>2</sub>PO<sub>4</sub> was dissolved in 800 mL Milli-Q water, and pH adjusted to pH 7.4. The final volume was adjusted to 1 L prior to autoclaving.

#### 2.1.8.7. TB Phosphate Buffer (0.17 M KH<sub>2</sub>PO<sub>4</sub>, 0.72M K<sub>2</sub>HPO<sub>4</sub>)

23.1 g KH<sub>2</sub>PO<sub>4</sub>, and 125.4 g K<sub>2</sub>HPO<sub>4</sub> was dissolved in dH<sub>2</sub>O, and made up to a final volume of 1 L, prior to autoclaving.

#### 2.1.8.8. 1 M HEPES Buffer (pH 7.0)

0.26 g HEPES (free acid) was dissolved in 90 mL of dH<sub>2</sub>O and pH adjusted to pH 7.0 with 10 M NaOH. The final volume was adjusted to 100 mL, and autoclaved.



### 2.1.9. Enzymes

#### 2.1.9.1. 10 mg mL<sup>-1</sup> Lysozyme

500 mg lysozyme (Merck, L3790) was dissolved in 50 mL dH<sub>2</sub>O, filter sterilized, aliquoted, and stored at -20°C.

#### 2.1.9.2. Restriction Enzymes

All restriction enzymes, unless stated otherwise, were provided by New England Biolabs (NEB, Massachusetts USA).

## 2.2. Methods

### 2.2.1. *Escherichia coli* Propagation and Maintenance

#### 2.2.1.1. Standard Growth Conditions

*Escherichia coli* was routinely grown in LB broth, with appropriate antibiotic(s), at 37°C with shaking (200 RPM) for 16-20 hr. If a temperature sensitive plasmid was present, the growth temperature was adjusted accordingly.

*E. coli* was routinely grown on LB agar, with appropriate antibiotic(s), at 37°C for 16-20 hr. If a temperature sensitive plasmid was present, the growth temperature, as well as incubation time, was adjusted accordingly. e.g. 30°C for 40-48 hr.

Final concentrations of antibiotics used were as follows: 50 µg mL<sup>-1</sup> kanamycin (<sub>kan50</sub>); 100 µg mL<sup>-1</sup> ampicillin (<sub>amp100</sub>); and 35 µg mL<sup>-1</sup> chloramphenicol (<sub>chlor35</sub>).

#### 2.2.1.2. Preparation of Electrocompetent Cells

The desired strain was inoculated for single colonies on LB agar, and incubated at 37°C for 12-16 hr. A single colony was inoculated into 10 mL of LB broth and incubated for 12-16 hr at 37°C with shaking (200 RPM). 2 mL of overnight culture was then inoculated into 200 mL of LB broth and incubated at 37°C with shaking, until an OD<sub>600</sub> of 0.4-0.6 was reached. The culture flask was placed on ice for 20 min, and cells then transferred to a pre-chilled 250 mL centrifuge bottles (Nalgene, 3141) and centrifuged at 6,000 x g for 10 min at 4°C (Beckman J2-21). The cell pellet was re-suspended in 100 mL of 10% glycerol, which was at a temperature of 4°C, and centrifuged again at 6,000 x g for 10 min at 4°C. The cell pellet was re-suspended in 100 mL 10% glycerol, which was at a temperature of 4°C, and centrifuged again, 6,000 x g for 10 min. The

cell pellet was re-suspended into 960  $\mu\text{L}$  10% glycerol, and 40  $\mu\text{L}$  aliquoted into 1.5 mL microcentrifuge tubes, which were flash frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ .

### 2.2.1.3. Electroporation Transformation

Either 40-60 ng of plasmid DNA, 5  $\mu\text{L}$  of ligation mixture, or 3  $\mu\text{L}$  of NEBuilder HiFi assembled product was added to 40  $\mu\text{L}$  of thawed electrocompetent cells and then transferred to a pre-chilled 2 mm electrode gap electroporation cuvette (Flowgen Bioscience, FBR-102). The transformation mixture was electroporated in an electroporator (Eppendorf 2510) at 2.5 V with a time constant 4 ms. Immediately after electroporation, 960  $\mu\text{L}$  of SOC media was added and cells were incubated at  $37^{\circ}\text{C}$  for 1 hr with shaking (200 RPM). If a temperature sensitive plasmid was transformed, cells were incubated at  $30^{\circ}\text{C}$  for 2 hr. 1000  $\mu\text{L}$  of transformation culture was plated onto LB agar containing appropriate antibiotics, with dilution ranging from  $10^0$  to  $10^{-4}$ , and incubated at  $37^{\circ}\text{C}$  for 12-16 hr or  $30^{\circ}\text{C}$  for 24-36 hr. Transformants were confirmed by PCR and sequence analysis (Eurofins Genomics).

### 2.2.1.4. Chemical Transformation

5  $\mu\text{L}$  of site directed mutation (SDM) assembled product, 5  $\mu\text{L}$  of ligation product, or 40-60 ng of plasmid DNA was added to 50  $\mu\text{L}$  of thawed NEB 5-alpha chemically competent cells (NEB, C2987), and placed on ice for 30 min. The cells were heat shocked at  $42^{\circ}\text{C}$  for 30 s and placed back on ice for 5 min; 950  $\mu\text{L}$  of room temperature SOC media was added, and incubated at  $37^{\circ}\text{C}$  for 1 hr with shaking (200 RPM). If a temperature sensitive plasmid was transformed, the cells were incubated at  $30^{\circ}\text{C}$  for 2 hr. Transformants were plated onto LB agar with appropriate antibiotics, with dilution ranging from  $10^0$  to  $10^{-4}$ , and incubated at  $37^{\circ}\text{C}$  for 12-16 hr or  $30^{\circ}\text{C}$  for

24-36 hr. Transformants were confirmed by PCR and sequence analysis (Eurofins Genomics).

#### 2.2.1.5. Long Term Storage in 25% Glycerol Stock

A single bacterial colony was inoculated into 5 mL of LB broth containing appropriate antibiotic(s) and incubated at 37°C with shaking (200 RPM) for 16-18 hr. 500 µL of the overnight culture was transferred into a 2 mL cryotube containing 500 µL of 50% glycerol, and the solutions mixed. The glycerol stock culture was stored at -80°C.

#### 2.2.1.6. Long Term Storage in Microbank™

A single bacterial colony was inoculated into 5 mL of LB broth containing appropriate antibiotics, and incubated at 37°C with shaking (200 RPM) for 16-18 hr. 500 µL of an overnight culture was transferred into a Microbank™ tube (Pro-Lab Diagnostics, Ontario Canada, PL.170), and incubated at room temperature for 2-3 min, according to the manufacturer's instructions. Excess liquid was then removed by aspiration, and the Microbank™ stock culture was stored at -80°C.

### 2.2.2. *Cupriavidus necator* Propagation and Maintenance

#### 2.2.2.1. Standard Growth Conditions

*Cupriavidus necator* was routinely grown in SOB media, containing appropriate antibiotic, at 30°C, with shaking (250 RPM) for 16-20 hr, in baffled flasks. *C. necator* was routinely grown on Nutrient Agar (NA) or LB agar, containing appropriate antibiotic, at 30°C for 24-48 hr.

Antibiotic final concentrations were used as follows; 200 µg mL<sup>-1</sup> kanamycin (kan<sub>200</sub>).

#### 2.2.2.2. Preparation of Electrocompetent Cells

*C. necator* was inoculated for single colonies on LB agar, and incubated at 30°C for 48 hr. 10 mL of SOB was inoculated with 4-5 colonies, serially diluted  $10^0$  to  $10^{-3}$  in SOB, and incubated at 30°C with shaking (250 RPM) for 16-18 hr. 2 mL of the culture with the lowest OD<sub>600</sub> (usually  $10^{-3}$ ) was inoculated into fresh 50 mL SOB, and incubated at 30°C with shaking (250 RPM) until the OD<sub>600</sub> reached 0.3. The culture flask was placed on ice for 20 min, and 2 mL of the *C. necator* culture was transferred into a pre-chilled, sterile 2 mL microcentrifuge tube and centrifuged at 6,000  $\times g$  for 5 min at 4°C (Sigma 1-16K). The cell pellet was re-suspended in 1 mL of sterile pre-chilled 1 mM HEPES buffer, and centrifuged at 6,000  $\times g$  for 5 min at 4°C. The cell pellet was again re-suspended in 1 mL of 1 mM HEPES, and was centrifuged at 6,000  $\times g$  for 5 min at 4°C. The cell pellet was re-suspended in 100  $\mu$ L of 1 mM HEPES buffer and was kept on ice until transformation. Electrocompetent *C. necator* was made fresh prior to transformation.

#### 2.2.2.3. Electroporation Transformation

400-500 ng of plasmid DNA was added to 100  $\mu$ L of electrocompetent *C. necator* cells and then transferred to a pre-chilled 2 mm electrode gap electroporation cuvette (Flowgen Bioscience, FBR-102). The transformation mixture was incubated on ice for 5 min. Cells were electroporated in an electroporator (Eppendorf 2510) at 2.5 V with a time constant of 4 ms. Immediately after electroporation, 900  $\mu$ L of SOC media was added and the cells were incubated at 30°C for 2 hr with shaking (250 RPM). 1000  $\mu$ L of transformation culture was plated onto NA<sub>kan50</sub> with the dilution range  $10^0$  to  $10^{-1}$ , and incubated at 30°C for 40-48 hr. A single colony of the transformant was inoculated onto NA<sub>kan200</sub>, and incubated at 30°C for 40-48 hr. Transformants were confirmed by PCR and sequence analysis (Eurofins Genomics).

#### 2.2.2.4. Long Term Storage in Microbank™

A single bacterial colony was inoculated into 5 mL of SOB broth and incubated at 30°C for 16-20 hr with shaking (250 RPM), overnight. 500 µL of an overnight culture was transferred into the Microbank™ tube (Pro-Lab Diagnostics) and incubated at room temperature for 2-3 min, according to manufacturer's instructions. Liquid was aspirated off, and the Microbank™ was stored at -80°C.

#### 2.2.3. DNA Manipulation and Purification

All DNA work was conducted using filter pipette tips (Starlab, TipOne® Filter Tips). Gloves were always worn to reduce DNase and RNase contamination of sample.

##### 2.2.3.1. Plasmid Purification

A single colony of *E. coli* containing the desired plasmid was inoculated into 10 mL of LB broth containing appropriate antibiotic(s) and incubated at 37°C for 12-16 hr with shaking (200 RPM). Growth conditions of 30°C were used if a temperature sensitive plasmid was present. 5 mL of the culture was centrifuged at 4,200 x *g* for 5 min, and supernatant discarded. Plasmid DNA was purified using a Monarch® Plasmid Mini Prep Kit (NEB, T1010) in accordance with manufacturer's instructions and eluted into 30 µL elution buffer (pre warmed at 55°C).

##### 2.2.3.2. PCR and DNA Purification

15 µL of a PCR product or 50 µL of a R.E digested product was purified using a Monarch® PCR and DNA clean up kit (NEB, T1030), in accordance with manufacturer's instructions, and eluted into 6 µL elution buffer (pre warmed at 55°C).

#### 2.2.3.3. Quantification of DNA

DNA concentration of a 1  $\mu\text{L}$  volume of a DNA sample was determined using a Nano Drop® 1000 spectrophotometer (Thermo Fisher), in accordance with manufacturer's instructions. Elution buffer (NEB, T1016) was used as the blank.

#### 2.2.3.4. DNA Sequencing

Plasmid DNA was diluted to a concentration of 75  $\text{ng } \mu\text{L}^{-1}$ , to a volume of 15  $\mu\text{L}$ , with HPLC grade  $\text{H}_2\text{O}$ . Purified PCR product was diluted to a concentration of 1-10  $\text{ng } \mu\text{L}^{-1}$ , to a volume of 15  $\mu\text{L}$ , with HPLC grade  $\text{H}_2\text{O}$ . 15  $\mu\text{L}$  of DNA was mixed with 2  $\mu\text{L}$  of the appropriate 10  $\mu\text{M}$  sequencing primer. The sample was sent for Sanger sequencing (TubeSeq Service, Eurofins Genomics). The returned DNA sequence data files were analysed using ClustalW (Kyoto University Bioinformatics Center, 2019), and the chromatogram analysed using SnapGene® viewer (GSL Biotech, 2019).

#### 2.2.3.5. Agarose Electrophoresis

A 1% agarose gel was made by adding 0.5 g of agarose (Sigma-Aldrich, A9539) to 50 mL of 1x TAE buffer, which was then dissolved by microwave heating. The agarose solution was left to cool to less than 50  $^{\circ}\text{C}$ , ethidium bromide was added to a final concentration of 0.2  $\mu\text{g mL}^{-1}$  and the gel poured into a casting tray. After setting, 5  $\mu\text{L}$  samples of the PCR product, or other DNA, was combined with 1  $\mu\text{L}$  of 6x loading dye (NEB, B7024) and loaded into the agarose gel well. Either 5  $\mu\text{L}$  of 100 bp ladder (NEB, N0551) or 1 kbp ladder (NEB, N0552) was loaded as a reference. The agarose gel was electrophoresed at 80 V for 40-60 min. DNA was visualized on a UV trans illuminator (BIO-RAD Gel Doc XR+).

#### 2.2.3.6. Bacterial gDNA Extraction, Heat Method

A single colony was inoculated into 50  $\mu$ L TE buffer, heated to 100°C for 10 min, and centrifuged at 13,000  $\times g$  for 1 min (Sigma 1-16K). The supernatant was aspirated and stored at -20°C.

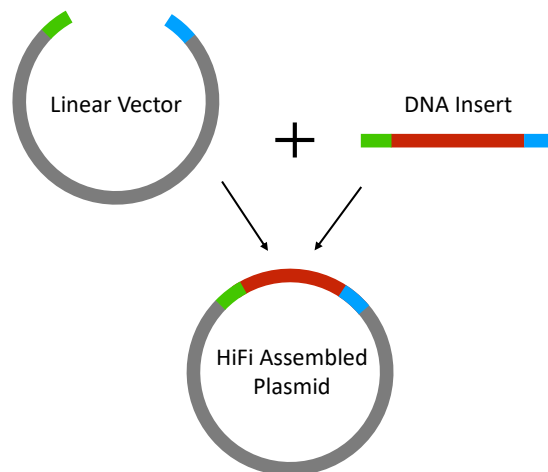
#### 2.2.3.7. Monarch® Genomic DNA Purification Kit Protocol

A single colony of the desired bacteria was inoculated into 5 mL of LB broth containing appropriate antibiotic(s) and incubated at 37°C for 12-16 hr with shaking (200 RPM). 1.5 mL of the overnight culture was pipetted into a 1.5 mL microcentrifuge tube (Eppendorf), centrifuged at 12,000  $\times g$  for 1 min (Sigma 1-16K), and supernatant removed. A Monarch® Genomic DNA Purification Kit (NEB, T3010) was used to extract gDNA, according to manufacturer's instructions and eluted in a final volume of 100  $\mu$ L with gDNA elution buffer.

#### 2.2.3.8. NEBuilder® HiFi DNA Assembly

If the insert that was to be cloned was <200 bp in size: 5  $\mu$ L of NEBuilder® HiFi DNA Assembly Master Mix 2x (NEB, E2621); 0.012 pmol ends of linearized plasmid vector; and 0.06 pmol ends of the insert were combined and made up to 20  $\mu$ L with sterile dH<sub>2</sub>O. If the insert was greater than >200 bp in size: 5  $\mu$ L of NEBuilder® HiFi DNA Assembly Master Mix 2x; 0.016 pmol ends of linearized plasmid vector; and 0.032 pmol ends of the insert were combined and made up to 20  $\mu$ L with HPLC grade H<sub>2</sub>O. The sample was mixed in a 0.2 mL PCR tube and placed in a thermocycler (BIO-RAD c1000) and heated to 50°C for 20 min (Figure 2.1). 3  $\mu$ L of the HiFi DNA assembled product was transformed into *E. coli* using the standard electrocompetent cell transformation protocol (2.2.1.3).





**Figure 2.1 Diagram of NEBuilder HiFi DNA Assembly**

Diagram representing NEBuilder HiFi DNA assembly with a single insert.

#### 2.2.3.9. Standard Restriction Enzyme Digest

All restriction enzymes (R.E) were supplied by New England Biolabs (NEB, Massachusetts, USA). 10 units of restriction enzyme, 1 µg of plasmid DNA, and 5 µL of 'Cut Smart' buffer 10x were combined and made up to 50 µL with HPLC grade H<sub>2</sub>O. Restriction enzyme digests were incubated at 37°C for 2-3 hr. For cloning work, 3 µL of Quick CIP (Calf Intestinal Phosphatase) (NEB, M0525) was added and incubated for a further 20 min. Quick CIP was inactivated by heating to 80°C for 2 min. Restriction enzyme inactivation was conducted either by heating the restriction enzyme digest tube at 65°C for 20 min, or by the addition of 50 µL of 6x loading dye (NEB, B7024) to the restriction digest. DNA was purified using a Monarch PCR & DNA clean up kit (NEB, T1030) (2.2.3.2), for downstream applications.

#### 2.2.3.10. Standard Ligation Protocol

2 µL of T4 DNA ligase buffer (10x), 50 ng of purified R.E digested vector, a 1:3 molar ratio of R.E digested insert, and 1 µL of T4 ligase (NEB, M0202) were combined and made up to 20 µL with HPLC grade H<sub>2</sub>O. The ligation mixture was incubated at 16°C for

14-16 hr and inactivated at 65°C for 10 min. 5 µL of the ligation mixture was used in the standard electrocompetent transformation protocol (2.2.1.3).

### 2.2.4. RNA Purification and Quantification

All RNA work was conducted using dedicated filter pipette tips (Starlab, TipOne® Filter Tips), and dedicated RNA only pipettes. Double gloves were worn to reduce RNase contamination of samples, with regular changing of gloves. Where possible, RNA work was conducted in biological safety cabinet, with UV treatment prior to use. Where not possible, RNA work was conducted in an RNA dedicated area, treated with RNaseZAP™ (Merk, R2020), with regular RNaseZAP™ treatment throughout the work, whilst also maintaining good aseptic procedures.

#### 2.2.4.1. RNA Extraction

RNA was extracted using a Monarch® Total RNA Miniprep Kit (NEB, T2010), according to manufacturer's instructions, using enzymatic lysis of cells (lysozyme), and elution from the column, using 50 µL nuclease free H<sub>2</sub>O. Enzymatic lysis was performed by re-suspending the bacterial pellet in 200 µL Lysozyme solution (1 mg mL<sup>-1</sup>).

#### 2.2.4.2. RNA Quantification

RNA was quantified using a Qubit 3 Fluorometer (Invitrogen), using the Qubit RNA BR (broad range) assay kit (Invitrogen, Q10210) according to manufacturer's instructions.

### 2.2.5. Polymerase Chain Reaction (PCR)

All PCR work was conducted using filter pipette tips (Starlab, TipOne® Filter Tips). Gloves were always worn to reduce DNase and RNase contamination of sample. Note, hot start polymerase was used, unless otherwise stated, thus PCR reaction was not set up on ice.

### 2.2.5.1. Standard PCR using Q5® Hot Start High-Fidelity 2x Master Mix

10  $\mu$ L of Q5® hot start high-fidelity 2x master mix (NEB, M0494), 1  $\mu$ L of forward primer (10  $\mu$ M), 1  $\mu$ L of reverse primer (10  $\mu$ M), and  $x$   $\mu$ L of template DNA (0.1-2 ng of plasmid DNA, or 50-100 ng gDNA), were added to a PCR reaction mixture, which was made up to 20  $\mu$ L, with nuclease free H<sub>2</sub>O, in a 0.2 mL PCR tube. PCR was conducted in a thermocycler (BIO-RAD c1000), using the standard thermocycle conditions (Table 2.5 Q5 Thermocycle Conditions).

Temp °C	Time (s)	
98	30	x 35
98	10	
$x$ Tm	15	
72	20 s / 1 kbp	
72	120	
4	hold	

**Table 2.5 Q5 Thermocycle Conditions**

### 2.2.5.2. Standard PCR using DreamTaq Hot Start PCR 2x Master Mix

10  $\mu$ L of DreamTaq hot start PCR master mix (2x) (Thermo Scientific, K9011), 1  $\mu$ L of forward primer (10  $\mu$ M), 1  $\mu$ L of reverse primer (10  $\mu$ M), and  $x$   $\mu$ L of template DNA (0.1-2 ng of plasmid DNA, 50-100 ng gDNA) were added to a PCR reaction mixture, which was made up to 20  $\mu$ L with nuclease free H<sub>2</sub>O was in a 0.2 mL PCR tube. PCR was conducted in a thermocycler (BIO-RAD c1000), using the standard thermocycle condition (Table 2.6 DreamTaq Thermocycle Conditions).

Temp °C	Time (s)	
95	60	
95	30	x 35
x Tm	15	
72	60s / 1kbp	
72	300	
4	hold	

**Table 2.6 DreamTaq Thermocycle Conditions**

### 2.2.6. Nitric Acid Soaking of PMP Erlenmeyer Flasks

Polymethylpentene (PMP) Erlenmeyer flasks (Nalgene, 4109-0250) were treated with nitric acid to remove all possible zinc. Nitric acid reacts with  $\text{Zn}^{2+}$  to form  $\text{Zn}(\text{NO}_3)_2$ , a crystalline form of zinc. Erlenmeyer flasks were washed three times in ultra-pure Milli-Q water and left to soak overnight with Milli-Q water. Erlenmeyer flasks were washed again three times in Milli-Q water. Flasks were then left to soak for 24-48 hr in 6% (V/V) nitric acid. Flasks were once again washed three times in Milli-Q water and autoclaved.

### 2.2.7. Nitric Acid Soaking of Magnetic Followers

Magnetic followers were washed three times in ultra-pure Milli-Q water and left to soak overnight with Milli-Q water. Followers were washed again three times in Milli-Q water and left to soak for 24-48 hr in 6% (V/V) nitric acid. Followers were washed three times in Milli-Q water and autoclaved.

### 2.2.8. Zinc Depleted Neidhardt's MOPS Minimal Media (NH) Preparation

To further reduce zinc contamination, chemical solution used in NH was made up in a specific way; this worked out the mass of  $\text{H}_2\text{O}$  needed to be added to a chemical

solution, rather than making the solution up to  $x$  L. This allows solution to be made directly in virgin PMP plasticware and eliminated the need for a measuring cylinder and a volumetric.

The mass of the chemical required was calculated; this was then multiplied by the density of the chemical, which gave the volume of the chemical. This chemical volume was subtracted from the intended final solution volume. This gave the mass of  $H_2O$  needed to be added, knowing that 1 mL = 1 g of  $H_2O$ .

$$H_2O (g) = Final\ volume - (mass * density)$$

$$H_2O (g) = Final\ volume - (g * g\ cm^{-3})$$

#### 2.2.8.1. 10x MOPS Stock – Chelex Treated

The following chemicals (Table 2.7 10x MOPS Composition) were combined in a nitric acid, treated 1 L Erlenmeyer flask and adjusted to pH 7.2, with 10 M NaOH.

Chemical	Grams (g)
MOPS	83.7
Tricine	7.17
NH <sub>4</sub> Cl Ammonium chloride	5.081
K <sub>2</sub> SO <sub>4</sub> Potassium sulphate	0.481
NaCl Sodium chloride	29.22
HPLC grade H <sub>2</sub> O	913.67

**Table 2.7 10x MOPS Composition**

50 g of Chelex® Resin 100 mesh (Sigma-Aldrich, C7901) was added and stirred at room temperature for 1 hr, using a nitric acid soaked magnetic follower. The 10x MOPS

solution was filter sterilized, using a 0.22  $\mu\text{m}$  filtration unit (Nalgene, NC526). 50 mL aliquots of the sterile Chelex treated 10x MOPS solution were stored in 50 mL PP centrifuge tubes (SARSTEDT) and frozen at  $-20^{\circ}\text{C}$ .

### 2.2.8.2. 30 % Glucose (% W/V) – Chelex Treated

150 g of ultra-pure glucose (Sigma-Aldrich, G7528) was dissolved in 402.29 g of HPLC grade  $\text{H}_2\text{O}$ , in a 1 L nitric acid treated Erlenmeyer flask. 25 g of Chelex<sup>®</sup> Resin 100 mesh was added, stirred at room temperature for 1 hr using a nitric acid soak magnetic follower, filtered using a 0.22  $\mu\text{m}$  filtration unit and stored at  $4^{\circ}\text{C}$ .

### 2.2.8.3. 0.132 M $\text{K}_2\text{HPO}_4$ – Chelex Treated

11.5 g of  $\text{K}_2\text{HPO}_4$  was dissolved in 495.29 g of HPLC grade  $\text{H}_2\text{O}$ , in a 1 L nitric acid treated Erlenmeyer flask. 25 g of Chelex<sup>®</sup> Resin 100 mesh was added, stirred at room temperature for 1 hr using a nitric acid soak magnetic follower, filter sterilized using a 0.22  $\mu\text{m}$  filtration unit and stored at  $4^{\circ}\text{C}$ .

### 2.2.8.4. 0.01 M $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$

0.14 g of  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  was dissolved in 49.926 g of HPLC grade  $\text{H}_2\text{O}$ , in a 50 mL PP centrifuge tube and filter sterilized using a 0.22  $\mu\text{m}$  PES filter (Sartorius, 16532). Fresh  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  was made prior to assembly of NH.

### 2.2.8.5. 2.5 M $\text{MgCl}_2$

25.375 g of  $\text{MgCl}_2$  was dissolved in 33.83 g of HPLC grade  $\text{H}_2\text{O}$  in a 50 mL PP centrifuge tube, filter sterilized using a 0.22  $\mu\text{m}$  filter (Sartorius) into a new 50 mL PP centrifuge tube and stored at room temperature.

#### 2.2.8.6. 0.02M $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$

0.147 g of  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  was dissolved in 99.84 g of HPLC grade  $\text{H}_2\text{O}$  in a 50 mL PP centrifuge tube, filter sterilized using a 0.22  $\mu\text{m}$  filter into a new 50 mL PP centrifuge tube and stored at room temperature.

#### 2.2.8.7. Micronutrient stock

9 mg of  $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ , 18 mg of  $\text{CoCl}_2$ , 6 mg of  $\text{CuSO}_4$ , and 40 mg of  $\text{MnCl}_2$  were dissolved and made up to 50 mL in HPLC grade  $\text{H}_2\text{O}$  and stored in a 50 mL PP sterile centrifuge tube (Sarstedt), at room temperature.

#### 2.2.8.8. Neidhardt's MOPS Minimal Media (NH) Assembly

The following components (Table 2.8 Neidhardt's Media) were combined directly in the filtration chamber of a 500 mL 0.22  $\mu\text{m}$  filtration unit (Nalgene), filtered, and stored at 4°C for up to 1 month.

Component	Volume
10x MOPS	50 mL
30% Glucose	5 mL
0.01 M $\text{FeSO}_4$	500 $\mu\text{L}$
0.02 M $\text{CaCl}_2$	12.5 $\mu\text{L}$
2.5 M $\text{MgCl}_2$	105 $\mu\text{L}$
Micronutrient stock	10 $\mu\text{L}$
0.132 M $\text{K}_2\text{HPO}_4$	5 mL
50x MEM amino acids solution (Gibco)	10 mL
HPLC grade $\text{H}_2\text{O}$	423.94 mL

**Table 2.8 Neidhardt's Media Composition**

### 2.2.9. pJI100 Plasmid Series Construction

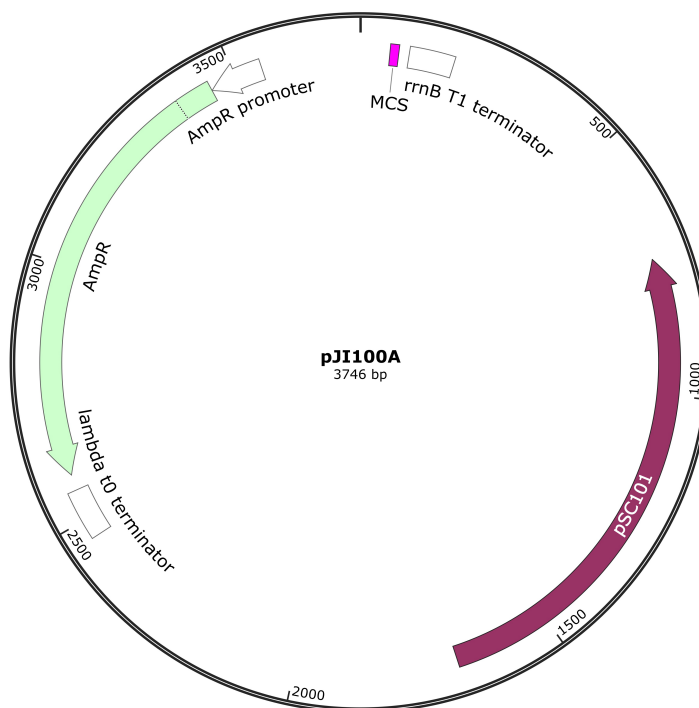
#### 2.2.9.1. pJI100 Plasmid Construction

pJI100, a low copy number expression plasmid was designed, using the backbone of pLUX (Burton *et al.*, 2010). A standard Q5<sup>®</sup> Hot Start Polymerase protocol (2.2.5.1) was conducted using primers pJI100\_F and pJI100\_R, with a T<sub>m</sub> of 72°C and extension time of 2 min, to amplify the *pSC101*, *rrnB*, and *kan<sup>R</sup>* elements of pLUX, whilst introducing a BamHI site at both 5' and 3' end. The PCR product was purified using a Monarch<sup>®</sup> PCR & DNA clean up kit (2.2.3.2). The purified PCR product was digested using BamHI-HF, using the standard restriction digest protocol (2.2.3.9), and purified once more, using a Monarch<sup>®</sup> PCR & DNA clean up kit. Self-ligation was conducted with T4 DNA ligase, and 0.006 pmol of the BamHI digested PCR product (2.2.3.10). 5 µL of the ligation mixture was transformed into *E. coli* MG1655, using standard electroporation protocol (2.2.1.3).



### 2.2.9.2. pJI100A plasmid construction

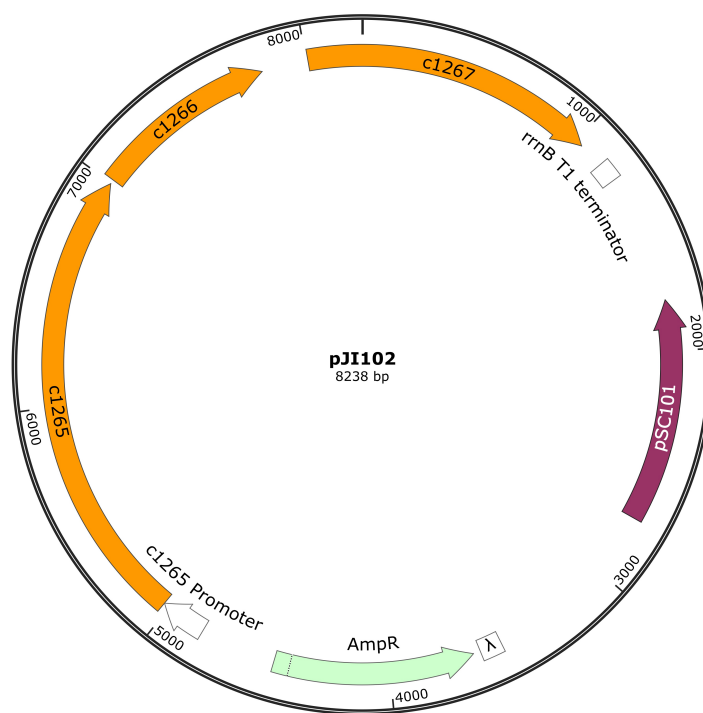
An ampicillin resistant alternative of pJI100 was created, pJI100A (Figure 2.2). A standard Q5<sup>®</sup> PCR protocol (2.2.5.1) was conducted to amplify pJI100 upstream and downstream of *kan<sup>R</sup>*, using primers pJI100A\_lin\_F and pJI100A\_lin\_R, with a *T<sub>m</sub>* of 69°C and an extension time of 2 min. *amp<sup>R</sup>* was amplified from pDOC-G (Lee *et al.*, 2009), using the primers pJI100A\_amp\_F and pJI100A\_amp\_R, with a *T<sub>m</sub>* of 66°C and an extension time of 40 s. Both PCR products were purified using a Monarch<sup>®</sup> PCR & DNA clean up kit (2.2.3.2). The standard NEBuilder<sup>®</sup> HiFi DNA assembly protocol (2.2.3.8) was used, with 0.016 pmol of linear pJI100 DNA, and 0.032 pmol of the *amp<sup>R</sup>* DNA insert. 1 µL of HiFi assembled product was transformed into *E. coli* MG1655, using standard electroporation protocol (2.2.1.3).



**Figure 2.2 Plasmid Map of pJI100A**

### 2.2.9.3. pJI102 – *c1265-7* Expression Plasmid

*c1265-7* was cloned into the pJI100A expression plasmid, creating pJI102 (Figure 2.3). *c1265-7* was amplified from *E. coli* CFT073 gDNA, extracted using a Monarch genomic extraction kit (2.2.3.7), with primers pJI102\_c12657\_F and pJI102\_c12657\_R, using a standard Q5® PCR protocol (2.2.5.1), with a  $T_m$  of 59°C and an extension of 2 min. pJI100A was linearized using the standard Q5® PCR protocol, with primers pJI102\_lin\_F and pJI102\_lin\_R, with a  $T_m$  of 72°C and an extension time of 90 s. The standard NEBuilder® HiFi DNA assembly protocol (2.2.3.8) was used, with 0.016 pmol of linear pJI100A and 0.032 pmol of *c1265-7* insert DNA. 1 µL HiFi assembled product was transformed into *E. coli* MG1655, *E. coli* MG1655  $\Delta znuA$ , and *E. coli* MG1655  $\Delta znuBC$ , using standard electroporation protocol (2.2.1.3).



**Figure 2.3 Plasmid Map of pJI102**

## 2.2.10. pJI300 Reporter Plasmid Construction

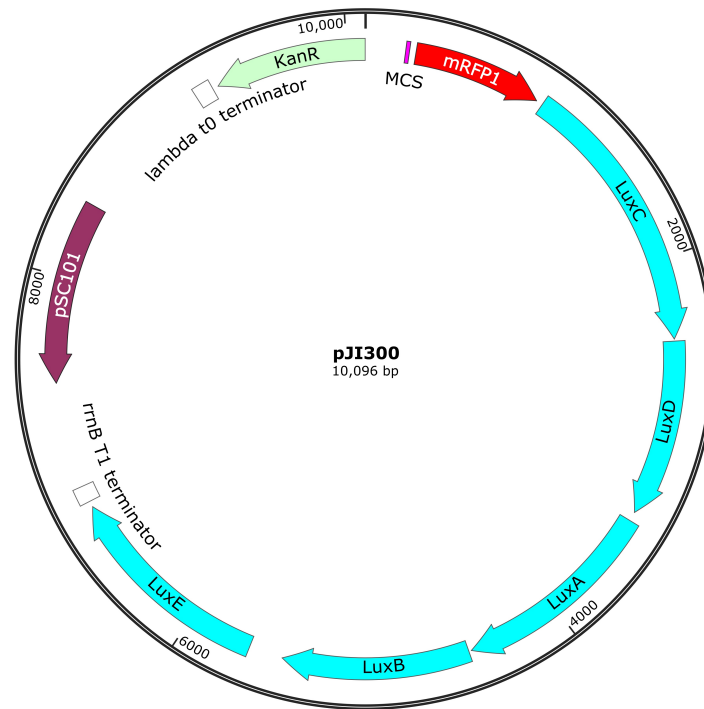


Figure 2.4 Plasmid Map of pJI300

## 2.2.10.1. Amplification and Linearization of pLUX

pLUX (Burton *et al.*, 2010) was linearized adjacent to the multiple cloning site (MCS). Linearized DNA included *luxCDABE*, *rrnB* T1, *pSC101*, and *kan<sup>R</sup>*. Linearization and amplification of pLUX was conducted using the standard Q5 polymerase PCR protocol (2.2.5.1), using primers pJI300\_Lin\_F and pJI300\_Lin\_R, with a  $T_m$  of 72 °C, and an extension time of 6:40 min. A standard restriction enzyme digest (2.2.3.9) was performed on the PCR product with *DpnI* and purified using a Monarch® PCR & DNA Cleanup Kit (2.2.3.2).

2.2.10.2. Amplification of *mRFP1*

*mRFP1* was amplified from pUltraRFP-KM (Mavridou *et al.*, 2016), using a standard Q5 polymerase PCR protocol (2.2.5.1), using the primers pJI300\_mRFP\_F and pJI300\_mRFP\_R, with a  $T_m$  of 63°C, and an extension time of 60 s. Primers were

designed to introduce a synthetic RBS (Ribosomal Binding Site) upstream of *mRFP1* [5'- AAGGAGTTTAAAAAACAAT-3']. The PCR product was purified using a Monarch® DNA & PCR Cleanup Kit (2.2.3.2).

### 2.2.10.3. pJI300 Plasmid HiFi Assembly and Transformation

Purified *mRFP1* DNA and the linearized pLUX were assembled using the standard NEBuilder® HiFi DNA Assembly (2.2.3.8), with 70 ng of linearized pLUX and 11 ng of *mRFP1*. 1 µL of the assembled product was transformed into *E. coli* MG1655, using the standard electroporation protocol (2.2.1.3).

### 2.2.11. pJI300 Plasmid Series Promoter Inserts

pJI300 (Figure 2.4) was digested with restriction enzymes BamHI-HF and XhoI, using the standard restriction enzyme digest protocol (2.2.3.9). Digested product was purified using a Monarch® DNA & PCR Cleanup Kit (2.2.3.2). DNA concentration was determined using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) (2.2.3.3).

Gene promoters were amplified to include roughly 100 bp upstream of the +1 site. Standard Q5® polymerase PCR protocol (2.2.5.1) was used, with an extension time of 20 s and  $T_m$ , found in Table 2.3 (Table of primers). PCR products were purified using the Monarch® PCR & DNA Cleanup Kit (2.2.3.2). DNA concentration was determined using a Nanodrop 1000 spectrophotometer (2.2.3.3).

Amplified promoter DNA was cloned into digested pJI300 plasmid, using the standard NEBuilder® HiFi Assembly protocol (2.2.3.8), with 0.012 pmol of R.E digested pJI300 and 0.06 pmol of insert DNA. Plasmids were transformed into *E. coli* MG1655, via the standard electroporation protocol (2.2.1.3) and plated onto LB<sub>kan50</sub>.

Promoter constructs were designated names pJI301-pJI308:  $P_{zntA}$ ;  $P_{trc}$ ;  $P_{znuA}$ ;  $P_{znuCB}$ ;  $P_{zinT}$ ;  $P_{pliG}$ ;  $P_{ykgM}$ ;  $P_{c1265}$  respectively. See Table 2.2 Table of Plasmids.

### 2.2.12. pJI300 Promoter Insert Confirmation via Colony PCR

10  $\mu$ L of Q5<sup>®</sup> hot start high-fidelity 2x master mix (NEB), 1  $\mu$ L of 204\_seq\_F primer (10  $\mu$ M), 1  $\mu$ L of rfp\_seq\_R primer (10  $\mu$ M), and 8  $\mu$ L of H<sub>2</sub>O was mixed in a PCR tube. A single colony of freshly transformed bacteria was selected using a sterile toothpick and inoculated into the PCR tube, containing the PCR reagent mix. PCR was conducted in a thermocycler (BIO-RAD c1000), using the standard thermocycle condition in Table 2.9 Q5 Colony PCR Thermocycle Conditions). The presence and size of the PCR product was confirmed through gel electrophoresis (2.2.3.5). An empty vector (pJI300) would produce a PCR product of 379 bp; a correct insert would produce a larger band, equivalent to the insert size (379 bp +  $x$  bp).

Temp °C	Time (s)	
98	30	x 35
98	10	
62	15	
72	20	
72	300	
12	hold	

**Table 2.9 Q5 Colony PCR Thermocycle Conditions**

### 2.2.13. $P_{zntA}$ (pJI301) Site Directed Mutagenesis

The Q5<sup>®</sup> Site-Directed Mutagenesis Kit (NEB, E0554) was used to incorporate specific mutations into  $P_{zntA}$  of pJI301. Primers were designed, using the NEBase Changer (NEB, 2018), to introduce specific mutations. 10  $\mu$ L of Q5<sup>®</sup> hot start high-fidelity 2x master

mix (NEB), 1  $\mu$ L of forward primer (10  $\mu$ M), 1  $\mu$ L of reverse primers (10  $\mu$ M), and 2 ng of pJI301 were combined in a 0.2 mL PCR tube, and made up to 20  $\mu$ L, with nuclease free H<sub>2</sub>O. Primers for SDM are found in Table 2.3. PCR was conducted in a thermocycler (BIO-RAD c1000), using the thermocycle conditions in Table 2.10 Q5 SDM Thermocycle Conditions).

Temp °C	Time (s)	
98	30	
98	10	x 25
x Tm	15	
72	480	
72	600	
4	hold	

**Table 2.10 Q5 SDM Thermocycle Conditions**

Kinase, Ligase & *DpnI* (KLD) treatment was conducted by mixing: 1  $\mu$ L of PCR product; 5  $\mu$ L of 2x KLD reaction buffer; 1  $\mu$ L of 10x KLD enzyme mix; and 2  $\mu$ L of dH<sub>2</sub>O in a 0.2 mL PCR tube, and incubated at room temperature or 5 min. 5  $\mu$ L of the KLD product was transformed into *E. coli* NEB 5-alpha chemical competent cells, using standard chemical competent cell transformation (2.2.1.4).

#### 2.2.14. Random Mutagenesis of P<sub>zntA</sub>

Error prone PCR was designed to incorporate random mutations into the P<sub>zntA</sub> in pJI301 (Ausubel *et al.*, 2003).

##### 2.2.14.1. Random Mutation of P<sub>zntA</sub> via Error Prone PCR

The solutions found in Table 2.11 Error Prone PCR Composition) were combined in a 0.2 mL PCR tube. PCR was conducted in a thermocycler (BIO-RAD c1000), using the

appropriate thermocycle conditions in Table 2.12 Error Prone PCR Thermocycle). The presence and size of the PCR product was confirmed via electrophoresis on a 1% agarose gel (2.2.3.5).

Component	Volume ( $\mu\text{L}$ )
Taq Buffer 10x (NEB)	10
$\text{MgCl}_2$ 50 mM	11
dCTP 25 mM	4
dTTP 25 mM	4
dATP 5 mM	4
dGTP 5 mM	4
pJI301(zntA)_XhoI_F 100 $\mu\text{M}$	2
pJI301(zntA)_BamHI_R 100 $\mu\text{M}$	2
pJI301 2 ng $\text{mL}^{-1}$	1
Hot Start Taq 5 U $\text{mL}^{-1}$ (NEB, M0495S)	1
Nuclease free $\text{H}_2\text{O}$	57

**Table 2.11 Error Prone PCR Composition**

Temp $^{\circ}\text{C}$	Time (s)	
95	60	
95	30	x 22
x Tm	60	
72	180	
72	300	
4	hold	

**Table 2.12 Error Prone PCR Thermocycle**

#### 2.2.14.2. Cloning of Error Prone PCR Insert

The PCR product was digested with restriction enzymes XhoI and BamHI-HF, using the standard restriction enzyme digest protocol (2.2.3.9). The digested insert was ligated with R.E digested pJI300 (BamHI-HF and XhoI), using the standard ligation protocol (2.2.3.10). 5  $\mu$ L of the ligation mixture was transformed into *E. coli* MG1655, using standard electroporation transformation protocol (2.2.1.3) and plated onto LB<sub>kan50</sub> + 200  $\mu$ M ZnSO<sub>4</sub> agar.

#### 2.2.14.3. Bioluminescence Optical Imaging of Error Prone PCR Clones

Transformants on LB<sub>kan50</sub> + 200  $\mu$ M ZnSO<sub>4</sub> were imaged for their bioluminescence. Plates were placed on the stage of an optical imager (Photon IMAGER – Biospace Lab) and an optical image taken. Bioluminescence was recorded via photon count, recording photon emission for between 30 s to 60 s. The Photon image was overlaid on the optical image, using software M3Vision (Biospace Lab). Overlay image was used to determine transformants with altered promoter activity (photon emissions), compared to a control *wt* pJI301.

#### 2.2.14.4. Analysis of Error Prone PCR Promoter Activity

Transformants with alterations in bioluminescence were inoculated onto fresh LB<sub>kan50</sub> and incubated at 37°C, for 16-18 hr. End point fluorescence assays were conducted, using the standard end point fluorescence protocol (2.2.18).

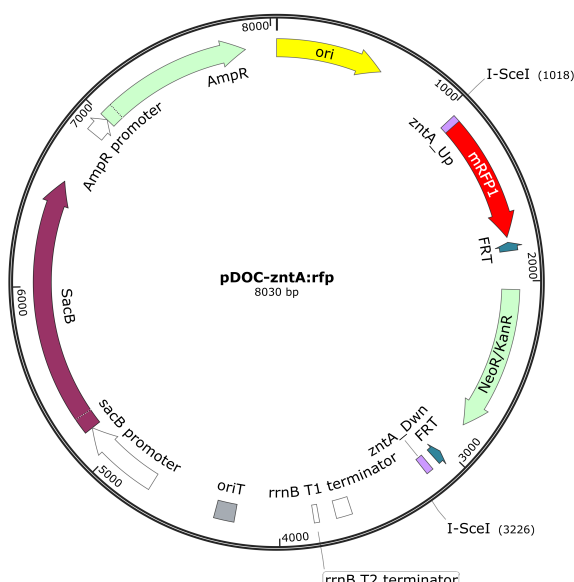
End point fluorescence compared potential mutant (pJI301) to the *wt* pJI301. Samples which had either a significant increase, or decrease, had the plasmid extracted and were sent for DNA sequencing (Eurofins) (2.2.3.4). ClustalW was used to align the sequenced DNA and compared to *wt* P<sub>zntA</sub>.



### 2.2.15. Gene Coupling of *mRFP1* to *zntA* by Gene Doctoring

*mRFP1* was fused to the C-terminal end of *zntA*, using the Gene Doctoring method (Lee *et al.*, 2009). *mRFP1* was chosen for fluorescence gene coupling, rather than GFP, therefore the same fluorescent protein is present in both the reporter plasmid construct (pJI300) and the chromosomal fusion. Note, this method also inserts a *kan<sup>R</sup>* cassette into the chromosome, for selection purposes.

#### 2.2.15.1. Construction of pDOC-*zntA:rfp*



**Figure 2.5 Plasmid Map of pDOC-*zntA:rfp***

Plasmid map of pDOC-*zntA:rfp* showing relevant genes. Whole plasmid sequence in appendix.

Primers were designed to amplify and linearize pDOC-C, amplify the *kan<sup>R</sup>* cassette and *FRT* sites (Flp-*FRT* recombinase site, for site directed recombination) from pDOC-G, and amplify *mRFP1* from pUltraRFP-KM. At the same time, incorporating 40 bp upstream sequence and 40 bp of homologous *zntA* sequence at the desired insertion site (Figure 2.5).

pDOC-C was linearized using primers: pDOC\_zntA:rfp\_Lin\_F and pDOC\_zntA:rfp\_Lin\_R using standard Q5 polymerase PCR protocol (2.2.5.1), with a  $T_m$  of 65°C with an extension of 6 min. The *kan<sup>R</sup>* cassette was amplified from pDOC-G using primers pDOC\_zntA:rfp\_KanR\_F and pDOC\_zntA:rpf\_KanR\_R using the standard Q5 polymerase PCR protocol (2.2.5.1) with a  $T_m$  of 72°C with an extension of 1 min. *mRFP1* was amplified from pUltraRFP-KM using primers pDOC\_zntA:rfp\_RFP\_F and pDOC\_zntA:rfp\_RFP\_R using standard Q5 polymerase protocol (2.2.5.1), with a  $T_m$  of 63°C and an extension of 1 min.

PCR products were verified by gel electrophoresis (2.2.3.5) and purified using a Monarch® PCR & DNA Cleanup Kit (2.2.3.2). The standard NEBuilder HiFi assembly protocol (2.2.3.8) was used with purified PCR products; 73 ng of linearized pDOC-C DNA, 18 ng of *kan<sup>R</sup>* DNA, and 35 ng of *mRFP1* DNA was combined and assembled. 3 µL of the assembly reaction was transformed into *E. coli* MG1655 using the standard electroporation protocol (2.2.1.3) and plated onto LB<sub>kan50&amp;100</sub>. The plasmid was purified and verified via Sanger sequencing (2.2.3.4), using the sequencing primers pDOC-C\_Seq\_F and pDOC-C\_Seq\_R.

#### 2.2.15.2. Chromosomal Fusion of *mRFP1* to *zntA*

pDOC-zntA:rfp and pACBSR were transformed into *E. coli* MG1655, using the standard electrocompetent transformation protocol (2.2.1.3) and plated into LB<sub>CAK</sub> (LB + Chlor 35 µg L<sup>-1</sup>, Amp 100 µg L<sup>-1</sup>, and Kan 50 µg L<sup>-1</sup>). Single colonies were patch plated onto both LB<sub>CAK</sub> and LB<sub>CAK</sub> + 5% sucrose. A single sucrose sensitive colony was inoculated into 2 mL of LB<sub>CAK</sub> + 1% glucose and incubated at 37°C with shaking (200 RPM) for 2 hr. The bacterial culture was transferred into a 1.5 mL microcentrifuge tube and centrifuged at 13,000 x *g* for 1 min (Sigma 1-16K), and the supernatant disposed of. The bacterial pellet was re-suspended in 1 mL of LB, pelleted again at 13,000 x *g* for

1 min, re-suspended in 1 mL LB + 1% arabinose, and incubated at 37°C with shaking for 5 hr. The bacterial culture was serial diluted ( $10^{-1}$  to  $10^{-4}$ ) and plated on LB<sub>kan50</sub> + 5% sucrose and incubated at 37°C for 16-20 hr. Single colonies were patch plated onto LB<sub>kan50</sub>, LB<sub>chlor35</sub>, and LB<sub>amp100</sub>. An Amp<sup>S</sup>, Kan<sup>R</sup>, and Chlor<sup>R</sup> colony suggested loss of pDOC-zntA:rfp, chromosomal fusion of *zntA:rfp*, insertion of *::kan<sup>R</sup>*, and residual helper plasmid pACBSR.

### 2.2.15.3. pACBSR Plasmid Curing

A single colony of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* (Kan<sup>R</sup>, Amp<sup>S</sup> and Chlo<sup>R</sup>) was inoculated into 5 mL of LB<sub>kan50</sub>, and incubated at 43°C for 12-16 hr, with shaking (200 RPM). Dilutions from  $10^{-1}$  to  $10^{-4}$  were plated onto LB<sub>kan50</sub>, and incubated at 42°C for 16hr. Single colonies were patch plated onto LB<sub>kan50</sub> and LB<sub>chlor35</sub> agar. Kan<sup>R</sup> only colonies suggest that pACBSR has been cured, and *zntA:rfp::kan<sup>R</sup>* was chromosomally fused.

### 2.2.16. Temporal Luminescence Assays

#### 2.2.16.1. Overnight Culture Set Up

pJI300 (empty vector) was used as a control to account for background fluorescence. *E. coli* MG1655 transformed with pJI300 and pJI300 promoter constructs were inoculated for singles on LB<sub>kan50</sub> agar, and incubated at 37°C for 12-16 hr. A single colony was inoculated into 5 mL of desired liquid growth media (LB or NH) + 50 µg mL<sup>-1</sup> kanamycin, with three biological repeats. Cultures were incubated at 37°C for 12-16 hr with shaking (200 RPM).

### 2.2.17. Luminescence Assay Measurement

Assay conditions were set up in 1.5 mL microcentrifuge tubes, with 1 mL LB<sub>kan50</sub> or NH<sub>kan50</sub> with desired conditions including, for example, 0-400 µM ZnSO<sub>4</sub>, 20 µM TPEN. These were set up with a 1:50 ratio, with three biological repeats.

The outer wells of a 96 well 350 µL black, clear bottom microtiter plate (Porvair, 215006) were filled with 300 µL H<sub>2</sub>O, to reduce moisture loss over time. 200 µL of blank media was pipetted into three wells, to serve as an absorbance blank. 200 µL of pJI300 and the pJI300 constructs were transferred into a 96 well microtiter plate, with three biological repeats. The inoculated microtiter plate was placed in a pre-warmed (37°C) microtiter plate reader (TECAN GENious pro). Measurements were taken every 10 min, for 16 hr, with 10 s of shaking (200 RPM), prior to each measurement. Absorbance was measured at 600 nm. Fluorescence was measured with an excitation of 580 nm and emission of 620 nm Luminescence was measured at 100 ms integration.

Data was analysed using (Equation 2.1), where: *RLUs* = relative luminescence unit of sample; *RLUb* = relative luminescence unit of blank; *ODs* = absorbance of sample; and *ODb* = absorbance of blank. This equation ensures values are ≥ 0. Values which would be below zero, become zero.

$$\text{Luminescence} = \frac{\text{IF} \left( ((\text{RLUs} - \text{RLUb}) < 0), 0, (\text{RLUs} - \text{RLUb}) \right)}{\text{IF} \left( ((\text{ODs} - \text{ODb}) < 0), 1, (\text{ODs} - \text{ODb}) \right)}$$

**Equation 2.1 Luminescence Background Correction Equation**

### 2.2.18. End Point Fluorescence Assay

#### 2.2.18.1. Overnight Culture Set Up

pJI300 (empty vector) was used as a control to account for background fluorescence. *E. coli* MG1655, transformed with pJI300 and pJI300 constructs, were inoculated for singles on LB<sub>kan50</sub> agar. A single colony was inoculated into 5 mL of liquid growth media + 50 µg L<sup>-1</sup> kanamycin, with three biological repeats. Media was supplemented with various ZnSO<sub>4</sub> concentrations, 20 µM TPEN, or other transition metals. Cultures were incubated at 37°C for 12-16 hr with shaking (200 RPM). [Note, this experiment was also conducted with *C. necator*. The setup is identical, except liquid broth and agar was supplemented with 200 µg L<sup>-1</sup> kanamycin and grown at 30°C].

#### 2.2.18.2. Endpoint Florescence Measurement

200 µL of overnight culture was transferred to a 96 well 350 µL black, clear bottom microtiter plate (Porvair, 215006), along with sterile media as a blank control, and pJI300 culture as a florescent background control. The 96 well microtiter plate was placed into a microtiter plate reader (TECAN GENios pro). Absorbance was measured at 600 nm. Fluorescence was measured with an excitation of 580 nm and emission of 620 nm, with an integration time of 40 µs and the gain set to extended dynamic range. Data was analysed using (Equation 2.2) where: *RFUs* = relative fluorescence unit of sample; *RFUb* = relative fluorescence unit of blank; *ODs* = absorbance sample; and *ODb* = absorbance blank. This equation ensures values are ≥ 0. Values which would be below zero, become zero.

$$\text{Fluoresence} = \frac{\text{IF}\left(\left((\text{RFUs} - \text{RFUb}) < 0\right), 0, (\text{RFUs} - \text{RFUb})\right)}{\text{IF}\left(\left((\text{ODs} - \text{ODb}) < 0\right), 1, (\text{ODs} - \text{ODb})\right)}$$

**Equation 2.2 Fluorescence Background Correction Equation**

### 2.2.19. BacLight LIVE:DEAD Assay

#### 2.2.19.1. 2x Staining Solution SYTO9/Propidium Iodide

6  $\mu$ L of 3.34 mM of SYTO9 dye and 6  $\mu$ L of 20 mM Propidium Iodide (PI) was combined with 2 mL of dH<sub>2</sub>O. Components provided as part of the LIVE/DEAD BacLight Viability Assay Kit (Thermo Fisher, L7012).

#### 2.2.19.2. BacLight TPEN Toxicity Test

*E. coli* MG1655 was inoculated for single colonies on LB agar, and incubated at 37°C for 16-20 hr. A single colony was inoculated into 5 mL of LB broth, with three biological repeats, and incubated at 37°C with shaking (200 RPM) for 16-20 hr. Each biological repeat was sub-cultured into a further 4 x 20 mL of LB, totalling 16 x 20mL, and incubated at 37°C with shaking (200 RPM), until mid-log phase (OD<sub>600</sub> 0.4) was reached. Each biological repeat was inoculated with various concentrations of TPEN (0, 10, 20 & 40  $\mu$ M), and incubated for a further 2 hr. 1 mL of culture was transferred into 1.5 mL microcentrifuge tubes and centrifuged at 13,000  $\times g$  (Sigma 1-16K) for 1 min. The supernatant was disposed of and re-suspended into 1 mL 0.85% NaCl. A negative control (dead) of 1 mL of 0  $\mu$ M TPEN was centrifuged at 13,000  $\times g$  for 1 min, supernatant disposed, and cell pellet re-suspended in 1 mL of 70% isopropanol. Samples were incubated at room temperature with shaking (400 RPM), in a shaking heat block (Eppendorf, Thermo Mixer Comfort), for 30 min. The negative control sample (in 70% isopropanol) was centrifuged at 13,000  $\times g$  for 1 min and re-suspended in 1 mL 0.85% NaCl.

A standard curve was made by combining various concentrations of live cells (0  $\mu$ M TPEN), and dead cells (0  $\mu$ M TPEN, 70% isopropanol treated). The ratio of live:dead used was as follows: 1:0; 4:1; 1:1; 1:4; and 0:1.

100  $\mu$ L of 2x staining solution was combined with 100  $\mu$ L of sample in a 96 well 350  $\mu$ L black, clear bottom microtiter plate (Corning, 215006), placed in a Tecan GENios Pro, and incubated at room temperature (18-22°C), for 15 min. Fluorescence for live cells (SYTO9) was recorded at em/ex 485/535 nm, and fluorescence for dead cells (PI) was recorded at em/ex 485/620 nm.

### 2.2.19.3. BacLight Zinc ( $\text{ZnSO}_4$ ) Toxicity Test

*E. coli* MG1655 was inoculated for single colonies on LB agar and incubated at 37°C, for 16-20 hr. A single colony was inoculated into 5 mL of LB broth, with three biological repeats, and incubated at 37°C with shaking (200 RPM), for 16-20 hr. Each biological repeat was sub-cultured into 6 x 20 mL of LB broth, totalling 18 x 20 mL, and incubated at 37°C with shaking (200 RPM), until mid-log phase ( $\text{OD}_{600}$  0.4) was reached. Each biological repeat was inoculated with various concentrations of  $\text{ZnSO}_4$  (0, 200, 400, 600, 800, and 1000  $\mu$ M) and incubated for a further 2 hr. 1 mL of culture was centrifuged at 13,000 x *g* for 1 min, the supernatant disposed and the cell pellet re-suspended in 1 mL of 0.85% NaCl. A negative control (dead) of 1 mL of 0  $\mu$ M  $\text{ZnSO}_4$  was centrifuged at 13,000 x *g* for 1 min, the supernatant disposed and the cell pellet and re-suspended in 1 mL 70% isopropanol. The samples were incubated at room temperature with shaking (400 RPM), in a shaking heat block, for 30 min. The negative control sample (in 70% isopropanol) was pelleted at 13,000 x *g* for 1 min and re-suspended into 1 mL 0.85% NaCl.

A standard curve was made by combining various concentrations of live cells (0  $\mu$ M ZnSO<sub>4</sub>), and dead cells (0  $\mu$ M ZnSO<sub>4</sub> , 70% isopropanol treated). The ratio of live:dead used was as follows: 1:0; 4:1; 1:1; 1:4; and 0:1.

100  $\mu$ L of 2x staining solution was combined with 100  $\mu$ L of sample in a 96 well 350  $\mu$ L black, clear bottom microtiter plate (Porvair, 215006), placed in a Tecan GENios Pro, and incubated at room temperature (18-22°C) for 15 min. Fluorescence for live cells (SYTO9) was recorded at em/ex 485/535 nm, and fluorescence for dead cells (PI) was recorded at em/ex 485/620 nm.

### 2.2.20. Flow Cytometry

#### 2.2.20.1. *E. coli* zntA:rfp::kan<sup>R</sup> Culture Set Up

A single colony, with three biological repeats, of both *E. coli* MG1655 *wt* and *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* was inoculated into 5 mL liquid broth (LB or NH) (+ 50  $\mu$ g mL<sup>-1</sup> kanamycin for *zntA:rfp::kan<sup>R</sup>*), and incubated at 37°C for 16-18hr, with shaking (200 RPM). Cultures were then sub-cultured with a 1:50 dilution into 5 mL of fresh liquid broth (LB or NH) (+ 50  $\mu$ g mL<sup>-1</sup> for *zntA:rfp::kan<sup>R</sup>*). *E. coli* MG1655 *wt* was sub-cultured in liquid broth (LB or NH) only. Whereas *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* was sub cultured into liquid broth (LB or NH) with added TPEN or ZnSO<sub>4</sub> at various concentrations. Samples were incubated at 37°C with shaking (200 RPM) for 16 hr. 100  $\mu$ L of bacterial culture was centrifuged at 13,000 x *g* for 1 min (Sigma 1-16K), supernatant disposed, and re-suspended in 1000  $\mu$ L 4% formaldehyde in PBS. Samples were stored on ice until analysed using flow cytometry.



#### 2.2.20.2. Plasmid Culture Set Up

A single colony, with three biological repeats, of *E. coli* MG1655 transformed with pJI300 and pJI301 were inoculated into 5 mL of LB<sub>kan50</sub> and incubated at 37°C for 16 hr, with shaking (200 RPM). Cultures were then sub-cultured with a 1:50 dilution into 5 mL of fresh LB<sub>kan50</sub> with; 0, 100, 200 or 400 µM ZnSO<sub>4</sub>, or 20 µM TPEN. Samples were incubated at 37°C with shaking (200 RPM) for 16-18 hr. 100 µL of bacterial culture was centrifuged at 13,000  $\times g$  for 1 min (Sigma 1-16K), supernatant disposed, and re-suspended in 1000 µL 4% formaldehyde in PBS. Samples were stored on ice until analysed using flow cytometry.

#### 2.2.20.3. Flow Cytometry Analysis

Samples were analysed using the flow cytometer (Beckman Coulter Astrios EQ). Samples were excited with a green (561 nm) laser with emission filter (614/20 nm). Forwards scatter (FCS), side scatter (SSC), and emission (604-624 nm) was recorded.

#### 2.2.20.4. Flow Cytometry Data Analysis (Kaluza)

Flow cytometry data was analysed using Kaluza Analysis Software (Beckman Coulter). Four gates (used to exclude data points and set value limit) were used for the flow cytometry analysis procedure in this work: time (rate of injection); whole cells; single cells; and fluorescent noise.

Histogram plots were created comparing; gated fluorescence (ex 651 nm, em 604-624 nm) against events/counts. Overlay plots were performed to compare samples, and controls.

#### 2.2.20.5. Bayesian Background Subtraction Analysis of Flow Cytometry Data

Kooperberg *et al.*, (2002) developed a Bayesian background subtraction approach to analysis data for microarray data; this was adapted for use in flow cytometry data. This analysis provides an estimation of the true foreground. The Bayesian background subtraction approach essentially; subtracts the background data from the foreground data by constructing a posterior distribution of the true foreground data (given the observed foreground and background data), knowing that the true foreground cannot be negative.

The original Kooperberg *et al.*, (2002) model was adapted, and altered slightly to work with flow cytometry data; designated the “Ingram” function (Figure 2.6). The ‘Supply’

function was applied to the Ingram function, allowing all foreground data points (f) to be run through the code, returning the estimation of the true foreground data.

```
Ingram = function(f) {
  SDS = 4
  sf = sb/10
  sd = sqrt(sf*sf+sb*sb)

  #Range
  low = max(0,as.integer(f-xb-sds*sd))
  hi = as.integer(f-xb+sds*sd+1)
  x = seq(low,hi)
  x

  #kooper
  y = dnorm((f-x-xb)/sd)*pnorm(((f-x)*sb*sb+xb*sf*sf)/(sf*sb*sd))
  y

  #Backsub
  Z = sum(x*y)/sum(y)
}
```

f = every foreground data point  
 xb = mean background  
 sf = standard deviation foreground  
 sb = standard deviation background

#### Figure 2.6 Ingram Function

R code of 'Ingram' function. Bayesian approach to substitution of background fluorescence from foreground fluorescence of Flow Cytometry data. Based on work by Kooperberg *et al.*, (2002). Code adapted from Dov Stekel

### 2.2.21. C1265-7 Phenotype Growth Experiments

#### 2.2.21.1. C1265-7 Phenotype Shake-Flask Growth Experiment

A single colony of freshly grown *E. coli* CFT073 (wt,  $\Delta znuABC$ ,  $\Delta c1265-7$ , and  $\Delta c1265-7 \Delta znuABC$ ) were inoculated into 5 mL of NH media, with three biological repeats, and incubated at 37°C with shaking (200 RPM) for 16-18 hr. The bacterial cultures were sub-cultured into 35 mL of NH media, at a 1:100 ratio, and incubated at 37°C with shaking (200 RPM). OD<sub>600</sub> readings were taken every hour for 10 hr, and a final reading at 24 hr.

### 2.2.21.2. C1265-7 Phenotype Microplate Growth Experiment

A single colony of freshly grown *E. coli* CFT 073 (*wt*,  $\Delta znuABC$ ,  $\Delta c1265-7$ , and  $\Delta c1265-7 \Delta znuABC$ ) were inoculated into 5 mL of NH media, with three biological repeats, and incubated at 37°C with shaking (200 RPM) for 16-18 hr. The bacterial cultures were diluted 1:50 in 1 mL NH media. 200  $\mu$ L of culture was transferred in to a 96 well 350  $\mu$ L black, clear bottom microtiter plate (Porvair, 215006), and placed into a pre warmed (37°C) TECAN GENios Pro. Absorbance<sub>600</sub> (600 nm) was recorded every 10 min for 960 min, with 10 s shaking (200 RPM) prior to each read.

### 2.2.21.3. C1265-7 Phenotype BacTiter-Glo® Cell Viability Assay

*E. coli* MG1655 (*wt*,  $\Delta znuA$ , and  $\Delta znuCB$ ) were transformed with either pJI100A (empty vector) or pJI102 (*c1265-7* expression vector). A single colony, with three biological repeats, was inoculated into 5 mL of NH<sub>amp100</sub>, and incubated at 37°C with shaking (200 RPM) for 16-18hr. The culture was sub-cultured with a 1:100 ratio into 25 mL of NH<sub>amp100</sub> and incubated at 37°C with shaking (200 RPM).

Luminescence readings were taken every hour, for 9 hr. BacTiter-Glo (Promega, G8230) reagent was made up as per manufactures instructions. 100  $\mu$ L of bacterial culture was combined with 100  $\mu$ L BacTiter-Glo reagent in a 96 well 350  $\mu$ L black, clear bottom microtiter plate (Porvair, 215006), and placed in a pre-warmed (22°C) TECAN GENios Pro. Luminescence readings were taken after 5 min incubation at 22°C, with a luminescence integration time of 250 ms.

### 2.2.22. ICP-MS Metal Analysis of Media

#### 2.2.22.1. ICP-MS Analysis - Preparation of LB and NH

A  $10^{-1}$  dilution of LB broth was needed to achieve an elemental molar concentration within the detectable range for ICP-MS analysis. 1 mL of LB broth was diluted by adding 9 mL of 2.22% nitric acid, to make a final concentration of 2% nitric acid

294  $\mu$ L of 68% Nitric acid was added to 10 mL of NH media, to make a final concentration of 2% nitric acid. No dilution of NH was needed to achieve an elemental concentration for the detectable range for ICP-MS analysis.

#### 2.2.22.2. ICP-MS Analysis - Preparation of *E. coli* Growth Curve in LB

A single colony of *E. coli* MG1655, with three biological repeats, was inoculated into 5 mL of LB broth, and incubated at 37°C for 16 hr with shaking (200 RPM). The bacterial culture was diluted  $10^{-3}$  in 50 mL of LB and incubated at 37°C with shaking (200 RPM). 1 mL was removed from each biological repeat, and OD<sub>600</sub> was determined using a spectrophotometer. Further OD<sub>600</sub> readings were taken hourly (0-7 hr and 22 hr).

2 mL of culture was extracted simultaneously as the OD<sub>600</sub> readings and transferred to a 2 mL microcentrifuge tube (Eppendorf). The sample was centrifuged at 20,000 x *g* for 3 min (Sigma 1-16K), the supernatant was removed and transferred to a new 2 mL 1.5 mL microcentrifuge tube. The supernatant was centrifuged again at 20,000 x *g* for 3 min to ensure all cells removed. 1 mL of the supernatant was transferred to a 15 mL PP centrifuge tube, combined with 9 mL of 2.22% nitric acid, and filtered through a 0.22  $\mu$ m PES (Sartorius) filter into a new 15 mL PP centrifuge tube.

### 2.2.22.3. ICP-MS Analysis

Samples were sent off for trace element analysis using ICP-MS. Analysis was conducted by Saul Vazquez Reina at The University of Nottingham, School of Bioscience, using a Thermo-Fisher Scientific iCAP-Q equipped with CCTED (collision cell technology with energy discrimination).

### 2.2.23. RT-qPCR to Determine Relative Abundance of *zntA*

RT-qPCR refers to Reverse Transcriptase quantitative Polymerase Chain Reaction.

#### 2.2.23.1. Control Reference Gene *rrsA*

*rrsA* was chosen as the reference gene for comparison of mRNA abundance. *rrsA* has previously been shown to be a suitable reference gene for similar RT-qPCR experiments (Zhou *et al.*, 2011; Peng *et al.*, 2014; Xu *et al.*, 2019).

#### 2.2.23.2. Preparation of Samples

*E. coli* MG1655 was inoculated for single colonies on LB agar, and incubated at 37°C for 16-18 hr. A single colony was inoculated into 10 mL of NH, with three biological repeats, and incubated at 37°C with shaking (200 RPM) for 16-18 hr. Each biological repeat was sub-cultured into 20 mL of; NH, NH + 200 µM ZnSO<sub>4</sub>, and NH + 400 µM ZnSO<sub>4</sub>, at a 10<sup>-3</sup> dilution, and incubated at 37°C with shaking (200 RPM). Samples were taken at 16 hr incubation. Samples were standardized to an OD<sub>600</sub> unit of 1. Samples were pelleted at 13,000 *x g* for 1 min, re-suspended in 1 mL of RNA*later* and stored at 4°C until RNA extraction. Total RNA was extracted using a Monarch® Total RNA Miniprep Kit (2.2.4.1) and diluted to 1 ng µL<sup>-1</sup> using nuclease free H<sub>2</sub>O.

### 2.2.23.3. RT-qPCR Using The Luna® Universal Probe One-Step RT-qPCR Kit

RT-qPCR was conducted for each biological repeat, with two technical repeats. *rrsA* and *zntA* were analysed for all three growth conditions. Luna® Universal Probe One-Step RT-qPCR Kit (NEB, E3006) components were thawed at room temperature, re-suspended by vortexing (Vortex-Genie 2) and placed on ice. 10 µL of Luna universal probe one-step reaction mix (2x), 2 µL of Luna warm start RT enzyme mix (20x), 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 0.5 µL probe (10 µM), and template RNA (1 ng) were combined, and made up to a final volume of 20 µL using nuclease free H<sub>2</sub>O in a 0.2 mL PCR tube (Xtra-clear, I1402-8200). RT-qPCR was conducted in a real-time PCR cycler (Qiagen Rotor-Gene Q) using thermocycle conditions in Table 2.13 qPCR Thermocycle Conditions) with green fluorescence (FAM).

Temp °C	Time (s)	
55	600	
95	60	
95	10	
60	30 (+ fluorescence read) [FAM, 495/520 nm]	X 50

**Table 2.13 qPCR Thermocycle Conditions**

### 2.2.23.4. RT-qPCR Calibration Curve

A standard curve was set up to analysis the efficiency of the qPCR. An RNA extract of *E. coli* MG1655 was diluted 10-fold from 10<sup>0</sup> to 10<sup>-3</sup>. An efficiency of 90-110% is widely acceptable. Deviation from 100 % efficiency, within this range can be due to pipetting error in the RNA dilutions or in qPCR set up.

### 2.2.23.5. Comparative C<sub>q</sub> Analysis ( $2^{-\Delta\Delta C_q}$ )

The comparative C<sub>q</sub> analysis (sometimes known as C<sub>t</sub>) is a common and powerful analysis to show relative gene expression. This can be used to determine relative gene expression of different genes, or the same gene under different conditions.

C<sub>q</sub> is the value of the cycle number in which the fluorescence intensity exceeds an arbitrarily appointed value. The Comparative C<sub>q</sub> makes two assumptions; that the efficiency of gene amplification is close 100% (90-110%), and that the gene of interest and internal reference gene have similar efficiencies. To be able to determine comparative C<sub>q</sub>, four values are needed, outlined in Table 2.14 RT-qPCR Samples & Controls). To determine the fold change in expression, the  $2^{-\Delta\Delta C_q}$  equation is used (Equation 2.3).

C <sub>T</sub> values	Condition A (Treated)	Condition B (Untreated)
Reference gene ( <i>rrsA</i> )	A <sub>Ref</sub>	B <sub>Ref</sub>
Target gene of interest ( <i>zntA</i> )	A <sub>Gene</sub>	B <sub>Gene</sub>

**Table 2.14 RT-qPCR Samples & Controls**

$$\text{Fold Change} = 2^{-\Delta\Delta C_q} = 2^{\Delta C_q - \Delta C_{ref}} = 2^{(A_{gene} - A_{ref}) - (B_{gene} - B_{ref})}$$

**Equation 2.3  $2^{-\Delta\Delta C_q}$**



## Chapter 3

# Fundamental Understanding of the ZntR Transcriptional Regulation of the *zntA* Promoter ( $P_{zntA}$ )

**P<sub>zntA</sub>**      -35                                  -10                                  +1  
TTGACTCTGGAGTC GACTCCAGAGTGTATCCTTCGGTTA

Promoter sequence of *zntA*. **Bold** indicates either -35 or -10 element. Underline indicates ZntR binding site, space indicates separation of the 11-11 bp palindromic repeat. **Green** indicates +1 start site.

108

equating to two free  $Zn^{2+}$  cation per cell (Rensing, Mitra and Rosen, 1997; Outten and O'Halloran, 2001).

Previous data has also suggested that there may be a tuneable element to  $P_{zntA}$ , that translational gene expression may relate to zinc concentration (Brocklehurst *et al.*, 1999; Outten and O'Halloran, 2001).

#### 3.1.1. Aims and Objectives

The main objective of this chapter is to further understand the adaptive transcriptional regulation of the ZntR regulated  $P_{zntA}$ , with the intention to utilise  $P_{zntA}$  as a novel zinc inducible system.

$P_{zntA}$  is known to be tuneable, somewhat, in response to zinc. This chapter aims to further develop understanding of the tuneable response of  $P_{zntA}$  to zinc, using both end point and temporal assays. Temporal expression allows an understanding of how the gene is regulated over time, and how this relates to bacterial growth phases. This understanding of zinc adaptive response to zinc is essential to understand if  $P_{zntA}$  can be used as a zinc inducible promoter.

## 3.2. Results

### 3.2.1. Nomenclature

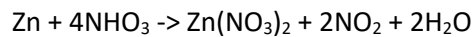
The promoter of a gene will be written using the following notation; ' $P_{gene}$ '. Throughout this work, data will be shown as  $P_{gene}$  rather than the plasmid containing the promoter, for ease; plasmids and promoter inserters can be seen in (Table 2.2 Table of Plasmids).

### 3.2.2. Validation of Zinc Depleted Neidhardt's MOPS Minimal Media (NH)

This work aims to develop a new novel promoter system based on the zinc homeostasis system of *E. coli*. For this, it is important to have a defined minimal media in which to conduct assays. Neidhardt's MOPS minimal media was chosen, as it has been used in previous zinc depleted studies (Clayton, 2012; Wang *et al.*, 2012). Neidhardt's MOPS minimal media was chosen due to its similar growth rates to traditional complex bacterial growth medias (Neidhardt, Bloch and Smith, 1974). Due to Neidhardt's media being a chemically defined media, it will allow the reduction of zinc more easily than a media containing complex biologically derived components such as yeast extract and peptone.

To make a zinc depleted media, a few alterations were made to the original recipe for Neidhardt's MOPS minimal media, developed by Neidhardt (Neidhardt, Bloch and Smith, 1974). Notably,  $ZnSO_4$  was omitted from the metal micronutrient stock. Further to this, to improve growth rate, MEM (minimum essential medium) amino acid solution was also added. MEM contains all essential amino acids apart from L-glutamine. Early experiment showed that, without additional amino acids, the growth rate was significantly slower than in LB, and  $OD_{600}$  did not reached comparable levels (data not shown). Henceforth, zinc depleted Neidhardt's MOPS minimal media will be known as NH.

To reduce zinc contamination from plasticware, either new virgin PP plastic was used, or if not possible, 6% nitric acid treated PMP was used. No glass was used in the making of NH media, or downstream experiments. 6% nitric acid treatment was used; nitric acid reacts with zinc to produce zinc nitrate, which can be removed from plasticware via washing with ultra-pure Mili-Q water.



All non-metal solutions were treated with 10% Chelex 100 resin for 1 hr (with stirring), prior to filter sterilization. Chelex 100 is a styrene divinylbenzene copolymer containing paired iminodiacetate ions (Bio-Rad, 2020). Chelex 100 resin binds polyvalent metal ions and other metal ions with strong affinity. According to Bio-Rad, (2020) Chelex 100 *“scavenges metal contaminants without altering the concentration of nonmetal ions”*. Therefore, Chelex 100 Resin was used to remove metal ions from non-metal solutions, removing environmental zinc from water and the chemicals. Chelex 100 resin has previously been used to remove metal ions from solutions (Knapp *et al.*, 1987; Clayton, 2012).

Highest grade and purity chemicals were used to reduce zinc contamination. HPLC grade  $\text{H}_2\text{O}$  was purchased and used to make NH media. To further reduce zinc contamination, filter tips were used through NH work. All solutions and NH were filter sterilized into virgin PP plastic.

### 3.2.3. ICP-MS Analysis of Growth Media

Two growth media are used predominantly throughout this work, Lysogeny Broth (LB) (Bertani, 1951), and zinc depleted Neidhardt’s MOPS minimal media (NH). As this work is centred around promoter activity, with zinc used as an inducer, it is essential that the zinc concentration of these two media is known.

ICP-MS (Inductively Coupled Plasma - Mass Spectrometry) was used for elemental analysis of the two main growth media. Table 3.1 shows the elemental concentrations of both media, as determined via ICP-MS. Two-tailed, unpaired, t-test concluded that all elements tested, apart from Cu, have a significant difference in concentrations between the two media. Zn is over 100x lower in NH than in LB.

Element (mol L <sup>-1</sup> )	LB	NH	Significant Difference Between LB and NH
Na	$9.65 \times 10^{-2}$	$7.40 \times 10^{-1}$	***
Mg	$1.68 \times 10^{-4}$	$9.40 \times 10^{-3}$	***
P	$4.27 \times 10^{-3}$	$1.67 \times 10^{-2}$	****
K	$8.83 \times 10^{-3}$	$2.74 \times 10^{-2}$	****
Ca	$8.54 \times 10^{-5}$	$1.80 \times 10^{-5}$	****
Mn	$2.48 \times 10^{-7}$	$9.67 \times 10^{-7}$	****
Fe	$4.39 \times 10^{-6}$	$9.54 \times 10^{-5}$	****
Co	$1.76 \times 10^{-7}$	$3.13 \times 10^{-7}$	****
Ni	$5.22 \times 10^{-8}$	$9.94 \times 10^{-8}$	*
Cu	$2.47 \times 10^{-7}$	$2.21 \times 10^{-7}$	ns
Zn	$1.22 \times 10^{-5}$	$1.15 \times 10^{-7}$	****
Ag	$4.15 \times 10^{-10}$	$2.89 \times 10^{-10}$	*

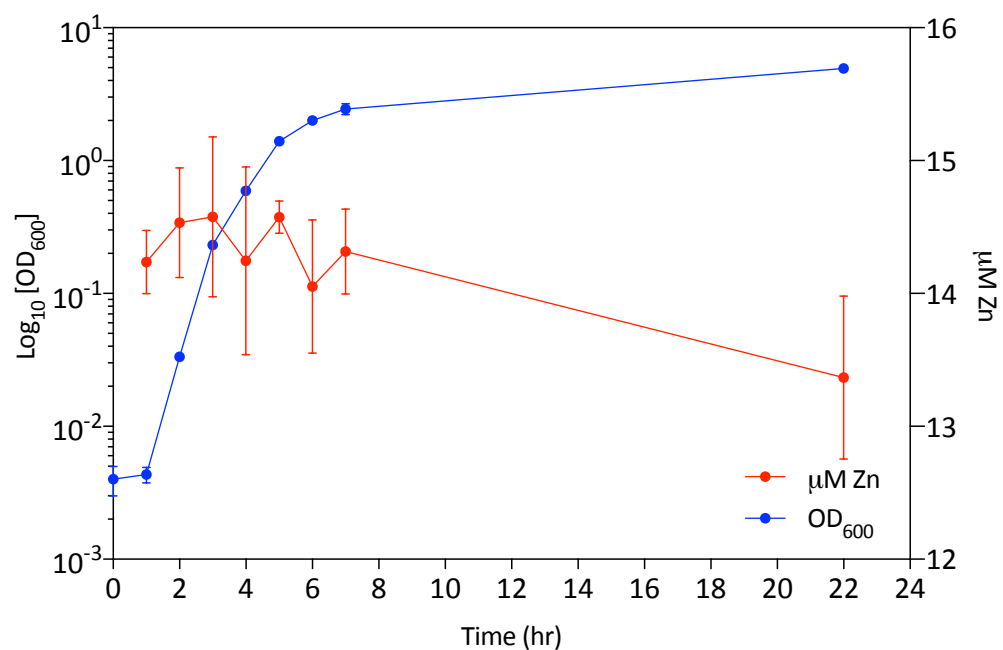
**Table 3.1 Molar Concentration of Different Growth Media**

Average elemental quantifications of LB and NH media through ICP-MS analysis. three technical repeats. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test)

Zinc is abundant and is found almost everywhere; removal of zinc from media is a hard and tedious task. Results show 106x reduction in zinc in NH compared to LB. These results show that NH is good and reliable media to use from zinc depleted conditions.

## 3.2.3.1. ICP-MS Analysis of Zinc in Media During Bacterial Growth

*E. coli* MG1655 was grown for 24 hr in LB media, with samples taken periodically and sent for ICP-MS analysis. This quantified the zinc concentration in the media over time and compared it to the *E. coli* growth phase. Figure 3.2 shows that between 0-24 hr there is no significant difference between zinc concentrations. A simple linear regression was used to analyse Zn concentration over time. This returned results showing a slight negative slope, (-0.052), with an  $R^2$  value 0.372. These results indicate that when *E. coli* is grown in LB, there is no significant reduction in overall zinc concentration.



**Figure 3.2 Molar Concentration of Zinc During *E. coli* Growth**

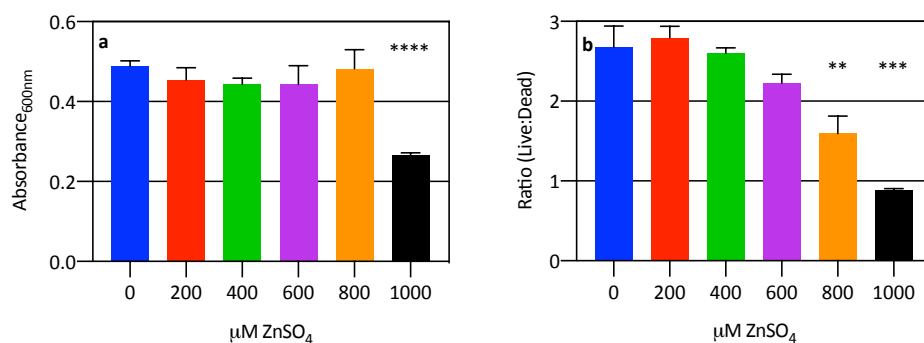
*E. coli* MG1655 was inoculated into 50 mL LB, with three biological repeats, and grown at 37°C with shaking (200 RPM). OD<sub>600</sub> and zinc concentration (ICP-MS) analysis was conducted periodically. OD<sub>600</sub>: 1 mL of culture was removed, and OD<sub>600</sub> was determined using a spectrophotometer. ICP-MS: 2 mL culture was pelleted, supernatant was diluted 1:10 with 2.22% nitric acid solution. Sample was analysed using ICP-MS to quantify available zinc in the media. Error bars indicate SD.

#### 3.2.4. BacLight Toxicity Tests

Throughout this thesis, both TPEN and  $\text{ZnSO}_4$  are used to evaluate the gene regulation of Zur and ZntR regulated genes. TPEN is a cell permeable, zinc chelator, which reduces the available zinc within a solution, as well as within the cell (Cho *et al.*, 2007). TPEN was chosen over other chelators such as EDTA, as TPEN is a specific zinc chelator. EDTA has been shown to chelate multiple transition metals, including: copper; nickel; iron; chromium; cobalt; manganese; palladium; and iridium (Zaitoun and Lin, 1997).  $\text{ZnSO}_4$  is used to increase zinc concentrations in the media. Initial toxicity tests were conducted to determine the safe concentration of  $\text{ZnSO}_4$  and TPEN for *E. coli* MG1655. The LIVE/DEAD BacLight viability assay kits (Thermo Fisher, L7012) were used; this kit provides more detail than a standard toxicity test, which, often solely looks at  $\text{OD}_{600}$  values after 16 hr growth, under different conditions. The BacLight kit allows one to look at the ratio of live to dead cells, rather than just optical density. This improved depth of analysis can be seen in Figure 3.3. Data shows that there is only a significant reduction in growth ( $\text{Absorbance}_{600}$ ) with the addition of 1000  $\mu\text{M}$   $\text{ZnSO}_4$ , whereas when looking at the ratio, there is a significant reduction of live cells with the addition of 800  $\mu\text{M}$   $\text{ZnSO}_4$ .



ZnSO<sub>4</sub> and TPEN were added during mid-log phase, as this is more indicative of how inducers are used. It is worth noting that bacteria were grown to an OD<sub>600</sub> of 0.4, grown for 2 hr, then Absorbance<sub>600</sub> was recorded. OD<sub>600</sub> was recorded in a spectrophotometer, whereas absorbance was recorded in a 96 well plate reader. This different measurement of growth gives the impression that no growth was observed between 2 to 4 hr time point, which is not necessarily true.



**Figure 3.3 Zinc Toxicity on *E. coli* During Mid Log Growth Phase**

*E. coli* MG1655 was inoculated into 20 mL LB broth, with three biological repeats, and incubated at 37°C with shaking (200 RPM) until OD<sub>600</sub> 0.4 was reached. Increasing concentrations of ZnSO<sub>4</sub> was added and incubated for a further 2 hr. LIVE/DEAD BacLight kit was used to determine live:dead ratio. Live fluorescence (ex/em 485/535), dead fluorescence (ex/em 485/620), and absorbance (600 nm) was recorded on a TECAN GENious Pro. (a) Absorbance<sub>600</sub> (b) ratio of live:dead cells. \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test) comparison to 0  $\mu\text{M ZnSO}_4$ .

#### 3.2.4.1. Zinc Toxicity Test

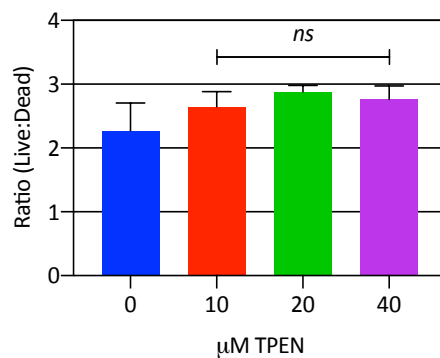
Figure 3.3a shows the effect of ZnSO<sub>4</sub> on OD<sub>600</sub> of *E. coli* grown with various concentrations of ZnSO<sub>4</sub>. Data shows that the addition of 1000  $\mu\text{M ZnSO}_4$ , significantly reduces the Absorbance<sub>600</sub>, compared to the control (0  $\mu\text{M ZnSO}_4$ ). At 1000  $\mu\text{M ZnSO}_4$ , there is a toxic effect on *E. coli*, reducing growth.

Figure 3.3b shows the effect of ZnSO<sub>4</sub> on live:dead ratio, as determined via a SYTO9 and Propidium Iodide (PI) dye, with increased concentrations of ZnSO<sub>4</sub>. This shows that with the lower addition of 800  $\mu\text{M ZnSO}_4$ , there is a negative effect on live:dead cell ratio, with an increase in dead cells, compared to live cells. With the addition of

600  $\mu\text{M}$   $\text{ZnSO}_4$ , there is no significant reduction in the live:dead ratio, however, there does seem to be a slight reduction in live:dead ratio.

#### 3.2.4.2. TPEN Toxicity Test

Figure 3.4 shows the effect of the addition of TPEN, on *E. coli* during mid log phase. This figure shows that with the addition of up to 40  $\mu\text{M}$  TPEN, there is no negative side effect on growth. With all three addition of TPEN (10  $\mu\text{M}$ , 20  $\mu\text{M}$ , and 40  $\mu\text{M}$ ) there is no significant difference in live:dead compared to the control of 0  $\mu\text{M}$  TPEN.



**Figure 3.4 TPEN Toxicity on *E. coli* During Mid Log Growth Phase**

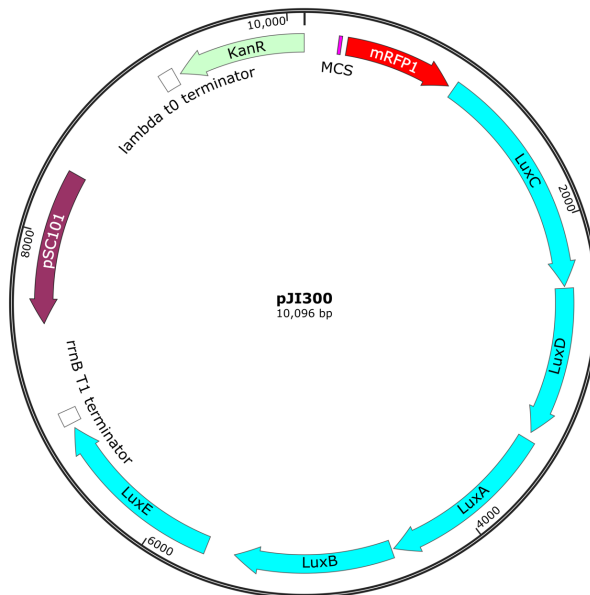
*E. coli* MG1655 was inoculated into 20 mL LB broth, with three biological repeats, and incubated at 37°C with shaking (200 RPM) until  $\text{OD}_{600}$  0.4 was reached. Various concentrations of TPEN was added and incubated for a further 2 hr. LIVE/DEAD BacLight kit (Invitrogen) was used to determine live:dead ratio. Live fluorescence 485/535, dead fluorescence 485/620, and absorbance (600 nm) was recorded on a TECAN GENious Pro. *ns*  $p > 0.01$  (two-tailed, unpaired, t-test) comparison to 0  $\mu\text{M}$  TPEN.

#### 3.2.5. pJI300: A Dual Reporter Plasmid

A dual reporter plasmid (pJI300) was designed and created for use throughout this work (2.2.10). pJI300 contains both *mRFP1* and *luxCDABE* reporter elements, as well as *pSC101*, *Kan<sup>R</sup>*, *rrnB*, *T1* terminator, and a multiple cloning site (MCS), with XhoI, SmaI and BamHI sites (Figure 3.5). Plasmid sequence in appendix (9.1.4).

mRFP1 was chosen, as this is a commonly used reporter element, due to its ease of use, and will be used for end point assays. mRFP1 matures 10x faster than its

predecessor (DsRed) and shows similar brightness in cells. Further to this, mRFP1 has improved cell penetration. mRFP1 has a maximum excitation of 584 nm and a maximum emission at 607 nm.



**Figure 3.5 Plasmid Map of pJI300**

Plasmid Map of pJI300 (empty vector) showing relevant genes

*luxCDABE*, the bioluminescence operon from *Photobacterium luminescens*, was chosen as it contains all genes for the bioluminescence; *luxAB* (luciferase) and *luxCDE* (fatty acid reductase complex). *luxCDABE* has a short half-life, and quick maturation; this allows one to look at temporal gene regulation. Luciferase catalyses reduction of the fatty acid reductase complex, which leads to emission of light (490 nm) (Meighen, 1991).

The stringently controlled pSC101 was chosen to reduce variation between cells. pSC101, regulated by RepA, replicates to around five copies per cell (Manen and Caro, 1991; Lee *et al.*, 2011).

Kanamycin resistance,  $kan^R$ , is used in the plasmid. Kanamycin is an aminoglycoside bactericidal antibiotic, which inhibits protein synthesis by bind to the 16S rRNA in the 30S subunit.  $kan^R$  encode the enzyme aminoglycoside-3'-phosphotransferase (APH(3')). APH(3') phosphorylates the 3'-hydroxly group of kanamycin, causing steric hinderance of the kanamycin binding to the 16S rRNA (Kotra, Haddad and Mobashery, 2000; ScienceDirect, 2008).

#### 3.2.5.1. Construction and Validation of pJI300

pJI300 was constructed using NEBuilder DNA Assembly kit. pLUX was linearized 1 bp downstream of the MCS, *mRFP1* amplified from pUltra-RFP, and the NEBuilder HiFi DNA Assembly kit was used to clone these two products together.

To ensure the plasmid pJI300 was constructed correctly, primers seq\_204\_F and luxC\_seq\_R were used to sequence the construct. ClustalW was used to align the expected sequence, and the returned sequence data showed correct insertion (data not shown).

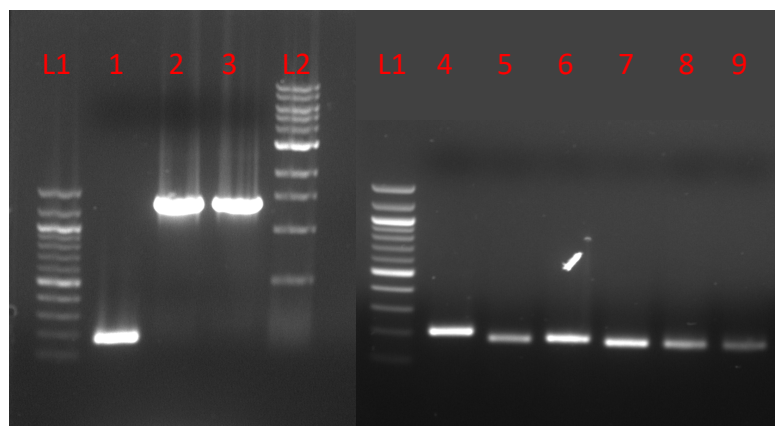
#### 3.2.5.2. pJI300 Promoter Series: Construction and Validation

Eight different promoters were cloned into the dual reporter plasmid pJI300. These promoters included the ZntR regulated  $P_{zntA}$ , and the Zur regulated promoters:  $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{kygM}$ , and  $P_{c1265}$  (Zur regulated constructs used later, Chapter 4). Additionally, a control promoter, the LacI regulated  $P_{trc}$ , was cloned into pJI300.

The construction of all eight promoter plasmids was conducted in the same way. Firstly, pJI300 was linearized by restriction enzyme digestion, using XhoI and BamHI. Digested plasmid concentration was determined using a nanodrop spectrophotometer and diluted to  $75 \text{ ng } \mu\text{L}^{-1}$ .

Promoters were amplified using NEB Q5 Polymerase, with primers designed using NEBuilder Assembly Tool (NEB, 2019). Primers were designed to amplify the promoter, as well as homologous DNA to pJI300 at the point where it was linearized (XhoI and BamHI digestion), to allow homologous recombination. PCR products were visualised on a 2% gel (Figure 3.6). Figure 3.6 shows PCR products of all promoter inserts at the expected size. These PCR products were then assembled with linearized pJI300 and transformed into *E. coli* MG1655 via electroporation.

Colony PCR was conducted on transformants for quick confirmation of correct promoter plasmid assembly. Positive samples had their plasmid extracted and were sent for Sanger sequencing, using primers 204\_seq\_R and rfp\_seq\_R. Sequence showed correct insertion (data not shown).



**Figure 3.6 PCR Product Gel Image for Promoter Inserts**

2% Gel image of PCR product of promoter inserts. L1 100 bp ladder, L2 1 kbp ladder, 1  $P_{zntA}$ , 2-3  $P_{trc}$ , 4  $P_{znuA}$ , 5  $P_{znuCB}$ , 6  $P_{zinT}$ , 7  $P_{pliG}$ , 8  $P_{ykgM}$ , 9  $P_{c1265}$ .

### 3.2.6. $P_{zntA}$ Mutants: Construction and Validation

Site directed mutagenesis (SDM) was used to introduce specific mutations into  $P_{zntA}$ , with the aim to elicit different responses to zinc induction and zinc depletion, compared to the wild type. NEB Q5 Site-Directed Mutagenesis kit was used to incorporate the desired mutations directly into pJI301.

Three mutation approaches were used to design the  $P_{zntA}$  mutation. The first approach mutated every nucleotide between the -10 and -35 element. These mutations were designed to cause the greatest physical change in the DNA, mutating from a purine to a pyrimidine; A $\leftrightarrow$ C and T $\leftrightarrow$ G. The second approach designed mutations in the: -35 element; -10 element; -14 TG motif; and -18 element, increasing their sequence consensus. The third approach introduced additional nucleotide between -25 and -24, increasing the spacing between the -10 and -35 elements. These additional nucleotides were added at the site which separates the ZntR-DNA binding of the two monomers (Figure 1.17). Figure 3.7 shows all mutations in  $P_{zntA}$ . [Note, due to the COVID-19 global pandemic, all mutations were created but not all were tested]. SDM promoter plasmids were purified and sent for sequencing; all showed sequence for desired SDM mutation (data not shown).

<i>P<sub>zntA</sub></i>	-35	-10	+1
<i>Wt</i> *	<b>TTGACT</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-36_-33>TACA)	<b>TTTACA</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-36_-33>TACA, -9_-8>AA) *	<b>TTTACA</b> CTGGAGTCGACTCCAGAGT <b>GTATAA</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-35A<C)	<b>TTG</b> <b>CCT</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-34C<A)	<b>TTGA</b> <b>AT</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-33T<G)	<b>TTGAC</b> <b>G</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-33T>A)	<b>TTGAC</b> <b>A</b> CTGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-32C<A)	<b>TTGACT</b> <b>A</b> TGGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-31T<G)	<b>TTGACT</b> <b>C</b> GGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-30G<T)	<b>TTGACT</b> <b>CT</b> GGAGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-29G<T)	<b>TTGACT</b> <b>CTGT</b> AGTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-28A<C)	<b>TTGACT</b> <b>CTGG</b> <b>C</b> GTCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-27G<T)	<b>TTGACT</b> <b>CTGGA</b> <b>T</b> TCGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-26T<G)	<b>TTGACT</b> <b>CTGGAG</b> <b>G</b> CGACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-25^-24insT) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CT</b> GACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-25^-24insTG) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CTG</b> GACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-25C<A)	<b>TTGACT</b> <b>CTGGAGT</b> <b>A</b> GACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-24G<T)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CT</b> ACTCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-23A<C)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CG</b> <b>C</b> TCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-22C<A)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGA</b> <b>A</b> TCCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-21T<G)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGAC</b> <b>G</b> CCAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-20C<A)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>A</b> CAGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-19C<A)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CA</b> AGAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-18A>G) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCT</b> GAGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-18A<C)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CC</b> <b>G</b> AGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-17G<T)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCA</b> <b>T</b> AGT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-16A<C)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAG</b> <b>C</b> GT <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-15G<T)	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAGA</b> <b>TT</b> G <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-14T>G) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAGAG</b> <b>GG</b> T <b>GTATCCT</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-12_-7>GTCAAA) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAGAGT</b> <b>G</b> <b>GTCAAA</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-11_-7>TGACA) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAGAGT</b> <b>GT</b> <b>TGACA</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-9_-8>AA) *	<b>TTGACT</b> <b>CTGGAGT</b> <b>CGACT</b> <b>CCAGAGT</b> <b>GTAT</b> <b>AA</b> TCGGTTA		
<i>P<sub>zntA</sub></i> (-9_-8>AA, -33T>A) *	<b>TTGAC</b> <b>A</b> CTGGAGTCGACTCCAGAGT <b>GTAT</b> <b>AA</b> TCGGTTA		

Figure 3.7 Sequence Alignment of *P<sub>zntA</sub>* Mutations

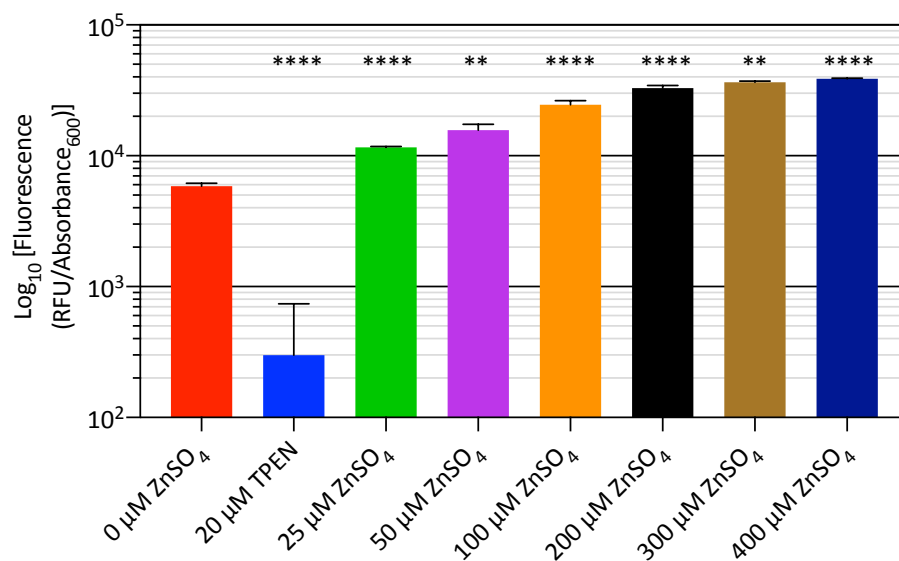
**Bold** indicates -35 or -10 element. **Red** indicates mutation. Underline indicates ZntR binding site. **Green** indicates +1 start site. '\*' indicates mutants which had an assay conducted. Standard mutation nomenclature naming was used.

### 3.2.7. End Point Fluorescence Assays in Lysogeny Broth (LB)

*E. coli*, containing the desired promoter plasmid, was grown in LB with various  $\text{ZnSO}_4$  and IPTG conditions for 16 hr, and then fluorescence and absorbance<sub>600</sub> was recorded, using a TECAN GENios Pro. This data shows the accumulation of fluorescent protein, which can be used as a proxy for promoter activity. As mRFP1 has a half-life of over 24 hr, there should be little to no degradation of fluorescent proteins during the time course of the assay.

Figure 3.8 shows the fluorescence of *P<sub>zntA</sub>* in various concentrations of  $\text{ZnSO}_4$ , as well as 20  $\mu\text{M}$  TPEN. Firstly, the data shows that with the addition of 20  $\mu\text{M}$  TPEN the

promoter activity is significantly reduced, with over 10-fold reduction in fluorescence, compared to LB only. As LB is not zinc depleted, with a concentration of approximately 12.2  $\mu\text{M}$ , a zinc chelator was used to reduce zinc concentration in the media. This data suggests that TPEN is an apt chelator for this purpose, and as previously shown in Figure 3.4, 20  $\mu\text{M}$  TPEN does not cause a toxic effect. Secondly, one can see with the addition of only 25  $\mu\text{M}$   $\text{ZnSO}_4$ , there is a significant increase in fluorescence compared to LB only. Maximal expression is observed with the addition of 400  $\mu\text{M}$   $\text{ZnSO}_4$ , which elicits almost a 10-fold increase in fluorescence compared to LB only. When comparing the increase of fluorescence from LB + 20  $\mu\text{M}$  TPEN to LB + 400  $\mu\text{M}$   $\text{ZnSO}_4$ , there is over 100-fold increase in fluorescence.



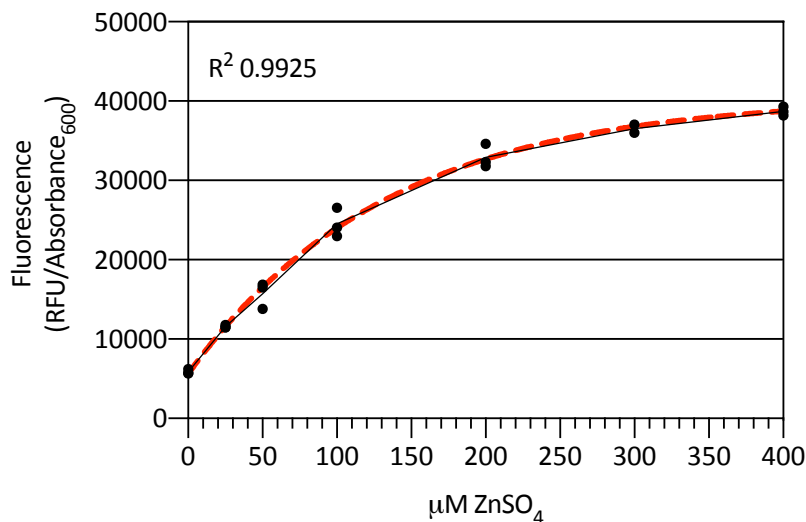
**Figure 3.8 End Point Fluorescence Assay of ZntR Regulated  $P_{zntA}$  in LB with Increasing  $\text{ZnSO}_4$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{LB}_{kan50}$  with increasing  $\text{ZnSO}_4$  concentration and TPEN, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test) comparison to LB only (0  $\mu\text{M}$   $\text{ZnSO}_4$ ).

Data in Figure 3.8 suggests that there is a relationship between increasing  $\text{ZnSO}_4$  concentration and fluorescence. A scatter plot was created to look at the relationship between fluorescence and  $\text{ZnSO}_4$  concentrations (Figure 3.9). A non-linear regression



analysis was conducted on this data, to determine if there is any statistical relationship between fluorescence and  $\text{ZnSO}_4$  concentrations. The non-linear regression returned an  $R^2$  value of 0.9925, suggesting that there is a strong relationship between fluorescence and increasing  $\text{ZnSO}_4$ .



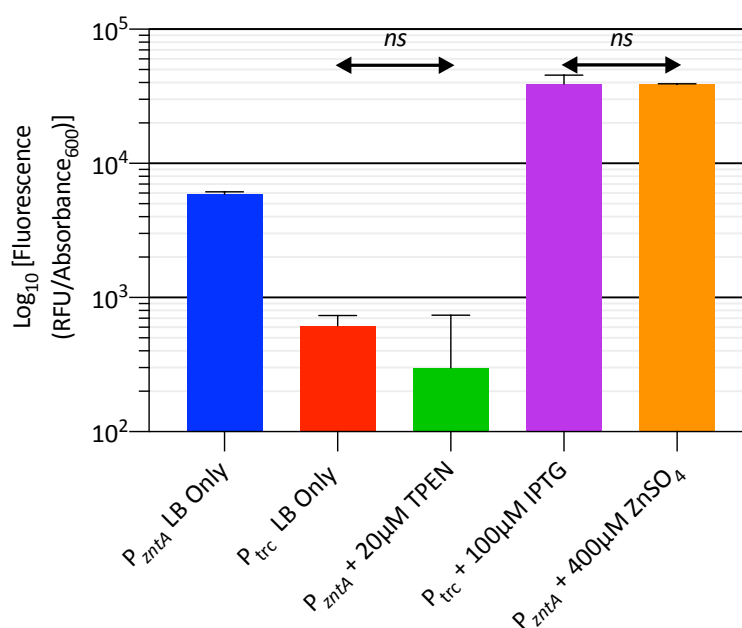
**Figure 3.9 Regression Analysis; End Point Fluorescence of  $P_{zntA}$  in LB**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{LB}_{kan50}$  with increasing  $\text{ZnSO}_4$  concentration, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Dots indicate individual replicates. Solid black line indicated the mean. Red dashed line indicated non-linear regression model (Exponential plateau). Goodness of fit  $R^2$  0.9925

For  $P_{zntA}$  to be considered as a potential novel expression system, it is useful to compare it to an inducible promoter system already used.  $P_{trc}$  was chosen as a control for this purpose.  $P_{trc}$  is a strongly expressed promoter, with low level of basal expression, under the regulation of LacI, and inducible by lactose or the allosteric IPTG.  $P_{trc}$  is a fusion promoter combining the -35 of  $P_{trp}$ , and the -10 of  $P_{lacUV5}$ . (de Boer, Comstock and Vasser, 1983; Brosius, Erfle and Storella, 1985).

Figure 3.10 shows the comparison of fluorescence of  $P_{zntA}$  and  $P_{trc}$ ; both were induced with their respective inducer, 400  $\mu\text{M ZnSO}_4$  or 100  $\mu\text{M IPTG}$ . In addition,  $P_{zntA}$  was repressed, using 20  $\mu\text{M TPEN}$ . The maximum promoter activity of  $P_{zntA} + 400 \mu\text{M ZnSO}_4$ ,

and  $P_{trc} + 100 \mu\text{M IPTG}$  show similar promoter activity levels. There is no significant difference between the maximum expression levels of the two promoters (two-tailed, unpaired, t-test). Further to this, when comparing basal levels of  $P_{zntA} + 20 \mu\text{M TPEN}$ , and  $P_{trc}$  in LB only, there is again no significant difference between basal levels. However, if comparing  $P_{zntA}$  in LB only and  $P_{trc}$ , there is a significantly higher fluorescence observed with  $P_{zntA}$ , with almost 10-fold higher fluorescence than  $P_{trc}$ .



**Figure 3.10 End Point Fluorescence Assay of  $P_{zntA}$  and  $P_{trc}$  in LB Induced or Uninduced**

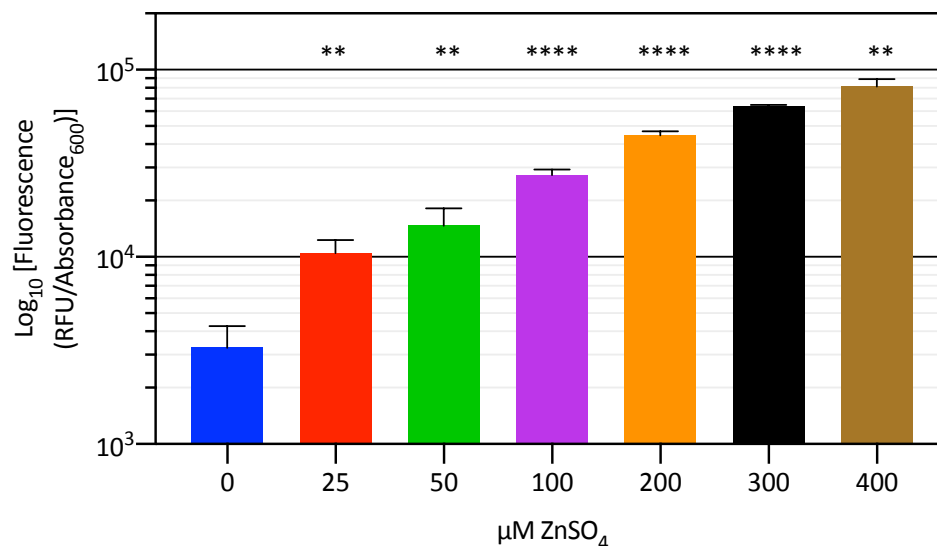
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> with an inducer ( $\text{ZnSO}_4$  for  $P_{zntA}$ , and IPTG for  $P_{trc}$ ) or a repressor (TPEN for  $P_{zntA}$ ), with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. ns = not significant (two-tailed, unpaired, t-test).

### 3.2.8. End Point Assays in NH Broth

Following on from the end point assays in LB broth, end point assays were conducted in the zinc depleted, defined minimal media NH. This allows one to look at promoter activity in minimal zinc conditions. As this is a zinc depleted media, TPEN is not required. When using TPEN in combination with NH, the zinc level is reduced further,

putting a high burden on bacterial growth, often leaving to no, or very minimal growth (data not shown).

Figure 3.11 shows the fluorescence of  $P_{zntA}$  in NH, with increasing concentrations of  $ZnSO_4$ . With the addition of 25  $\mu M$   $ZnSO_4$ , there is a significant increase in fluorescence, which is similar to that observed in LB (Figure 3.8). Further to this, maximum fluorescence is observed with the addition of 400  $\mu M$   $ZnSO_4$ , again similar to that observed in LB (Figure 3.8).

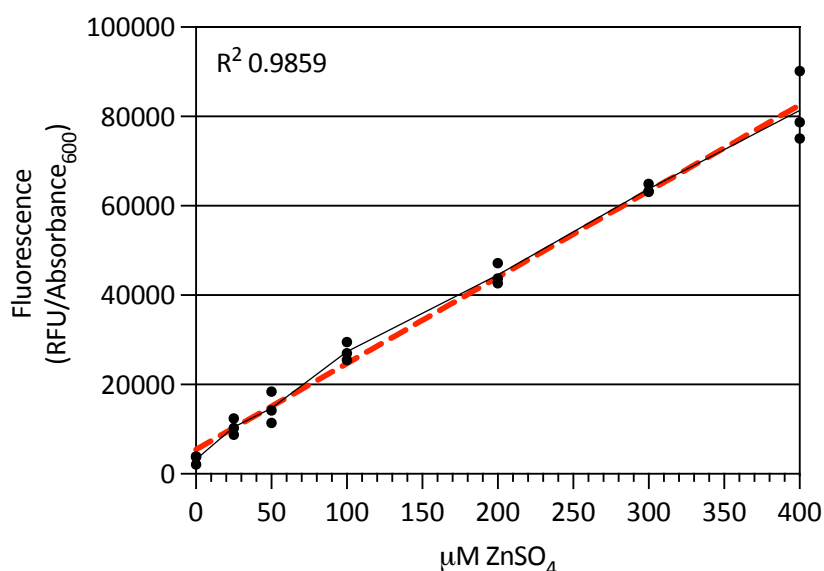


**Figure 3.11 End Point Fluorescence Assay of ZntR Regulated  $P_{zntA}$  in NH with Increasing  $ZnSO_4$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $NH_{kan50}$  with increasing  $ZnSO_4$  concentration, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test) comparison to NH only (0  $\mu M$   $ZnSO_4$ ).

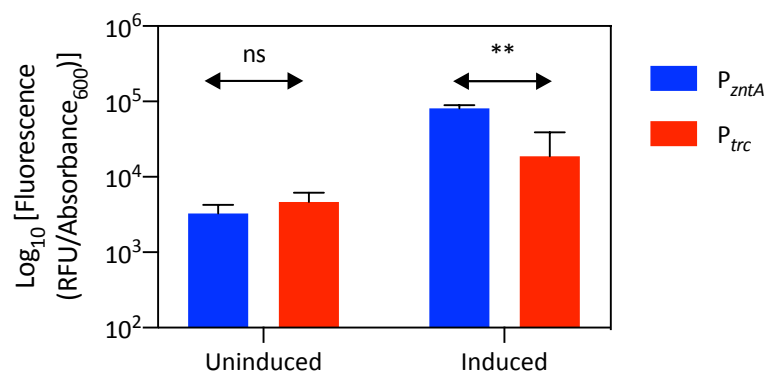
Figure 3.12 shows the correlation between increasing  $ZnSO_4$  and fluorescence, when  $P_{zntA}$  is expressed in NH media. A linear regression analysis was conducted on the data, which shows strong correlation between  $ZnSO_4$  concentration and fluorescence, with an  $R^2$  value of 0.9859, and a slope of +5.284. This shows that there is a strong linear relationship between increasing  $ZnSO_4$  concentration and fluorescence.

When comparing induced and uninduced promoters  $P_{zntA}$  and  $P_{trc}$  in NH, there is no significant difference in basal (uninduced) fluorescence (Figure 3.13). However, there is significantly higher fluorescence in induced  $P_{zntA}$  than in induced  $P_{trc}$ .



**Figure 3.12 Regression Analysis: End Point Fluorescence of  $P_{zntA}$  in NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{NH}_{\text{kan50}}$  with increasing  $\text{ZnSO}_4$  concentration, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Black dots indicate individual replicates. Solid black line indicated the mean. Red dashed line indicated linear regression model. Goodness of fit  $R^2$  0.9859.



**Figure 3.13 End Point Fluorescence Assay of  $P_{zntA}$  and  $P_{trc}$  in NH**

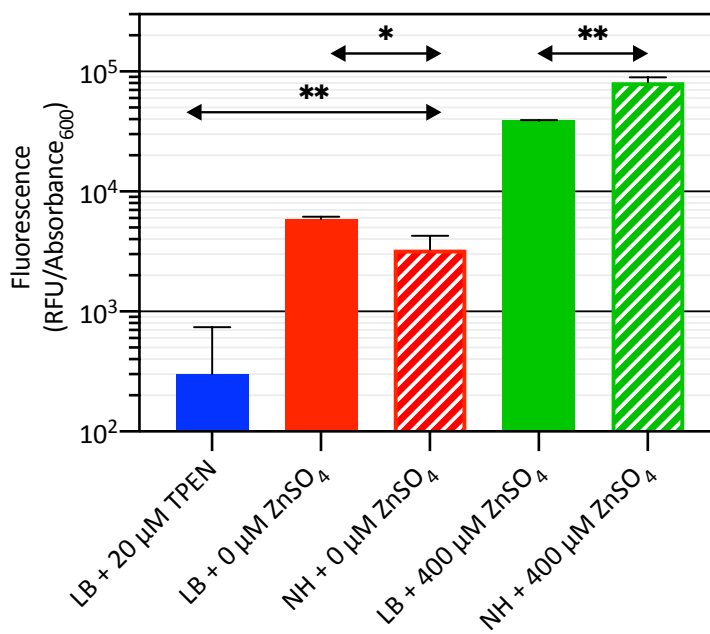
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $NH_{kan50}$  with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Uninduced NH only. Induced  $P_{zntA}$  (+ 400  $\mu$ M  $ZnSO_4$ ). Induced  $P_{trc}$  (+ 100  $\mu$ M IPTG). Error bars indicated S.D. \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test).

### 3.2.9. Comparison of End Point Assays in LB and NH

When comparing the fluorescence of  $P_{zntA}$  in LB and NH, there are disparities in the fluorescence data. Figure 3.14 shows the comparison of fluorescence of  $P_{zntA}$  grown in both LB and NH. Looking at induced  $P_{zntA}$  (400  $\mu$ M  $ZnSO_4$ ), there is a significant difference in fluorescence between expression in LB and NH, with NH showing significantly higher fluorescence. Looking at the uninduced expression, there is a significantly higher fluorescence in LB (LB only), compared to NH, leading to higher basal expression in LB. However, when  $P_{zntA}$  is expressed in LB with 20  $\mu$ M TPEN, this reduces the basal level to 10x lower than NH only, with strong significant differences between the two.

Table 3.2 Shows the fold-increase of the promoter activity of  $P_{zntA}$  and  $P_{trc}$  in basal and induced conditions. The greatest increase in fluorescence can be observed with  $P_{zntA}$  in LB; from LB + 20  $\mu$ M TPEN to LB + 400  $\mu$ M  $ZnSO_4$ , with an average of 129x increase in fluorescence. NH shows a lower increase of 24x, comparing NH only to NH + 400  $\mu$ M  $ZnSO_4$ . The second lowest variation in promoter activity is seen with  $P_{zntA}$ , between LB

only and LB + 400  $\mu$ M, with 6.6x increases in fluorescence. Interestingly, when looking at the fluorescence increase of  $P_{trc}$  from LB only to LB + 100  $\mu$ M IPTG, there is a 63x increase. The lowest increase is observed with  $P_{trc}$  is from NH to NH + 100  $\mu$ M IPTG.



**Figure 3.14 End Point Fluorescence Assay Comparison of  $P_{zntA}$  in Both LB and NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $NH_{kan50}$  or  $LB_{kan50}$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test) comparison between arrows.

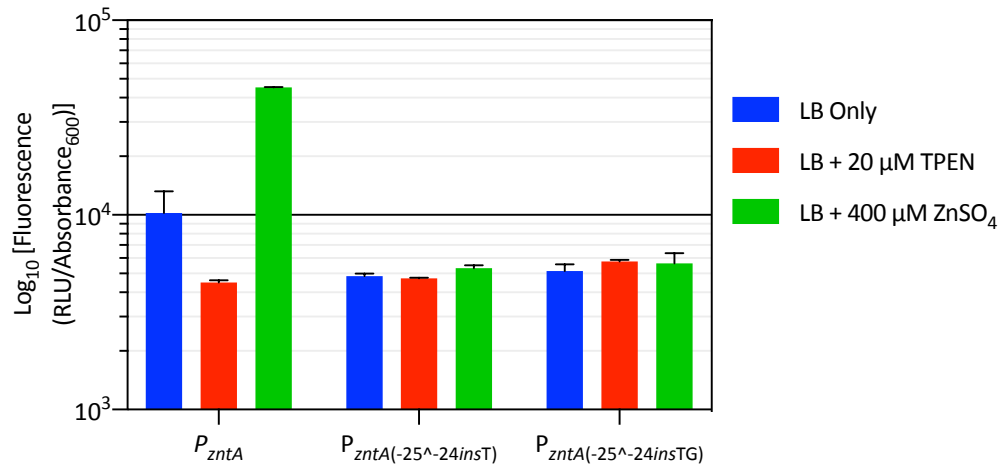
Promoter	Basal Media	Induced Media	Fluorescence Increase ( $x$ -fold)
$P_{zntA}$	LB only	LB + 400 $\mu$ M $ZnSO_4$	6.60
	LB + 20 $\mu$ M TPEN	LB + 400 $\mu$ M $ZnSO_4$	129.23
	NH only	NH + 400 $\mu$ M $ZnSO_4$	24.93
$P_{trc}$	LB only	LB + 100 $\mu$ M IPTG	63.38
	NH only	NH + 100 $\mu$ M IPTG	4.02

**Table 3.2 Fold Increase of Promoter Activity**

Fold increase in fluorescence (promoter activity) of  $P_{zntA}$  and  $P_{trc}$  in different basal and induced conditions.

### 3.2.10. End Point Assay of $P_{zntA}$ Mutants in LB Broth

To determine the effect of increasing the spacer between the -35 and -10 of  $P_{zntA}$ , at the intersect between the two ZntR monomer DNA binding, two  $P_{zntA}$  mutations were made to increase the spacer by either 1 bp or 2 bp;  $P_{zntA(-25\text{^-}24insT)}$  and  $P_{zntA(-25\text{^-}24insTG)}$ . An end point fluorescence assay was conducted using these mutants, using  $P_{zntA}$  *wt* as a control, and expressed in LB only, LB + 20  $\mu$ M TPEN, and LB + 400  $\mu$ M  $ZnSO_4$  (Figure 3.15).  $P_{zntA}$  *wt* shows a decrease in expression with 20  $\mu$ M TPEN and an increased in expression with 400  $\mu$ M  $ZnSO_4$ , as expected. However, when looking at  $P_{zntA(-25\text{^-}24insT)}$  and  $P_{zntA(-25\text{^-}24insTG)}$  there is little to no variation between LB only; 20  $\mu$ M TPEN or 400  $\mu$ M  $ZnSO_4$ . Further, expression levels in all three conditions are similar to repressed  $P_{zntA}$  *wt* in LB + 20  $\mu$ M TPEN.



**Figure 3.15 End Point Fluorescence Assay of  $P_{zntA}$  Mutants: Extended Spacer**

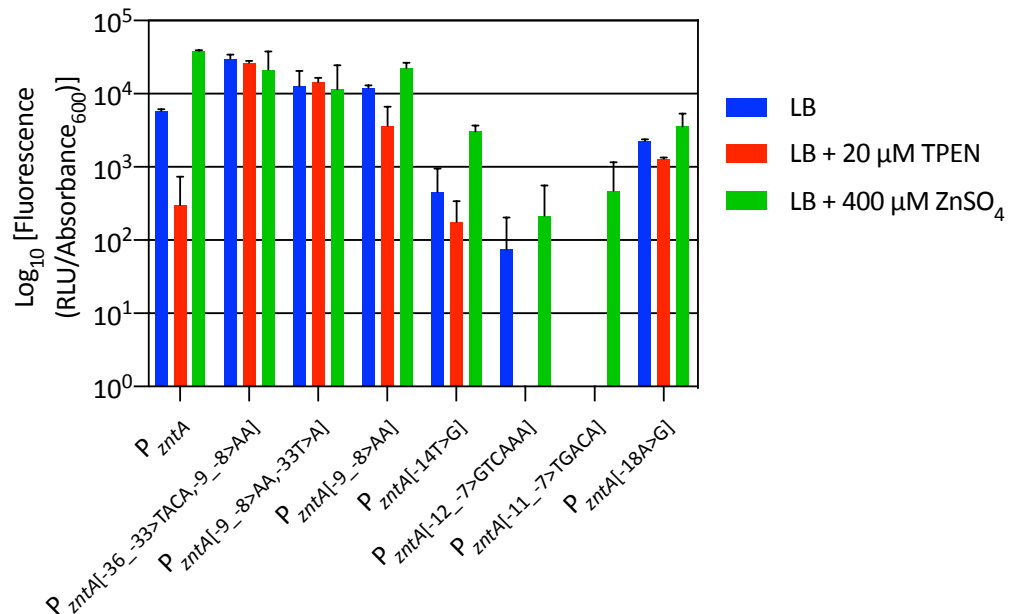
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> with or without 20  $\mu$ M TPEN, or 400  $\mu$ M  $ZnSO_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro.

Additional  $P_{zntA}$  mutations were created, to either increase or decrease their sequence alignment to the consensus promoter sequence.  $P_{zntA(-36_-33>TACA,-9_-8>AA)}$  improved the consensus to both -35 and -10.  $P_{zntA(-18A>G)}$  improved the -18 element consensus.  $P_{zntA(-9_-8>AA)}$  improved the -10 element consensus.  $P_{zntA(-14T>G)}$  improved the TG motif consensus.  $P_{zntA(-9_-8>AA, -33T>A)}$  lead to perfect alignment and sequence of both the -35 and -10 elements.  $P_{zntA(-11_-7>TGACA)}$  disrupted the -10 element by replacing the sequence with the consensus -35 sequence. Finally,  $P_{zntA(-12_-7>GTCAAA)}$  improved the consensus to the -10 element of  $\sigma^{24}$ . A study by Egler *et al.*, (2005) demonstrated that  $\sigma^{24}$  was needed for full resistance to  $Zn^{2+}$ , suggesting that  $\sigma^{24}$  may play a role in *zntA* expression. However, the general understanding of  $P_{zntA}$  is that it is under the regulation of  $\sigma^{70}$ , and if nothing else, this mutation will decrease the consensus to the  $\sigma^{70}$  -10 consensus sequence.

Figure 3.16 shows  $P_{zntA}$  mutations expressed in LB only, LB + 20  $\mu$ M TPEN and LB + 400  $\mu$ M  $ZnSO_4$ . The mutations  $P_{zntA(-36_-33>TACA,-9_-8>AA)}$  and  $P_{zntA(-9_-8>AA, -33T>A)}$  show fluorescence which is similar to induced  $P_{zntA}$  wt in LB + 400  $\mu$ M  $ZnSO_4$  and shows no



variation between the three conditions. This indicates that these two mutations have an increased promoter activity, however, at the same time removed the control element of the promoter. This could possibly suggest that ZntR no longer plays a role in gene regulation.  $P_{zntA(-9_-8>AA)}$  and  $P_{zntA(-18A>G)}$  show little variation between the three conditions. This could suggest these mutations reduced the effect of ZntR, but it still plays a minor role in gene regulation.  $P_{zntA(-14T>G)}$  shows a decrease in fluorescence with TPEN and an increase with  $ZnSO_4$ , however, the variation between repressed and induced fluorescence is less than that of  $P_{zntA}$  *wt*.  $P_{zntA(-12_-7>GTCAAA)}$  shows little variation between LB only and LB + 400  $\mu M$   $ZnSO_4$ , however, interestingly in LB + 20  $\mu M$  TPEN, no fluorescence is detected. Induced  $P_{zntA(-12_-7>GTCAAA)}$  shows similar fluorescence to that of  $P_{zntA}$  *wt* in LB + 20 TPEN, suggesting that the maximum expression is comparatively low.  $P_{zntA(-11_-7>TGACA)}$  shows the most interesting results, showing no detectable fluorescence when expressed in LB only or LB + 20  $\mu M$  TPEN. Fluorescence is only observed when expressed with 400  $\mu M$   $ZnSO_4$ . However, the fluorescence of  $P_{zntA(-11_-7>TGACA)}$  in LB + 400  $\mu M$  shows no significant difference to  $P_{zntA}$  *wt* in LB + 20  $\mu M$  TPEN (two-tailed, unpaired, t-test). This suggests that the mutant promoter  $P_{zntA(-11_-7>TGACA)}$  is still under the regulation of ZntR. However, the mutation reduced the consensus -10 element, leading to reduced basal expression, and now expression is only observed at low level, with the addition of  $ZnSO_4$ , when ZntR is in the active state.



**Figure 3.16 End Point Fluorescence Assay of  $P_{zntA}$  Mutants: Consensus Promoter Elements**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> with or without 20  $\mu$ M TPEN, or 400  $\mu$ M ZnSO<sub>4</sub>, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro.

### 3.2.11. RT-qPCR

For an inducible promoter, it is vital that the observed variations in promoter activity are due to transcriptional variations, rather than translational variations. Reverse Transcription-quantitative PCR (RT-qPCR) was used to determine two outcomes. Firstly RT-qPCR was used to determine if the correlation between increasing promoter activity (as previously measured by RFP end point fluorescence assay) and increasing zinc concentration can be measured at the transcriptional level, and thus confirm that  $P_{zntA}$  regulation is, at least partially, controlled by transcriptional regulation. Secondly, to determine if the RFP end point fluorescence assays are a reliable and informative proxy of transcriptional regulation of promoter activity.

This study used hydrolysis probes, also known as TaqMan. Hydrolysis probes bind to DNA between a primer set. These probes consist of a fluorophore at the 5' end, and a quencher at the 3' end. On successful amplification on DNA, endonuclease activity of

Taq polymerase cleaves the fluorophore from the probe, causing fluorescence. A qPCR thermocycler measures the fluorescence at the end of each PCR cycle, quantifying PCR product concentration. As the product is amplified exponentially each cycle, the fluorescence intensity increases in direct relationship with the PCR product concentration (Holland *et al.*, 1991). This can be coupled with reverse transcription, which allows qPCR to be conducted on RNA. RNA is extracted from a sample, and is reverse transcribed to cDNA, which is then used as a template for qPCR. A  $C_q$  (also known as  $C_t$ ) value is given to each reaction.  $C_q$  is the value of the cycle number in which the fluorescent intensity exceeds an arbitrarily appointed value.

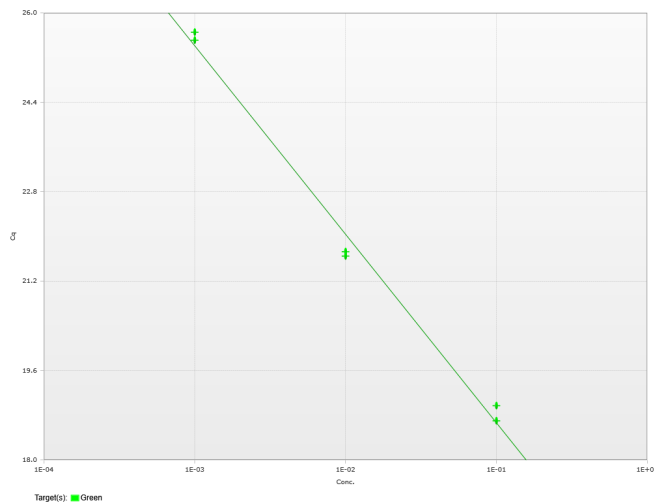
#### 3.2.11.1. RT-qPCR Validation: Calibration Curve

A calibration curve is used to determine the efficiency of the qPCR reaction; theoretically, the PCR product should double every cycle. A known concentration of nucleic acid is diluted  $10^{-1}$  to  $10^{-3}$ , which is then amplified through the qPCR thermocycle, and a  $C_q$  value determined. As PCR products double every cycle, with a 10-fold dilution of template nucleic acid, there should be a  $C_q$  value difference of 3.322 between each log dilution.  $C_q$  values are plotted against nucleic acid dilution, and linear regression analysis is conducted to determine the  $R^2$  value. Efficiency between 90-110% is seen to be within the acceptable range. Variations in efficiency are often accepted due to human errors, such as pipetting errors (Bustin *et al.*, 2009).

A calibration curve was conducted for each primer/probe set used in this work (*rrsA*, and *zntA*). *rrsA* is shown below, as an example. *zntA* showed efficiency between 90-110% (data not shown).

1 ng of total RNA extract was diluted 10-fold from  $10^{-1}$  to  $10^{-3}$ . RT-qPCR was set up using primers and probe for the reference gene *rrsA*. RT-qPCR was conducted, and  $C_q$

values were determined. Figure 3.17 shows  $C_q$  value against log dilution of total RNA. Linear regression analysis was conducted and returned an  $R^2$  value of 0.991, and a qPCR efficiency of 98% (determined using Q-Rex software, Qiagen).



**Figure 3.17 RT-qPCR Calibration Curve of *rrsA***

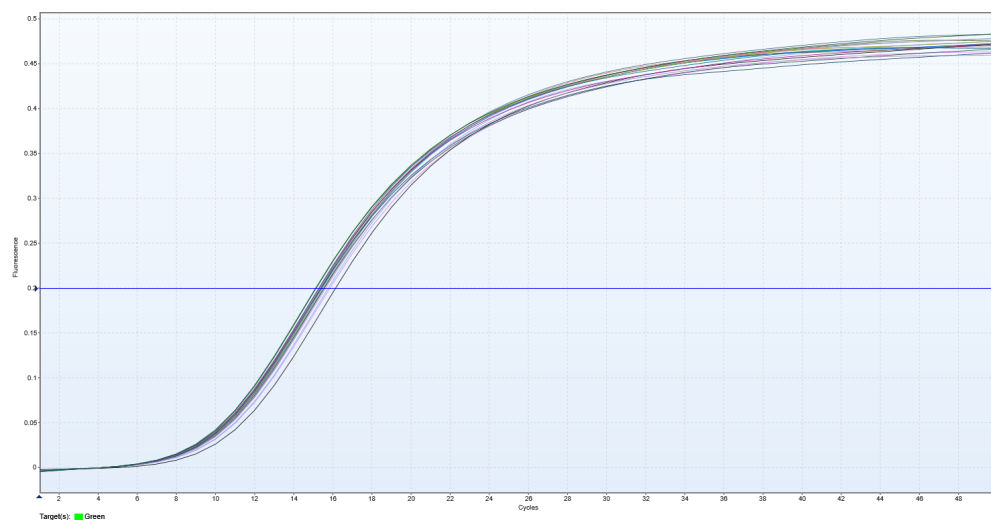
Calibration curve of *rrsA* reference gene. *E. coli* MG1655 was grown with aerobic incubation at 37°C with shaking (200 RPM) in NH. Total RNA extracted after 16 hr incubation. RT-qPCR conducted with primers (*rrsA*\_qPCR\_F, *rrsA*\_qPCR\_R) and probe (*rrsA*\_qPCR\_Probe), using template RNA (1 ng) dilutions from  $10^{-1}$  to  $10^{-3}$ , and two technical repeats. Green fluorescence was recorded after each cycle, and  $C_q$  value appointed.  $R^2$  0.99, efficiency 98% determined through Q-Rex software (Qiagen, Hilden Germany)

### 3.2.12. RT-qPCR: mRNA Variation of *zntA*

To determine if there is a transcriptional variation of *zntA* expression in various  $ZnSO_4$  conditions, RT-qPCR was used. The experimental set up involved growing *E. coli* MG1655 in NH media with either 0, 200 or 400  $\mu M$   $ZnSO_4$  for 16 hr, with three biological repeats. A 16 hr time point was chosen, as it correlates with the 16 hr time point reading of the RFP end fluorescence assay. [It is worth noting that an experiment was planned which conducted RT-qPCR readings at hourly intervals between 0-4 hr to correlate with the luminescence temporal assays. However, due to the COVID-19 global pandemic, my experimental lab time was cut short, and this experiment was never conducted]. Total RNA was extracted, RNA concentration determined, and

diluted to  $1 \text{ ng } \mu\text{L}^{-1}$ . RT-qPCR was performed on all samples using both *zntA* and *rrsA* primer/probe sets, with two technical repeats.

Firstly, *rrsA* was validated as a reference gene. Figure 3.18 shows fluorescence of *rrsA*, against cycle number; data plotted included all three conditions (0, 200, 400  $\mu\text{M}$   $\text{ZnSO}_4$ ), and three biological and two technical repeats. This graph shows the RT-qPCR plot lines are tight together, with little variation, indicating that the zinc concentration used has had no effect on the mRNA concentration of *rrsA*. There is less than 0.5  $C_q$  value difference between samples. *rrsA* can be used as a reference gene for this study, in the context of  $\text{ZnSO}_4$  conditions; further validation would be needed if to be used in other conditions. *rrsA* has also previously been used as a reference gene in similar studies (Zhou *et al.*, 2011; Peng *et al.*, 2014; Xu *et al.*, 2019).

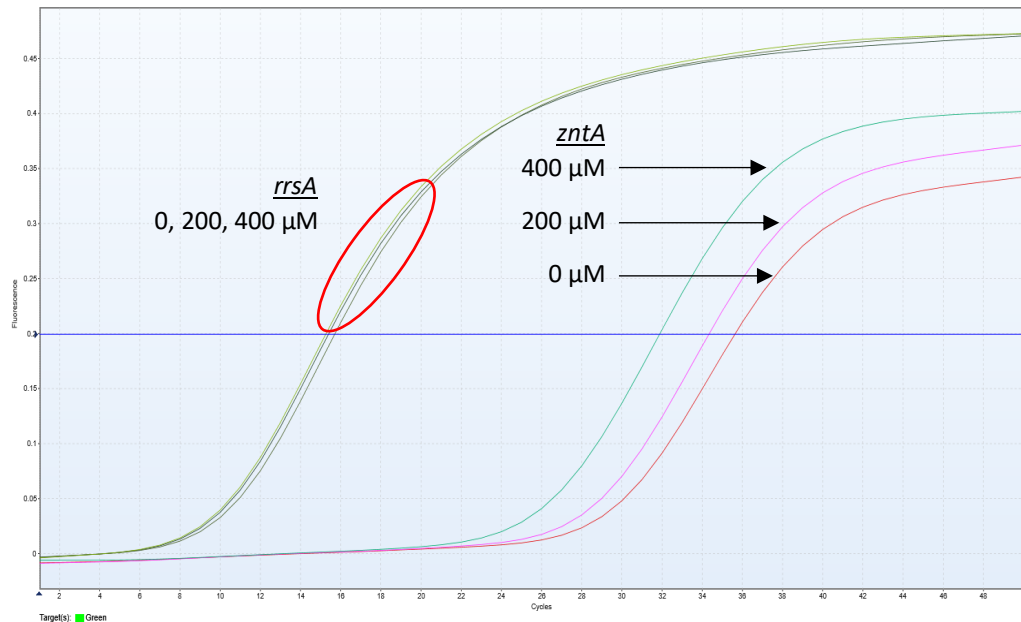


**Figure 3.18 RT-qPCR Plot of *rrsA***

*E. coli* MG1655 aerobically incubation at  $37^\circ\text{C}$  with shaking (200 RPM) in NH + (0, 200, or 400  $\mu\text{M}$   $\text{ZnSO}_4$ ), with three biological repeats. Total RNA extracted after 16 hr incubation. RT-qPCR conducted with primers (*rrsA*\_qPCR\_F, *rrsA*\_qPCR\_F, *rrsA*\_qPCR\_Probe), with two technical repeats, recording green fluorescence after each cycle. Solid blue line indicates  $C_q$  threshold.

The average of the three biological repeats for each sample were plotted on a single graph (Figure 3.19). This figure shows that all samples have a smooth sigmoidal curve, which indicates correct RT-qPCR performance. There is variation in *rrsA* and a much

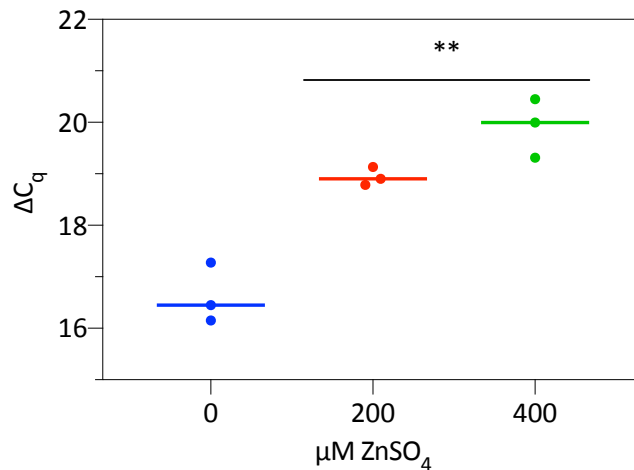
greater variation in *zntA* samples. It is clear that there is variation in *zntA* mRNA between three  $\text{ZnSO}_4$  concentrations; that there is more mRNA in 400  $\mu\text{M}$ , than in 200  $\mu\text{M}$  or 0  $\mu\text{M}$   $\text{ZnSO}_4$ .



**Figure 3.19 qPCR Plot for of *zntA* and *rrsA***

*E. coli* MG1655 aerobically incubation at 37°C with shaking (200 RPM) in NH + (0, 200, or 400  $\mu\text{M}$   $\text{ZnSO}_4$ ), with three biological repeats. Total RNA extracted after 16 hr incubation. RT-qPCR conducted with primers for *rrsA* (*rrsA\_qPCR\_F*, *rrsA\_qPCR\_F*, *rrsA\_qPCR\_Probe*) and *zntA* (*zntA\_qPCR\_F*, *zntA\_qPCR\_R*, *zntA\_qPCR\_Probe*), with two technical repeats, recording green fluorescence after each cycle. Solid blue line indicates  $C_q$  threshold.

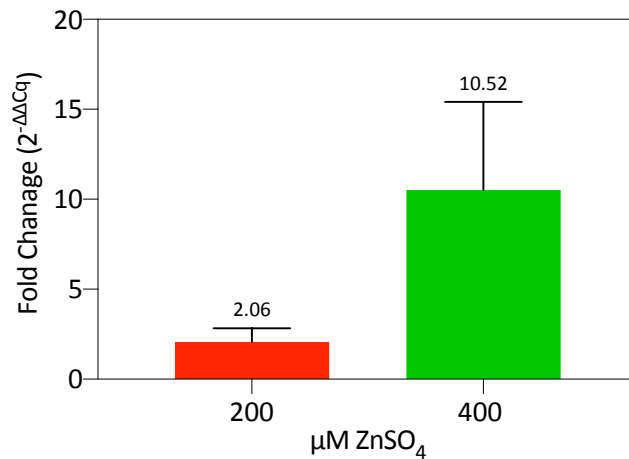
$C_q$  values were determined; technical repeats for each biological were averaged, leaving three biological repeats for each condition;  $\Delta C_q$  values were determined (*zntA-rrsA*) and plotted. Figure 3.20.  $\Delta C_q$  values shows there is a significant difference in  $C_q$  values with the addition of  $\text{ZnSO}_4$  at both 200  $\mu\text{M}$  and 400  $\mu\text{M}$ , compared to NH only (0  $\mu\text{M}$   $\text{ZnSO}_4$ ). The increased mRNA concentration in relationship with the increase of  $\text{ZnSO}_4$  corroborates the data seen in Figure 3.8 and Figure 3.11.



**Figure 3.20  $\Delta C_q$  Plot for mRNA of *zntA***

$C_q$  values for *zntA* and *rrsA* from RT-qPCR plots.  $\Delta C_q$  (*zntA-rrsA*) was plotted, showing biological repeats and mean. \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test) comparison to un-induced (0  $\mu\text{M ZnSO}_4$ ).

To further analyse the RT-qPCR data, the comparative  $2^{-\Delta\Delta C_q}$  method was used. This analyses the fold change in expression of a gene compared to a control, considering the reference gene as a standard. Figure 3.21 shows the fold change of mRNA concentration of the gene *zntA*. When induced with 200  $\mu\text{M ZnSO}_4$  *zntA* is over 2-fold higher than when grown in NH only (0  $\mu\text{M ZnSO}_4$ ). When induced with 400  $\mu\text{M ZnSO}_4$ , there is over 10-fold increase in mRNA concentration, compared to when grown in NH only. Again, this corroborates the data observed in Figure 3.8 and Figure 3.11. This further strengthens the argument that the fluorescence variation, recorded through fluorescence assay, is due to transcriptional regulation.



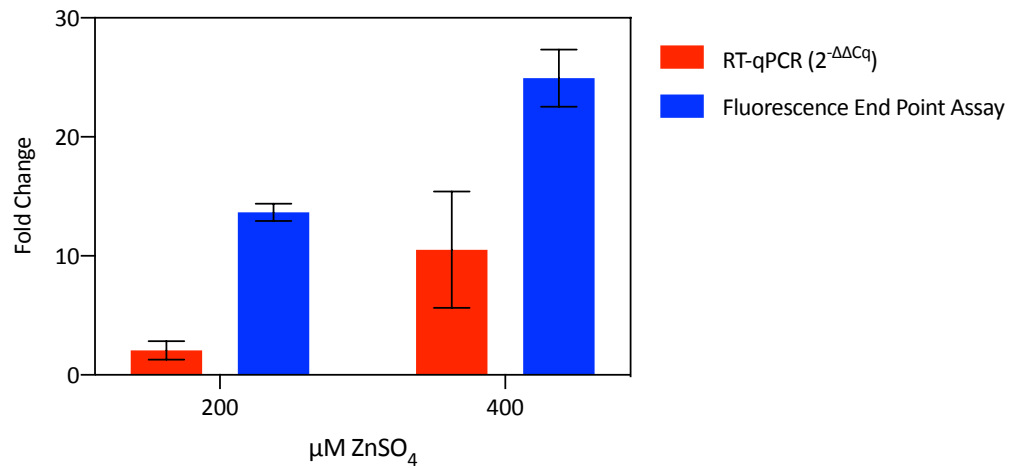
**Figure 3.21 Comparative  $C_q$  ( $2^{-\Delta\Delta C_q}$ ) Plot for mRNA of *zntA***

$C_q$  values previously determined for *zntA* and *rrsA* from qPCR plots.  $2^{-\Delta\Delta C_q}$  mean was plotted, using three biological repeats. Mean fold change shown above column.

To directly compare and contrast the experimental methods of *zntA* gene regulation (RT-qPCR and end point fluorescence assays), fold change was calculated for both experimental data sets. The fold change ( $2^{-\Delta\Delta C_q}$ ) of RT-qPCR was previously calculated in Figure 3.21, and the fold change was calculated with end point fluorescence assay data from Figure 3.11, and plotted onto a single graph (Figure 3.22). Both experiments show a significant increase in fold change from 200  $\mu\text{M}$  to 400  $\mu\text{M}$   $\text{ZnSO}_4$  (two-tailed, unpaired, t-test). If one looks at the increase (or difference) in fold change from 200  $\mu\text{M}$  to 400  $\mu\text{M}$   $\text{ZnSO}_4$  for each individual experiment ([RT-qPCR =  $10.52 - 2.06 = 8.46$ ] [end point fluorescence =  $24.93 - 13.65 = 11.28$ ]), there is no significant difference in these increases (two-tailed, unpaired, t-test). One could also work out the fold change between 200  $\mu\text{M}$  and 400  $\mu\text{M}$   $\text{ZnSO}_4$  of each experiment [(RT-qPCR =  $10.52 / 2.06 = 5.11$ ] [end point fluorescence =  $24.93 / 13.65 = 1.83$ ]), which does show a significant difference (two-tailed, unpaired, t-test). This data shows that the increase in fold change of the two experiments increases the same. However, the fold change increase in RT-qPCR is relatively larger than end point fluorescence, as the fold change with 200  $\mu\text{M}$   $\text{ZnSO}_4$  in RT-qPCR is lower than the fold change with 200  $\mu\text{M}$   $\text{ZnSO}_4$  in end



point fluorescence. This data does suggest a relationship between RT-qPCR data and end point fluorescence data, which support the use of mRFP1 end point fluorescence assay as a meaningful proxy of transcriptional gene regulation.



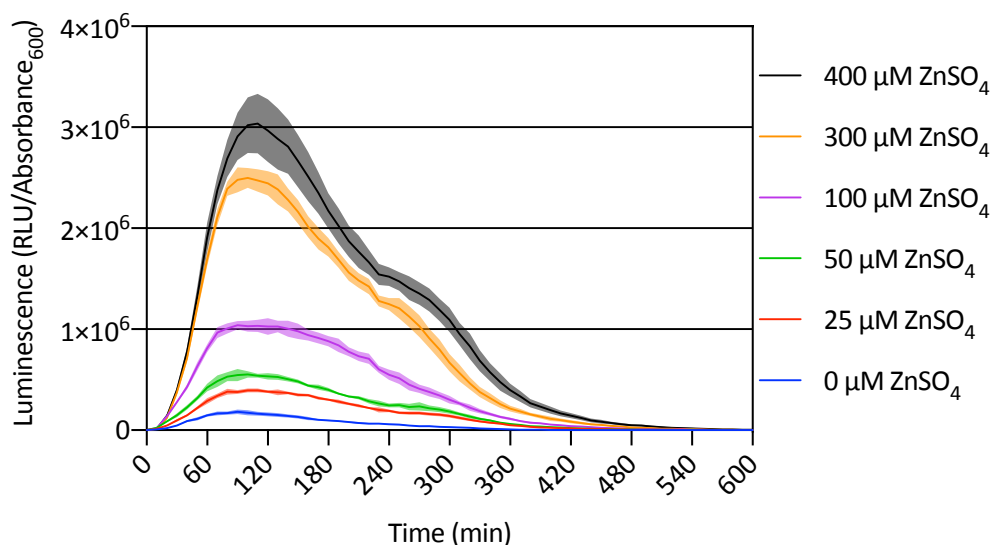
**Figure 3.22 Comparison of Fold Change; RT-qPCR vs End Point Fluorescence**

Fold change in *zntA* expression, as recorded by RT-qPCR and end point fluorescence assay. RT-qPCR fold change data Figure 3.21. End point fluorescence fold change data from Figure 3.11.

### 3.2.13. Temporal Luminescence Assay in NH Broth

To further understand the promoter activity of  $P_{zntA}$ , temporal luminescence assays were conducted. This allows further understanding of the promoter activity in relationship with bacterial growth phase.

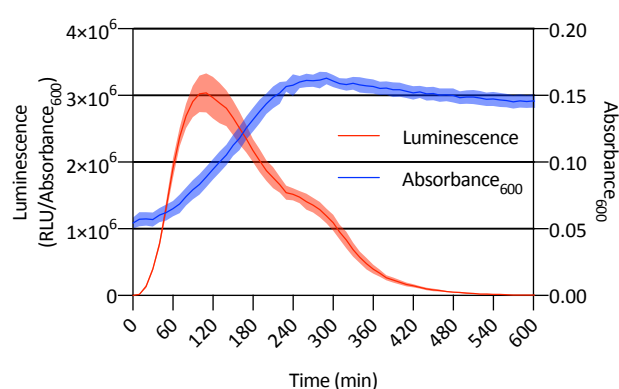
Using the  $P_{zntA}$  plasmid pJI301, *E. coli* was grown in NH, in various  $\text{ZnSO}_4$  concentrations, for 10 hr in a 96 well plate reader. Absorbance, fluorescence, and luminescence readings were taken every 10 min (Figure 3.23). Figure 3.23 shows that, like fluorescence assays, there is a relationship between increasing  $\text{ZnSO}_4$  concentrations and promoter activity, as measured by luminescence or fluorescence. Luminescence is highest when induced with 400  $\mu\text{M}$   $\text{ZnSO}_4$ , and lowest with no induction (0  $\mu\text{M}$   $\text{ZnSO}_4$ ). At maximum promoter activity, which is observed at around 120 min for all added  $\text{ZnSO}_4$  concentrations, there is a clear relationship between increasing luminescence and  $\text{ZnSO}_4$ .



**Figure 3.23 Temporal Luminescence Assay of ZntR Regulated  $P_{zntA}$  in NH Media**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in  $NH_{kan50}$  with increasing  $ZnSO_4$  concentrations, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader.

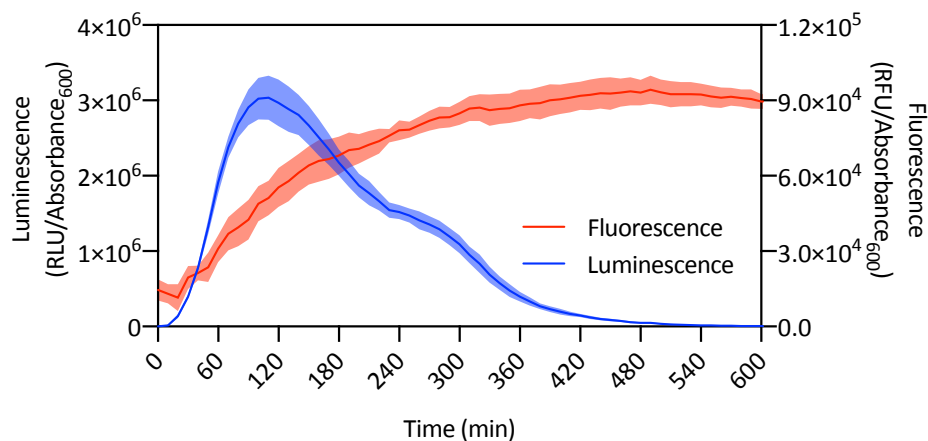
Using the example of  $P_{zntA}$  in NH + 400  $\mu M$   $ZnSO_4$ , an overlay graph was produced, showing absorbance and luminescence (Figure 3.24). From this overlay graph, it can be observed that maximal expression is reached during exponential growth phase, and as stationary phase occurs, regulation of  $P_{zntA}$  drops off.



**Figure 3.24 Temporal Luminescence Assay of  $P_{zntA}$  In NH + 400  $\mu M$   $ZnSO_4$ : Absorbance and Luminescence**

Temporal expression of  $P_{zntA}$  in NH + 400  $\mu M$   $ZnSO_4$ . Left Y-axis shows luminescence (RLU/Absorbance<sub>600</sub>) in red. Right Y-axis shows Absorbance<sub>600</sub> in blue.

Figure 3.25 shows both fluorescence and luminescence, using the example  $P_{zntA}$  in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ . This graph shows that maximum fluorescence is observed at around 480 min, which is around 360 min after the maximum luminescence is observed. At 480 min, fluorescence stops increasing, and starts to fall, likely due to RFP protein degradation. Also, at 480 min, luminescence appears to be zero, indicating that promoter activity is no longer occurring. Figure 3.24 shows that at 480 min, the bacteria have been in station phase for around 3 hr already, suggesting that promoter activity stopped before 480 min; the increasing fluorescence observed can be due to maturation time of RFP, which falsely suggests that promoter activity is occurring, where, in fact, it is not.

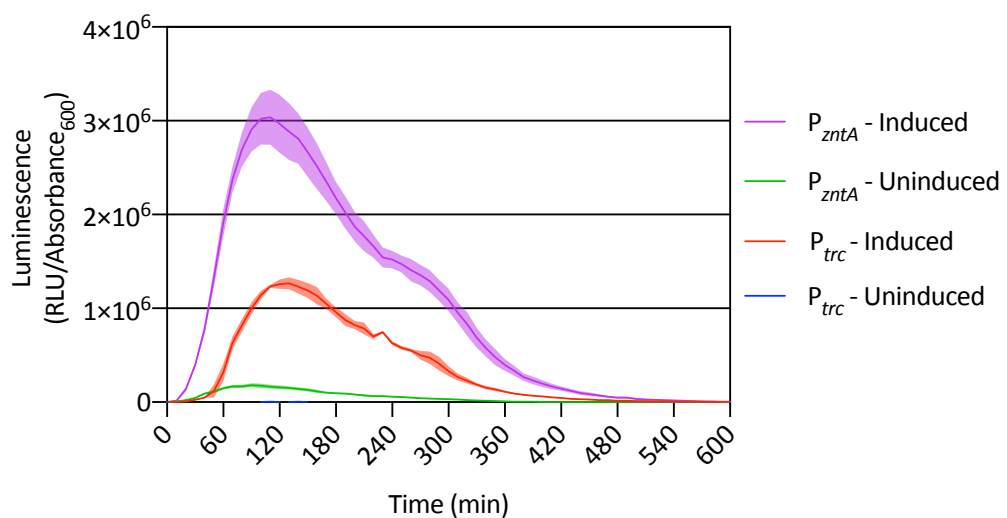


**Figure 3.25 Temporal and Accumulative Assay of  $P_{zntA}$  in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$**

Dual reporter: Luminescence and fluorescence assay of  $P_{zntA}$  in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ , over 16 hr. Left Y-axis shows Luminescence (RLU/Absorbance<sub>600</sub>) in blue. Right Y-axis shows Fluorescence (RFU/Absorbance<sub>600</sub>) in red.

From Figure 3.23 it can be said that maximum luminescence is  $P_{zntA}$  in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ , and minimum luminescence is  $P_{zntA}$  in NH only (0  $\mu\text{M}$   $\text{ZnSO}_4$ ). With this considered, luminescence of  $P_{zntA}$  (0 and 400  $\mu\text{M}$   $\text{ZnSO}_4$ ) was plotted alongside the control promoter  $P_{trc}$  (0 and 100  $\mu\text{M}$  IPTG) (Figure 3.26). From comparison of  $P_{trc}$  and  $P_{zntA}$ , induced  $P_{zntA}$  shows a higher maximum luminescence at all time points from 0 min to 600 min, compared to induced  $P_{trc}$ . However, the uninduced luminescence of  $P_{zntA}$

is also higher than uninduced  $P_{trc}$  at all time points. Again, this corroborates the previous end point fluorescence data; that  $P_{zntA}$  has a higher basal level of expression than  $P_{trc}$ . Further, the time point with the maximum luminescence of both  $P_{zntA}$  and  $P_{trc}$  is approximately the same, at 120 min, and is followed by a drop in luminescence. As the drop-in luminescence is observed in both  $P_{zntA}$  and  $P_{trc}$ , this drop is unlikely an artefact of this  $P_{zntA}$  promoter dynamics, and most likely due to the decrease of luciferase substrate, FMNH<sub>2</sub> and RCHO.



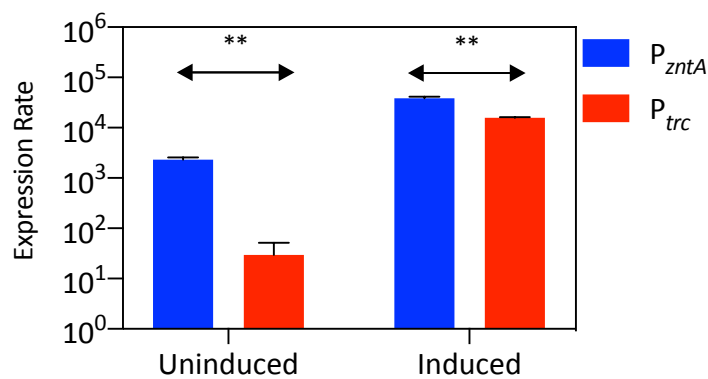
**Figure 3.26 Temporal Luminescence Assay of  $P_{zntA}$  and  $P_{trc}$  in NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in NH<sub>kan50</sub>, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader. Induced  $P_{zntA}$ , 400  $\mu$ M ZnSO<sub>4</sub>. Induced  $P_{trc}$ , 100  $\mu$ M IPTG.

To provide further information on promoter activity, beyond that of end point fluorescence assays, the maximum rate of expression (RLU/Absorbance<sub>600</sub> min<sup>-1</sup>) was analysed from luminescence data. Each sample had the maximum rate of expression evaluated, using a simple linear regression analysis. Statistical tests were then conducted on the rate of expression, as a means of comparison.

Figure 3.27 shows the average rate of expression for  $P_{zntA}$  and  $P_{trc}$  induced and uninduced. All have an R<sup>2</sup> value or 0.95 of greater, besides uninduced  $P_{trc}$  which had

an average  $R^2$  of 0.56. This low  $R^2$  of uninduced  $P_{trc}$  is likely due to the fact the expression rate is so low; there is greater variation in data due to noise (data not shown). Induced  $P_{zntA}$  shows a significantly higher expression rate compared to induced  $P_{trc}$ . However, uninduced  $P_{zntA}$  also has a significantly higher expression rate than uninduced  $P_{trc}$ .



**Figure 3.27 Luminescence Expression Rate of  $P_{zntA}$  and  $P_{trc}$  in NH**

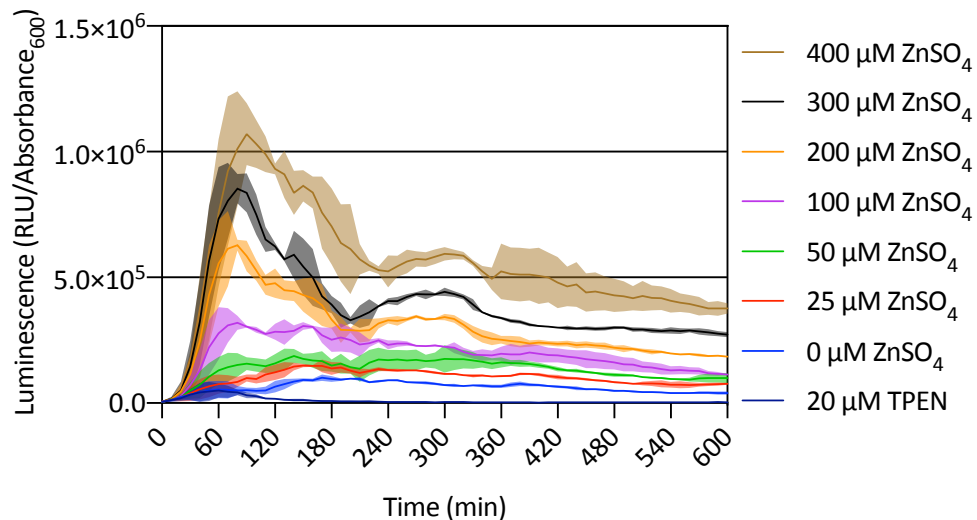
Expression rate (RLU/Absorbance<sub>600</sub> min<sup>-1</sup>) of  $P_{zntA}$  and  $P_{trc}$  in NH with or without an inducer. Induced  $P_{zntA}$ , 400  $\mu$ M ZnSO<sub>4</sub>. Induced  $P_{trc}$ , 100  $\mu$ M IPTG. Error bars indicated S.D. \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test)

The combination of end point fluorescence data and temporal luminescence data of  $P_{zntA}$  suggests that induced  $P_{zntA}$  has a significantly higher promoter activity than induced  $P_{trc}$ . However, the uninduced (or basal)  $P_{zntA}$  promoter activity is the same in end point assays, but significantly higher in temporal assays, compared to uninduced  $P_{trc}$ .

### 3.2.14. Temporal Luminescence in LB

Following on from the luminescence assays of  $P_{zntA}$  in NH, luminescence assays were conducted in LB media. Similar to the luminescence results in NH (Figure 3.23) maximum luminescence is observed in LB, with the addition of 400  $\mu$ M ZnSO<sub>4</sub>, and the minimum expression with 20  $\mu$ M TPEN (Figure 3.28). Further to this, there seems to

be a relationship between luminescence and  $\text{ZnSO}_4$  concentrations: increasing  $\text{ZnSO}_4$  concentrations increases luminescence.



**Figure 3.28 Temporal Luminescence Assay of ZntR regulated  $P_{zntA}$  in LB**

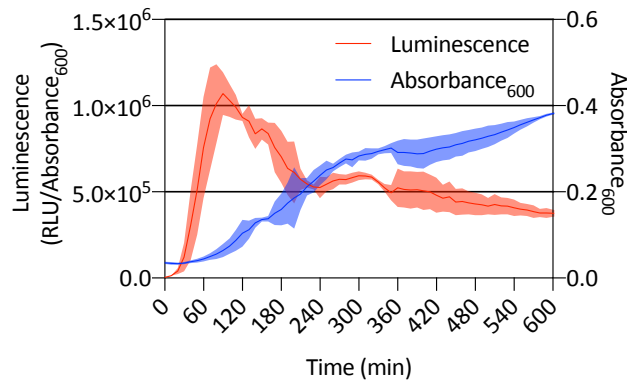
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in  $\text{LB}_{\text{kan50}}$  with increasing  $\text{ZnSO}_4$  concentrations, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader.

Using  $P_{zntA}$  + 400  $\mu\text{M}$   $\text{ZnSO}_4$  as an example, luminescence was compared to absorbance.

Figure 3.29 shows that maximum luminescence is observed during exponential growth phase. There is a reduction in growth rate at around 360 min, which also coincides with the plateau of luminescence. Interestingly, unlike the luminescence assay conducted in NH media (Figure 3.24), when grown in LB, luminescence is relatively stable after 240 min and is still expressing, whereas in NH, after 540 min, there appears to be little, or no luminescence.

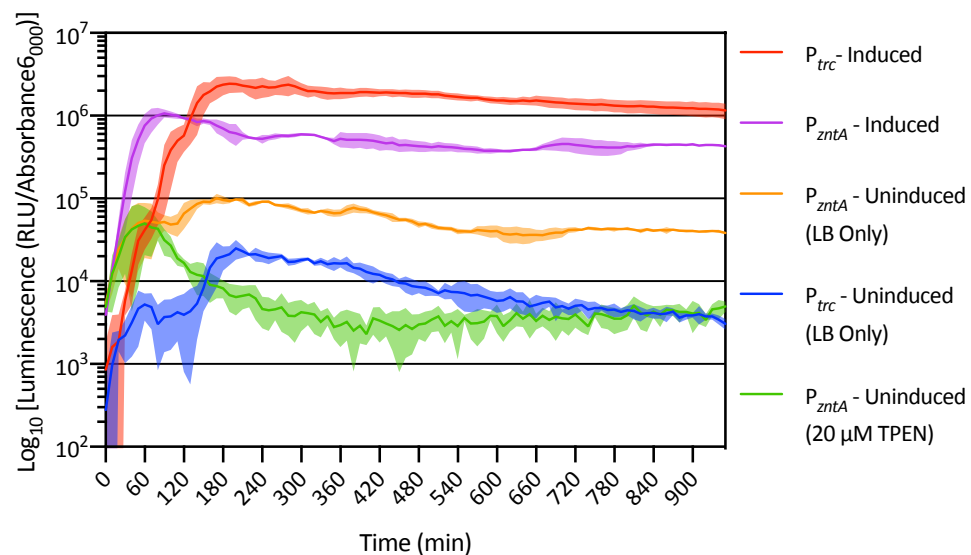
Again, it is useful to compare the promoter activity of  $P_{zntA}$  to the standard inducible promoter,  $P_{trc}$  (Figure 3.30) [note a  $\text{Log}_{10}$  graph was plotted as the variation between samples could not be seen on a linear graph]. Figure 3.30 shows that induced  $P_{trc}$  has a higher level of luminescence, compared to induced  $P_{zntA}$ , and that both induced promoters follow a similar expression pattern - an initial peak, followed by a constant

expression level. It is worth noting that  $P_{zntA}$  appears to reach maximum luminescence faster than  $P_{trc}$ , and that  $P_{zntA}$  stabilises after 240 min, whereas  $P_{trc}$  has a maximum expression, approximately 120 min later than  $P_{zntA}$ .



**Figure 3.29 Temporal Luminescence Assay of  $P_{zntA}$  in LB + 400  $\mu$ M  $ZnSO_4$ : Absorbance and Luminescence**

Temporal luminescence assay of  $P_{zntA}$  in LB + 400  $\mu$ M  $ZnSO_4$ . Left Y-axis shows luminescence (RLU/Absorbance<sub>600</sub>) in red. Right Y-axis shows absorbance<sub>600</sub> in blue.

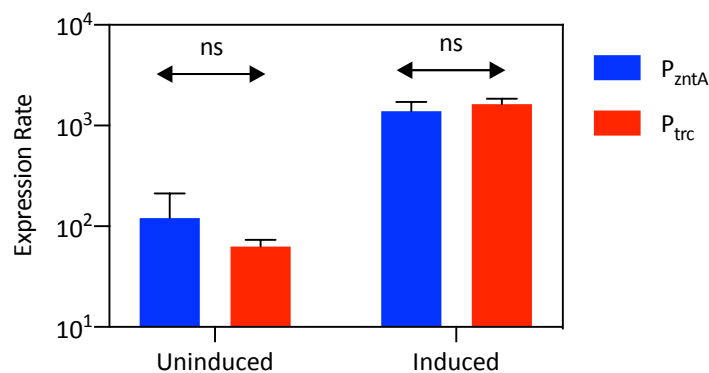


**Figure 3.30 Temporal Luminescence of  $P_{zntA}$  and  $P_{trc}$  in LB**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in LB<sub>kan50</sub>, with three biological repeats. Absorbance<sub>600</sub> and luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader. Induced  $P_{zntA}$ , 400  $\mu$ M  $ZnSO_4$ . Induced  $P_{trc}$ , 100  $\mu$ M IPTG. Uninduced either LB only or LB + 20  $\mu$ M TPEN.

Uninduced  $P_{zntA}$  (20  $\mu$ M TPEN) has the lowest constant luminescence, but has a relatively high peak expression at 60 min. Uninduced  $P_{trc}$  has a lower peak luminescence at 180 min, but has mostly higher stable luminescence, compared to uninduced  $P_{zntA}$  (20  $\mu$ M TPEN). Besides the higher peak at 60 min, uninduced  $P_{zntA}$  (20  $\mu$ M TPEN) shows the lowest luminescence after 150 min. When looking at uninduced  $P_{zntA}$  (LB only), luminescence is higher at all time points, compared to uninduced  $P_{trc}$  and uninduced  $P_{zntA}$  (20  $\mu$ M TPEN). Interestingly, the peak luminescence of uninduced  $P_{zntA}$  (LB only) and uninduced  $P_{zntA}$  (20  $\mu$ M TPEN) are observed at around 60 min and show similar luminescence levels. However, the luminescence of uninduced  $P_{zntA}$  (LB only) remains approximately at the maximum luminescence for the remainder of the assay, whereas uninduced  $P_{zntA}$  (20  $\mu$ M TPEN) luminescence drops over 10-fold.

When comparing expression rates of  $P_{zntA}$  and  $P_{trc}$  in LB, there is no significant difference between the induced states (Figure 3.31). There is also no significant difference between uninduced  $P_{trc}$  and uninduced  $P_{zntA}$  (20  $\mu$ M TPEN).



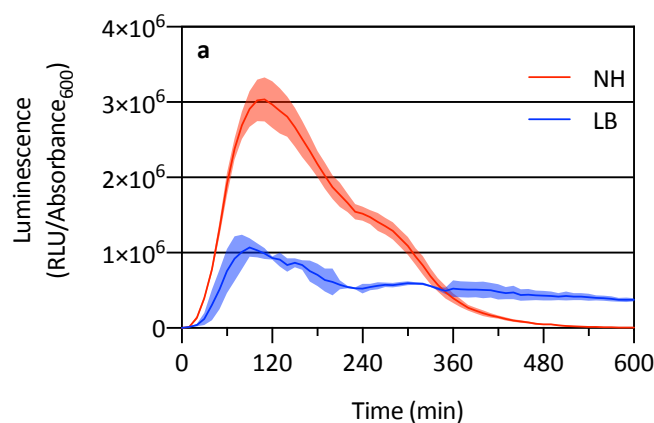
**Figure 3.31 Luminescence Expression Rate of  $P_{zntA}$  and  $P_{trc}$  in LB**

Expression rate (RLU/Absorbance<sub>600</sub> min<sup>-1</sup>) of  $P_{zntA}$  and  $P_{trc}$  in LB with or without inducer. Induced  $P_{zntA}$  (+ 400  $\mu$ M ZnSO<sub>4</sub>). Induced  $P_{trc}$  (+ 100  $\mu$ M IPTG). Uninduced. Error bars indicated S.D. ns  $p \geq 0.05$  (two-tailed, unpaired, t-test).



### 3.2.15. Temporal Luminescence: Comparison of LB and NH

To further compare the luminescence, induced  $P_{zntA}$  (400  $\mu\text{M}$   $\text{ZnSO}_4$ ) luminescence in LB and NH were compared (Figure 3.32). The first thing to notice when looking at Figure 3.32, is that luminescence in NH media gives a higher maximum value, however, expression dynamics in LB and NH are different. Luminescence in NH shows a much higher peak, at approximately the same time as LB (120 min), however, luminescence in NH drops significantly to a barely detectable level, due to the lack of luciferase substrate. In contrast, luminescence in LB shows a base line throughout the assay. This suggests that LB media allow production of  $\text{FNMH}_2$  and  $\text{RCHO}$  to occur for longer than in NH media. These variations between media could be attributed to the fact that NH is a minimal defined media, whereas LB is a rich basic media; NH does not have sufficient nutrients to support growth and bacterial function for as long as LB. This lack of nutrients may put the bacteria into a starvation mode, which reduces the output of luciferase substrates.



**Figure 3.32 Temporal Luminescence of Induced  $P_{zntA}$  in LB and NH**

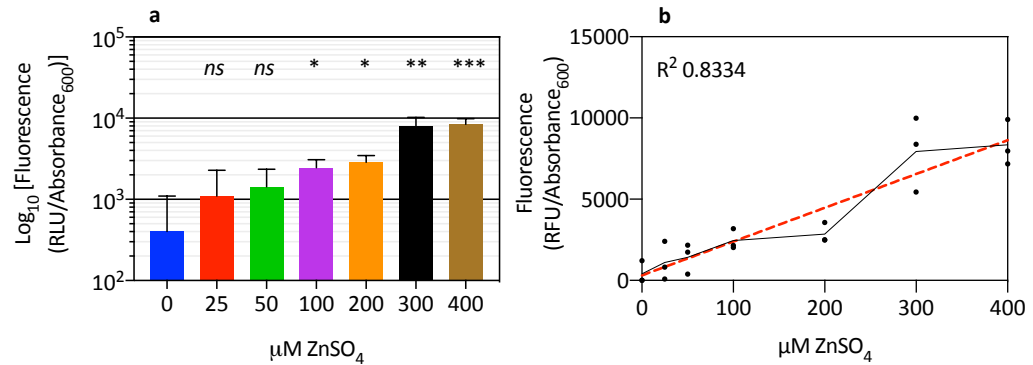
Temporal luminescence of induced  $P_{zntA}$  (400  $\mu\text{M}$   $\text{ZnSO}_4$ ) in either  $\text{NH}_{\text{kan50}}$  or  $\text{LB}_{\text{kan50}}$  media. NH data from (Figure 3.23). LB data from (Figure 3.28)

### 3.2.16. Alternative Bacteria

This chapter has focused on understanding the regulation of  $P_{zntA}$  in *E. coli*, however, it is useful to understand the regulation of  $P_{zntA}$  in other bacteria, to determine if a novel inducible promoter can be used in another genus. An end point fluorescence assay of  $P_{zntA}$  was conducted in *Cupriavidus necator* H16, grown in LB, with increasing concentration of  $ZnSO_4$  (Figure 3.33). Figure 3.33a shows that with the addition of  $\geq 100 \mu M$   $ZnSO_4$  there is a statistically significant increase in fluorescence of  $P_{zntA}$  in *C. necator*, compared to LB only ( $0 \mu M$   $ZnSO_4$ ). Figure 3.33b shows that there is a minor relationship between fluorescence and  $ZnSO_4$  concentration: increasing  $ZnSO_4$  concentration will increase fluorescence in a linear relationship, with a goodness of fit  $R^2$  0.8334. The relationship between gene expression and  $ZnSO_4$  concentration is not as strong as observed with  $P_{zntA}$  in *E. coli*, in either LB or NH media (Figure 3.9 and Figure 3.11), which both have a  $R^2 \geq 0.95$ . There are large deviations from the mean in this data set, which account for the lower  $R^2$  value compared to *E. coli*.

Genome analyses of *C. necator* showed a putative annotated gene for *zntA*, however, no putative labelled gene is found for *zntR* (Pohlmann *et al.*, 2006). The promoter of *zntA* in *C. necator* shows a potential ZntR binding site, which shows 45% homology to the ZntR binding site of  $P_{zntA}$  in *E. coli*. Protein BLAST and nucleotide BLAST returned no results for a ZntR protein or *zntR* gene in *C. necator*. Protein Blast returned a result of MerR in *C. necator*, which showed 40.3% sequence identity to *E. coli* ZntR. Brocklehurst *et al.*, (1999) showed that ZntR of *E. coli* is induced by  $Hg^{2+}$ . *E. coli* ZntR contains five Cysteine residues, of which three are in identical locations in MerR (Khan *et al.*, 2002). Knowing that Cysteine residues are often linked to metal ion binding in enzymes, such as  $Zn^{2+}$  in zinc fingers (Klug and Rhodes, 1987) and  $Zn^{2+}$  in metallothionein (Chang, Liao

and Huang, 1998), it is plausible to assume that if *C. necator* does not encode *zntR*, MerR replaces the function and regulates *zntA* during increased zinc conditions.



**Figure 3.33 End Point Assay of ZntR Regulated  $P_{zntA}$ : Expressed in *C. necator***

Plasmid constructs transformed and expressed in *C. necator* H16 with aerobic incubation at 30°C with shaking (250 RPM) in  $\text{LB}_{\text{kan200}}$  with increasing  $\text{ZnSO}_4$  concentration, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. (a) bar chart. Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$  (two-tailed, unpaired, t-test) comparison to un-induced (0  $\mu\text{M ZnSO}_4$ ). (b) scatter plot. Dots indicate individual replicates. Solid black line indicates the mean. Red dashed line indicates linear regression model. Goodness of fit  $R^2 = 0.8334$ .

### 3.3. Discussion and Conclusion

#### 3.3.1. Growth Media Validation

Zinc depleted Neidhardt's MOPS minimal media (NH) has proved an invaluable tool for understanding the promoter activity of  $P_{zntA}$ . ICP-MS analysis showed that the zinc concentration in NH is 100-fold less than in LB media. NH allows one to conduct assays without the need for zinc chelators; these chelators may have a negative affect and chelator or metals as well as zinc. NH media has previously been used as a zinc depleted media in similar studies, and showed comparable growth rates of other traditional media (Neidhardt, Bloch and Smith, 1974; Clayton, 2012).

Assays were also conducted in standard, commonly used media, Lysogeny Broth (LB). To reduce the available zinc in LB media, TPEN was used. TPEN can permeate bacterial cell membranes and thus reduce internal zinc concertation, as well as external. TPEN has been used as a zinc chelator in previous studies (Sigdel, Easton and Crowder, 2006; Graham *et al.*, 2009; Clayton, 2012). A toxicity test was initially conducted to look at the concentration of TPEN, which would cause toxicity on *E. coli*. Results showed that up to 40  $\mu\text{M}$  TPEN does not cause a toxic effect on *E. coli*, thus the desired concentration of 20  $\mu\text{M}$  TPEN is safely below toxic levels. It is worth noting that high specificity of zinc binding of TPEN has been debated. Sigdel, Easton and Crowder, (2006) suggested that TPEN binds to  $\text{Cd}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$  more tightly than it does to  $\text{Zn}^{2+}$ . However, generally TPEN is regarded as a highly specific zinc chelator.

A zinc toxicity test was also conducted; results showed that with the addition of up to 600  $\mu\text{M}$   $\text{ZnSO}_4$  it does not cause a toxic effect on *E. coli*. Most assays throughout this experiment used the addition 400  $\mu\text{M}$   $\text{ZnSO}_4$  as the maximum zinc concentration, which is safely under toxic levels.

For a strong inducible promoter, it is essential that the inducer is not metabolized. Looking at the example of the *lac* operon, lactose was replaced with IPTG as the inducer, due to the fact that IPTG cannot be metabolised, whereas lactose can (Reznikoff, 1992). The promoter  $P_{zntA}$  is regulated by ZntR, which in turn is activated by zinc. Obviously, zinc cannot be metabolised or broken down. However, it is essential that the level of zinc concentrations in the media does not significantly reduce over time, due to it being utilised by bacteria for growth and cellular functions. Zinc concentrations were determined from the media of *E. coli* liquid culture over 24 hr; results indicated that the zinc concentration in the media did not significantly reduce over 24 hr of bacterial growth.

### 3.3.2. End Point Fluorescence Assays

End point fluorescence assays in both LB and NH media showed a strong relationship between fluorescence (a proxy for promoter activity) and increased  $ZnSO_4$  concentrations. There was a stronger relationship in NH media than LB. This relationship between increased promoter activity and increased inducer concentration gives promise for  $P_{zntA}$  to be used as a tuneable zinc inducible promoter.

This data is in line with previously published data (Brocklehurst *et al.*, 1999). However, I believe the data I have collected and presented is much fairer and more accurately presented. Brocklehurst *et al.*, (1999) showed increasing luminescence with increasing zinc concentrations, however, due to the nature of the *luxCDABE* operon, using luminescence as an end point assay can be unreliable, partially due to the short half-life luciferase and the multiple factors which influence luminescence. Using a fluorescence-based end point assay, as used in this work, is a much more accurate way of depicting this data; the 24 hr + half-life of mRFP1 means that little degradation of fluorescent protein will be observed within the assay. Fluorescence will show the

accumulation of RFP, which can be used as a proxy for accumulated promoter activity. In this thesis LuxCDABE luminescence data was used to determine the rate of expression, as a statistical method of comparison. This provides further information on the regulation of  $P_{zntA}$ .

When comparing  $P_{zntA}$  to  $P_{trc}$  in LB, there is a significant increase in fluorescence of uninduced  $P_{zntA}$  (LB only), compared to uninduced  $P_{trc}$ . However, if the uninduced  $P_{zntA}$  is measured with the addition of 20  $\mu$ M TPEN, this reduces the basal level to that of uninduced  $P_{trc}$ . There is also no significant difference in induced  $P_{zntA}$  (400  $\mu$ M  $ZnSO_4$ ) and induced  $P_{trc}$  (100  $\mu$ M IPTG), when measured in LB. Moreover, the greatest fold increase in fluorescence is observed between uninduced  $P_{zntA}$  (20 $\mu$ M TPEN) and induced  $P_{zntA}$  (400  $\mu$ M  $ZnSO_4$ ), with 129-fold increase. When comparing  $P_{zntA}$  and  $P_{trc}$  in NH, induced  $P_{zntA}$  showed a significantly higher fluorescence than induced  $P_{trc}$ , and both showed no significant difference in their uninduced states. This comparison of end point fluorescence data suggests that induced  $P_{zntA}$  is comparable to  $P_{trc}$ , however, the basal level of expression of  $P_{zntA}$  is not optimal.

### 3.3.3. End Point Fluorescence Assay of Mutant $P_{zntA}$

Results showed that by increasing the spacer length of  $P_{zntA}$  it significantly reduced the fluorescence compared to  $P_{zntA}$  *wt* and reduced it to the same level as uninduced  $P_{zntA}$  *wt* (20  $\mu$ M TPEN). These results show that by increasing the spacer by either 1 or 2 bp, it removes the ZntR regulation of *zntA*; the addition of 20  $\mu$ M TPEN or 400  $\mu$ M  $ZnSO_4$  does not cause any variation in gene expression. It is possible that with the increase in spacer length, ZntR monomers bind to the 11 bp sequence, but cannot form successful bonds between each other, therefore when zinc binds to the monomers, no conformation change occurs. At the same time, the addition of 1 or 2 bp increases the spacer region to 21/22, which further reduces the spacer region from the optimal

17±1 bp (Harley and Reynolds, 1987), reducing successful  $\sigma^{70}$  interactions with the -10 and -35 element. A similar study was conducted by Brocklehurst *et al.*, (1999), in which the spacer region was shortened. Results demonstrated that shortening the spacer region by 1 bp reduced gene expression, but still showed ZntR regulation. However, reducing the spacer region by 2 bp completely removed the ZntR regulation. Reducing the spacer region by 2 bp brings the -35 and -10 element in perfect alignment with  $\sigma^{70}$ , and thus is now constitutently expressed, regardless of zinc concentrations and ZntR interaction. The results from this chapter on increasing the spacer region, and the results from Brocklehurst *et al.*, (1999) on decreasing the spacer region, give a clear understanding of this spacer region. This combined data show that the  $P_{zntA}$  spacer region is optimised, and that increasing or decreasing the spacer region by a single base, significantly reduces the ZntR regulation of the *zntA*.

Mutations which altered the promoter sequences' similarity to consensus sequences did not prove fruitful. If the -10 element was mutated to the consensus (5'-TATAAT-3'), it greatly improved gene expression in all three conditions, (LB only, 20  $\mu$ M TPEN, or 400  $\mu$ M  $\text{ZnSO}_4$ ) but with minor variation between the conditions. If mutations improved consensus to both -10 and -35 element, this increased gene expression to a similar level of induced  $P_{zntA}$  *wt*, but showed no ZntR regulation; there was no variation between the three conditions. Mutation in the TG motif [ $P_{zntA}(-14\text{T>G})$ ] reduced the basal level expression like  $P_{zntA}$  *wt* in LB + 20  $\mu$ M TPEN, however, when induced with 400  $\mu$ M  $\text{ZnSO}_4$ , fluorescence was also significantly reduced compared to  $P_{zntA}$  *wt*. The most interesting mutation was reducing the consensus of the -10 element [ $P_{zntA}(-11_{-7}\text{>TGACA})$ ]. The  $P_{zntA}(-11_{-7}\text{>TGACA})$  mutation, which disrupted the -10 element, showed no detectable fluorescence in either LB only or LB + 20  $\mu$ M TPEN; fluorescence was only observed in LB + 400  $\mu$ M  $\text{ZnSO}_4$ . This mutation drastically reduced the gene expression to a level which is undetectable in either uninduced LB only or LB + 20  $\mu$ M TPEN. Gene

expression is only observed as fluorescence, when induced with 400  $\mu\text{M}$   $\text{ZnSO}_4$ . Essentially, this mutation drastically reduced gene expression in all three conditions, and only when ZntR is activated does gene expression occur at a level which can be detected. Unfortunately, induced [ $P_{zntA(-11\_7>\text{TGACA})}$ ] showed no significant difference in fluorescence, compared to uninduced  $P_{zntA}$  (20  $\mu\text{M}$  TPEN) (two-tailed, unpaired, t-test). This mutation does reduce the basal level of expression, which is advantageous, however, the level of induced gene expression, I believe, is too low to be considered advantageous.

#### 3.3.4. mRNA Analysis of *zntA* Through RT-qPCR

Direct quantification of mRNA was conducted using RT-qPCR. The reference gene *rrsA* was first validated as a reliable reference gene for experiments in this thesis. *rrsA* showed little variation in mRNA concentration, which was not affected by increasing  $\text{ZnSO}_4$  concentrations. Tight RT-qPCR graphs, and little variation in  $C_q$  ( $\sim 0.5$ ) validated *rrsA*. *rrsA* has previously been validated as a reference gene in the context zinc RT-qPCR assays (Zhou *et al.*, 2011; Peng *et al.*, 2014; Xu *et al.*, 2019).

After *rrsA* validation, RT-qPCR was used to analyse mRNA of *zntA* in three different conditions, NH + 0, 200, or 400  $\mu\text{M}$   $\text{ZnSO}_4$ .  $\Delta C_q$  analysis showed that with the addition of both 200 and 400  $\mu\text{M}$   $\text{ZnSO}_4$ , there was a significant increase in the  $C_q$  value, which can be interpreted as an increase in *zntA* mRNA. Further analysis was conducted using the comparative  $2^{-\Delta\Delta C_q}$  method.  $2^{-\Delta\Delta C_q}$  analysis showed a 2-fold increase in *zntA* mRNA, with the addition of 200  $\mu\text{M}$ ; a 10-fold increase in *zntA* mRNA was observed with the addition of 400  $\mu\text{M}$   $\text{ZnSO}_4$ .

These RT-qPCR results support the hypothesis that the alterations in the fluorescence data, i.e increased expression observed with increasing  $\text{ZnSO}_4$  concentration, are due



to transcriptional variation, rather than solely translation variation. This is vital for an inducible promoter; regulation and variation should be due to transcription, rather than translation. RT-qPCR results cannot unequivocally say that translational variation does not play a role in  $P_{zntA}$  fluorescence assays, however, it can without doubt show that transcriptional variation plays a major role in  $P_{zntA}$  fluorescence assay variation. These RT-qPCR results show that the fluorescence assays are a good method for transcription analysis of promoter activity of  $P_{zntA}$  and could be a validation for use of the pJI300 plasmid.

When comparing *zntA* gene expression measure by either RT-qPCR or mRFP1 end point fluorescence assay, there seems to be a relationship between the two. To directly compare these two experimental datasets, fold change was calculated for each experiment, as a method to normalize the data for direct comparison. The difference between fold change in 200  $\mu\text{M}$  and 400  $\mu\text{M}$   $\text{ZnSO}_4$ , measured from RT-qPCR and fluorescence, showed no significant difference, suggesting a direct relationship between this data. This direct relationship can validate the use of mRFP1 as a reliable method to measure *zntA* gene expression, or at least as a proxy. However, the fold change in 200  $\mu\text{M}$   $\text{ZnSO}_4$  is larger in RT-qPCR, than in end point fluorescence. This could be due to the longer half-life of mRFP1 than mRNA, with residual mRFP1 affecting the data. Measuring gene regulation via reporter plasmids will always affect the data; the intrinsic process involved in transcription and translation will always be a factor, as will the involvement of plasmid copy number. This data, however, does suggest that mRFP1 based end point fluorescence assays are useful proxy for gene expression analysis.

However, an issue with the RT-qPCR experiment conducted was the chosen 16 hr time point. The 16 hr time point was initially chosen, as it coincides with the fluorescence

assay time point. As mentioned before, a more extensive experiment was planned which looked at *zntA* mRNA concentrations between 0-4 hr, but due to the COVID-19 global pandemic, this experiment was not conducted. Bacterial mRNA half-life is generally considered short, <60 min (Rauhut and Klug, 1999; Richards *et al.*, 2008), compared to 24 hr + for mRFP1. This means that if comparing  $C_q$  values directly against fluorescence, value would be misleading, as fluorescence would measure accumulation of RFP for the whole assay, whereas  $C_q$  values will only measure mRNA accumulation of approximately <60 min, prior to sample reading. It would be more accurate to record and compare  $C_q$  values at time point less than 4 hr, as luminescence assays suggest that after this time point gene expression appears to reduce significantly.

### 3.3.5. Temporal Luminescence Assays

Luminescence assays were conducted to allow one to look at temporal promoter activity, rather than an accumulation of activity. It is worth noting, that even though luminescence can be used for temporal data, you are still looking at a historical time point of promoter activity due to time involved in: transcription; translation; maturation; and reaction. However, the lag time between gene upregulation and the measurable luminescence is significantly less than when using mRFP1.

Similar to the results observed in fluorescence assays, luminescence temporal assays show a relationship between promoter activity and  $ZnSO_4$  concentrations. With increased  $ZnSO_4$  concentration, there was an increase in luminescence. Again, like the end point assays, this relationship can be observed in both NH and LB media. Luminescence provides another depth of analysis of promoter activity. Luminescence assays allow one to look at promoter activity and compare it to bacterial growth phases. Expression dynamics of  $P_{zntA}$  varied between NH and LB assays.  $P_{zntA}$

luminescence in NH had a higher maximum peak, but dropped to essentially no expression after 480 min. In contrast,  $P_{zntA}$  luminescence in LB showed a lower maximum peak, but showed constant expression for up to 10 hr. Both NH and LB showed that the peak of promoter activity is observed during the exponential phase of growth.

When comparing  $P_{zntA}$  and  $P_{trc}$  luminescence in LB, induced  $P_{trc}$  showed higher maximum luminescence than induced  $P_{zntA}$ ; uninduced  $P_{trc}$  showed lower maximum luminescence than uninduced  $P_{zntA}$  (LB only or 20  $\mu$ M TPEN). Interestingly, when comparing the expression rate, there was no significant difference between induced  $P_{zntA}$  and induced  $P_{trc}$ , or uninduced  $P_{zntA}$  (20  $\mu$ M TPEN) and uninduced  $P_{trc}$ . However, if looking at expression rate between  $P_{zntA}$  and  $P_{trc}$  in NH,  $P_{zntA}$  showed higher expression rate in both the uninduced and induced state.

### 3.3.6. Alternative Bacteria

The novel inducible promoter  $P_{zntA}$  was tested in *Cupriavidus necator* to determine if  $P_{zntA}$  shows the same controllable zinc induction as observed in *E. coli*. Results showed that  $P_{zntA}$  in *C. necator* can be induced with the addition of  $\geq 100 \mu$ M  $\text{ZnSO}_4$ . Further, there is a weak linear relationship between fluorescence and increasing  $\text{ZnSO}_4$  concentration. Genome analysis showed an annotated gene for *zntA*, but not the regulatory *zntR*. However, genome analysis did show an annotated gene, *merR*.  $P_{zntA}$  of *C. necator* showed a potential ZntR binding site, which has 45% similarity to the ZntR binding of  $P_{zntA}$  in *E. coli*. I hypothesise that if no *zntR* is encoded within *C. necator*, MerR replaces the role of ZntR, due to their similarity in function, structure, and cysteine residue locations (based on ZntR and MerR in *E. coli*). However, it is also possible that ZntR is present on the *C. necator* chromosome, but has not yet been annotated.

### 3.3.7. Concluding Remarks

Data thus far shows promise for  $P_{zntA}$  to be used as a novel, tuneable inducible promoter. Having similar or even higher expression than the IPTG inducible  $P_{trc}$  shows potential, but is let down by the high basal level of expression, mainly observed in LB. Unfortunately, site directed mutagenesis of  $P_{zntA}$  did not provide any novel promoters which were more favourable than  $P_{zntA}$  wt. RT-qPCR confirmed that  $P_{zntA}$  regulation is, at least partially, regulated at the transcriptional level.

### 3.3.8. Future work

Often, when using an inducible promoter, the inducer is added mid-exponential phase, rather than during lag phase, which is when it was added in assays conducted in this thesis. It would be useful to look at promoter activity, using both luminescence and RT-qPCR assays, of  $P_{zntA}$  when induced with  $ZnSO_4$  during mid-exponential phase. This could provide data which shows how long the lag time is between induction and increased promoter activity.

RT-qPCR was conducted at 16 hr, but as previously mentioned, due to the short half-life of RNA, it is not ideal as a direct comparison to 16 hr end point fluorescence assays. It would be useful to compare RT-qPCR data directly against luminescence and fluorescence data by taking RT-qPCR, luminescence, and fluorescence readings at regular intervals between 0-4 hr, as this was seen to be the most active for promoter activity from luminescence data. With all this data, one would be able to directly compare the fluorescence and luminescence data, against the direct mRNA quantification via RT-qPCR. This would provide a few valuable pieces of information; it would definitively show if fluorescence and luminescence assays in this study are a

good and reliable proxy for promoter activity. It would also provide direct information on gene regulation of  $P_{zntA}$ , which could be standalone data on  $P_{zntA}$  gene regulation.

## Chapter 4

# Fundamental Understanding of the Transcriptional Regulation of Zur Regulated Promoters

### 4.1. Introduction

The previous chapter (Chapter 3) discussed the ZntR regulation of *zntA*, which encodes the major zinc efflux protein in *E. coli*, ZntA, and which forms part of the zinc homeostasis system. This chapter will focus on Zur regulated genes, which makes up the zinc import system of the *E. coli* zinc homeostasis system.

The Zur (Zinc Uptake Regulator) transcription factor is a major TF which is involved in the regulation of genes associated with: zinc import; zinc acquisition; or freeing zinc from internal reservoirs. Zur is an active repressor, when internal zinc concentrations are optimal, zinc binds to Zur, which in turn binds to promoter DNA, causing repression of transcriptional via steric hinderance of RNAP. When internal free zinc concentrations become low, the zinc bound to Zur is released, and Zur disassociates from the promoter DNA, allowing RNAP to bind and initiate transcription (Patzner and Hantke, 1998; Outten and O'Halloran, 2001).

Zur has been shown to regulate six promoters in *E. coli*: *znuA*; *znuCB*; *zinT*; *pliG*; *ykgMO*; and *c1265-7* (Patzner and Hantke, 2000; Clayton, 2012; Gilston *et al.*, 2014). Zur binds to a consensus sequence (Zur box), originally described as a specific DNA sequence [5'-GAAATGTTATA-N-TATAACATTTTC-3'], which was discovered through a DNA-footprint assay of Zur regulated *znuABC* promoter (Patzner and Hantke, 2000). However, it has more recently been described, through Zur-DNA crystallography, as a motif; Purine-N-N-N-Pyrimidine (RNNNY), separated by 3 bp giving a repeated motif of RNNNYxxxRNNNY (Figure 4.1) (Gilston *et al.*, 2014).

#### 4.1.1. Aims and Objectives

This chapter's main objective is to further understand the transcriptional response of Zur regulated promoters to zinc, with the aim to develop a novel inducible or repressible promoter system.

Chapter 3 showed a strong relationship between zinc concentration in the growth media and gene expression of the ZntR regulated  $P_{zntA}$ . This chapter will also aim to determine if, like  $P_{zntA}$ , any of the six Zur regulated promoter show any relationship between zinc concentration and gene expression.



## 4.2. Results

### 4.2.1. Zur Regulated Promoters

Six Zur regulated promoters were chosen for this study:  $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$  (*ykgMO*); and  $P_{c1265}$  (*c1265-7*). A combination of literature, and previous transcriptomic data obtained in the Hobman research group was used to determine these six Zur regulated genes. *znuA*, *znuCB*, *zinT*, *pliG* and *ykgM* have previously been reported in literature to be regulated by Zur (Patzner and Hantke, 1998; Panina, Mironov and Gelfand, 2003; Graham *et al.*, 2009; Gabbianelli *et al.*, 2011; Gilston *et al.*, 2014). *c1265-7* is a potential novel Zur regulated gene, proposed by Clayton, (2012), through a combination of transcriptomic and bioinformatic analysis.

#### 4.2.1.1. Novel *c1265-7* Zur Regulated Operon

A novel Zur regulated gene operon *c1265-7*, found only chromosomally in UPEC *E. coli* strains, was discovered by Clayton, (2012). Transcriptomic data, performed by Clayton, (2012) showed that there was a significant fold increase in gene transcripts of the *c1265-7* operon; when comparing cells grown in a zinc depleted media (NH only) to cells grown in a zinc repleted media (NH + 1  $\mu$ M ZnSO<sub>4</sub>) (Table 4.1). The transcriptomic data also showed a significant fold increase in gene transcripts from *E. coli* CFT073  $\Delta$ *zur* compared to *E. coli* CFT073 *wt* when both were grown in the zinc repleted media (1  $\mu$ M ZnSO<sub>4</sub>). This data showed that *c1265-7* gene transcription is upregulated in a zinc depleted state, and that deletion of *zur* ( $\Delta$ *zur*) also upregulated transcription of *c1265-7*. This data strongly suggests that *c1265-7* is under Zur regulation. Further Patser bioinformatic analysis concluded that  $P_{c1265}$  has a high similarity to the Zur binding in *E. coli* CFT073 (Patser score = 19.9) (Clayton, 2012). The bioinformatics, and

transcriptomics data suggest that  $P_{c1265}$  may be a useful novel Zur regulated promoter and is worth further investigation.

Gene	Fold Change in <i>E. coli</i> CFT073 (Transcripts)	
	<i>wt</i> (+ zinc) → <i>wt</i> (- zinc)	<i>wt</i> (+ zinc) → $\Delta$ <i>zur</i> (+ zinc)
<i>c1265</i>	41.83	30.98
<i>c1266</i>	60.92	46.34
<i>c1267</i>	31.56	34.69

**Table 4.1 Transcriptomics data of *E. coli* CFT073**

Transcriptomic data (fold change) of the novel *c1265-7* gene expressed in *E. coli* CFT073 *wt* or  $\Delta$ *zur*. - zinc (NH only). + zinc (NH + 1  $\mu$ M ZnSO<sub>4</sub>). Data from (Clayton, 2012).

#### 4.2.1.2. Consensus Zur Binding

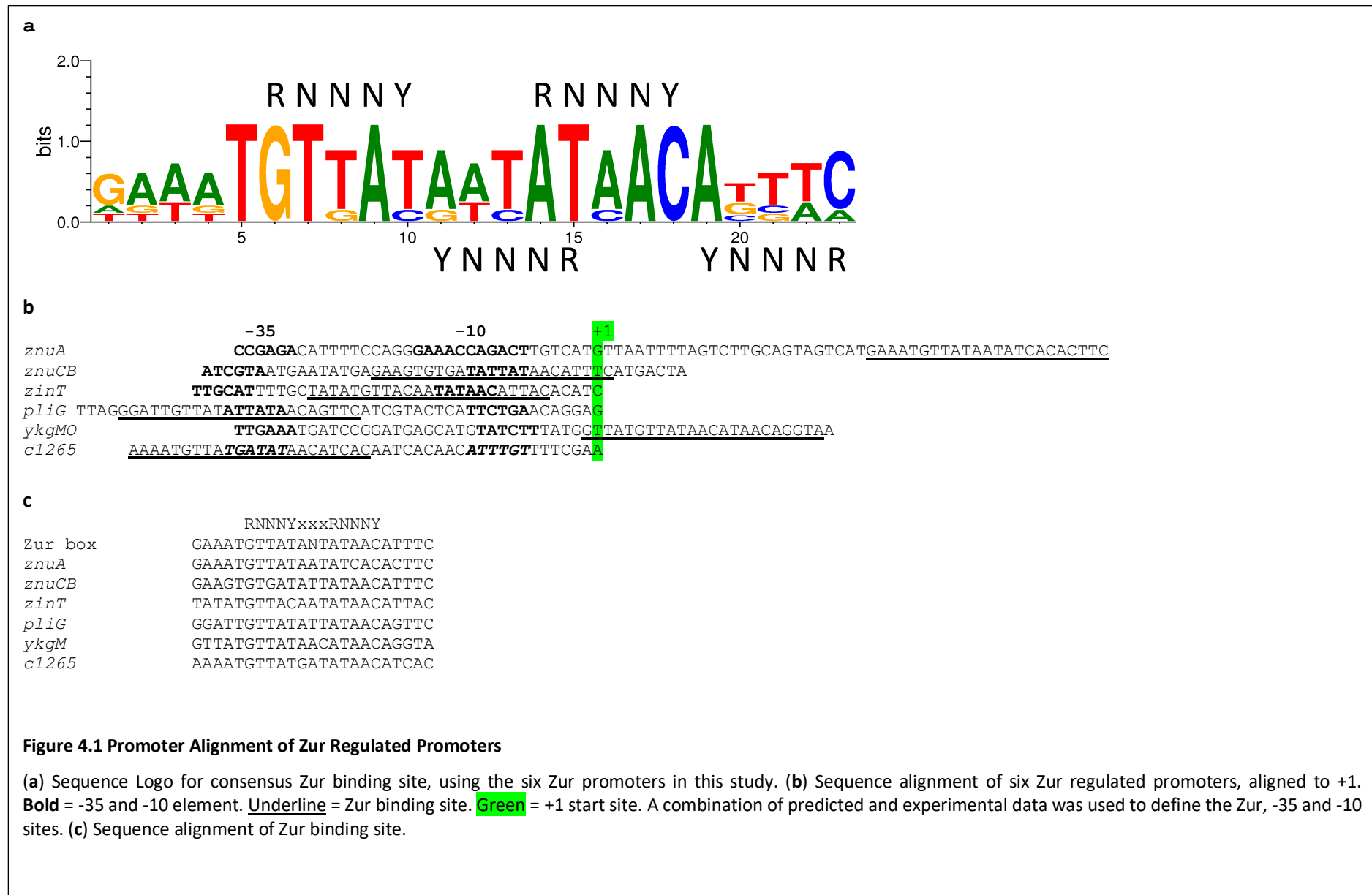
The Zur DNA binding consensus has been described two different ways in literature. The first to be described was the Zur box [5'-GAAATGTTATA-N-TATAACATTTTC-3'] (Patzer and Hantke, 2000). The Zur box was discovered through a DNA foot-printing assay which looked at the Zur protected area of the divergently expressed *znuABC*. Gilston *et al.*, (2014) determined a more accurate consensus sequence, using Zur-DNA crystallography to analyse the specific promoter DNA nucleotides which interact with Zur. This gave a motif consensus motif sequence, RNNNYxxxRNNNY.

Using the known and potential Zur binding sites of the six Zur regulated promoters in this thesis (Figure 4.1b), a consensus sequence logo was created (Figure 4.1a). From looking at the logo consensus, the six promoters used in this study ( $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$ , and  $P_{c1265}$ ) show perfect homology to the RNNNY-xxx-RNNNY motif pattern, which agrees with the findings of Gilston *et al.*, (2014) (Figure 4.1c). It can also be seen that there is strong homology to the Zur box consensus sequence, with the 'most common' nucleotide showing identical homology to the Zur box. Interestingly, when looking at the Zur binding position relative to the +1 start site, there is large variation

in Zur binding (Figure 4.1**b**). Zur binding can be observed both upstream and downstream of the +1 start site. When looking at the nucleotides outside of the RNNNYxxxRNNNY motif there is a noticeable increase in variation between the bp.

The potential Zur binding location in P<sub>c1265</sub> shows 78% sequence homology to the Zur box, and perfect homology to the RNNNYxxxRNNNY motif (Figure 4.1), strengthening the argument that c1265-7 is Zur regulated.

The 'N' of the RNNNY motif demonstrates highly conserved base pairs between the aligned Zur regulated promoters (Figure 4.1**a**). Crystallography data (Gilston *et al.*, 2014) demonstrated that each Zur monomer creates only two hydrogen bonds with purines, denoted as 'R' and 'Y' in the RNNNY motif (Figure 1.14**b**) (Gilston *et al.*, 2014). However, the (Zur<sub>2</sub>)<sub>2</sub>-DNA complex has over 100 contacts between amino acids and DNA.



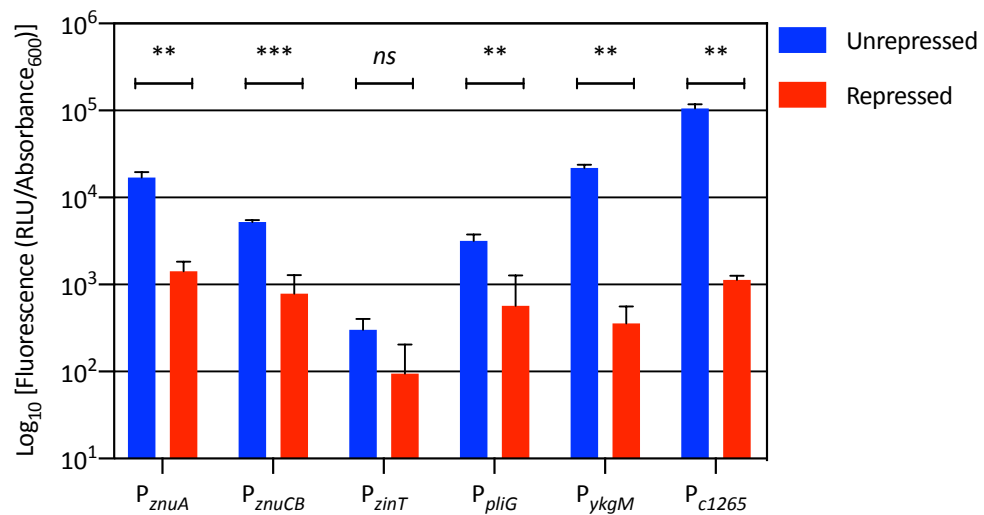
#### 4.2.2. End Point Fluorescence Assay In NH

To increase the promoter activity of Zur regulated genes, the internal free zinc concentration in the bacterial cell must be reduced. One way to do this is to grow bacteria in a zinc depleted media. Initial end point fluorescence assays were conducted in zinc depleted Neidhardt's MOPS minimal media (NH), using the pJI300 dual report plasmid with cloned Zur regulated promoter sequences. It is vital to understand promoter activity of these Zur regulated promoters both when unrepressed (zinc depleted) and repressed (zinc repleted). [Note, when talking about the ZntR regulated promoter  $P_{zntA}$ , the term induced and uninduced was used, as  $ZnSO_4$  was added as an inducer to increase promoter activity]. Zur regulated promoters use added  $ZnSO_4$  to repress promoter activity, therefore the term repressed and unrepressed is used when expressed in NH. Zur regulated promoters were expressed in NH only (unrepressed) and NH + 400  $\mu M$   $ZnSO_4$  (repressed)

Six Zur regulated promoters were cloned into the reporter plasmid, pJI300, and transformed into *E. coli* MG1655. *E. coli* was grown overnight (16 hr) in NH with various increasing concentration of  $ZnSO_4$ . After incubation, fluorescence and absorbance<sub>600</sub> was recorded using a TECAN GENios Pro. End point fluorescence assay data shows the accumulation of fluorescent protein.

Figure 4.2 shows fluorescence of Zur regulated promoters expressed in NH media, NH only (unrepressed) and NH + 400  $\mu M$   $ZnSO_4$  (repressed). It is worth noting the zinc concentration of NH only is 115 nM, and therefore not completely zinc depleted. The fluorescence of promoters in NH only (unrepressed) shows an obvious difference in fluorescence between lowest and highest expressed promoter,  $P_{zinT}$  and  $P_{c1265}$ , with over 350-fold difference between the two. A one-way ANOVA on the six Zur regulated

promoters in NH only returned a p-value of <0.0001, showing a strong significant variation between all the samples.



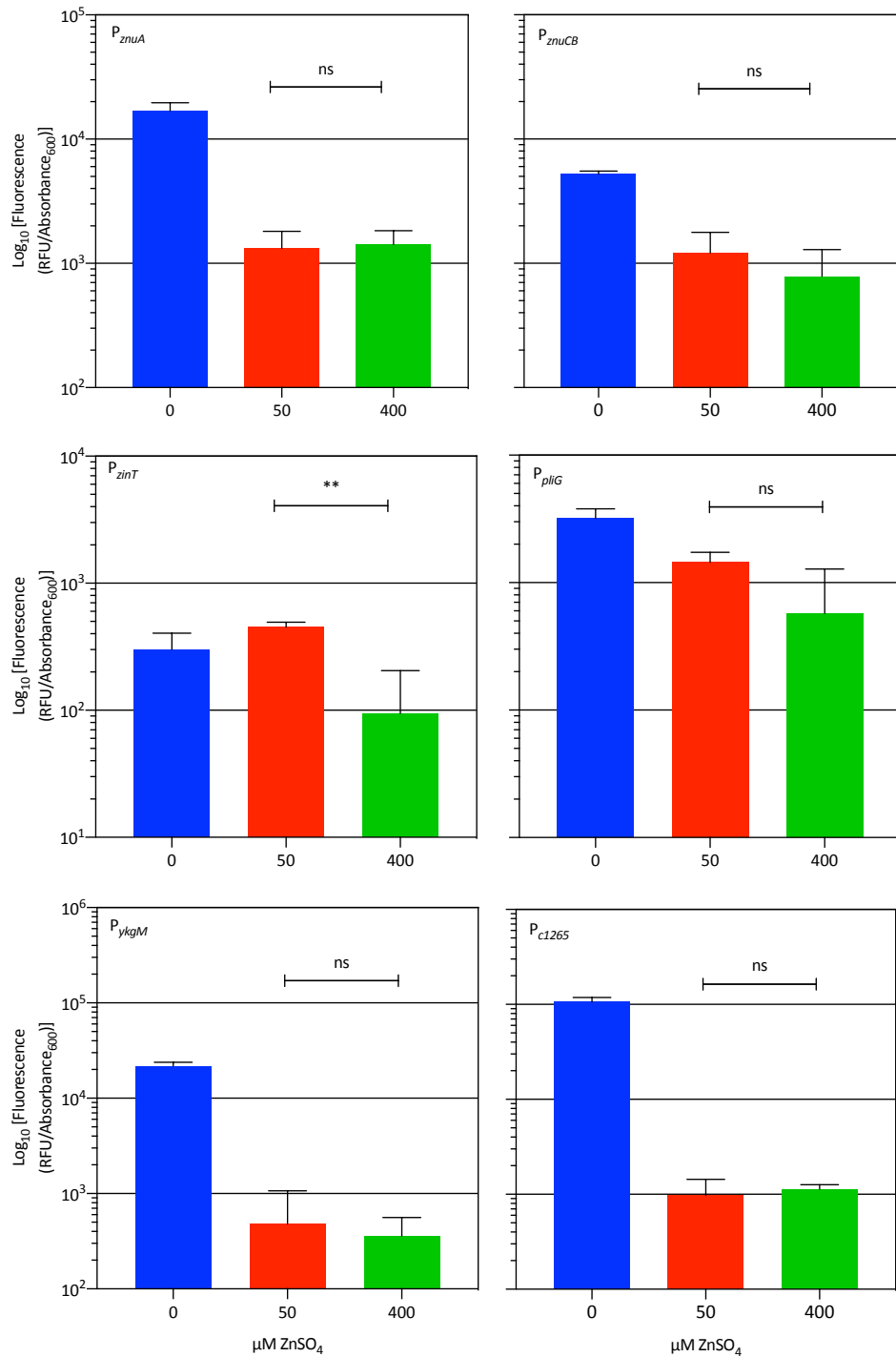
**Figure 4.2 End Point Fluorescence Assay of Zur Regulated Promoters in NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in NH<sub>kan50</sub> with or without 400 μM ZnSO<sub>4</sub>, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Unrepressed, NH only. Repressed, NH + 400 μM ZnSO<sub>4</sub>. Error bars indicated S.D. ns p ≥ 0.05, \*\* p ≤ 0.01, \*\*\* p ≤ 0.001, (two-tailed, unpaired, t-test) comparison of unrepressed and repressed.

When looking at the repressed and unrepressed fluorescence of each promoter, there is a significant difference between them, except for P<sub>zinT</sub>. P<sub>zinT</sub> is the only Zur regulated promoter with no significant difference in fluorescence between repressed and unrepressed. The largest difference in fluorescence is observed with P<sub>c1265</sub>, with over 90-fold increase in fluorescence between repressed and unrepressed. P<sub>c1265</sub> shows the highest unrepressed fluorescence, but not the lowest repressed. The lowest repressed fluorescence is observed in P<sub>zinT</sub>. However, the unrepressed fluorescence of P<sub>zinT</sub> is lower than all repressed Zur promoters.

There is a significant increase in fluorescence in all Zur regulated promoters, except P<sub>zinT</sub>, when levels are compared between NH and NH + 400 μM ZnSO<sub>4</sub>. Further tests were conducted to look at the impact of ZnSO<sub>4</sub> concentration on fluorescence, to see if there is a correlation between promoter activity and ZnSO<sub>4</sub> concentrations. Figure

4.3 shows the fluorescence of promoters in NH with 0, 50 or 400  $\mu\text{M}$  added  $\text{ZnSO}_4$ . All promoters, except  $P_{zinT}$ , show no significant difference in fluorescence with either 50  $\mu\text{M}$  or 400  $\mu\text{M}$  added  $\text{ZnSO}_4$ . This suggests that these five Zur regulated promoters are not tuneable within this zinc range, 50 to 400  $\mu\text{M}$  added  $\text{ZnSO}_4$ . The addition of less than 50  $\mu\text{M}$   $\text{ZnSO}_4$  to NH media may be needed to see variation in fluorescence. One can say that with 50  $\mu\text{M}$   $\text{ZnSO}_4$  repression of Zur promoters is observed. Looking at  $P_{zinT}$ , there is a significant difference in fluorescence between 50 and 400  $\mu\text{M}$   $\text{ZnSO}_4$ , however the fluorescence with the addition of 50  $\mu\text{M}$  is higher than with 0  $\mu\text{M}$   $\text{ZnSO}_4$ , which goes against our knowledge of Zur regulation. I would argue that the variations observed in  $P_{zinT}$  is not real and may be due to biological variation. Further to this, if the difference in fluorescence observed in  $P_{zinT}$  is true, there is half a  $\log_{10}$  difference between values, which if intended to be used as a repressible promoter, is not enough variation for it to be considered useful. It is further possible that  $P_{zinT}$  is poorly repressed, as it is already in the maximum repressed state; the 115 nM of zinc in the NH media causes maximum repression of  $P_{zinT}$ , and that the repressible range is different than other Zur regulated promoters.



**Figure 4.3 End Point Fluorescence Assay of Zur Regulated Promoters in NH with Increasing  $\text{ZnSO}_4$  Concentrations**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{NH}_{\text{kan50}}$  with 0, 50 or 400  $\mu\text{M}$   $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Blue bars indicate NH only, Red indicated NH + 50  $\mu\text{M}$   $\text{ZnSO}_4$ , Green indicates NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ . Error bars indicated S.D. ns  $p > 0.05$ , \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test).



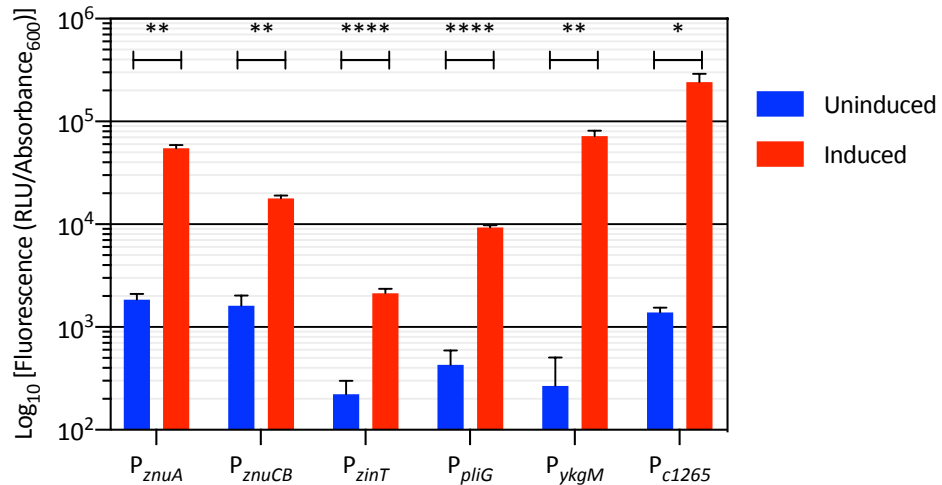
### 4.2.3. End Point Fluorescence Assays in LB

Conducting end point assays in NH helps to further understand the regulation of Zur regulated promoters, however it is useful to look at regulation in LB, a standard growth media, which contain a zinc concentration of 12.2  $\mu\text{M}$  as determined by ICP-MS (Table 3.1). When conducting assays in LB, TPEN is used to chelate zinc. The use of TPEN can be used as an inducer, as it reduces the biologically available zinc and thus increases promoter activity of Zur regulated promoters. When conducting Zur regulation assays in LB, the term induced (LB + 20  $\mu\text{M}$  TPEN) and uninduced (LB only) are used.

An end point fluorescence assay was conducted on all six Zur regulated promoters in both LB only (uninduced) and LB + 20  $\mu\text{M}$  TPEN (induced) (Figure 4.4). A one-way ANOVA analysis was conducted, which shows there is a statistically significant difference between the fluorescence values between the six promoters, true for both uninduced (LB only) and induced (LB + 20  $\mu\text{M}$  TPEN). In uninduced (LB only),  $P_{zinT}$  shows the lowest fluorescence, and  $P_{znuA}$ ,  $P_{znuCB}$  and  $P_{c1265}$  show no significant difference in fluorescence between each other (one-way ANOVA). When induced with 20  $\mu\text{M}$  TPEN  $P_{c1265}$  shows the highest level of fluorescence, and  $P_{zinT}$  shows the lowest level of fluorescence. Interestingly, when comparing the level of fluorescence NH only (Figure 4.2) and LB + 20  $\mu\text{M}$  TPEN (Figure 4.4), promoters show the same pattern of fluorescence; the most active promoter to the least active promoter [ $P_{c1265} > P_{ykgM} > P_{znuA} > P_{znuCB} > P_{pliG} > P_{zinT}$ ].

A two-tailed, unpaired, t-test was conducted on all promoter assay data, comparing uninduced to induced fluorescence. T-test confirmed that all promoters show a significant difference in fluorescence between the two conditions. It is worth noting that even though  $P_{c1265}$  has a large increase in fluorescence in the presence of TPEN, it has a low p-value, this is due to large variation of fluorescence in LB + 20  $\mu\text{M}$  TPEN.

This variation could be due to a heterogenous response to TPEN, or that a high fluorescence values, there is a greater variation in equipment readings.



**Figure 4.4 End Point Fluorescence Assay of Zur Regulated Promoters in LB**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> or LB<sub>kan50</sub> + 20 µM TPEN, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Uninduced (LB only), Induced (LB + 20 µM TPEN). Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test) comparison of induced to uninduced

The induced fluorescence of P<sub>zinT</sub> is not dissimilar to the uninduced P<sub>c1265</sub> fluorescence.

This low fluorescence, and variation in promoter activity could be due to the role of ZinT in zinc homeostasis. ZinT has been shown to be an auxiliary protein for ZnuABC, used to acquire zinc during zinc depleted growth conditions. Specifically ZinT is a periplasmic binding protein, with high specificity to zinc, which chaperones zinc to the periplasmic binding protein (ZnuA) part of the ZnuABC ATP-binding cassette (Kershaw, Brown and Hobman, 2007; Graham *et al.*, 2009; Petrarca *et al.*, 2010; Gabbianelli *et al.*, 2011). It is possible that (knowing ZinT is an auxiliary component of ZnuABC) *zinT* is less regulated than *znuABC*, as ZnuABC plays a more dominant role in zinc homeostasis than ZinT. ZnuABC requires ZinT, but if zinc is present at sufficient levels, zinc import via ZnuABC is independent of ZinT, and increased ZinT levels in the periplasm will not increase zinc import, only increasing ZnuABC will increase zinc

import. Or, that *zinT* expression is relative to *znuABC* expression, but lower expression of *zinT* is required, as ZnuACB plays the larger role in zinc homeostasis.

$P_{ykgM}$  has the largest increase in fluorescence from uninduced (LB only) to induced (LB + TPEN), and has the lowest uninduced fluorescence, save  $P_{zinT}$ . However,  $P_{c1265}$  has the highest induced (LB + TPEN) fluorescence, which is significantly higher than  $P_{ykgM}$  (p value 0.0046).

Table 4.2 shows the average fold increase in fluorescence from uninduced (LB only) to induced (LB + 20  $\mu$ M TPEN).  $P_{ykgM}$  shows the highest fold increase at 269-fold increase. This is followed by  $P_{c1265}$  with a 173-fold increase. Interestingly, if compared to the IPTG inducible promoter  $P_{trc}$ , both  $P_{ykgM}$  and  $P_{c1265}$  show higher fold increase in fluorescence.  $P_{trc}$  shows a fold increase of only 63x (Table 3.2). This shows promise that  $P_{ykgM}$  and  $P_{c1265}$  could be used as a TPEN inducible promoter.

The lowest variation is observed in  $P_{znuCB}$ . It is interesting that both  $P_{znuA}$  and  $P_{znuCB}$  show different responses considering that they are part of the same operon, and divergently regulated from a single central Zur binding site (Figure 1.16). Both  $P_{znuA}$  and  $P_{znuCB}$  show similar uninduced levels but show significant differences in induced fluorescence (P-value  $\leq 0.05$ , two-tailed, unpaired, t-test). ZnuABC is a Type-I ATP Binding Cassette (ABC transporter) which has the structure (AB<sub>2</sub>C<sub>2</sub>) (Patzner and Hantke, 1998; Locher, 2009). It would make sense that *znuCB* is more upregulated than *znuA* as, theoretically, ZnuCB requires twice the amount of proteins than ZnuA to form the ZnuABC complex.

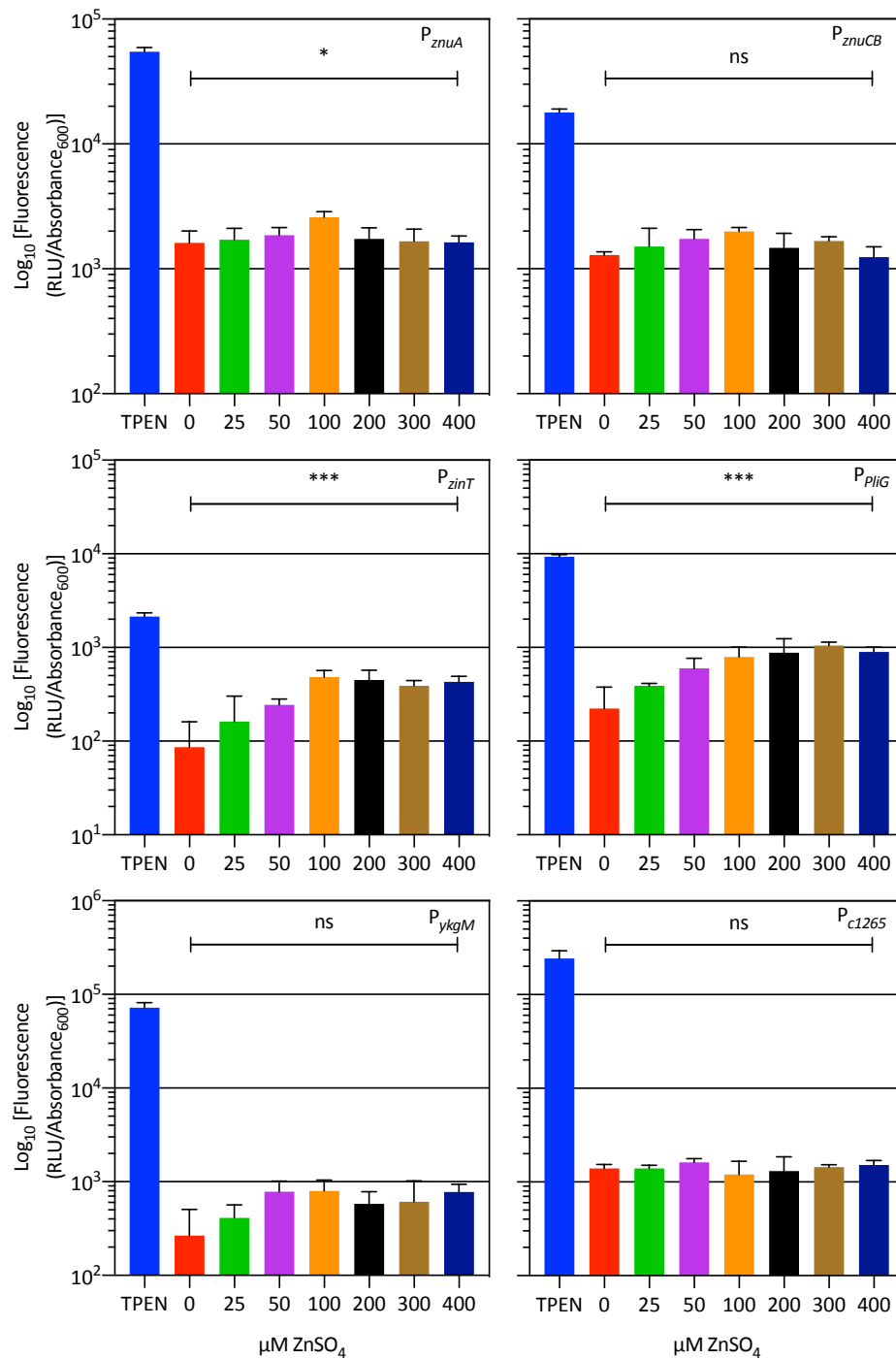
Promoter	Fluorescence increase (x)
P <sub>znuA</sub>	33.95
P <sub>znuCB</sub>	13.86
P <sub>zinT</sub>	24.76
P <sub>pliG</sub>	41.63
P <sub>ykgM</sub>	269.77
P <sub>c1265</sub>	173.58
P <sub>trc</sub> *	63.38

**Table 4.2 Fold Increase of Zur Promoter Activity in LB**

Fold increase in fluorescence of Zur regulated promoters expressed in LB. Increase fluorescence from uninduced (LB only) to induced (LB + 20  $\mu$ M TPEN). \* from LB only to LB + 100  $\mu$ M IPTG (data from Chapter 3).

Previous data in Figure 4.3 suggested that Zur regulated promoters are not tuneable using different concentrations of ZnSO<sub>4</sub>. A fluorescence assay was conducted with Zur regulated promoters in LB with increasing concentrations of added ZnSO<sub>4</sub> to see if increasing ZnSO<sub>4</sub> concentrations will further repress promoter activity (Figure 4.5). A larger range of ZnSO<sub>4</sub> concentrations were used for the assay compared to NH. One-way ANOVA was conducted to see if there is any variation in fluorescence between the samples, between 0-400  $\mu$ M added ZnSO<sub>4</sub>. ANOVA showed that P<sub>znuCB</sub>, P<sub>ykgM</sub>, and P<sub>c1265</sub> show no statistical significance in variation between different ZnSO<sub>4</sub> concentrations. This suggests that these Zur regulated promoters are not tuneable but are an on/off system, or that the zinc concentration in LB is enough for maximum suppression of these promoters. However, one-way ANOVA showed that P<sub>znuA</sub>, P<sub>zinT</sub>, and P<sub>pliG</sub> show statistically significant variation in fluorescence at the various ZnSO<sub>4</sub> concentrations. Looking at P<sub>zinT</sub> it suggests that increasing ZnSO<sub>4</sub> increases fluorescence, with the consensus understanding of Zur, this should not be true. It is possible that these variations are due to some other unknown factor which is involved

in  $P_{zinT}$  regulation, such as additional transcription factors. Figure 4.5 shows that in all cases,  $ZnSO_4$  concentrations (0-400  $\mu M$ ) compared to LB + 20  $\mu M$  TPEN, there is always a significant difference in these two conditions for all Zur regulated promoters (data not shown). This data suggests that the addition of  $ZnSO_4$  up to 400  $\mu M$  in LB does not further repress promoter activity. It also suggests that the zinc concentrations naturally present in LB is high enough to repressed promoter expression of Zur regulated genes.



**Figure 4.5 End Point Fluorescence Assay of Zur Regulated Promoters in LB with Increasing  $\text{ZnSO}_4$  Concentrations**

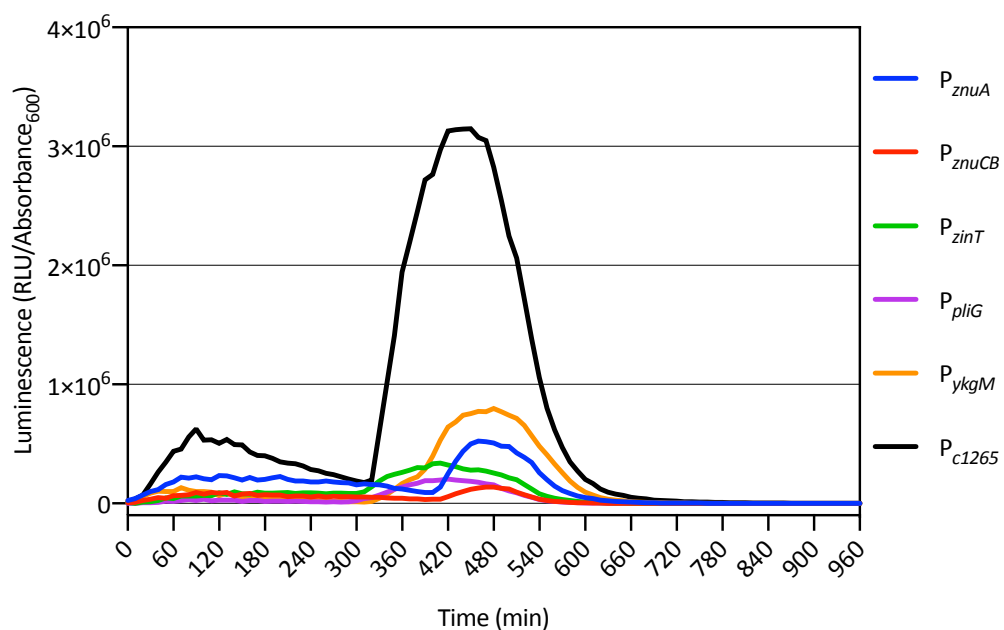
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{LB}_{kan50}$  with various concentrations of  $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Blue bars indicate LB + 20  $\mu\text{M}$  TPEN. Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$  (One-way ANOVA, between samples)

#### 4.2.4. Temporal Luminescence Assay in NH

Figure 4.6 shows the luminescence of the six Zur regulated promoters in NH only media. It is worth noting that error bars are not shown on this graph, for the simple reason that if plotted the graph becomes overcrowded and the data obscured. The experiment was conducted in triplicates, and all statistical tests will incorporate the three biological repeats. Figure 4.6 shows two distinct peaks of luminescence, first at around 120 min and a second at 480 min. Only the second peak at 480 min will be discussed, unless specified otherwise. The first peak is amplified by the normalization equation, which can skew the data, this will be discussed in more depth later.

It is clear that the most expressed promoter is  $P_{c1265}$ , followed by  $P_{ykgM}$  then  $P_{znuA}$ , which agrees with the end point fluorescence data seen in Figure 4.2 and Figure 4.4. Interestingly all six promoters have a similar 'shape' or promoter dynamic. They all have a maximum luminescence peak between 420 to 480 min, which is relatively short-lived. Luminescence is only seen to be high for less than 5 hr, and then drops to relatively low. After peak expression, the luminescence of all six Zur regulated promoter dropped on average, 1000-fold ( $10^{-3}$ ). This is not too dissimilar to the promoter dynamics of  $P_{zntA}$  and  $P_{trc}$  in NH (Figure 3.26).  $P_{zntA}$  and  $P_{trc}$  both show a rapid rate of increase to maximum luminescence and trails off to relatively low levels of expression. The six Zur regulated promoters have a similar promoter dynamic: a rapid rate of increased expression to maximum luminescence and tails off to relatively low levels. However,  $P_{zntA}$ ,  $P_{trc}$  and Zur regulated promoters expressed in LB show different dynamics than in NH, but similar dynamics to each other: rapid rate to maximum expression and then stabilised luminescence like maximum luminescence (Figure 4.8) (Zur regulated promoters in LB to be discussed in full details below). these data suggest that promoter dynamics largely depends on the growth media. NH media may cause

unwanted and misleading side effects on luminescence assays, which had not previously been thought or been discussed. As discussed in Chapter 3, it is possible that NH cannot support bacterial growth for as long as LB can; this could reduce production of the luciferase substrate, which leads to a decline in luminescence. For example, ZnuA is involved in zinc import during zinc depleted conditions; one would expect *znuA* to be expressed at a high and stabilised level in NH, rather than the peak which is observed. This peak could also be due to the internal concentration of zinc is now enough to cause repression of Zur, and as this coincides with stationary growth phase; *E. coli* are no longer replicated and thus no longer need to import zinc.



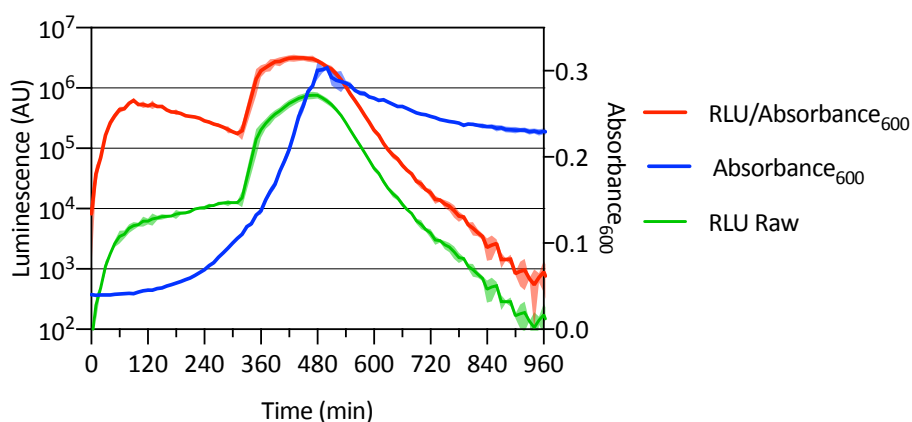
**Figure 4.6 Temporal Luminescence Assay of Zur Regulated Promoters in NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in  $NH_{kan50}$ , with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 16 hr in a TECAN GENios Pro plate reader. Error bars NOT shown due to overcrowding of data on graph

Using the example of  $P_{c1265}$  grown in NH only media, normalised luminescence, raw luminescence, and Absorbance<sub>600</sub> was plotted onto a single graph (Figure 4.7). A  $\log_{10}$  y-axis was used to more easily show the dual peak's value. As mentioned previously the first peak of  $P_{c1265}$  in Figure 4.6 is amplified using the normalisation equation. If one



looks at the raw luminescence, there are still two peaks in the promoter dynamics. Normalised luminescence gives the impression that luminescence drops after the first peak, where as raw luminescence shows a plateau in luminescence. The second, and major peak, shows similar promoter dynamics in both the normalized and raw luminescence data. The second peak also coincide with the mid log phase of bacterial growth, and the decline in luminescence can be seen when bacteria has reached stationary phase. The first peak can be observed in  $P_{c1265}$  in NH only and NH + 400  $\mu$ M  $ZnSO_4$ , but the second peak is only observed in NH only (Data not shown). Again, the promoter dynamics observed in NH are vastly different than in LB (discussed below), caused by the different composition of the two growth media used. However, it cannot be ruled out there that are other transcription factors, or catabolite repression during growth in NH, but not LB, which causes this distinctive dual peak of luminescence.



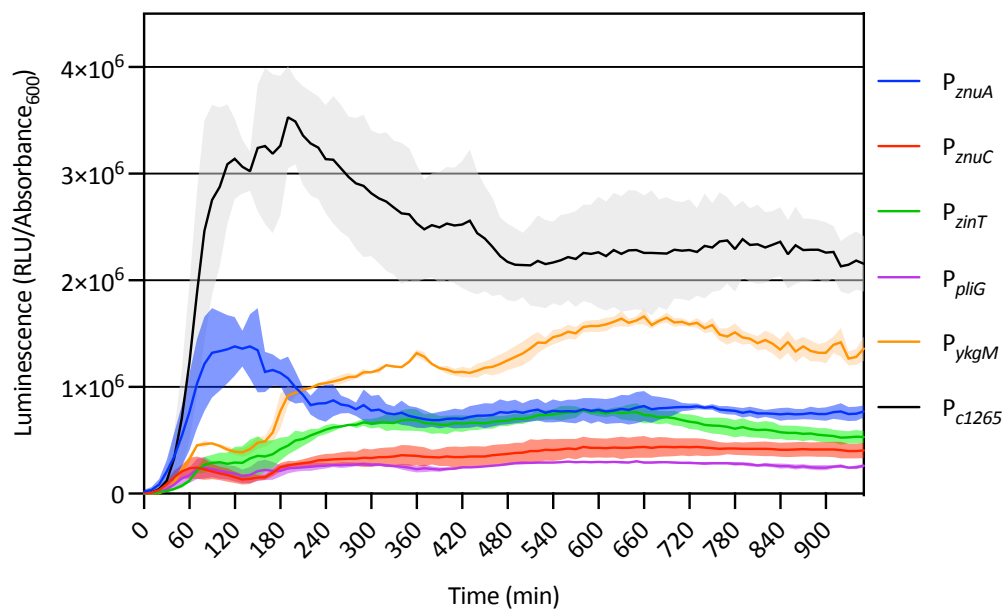
**Figure 4.7 Temporal Assay of  $P_{c1265}$  in NH; Luminescence and Absorbance**

Temporal expression of  $P_{c1265}$  in  $NH_{kan50}$  media. Left Y-axis shows  $\text{Log}_{10}$  luminescence (raw and normalized), right Y-axis shows Absorbance. Shaded area shows S.D.

#### 4.2.5. Temporal Luminescence Assays in LB

As with end point fluorescence assays, Zur regulated promoters can be ‘induced’ with the addition on TPEN. Temporal luminescence assays were conducted in LB for all six Zur regulated promoters.

Figure 4.8 shows temporal luminescence of induced Zur regulated promoters in LB + 20  $\mu$ M TPEN.  $P_{c1265}$ , like in NH, shows the highest peak luminescence. On average, throughout the time course of the assay,  $P_{c1265}$  shows higher luminescence than the other five Zur regulated promoters.  $P_{c1265}$  maximum luminescence is observed at between 200-220 min, luminescence then decreases, and stabilises.  $P_{ykgM}$  does not have a rapid increase in luminescence at the beginning of the assay, unlike  $P_{c1265}$  and  $P_{znuA}$ , however it steadily increases up until about 660 min where peak luminescence is observed.  $P_{ykgM}$  increases luminescence longer than any other Zur regulated promoter.  $P_{znuA}$  has a more rapid increase to maximum expression than  $P_{ykgM}$ , and even has higher luminescence for 180 min, where after  $P_{ykgM}$  has higher luminescence for the remainder of the assay. If compared to end point fluorescence assay in LB + 20  $\mu$ M TPEN (Figure 4.4)  $P_{ykgM}$  shows higher fluorescence than  $P_{znuA}$ . Even though  $P_{znuA}$  shows a higher expression rate than  $P_{ykgM}$ , the accumulation of  $P_{ykgM}$  is higher than  $P_{znuA}$ . This can easily be seen in end point assays, and can be deduced in temporal assay, as  $P_{ykgM}$  has a higher luminescence for most of the temporal assay.



**Figure 4.8 Temporal Luminescence Assay of Zur Promoters in LB + 20 µM TPEN**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in LB<sub>kan50</sub> + 20 µM TPEN, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 16 hr in a TECAN GENios Pro plate reader. Shaded areas show S.D.

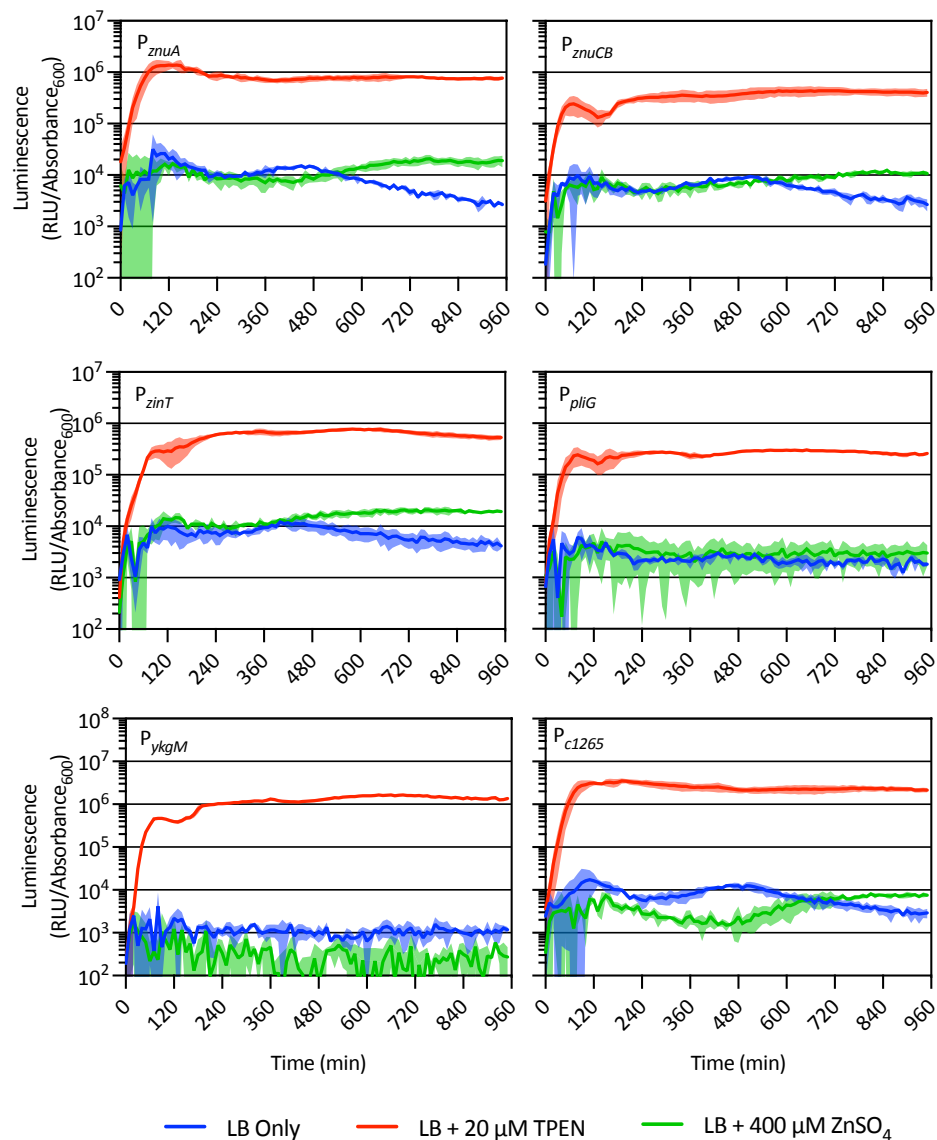
Induced (LB + 20 µM TPEN)  $P_{c1265}$  and  $P_{znuA}$  show similar promoter dynamics: a rapid increase followed by a slight decline to a plateau of luminescence. Whereas  $P_{ykgM}$  has a different promoter dynamic: a steadier increase of luminescence, which takes longer than  $P_{znuA}$  and  $P_{c1265}$  to reach maximal luminescence. This suggest that the promoter activity, and gene expression of these genes are different. Similarly, if we look at maximum expression from highest to lowest,  $P_{c1265} > P_{ykgM} > P_{znuA} > P_{zinT} > P_{znuC} > P_{pliG}$ , this is similar to that observed in NH, except  $P_{pliG}$  was higher than  $P_{znuC}$  (Figure 4.7). However, it is hard to compare expression of  $P_{ykgM}$  and  $P_{znuA}$  in LB, due to different promoter dynamics.

The data observed in Figure 4.8 is similar and agrees with the data seen in end point fluorescence assays in LB + 20 µM TPEN (Figure 4.4), that  $P_{c1265}$  is the highest expressed, followed by  $P_{ykgM}$ .

Again, to be able to understand promoter activity, with the desire to develop a new inducible promoter system, it is essential to look at the basal level of expression, i.e LB only, and further suppression with LB + 400  $\mu$ M ZnSO<sub>4</sub>. Figure 4.9 shows temporal luminescence of the six Zur regulated promoter either; uninduced (LB only), suppressed (LB + 400  $\mu$ M ZnSO<sub>4</sub>) or induced (LB + 20  $\mu$ M TPEN). Generally, there is very little difference between uninduced and suppressed promoter activity in all six Zur regulated promoters. In *P<sub>znuA</sub>*, *P<sub>znuCB</sub>*, and *P<sub>zinT</sub>*, after about 480 min luminescence seems to be slightly higher in LB + 400  $\mu$ M ZnSO<sub>4</sub> than LB only, but still over a 10-fold lower than induced. However, all six promoters show a distinct difference in luminescence in induced (LB + 20  $\mu$ M TPEN) and uninduced (LB only).

As Figure 4.9 shows there is rapid increase in luminescence to around 120 min, where it then plateaus, true for all six Zur regulated promoters. *P<sub>c1265</sub>* shows the highest plateau luminescence, and *P<sub>plIG</sub>* shows the lowest. Similar to the end point fluorescent assays (Figure 4.4), temporal assays suggest that *P<sub>ykgM</sub>* has the largest increase in promoter activity, from LB only to LB + 20  $\mu$ M TPEN, which is closely followed by *P<sub>c1265</sub>*.

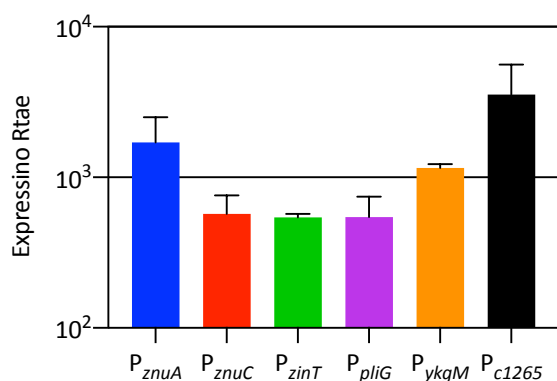
Interestingly, if one compares the temporal luminescence of *P<sub>c1265</sub>* in NH (Figure 4.7) and in LB (Figure 4.9), again it is clear that the promoter dynamics are different. In LB promoter activity plateaus after 120 min till the end of the assay. Whereas in NH there are two distinct peaks, which drop significantly. However, the peak in NH shows approximately the same maximum luminescence as the plateau in LB + 20  $\mu$ M (5x10<sup>6</sup>).



**Figure 4.9 Temporal Luminescence Assay of Zur Regulated Promoters in LB: Induced; Uninduced; and Suppressed**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in LB<sub>kan50</sub> only, LB<sub>kan50</sub> + 20 μM TPEN or LB + 400 μM ZnSO<sub>4</sub>, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader. Induced (LB + 10 μM TPEN), uninduced (LB only), Repressed (LB + 400 μM ZnSO<sub>4</sub>). Shaded areas show S.D.

Expression rate was determined for all six Zur regulated promoters in their induced state (LB + 20 μM TPEN) (Figure 4.10). The expression rates show that *P<sub>c1265</sub>* has the highest rate, followed by *P<sub>znuA</sub>* and then *P<sub>ykgM</sub>*. *P<sub>znuC</sub>*, *P<sub>zinT</sub>* and *P<sub>pliG</sub>* all show similar rates of expression, with a one-way ANOVA returned no significant variation ( $p \geq 0.05$ ).



**Figure 4.10 Luminescence Expression Rate of Zur Promoters in LB + 20  $\mu$ M TPEN**

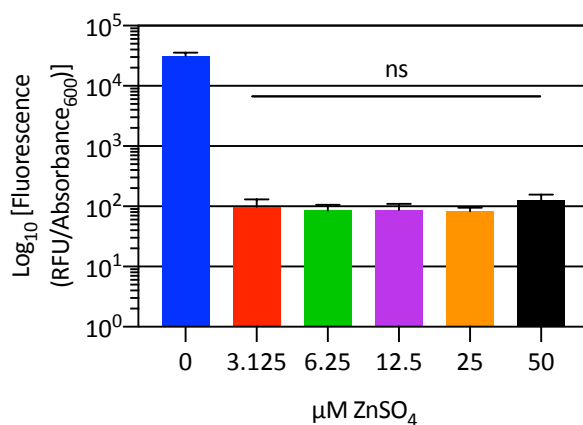
Expression rate (RLU/Absorbance<sub>600</sub> min<sup>-1</sup>) of Zur regulated promoters in LB + 20  $\mu$ M TPEN (induced). Error bars indicated S.D

This temporal luminescence data, especially for P<sub>c1265</sub> and P<sub>ykgM</sub>, shows the potential use for a novel TPEN inducible promoter. Both show strong a statistical difference between an induced and uninduced state. Further to this, when expressed in LB, promoter dynamics are favourable for an inducible promoter system. Promoter activity rapidly rises to maximum expression, and then stabilised at this high level. P<sub>c1265</sub> shows the highest level of expression rate out of the six Zur promoters.

#### 4.2.6. Further understanding of P<sub>c1265</sub>

To further understand the promoter activity of the novel Zur regulated gene operon, *c1265-7*, more tests were conducted. A test was conducted to determine what concentration of ZnSO<sub>4</sub> is needed to be added to NH media to reduce promoter activity. Figure 4.11 shows that with the addition of only 3.125  $\mu$ M ZnSO<sub>4</sub> in NH media, it significantly reduced the fluorescence. This addition of 3.125  $\mu$ M ZnSO<sub>4</sub> seems low but considering that ICP-MS analysis determine the zinc concentration of NH to be 115 nM (Table 3.1), this makes the addition of 3.125  $\mu$ M ZnSO<sub>4</sub> a 27-fold increase in

zinc. Further, there is no variation in fluorescence between samples with additional  $\text{ZnSO}_4$  from 3.125 to 50  $\mu\text{M}$   $\text{ZnSO}_4$ .



**Figure 4.11 End Point Fluorescence Assay of  $P_{c1265}$  in NH with Increasing  $\text{ZnSO}_4$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{NH}_{\text{kan50}}$  with increasing  $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. ns  $p > 0.05$  (One-way ANOVA)

### 4.3. Discussion and Conclusion

#### 4.3.1. End Point Fluorescence Assay

End point fluorescence assays conduction in NH only showed the theoretical maximum expression of Zur regulated promoters. The results in (Figure 4.2) showed that the most upregulated Zur regulated promoter, from repressed to unrepressed, is  $P_{c1265}$ , and the least is  $P_{zinT}$ . Results also showed that between the two states; repressed (NH + 400  $\mu$ M  $ZnSO_4$ ) and unrepressed (NH only)  $P_{c1265}$  showed the greatest variation (90x), followed by  $P_{ykgM}$ . There was very little variation observed with  $P_{zinT}$ .

Interestingly, when end point fluorescence assay was conducted in LB,  $P_{ykgM}$  showed the greatest variation (269x) between the two states: uninduced (LB only) and induced (LB + 20  $\mu$ M TPEN), followed by  $P_{c1265}$  (173x).  $P_{c1265}$  shows the highest maximum induced fluorescence, whereas  $P_{ykgM}$  showed the lowest uninduced fluorescence. However, when comparing the maximum fluorescence in LB and NH, the order of highest to lowest maximum fluorescence was the same;  $P_{c1265} > P_{ykgM} > P_{znuA} > P_{znuCB} > P_{pliG} > P_{zinT}$ . The end point fluorescence data in *E. coli* MG1655 agrees with transcriptional microarray data by Sigdel, Easton and Crowder, (2006) and Hensley *et al.*, (2012) which found *ykgM* to be the most up regulated gene in *E. coli* K-12 when grown in a zinc depleted state.

Results suggest that the zinc levels in LB are sufficient to repress all six Zur regulated promoter, and that the addition of up to 400  $\mu$ M  $ZnSO_4$  will not cause any further repression. A further assay, only conducted on  $P_{c1265}$ , tested the effect of lower concentrations of additional  $ZnSO_4$  on fluorescence (Figure 4.11). Results showed that  $P_{c1265}$  can be repressed with the addition of as little as 3.125  $\mu$ M  $ZnSO_4$ . Interestingly, 3.125  $\mu$ M  $ZnSO_4$  may seem low, but it is in fact 27x more zinc than in NH media only.



Result also showed that the addition of ZnSO<sub>4</sub> between 3.125 µM to 400 µM ZnSO<sub>4</sub> does not cause any significant difference in fluorescence, therefore suggesting that P<sub>c1265</sub> is not a tuneable promoter. However, it can only be said with certainty that P<sub>c1265</sub> is not tuneable between 3.125 µM to 400 µM ZnSO<sub>4</sub>, smaller concentrations of added zinc (<3.125 µM ZnSO<sub>4</sub>) may be needed to observe a tuneable effect.

Interestingly, zinc homeostasis has been shown to control internal zinc levels to femtomolar concentration, allowing approximately two Zn<sup>2+</sup> ions free internally (Outten and O'Halloran, 2001). Data from this chapter suggest that, unlike ZntR regulation, Zur regulation is not tuneable within the same range which ZntR shows a tuneable response to zinc concentration. However Outten and O'Halloran, (2001) demonstrated a tuneable response to both ZntR and Zur within the same zinc concentration range [Outten and O'Halloran, (2001) *Fig. 4*], this was not observed in data presented in this study. It is possible that a lower concentration range 0-3.125 µM zinc is needed to observe a tuneable response zinc in Zur regulation.

I hypothesise that the tight internal zinc concentration of *E. coli* is dominantly controlled by ZntA. Data in this chapter suggested that Zur does not alter its response to zinc within the same range that shows tuneable ZntR regulation, suggesting that Zur regulated genes will acquire zinc indiscriminately when in a zinc depleted state, and that ZntA is responsible for the actual internal concentration of zinc. I hypothesis that; Zur is an on / off tap for zinc import, importing zinc equal or greater than the desired level. Then ZntA will, with its tuneable response, export zinc to the desired internal femtomolar concentration.

### 4.3.2. Temporal Luminescence Assays

Temporal luminescence assay in NH showed that  $P_{c1265}$  has the highest luminescence, followed by  $P_{ykgM}$ , which is in line with the fluorescence assay in NH. However,  $P_{zinT}$  did not show the lowest luminescence, which is suggested from the fluorescence assays,  $P_{znuCB}$  showed the lowest peak in luminescence. The peak of expression in  $P_{c1265}$  can be observed during mid-log phase to early-stationary phase. All six Zur regulated promoters showed various expression levels, or peak luminescence.

Similarly, in LB temporal luminescence assays,  $P_{c1265}$  showed the highest peak luminescence, as well as highest plateau or stabilised luminescence when induced (LB + 20  $\mu$ M TPEN).  $P_{znuA}$  showed the second highest peak luminescence early in the assay which dropped to a lower plateau expression. Whereas  $P_{ykgM}$  had a low luminescence early in the assay, which continued to rise to the second highest plateau luminescence. All six Zur regulated promoters showed stable levels of expression throughout most of the assay. Either a slow or rapid increase to a plateau of stable luminescence level. This stable luminescence (gene expression) is vital for an inducible promoter system. Stable gene expression throughout an experiment is more desirable than a short-lived burst of expression. However, there may be instances where a short burst of expression is more desirable; all six promoters showed significant difference between an induced state (LB + 20  $\mu$ M TPEN) and an uninduced state (LB only).

Expression dynamics are drastically different between expression in NH media and LB media. Expression dynamics in NH show a large peak: rapid increase and rapid decrease in luminescence. Whereas expression dynamics in LB showed a rapid increase, followed by a plateau of stable luminescence. These variations in expression dynamics can also be seen in the ZntR regulated  $P_{zntA}$  in Chapter 3, (Figure 3.32). As this peak expression is observed in ZntR and Zur regulated promoters, it further

strengthens the arguments that NH media may not be ideal for luminescence assays, as the expression dynamics are not as one would expect. NH media may have a previously unknown side effect on luminescence expression. Due to the defined nature of NH, as previously mentioned, it is possible that NH cannot support growth for as long as LB can, consequently leading to the bacteria not being able to produce the luciferase substrate.

Results for  $P_{zinT}$  suggest that little to no regulation is observed in NH, but significant regulation is observed in LB media. As mentioned, there are differences in the expression dynamics between NH and LB media. It is possible that growth in NH causes suppression of the regulation of *zinT*, which is only observed in NH and not LB.

#### 4.3.3. Concluding remarks

$P_{zinT}$  does not seem to offer any potential use as an inducible promoter system. There is very little difference between an induced and uninduced state, or a suppressed or unrepresed state. Further to this, when induced in LB, it shows the slowest expression rate of all the six Zur regulated promoters.

$P_{c1265}$  shows strong promise for an on/off inducible promoter system;  $P_{c1265}$  showed the highest expression out of the six Zur regulated promoters tested.  $P_{c1265}$  showed the highest level of expression in both induced LB and NH.  $P_{c1265}$  shows significant increase in fluorescence between uninduced and induced (>100x). Luminescence assays conducted in induced LB (LB + 20  $\mu$ M TPEN) showed rapid induction, followed by stable and high expression throughout the whole assay (16 hr). As *c1265-7* is only found in UPEC *E. coli* strains, it is more advantageous as  $P_{c1265}$  will not be found chromosomally in lab strain *E. coli*.

Even though  $P_{c1265}$  shows the most promise for a novel TPEN inducible promoter system;  $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{plig}$  and  $P_{ykgM}$  also show promise. These other four promoters are useful if various levels of induction are required, i.e. in an operon. These are all regulated by Zur, and induced with TPEN, genes within a synthetic operon could be induced with the same inducer but elicit different levels of induction.

I propose a new hypothesis on zinc Homeostasis in *E. coli*. The ability of *E. coli* to control its internal zinc concentration to a femtomolar level is controlled, mostly, via ZntA. ZntA demonstrates tuneable gene regulation in response to  $ZnSO_4$  concentrations, whereas the six Zur regulated promoters showed no indication of tuneability. I hypothesize that Zur is an on/off tap for zinc import, importing zinc equal or greater than the desired level. Then, ZntA will, with its tuneable response, export zinc to the desired internal femtomolar concentration

#### 4.3.4. Future work

Unlike  $P_{zntA}$ , the six Zur regulated promoters have not been tested with RT-qPCR to determine if the luminescence and fluorescence change is observed due to transcriptional gene regulation. To improve the reliability of this data, RT-qPCR validations of all six Zur regulated promoter should be conducted. The same protocol used for  $P_{zntA}$  validation through RT-qPCR can be used, with *rrsA* as a reference gene. [This experiment was intended to part of this thesis. However, due to the 2020 COVID-19 global pandemic, national lockdown reduced my experiment lab time, and this experiment was not able to be conducted].

The data thus far shows the potential use of  $P_{c1265}$  and  $P_{ykgM}$  as a novel TPEN inducible (and zinc repressible) promoter system. However, to further develop our understanding of  $P_{c1265}$  and  $P_{ykgM}$  as a novel inducible promoter, further induction

assays would be beneficial. It would be useful to understand the expression dynamics of these promoters when induced during mid-log phase, rather than at lag phase as in all the above experiments.

To further strengthen my argument that ZntA controlled the internal zinc concentration, rather than the Zur regulated genes, it would be useful to look at all six Zur regulated promoters to determine at what level of zinc is needed to cause repressed promoter expression in NH.

## Chapter 5

# Understanding the Dynamic Response of the Transcriptional Regulation of *zntA*

### 5.1. Introduction

ZntA is a P-type ATPase, which, during increased intracellular zinc concentrations will export zinc from the cytoplasm to the periplasm. Expression of ZntA is controlled through transcriptional regulation via the ZntR transcription factor. ZntR is a MerR family transcriptional activator; essentially when internal zinc concentrations are high,  $\text{Zn}^{2+}$  binds to ZntR, causing ZntR to upregulate *zntA* transcriptional expression.

It has been hypothesised that during normal and increased environmental zinc concentrations, *E. coli* mounts a heterogeneous response in *zntA* transcriptional regulation (Takahashi *et al.*, 2015). A heterogeneous gene response, sometimes referred to as ‘bet-hedging’, is the mechanism in which single cells within a genetically identical population, and within a homogenous environment regulate a specific gene at different levels. This is a well-studied (both experimental and mathematically modelled) bacterial mechanism of survival (Raj and van Oudenaarden, 2008; Grote, Krysciak and Streit, 2015; Fritz, Walker and Gerland, 2019; Han and Zhang, 2020).

#### 5.1.1. Aim and Objectives

The main of this chapter was to further understand the stochastic response to the ZntR regulated  $P_{zntA}$  and determine if *zntA* shows a homogenous or heterogenous transcriptional gene regulation.

## 5.2. Results

### 5.2.1. Construction and Validation of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>*

*mRFP1* was chromosomally fused to the C-terminal end of ZntA using the gene doctoring method (Lee *et al.*, 2009), which uses  $\lambda$ -Red homologous recombination. Two plasmids were used; a helper plasmid (pACBSR) which contains the  $\lambda$ -Red genes, and the doctoring plasmid (pDOC-*zntA:rfp*) which contains the fusion gene with homologous DNA at the site of the chromosomal fusion.

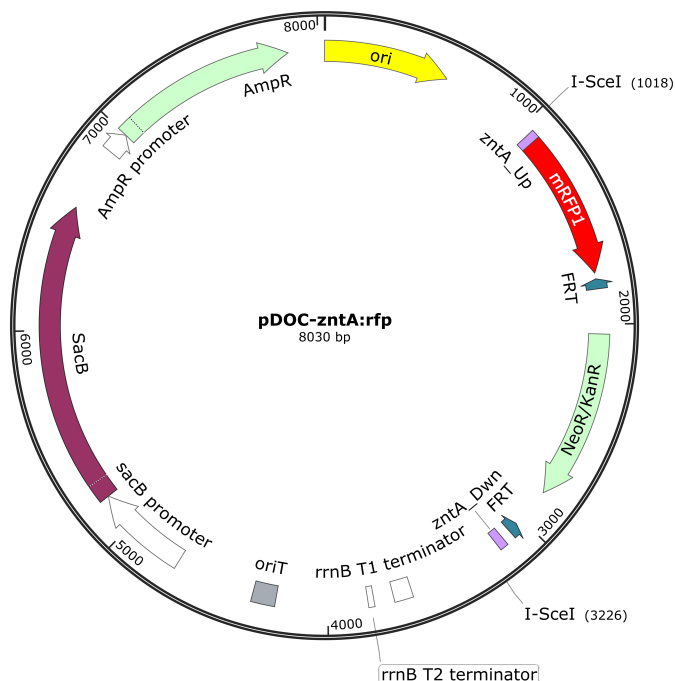
The *zntA:rfp::kan<sup>R</sup>* chromosomal fusion was created for the use in flow cytometry assays. By using a fluorescent protein sequence chromosomally fused to a gene, it removes the influence that plasmid copy number plays on results and mitigates translation variation. In theory, measured fluorescence is direct measure of ZntA. Further, it allows native gene expression to occur, from the native chromosomal position of the gene. mRFP1 was selected as the fluorescent protein, as this allows harmony between plasmid assays and chromosome fusion assays, as both use mRFP1. Note; “mRFP1” will be shortened to “rfp” throughout this chapter.

#### 5.2.1.1. pDOC-*zntA:rfp* Construction

Lee *et al.*, (2009) defined a process of chromosomal gene fusion which requires: a donor plasmid, pDOC-G (containing *gfp* and *kan<sup>R</sup>* with up and downstream *FRT* sites); a cloning plasmid, pDOC-C; and a helper plasmid, pACBSR (containing  $\lambda$ -red). In the initial step, the *gfp* and *kan<sup>R</sup>* genes are amplified from pDOC-G using PCR primers which incorporate homologous chromosomal DNA sequences for the desired insertion site, designed into the 5'-end of the primers. The *gfp* and *kan<sup>R</sup>* gene amplicons were cloned into the cloning vector, pDOC-C. However, as *mRFP1* was to be used, rather than *gfp*, it was easier to bypass the donor plasmid step, and use NEB DNA HiFi assembly to



directly construct the final plasmid (pDOC-zntA:rfp) containing; *kan<sup>R</sup>*, *mRFP1* and chromosomal homologous chromosomal DNA (Figure 5.1).



**Figure 5.1 Plasmid Map of pDOC-zntA:rfp**

Plasmid Map of pDOC-zntA:rfp showing relevant genes.

pDOC-C was linearized at the MCS using PCR amplification; *mRFP1* was amplified from pUltra-RFP incorporating into the amplicon 40 bp of DNA homologous to the chromosomal DNA directly upstream of the stop site (taa) of *zntA*; and *kan<sup>R</sup>* (with the up-stream and down-stream *FRT* sites) was amplified from pDOC-G whilst incorporating 40 bp of homologous chromosomal DNA directly down-stream of *zntA*. NEBuilder HiFi DNA assembly was used for plasmid construction. pDOC-zntA:rfp was sent for sequencing (using primers pDOC-C\_seq\_F and pDOC-C\_seq\_R). Sequence results showed that it was the correct construction (data not shown).

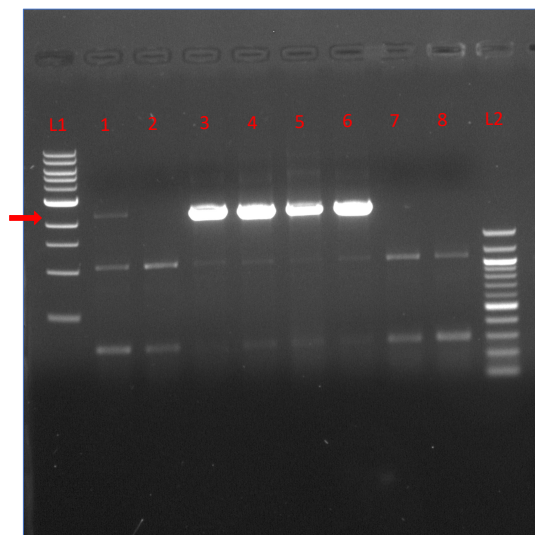
### 5.2.1.2. MG1655 *zntA:rfp::kan<sup>R</sup>* $\lambda$ -Red Induction and Validation

pDOC-*zntA:rfp* and pACBSR were transformed into *E. coli* MG1655 and  $\lambda$ -red was induced with arabinose. Patch plating on LB<sub>kan50</sub>, LB<sub>chlor35</sub>, and LB<sub>amp100</sub> was used to screen for chromosomal fusions. A colony which was sensitive to ampicillin, but resistant to kanamycin and chloramphenicol was chosen. Kan<sup>R</sup> with Amp<sup>S</sup> suggests chromosomal fusion, if Kan<sup>R</sup> and Amp<sup>R</sup> was observed it would suggest that pDOC-*zntA:rfp* was still present. Chlor<sup>R</sup> suggest that the helper plasmid pACBSR was still present.

An Amp<sup>S</sup>, Kan<sup>R</sup>, and Chlor<sup>R</sup> colony was selected, and was subject to plasmid curing at 43°C to remove the remaining helper plasmid (pACBSR). Single colonies were patch plated on to LB<sub>kan50</sub> and LB<sub>chlor35</sub> and incubated at 43°C. A colony with a Kan<sup>R</sup> and Chlor<sup>S</sup> phenotype suggested that pACBSR had been lost and a chromosomal fusion of *zntA:rfp::kan<sup>R</sup>* was present.

PCR screening was conducted on Kan<sup>R</sup>, Chlor<sup>S</sup> and Amp<sup>S</sup> colonies, using colony PCR with the primers *zntA:rfp\_scan\_F* and *zntA:rfp\_scan\_R*. *zntA:rfp\_scan\_F* is homologous to a region 159 bp upstream of the stop (att) and the site of potential fusion. *zntA:rfp\_scan\_R* is homologous to a region 139 bp downstream of *zntA* and the potential site of fusion. If no chromosomal fusion occurred, a PCR product of 298 bp would be observed; if a chromosomal fusion is present, a PCR product of 2259 bp would be observed. Eight potential colonies were screened for the *zntA:rfp::kan<sup>R</sup>* chromosomal fusion. 50% of the screened isolates showed successful insertion of *zntA:rfp::kan<sup>R</sup>* (Figure 5.2). Lanes 3 to 6 show the correct band size (~2259 bp), which suggests the correct gene fusion; these PCR products were purified and sent for Sanger sequencing using the primers *zntA:rfp\_scan\_F* and *zntA:rfp\_scan\_R*. Sequence results

were analysed to check for the correct insertion where data showed the correct fusion of *zntA:rfp::kan<sup>R</sup>* (data not shown).

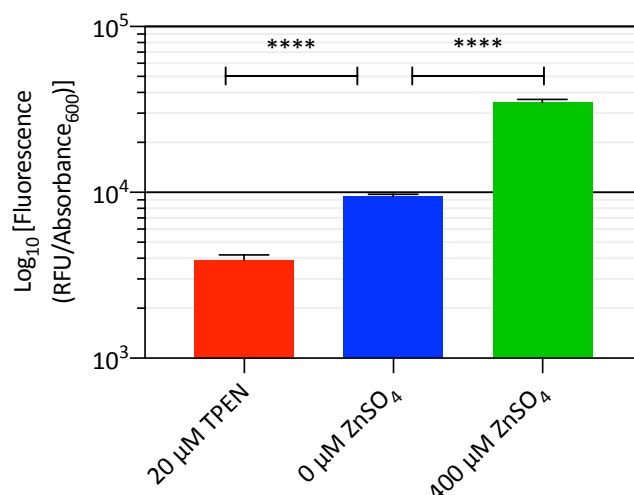


**Figure 5.2 PCR Product Gel Image for Conformation of *zntA:rfp::kan<sup>R</sup>* Fusion**

1% Gel image of PCR products from colony PCR to identify successful *zntA:rfp::kan<sup>R</sup>* fusions. L1 1 kbp ladder. L2 100 bp ladder. Lanes 1 to 8 isolates. Red arrow indicates expected band size for successful fusion (2259bp).

### 5.2.2. *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* Test of Validation

A simple test was conducted to validate the fluorescence of the chromosomal fusion of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>*. Knowing that *zntA* is upregulated in the presence of  $\text{ZnSO}_4$ , a standard overnight end point fluorescence assay was conducted, with *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* grown in LB only, LB + 20  $\mu\text{M}$  TPEN or LB + 400  $\mu\text{M}$   $\text{ZnSO}_4$ . Figure 5.3 shows there is a significant increase in fluorescence when *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* was grown in the presence of 400  $\mu\text{M}$   $\text{ZnSO}_4$  compared to LB only. There is also a significant reduction in fluorescence when *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* was grown in LB containing 20  $\mu\text{M}$  TPEN compared to cells grown in LB only. These results show that the chromosomal fusion of *zntA:rfp::kan<sup>R</sup>* has worked, and shows that the variations in fluorescence are in line with what one expects of *zntA* regulation. This has validated the use of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* for flow cytometry experiments.



**Figure 5.3 End Point Fluorescence Assay of Chromosomally Fused *zntA:rfp::kan<sup>R</sup>***

*E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> with either 20 µM TPEN or 400 µM ZnSO<sub>4</sub> added, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. \*\*\*\*  $p \leq 0.0001$  (two-tailed, unpaired, t-test).

### 5.2.3. Flow Cytometry

As previously described, mRFP1 was fused to the C-terminal end of ZntA, therefore fluorescence is a direct representation of ZntA abundance. Flow cytometry allows one to measure the fluorescence of a population of cells, by measuring the fluorescence of individual cells within the population. Standard end point fluorescence assays, as used in Chapter 3 and Chapter 4, measure the fluorescence of the whole population, returning an average fluorescence value. These standard fluorescence assays usually divide fluorescence over optical density (OD), to give a proxy of relative fluorescence units per cell. These assays do not provide information on cell to cell variation of fluorescence.

Previous work by Takahashi *et al.*, (2015) proposed that during normal and increased zinc concentrations, *E. coli* populations mount a heterogeneous response to zinc, regulating *zntA* at various levels within a genetically identical population and a homogenous environment. The theory behind this is that *E. coli* use heterogeneous

gene regulation as a survival mechanism, allowing cells within a population to more readily prepare for zinc shock. This hypothesis was proposed through a combination of experimental and mathematical modelling.

This work follows on from the hypothesis proposed by Takahashi *et al.*, (2015), and attempts to use flow cytometry to examine regulation of *zntA* in individual cells within a homogeneous environment. This should provide a greater understanding of the regulation of *zntA* and allow us to determine if there is a heterogeneous or homogeneous gene regulation of *zntA* in response to zinc.

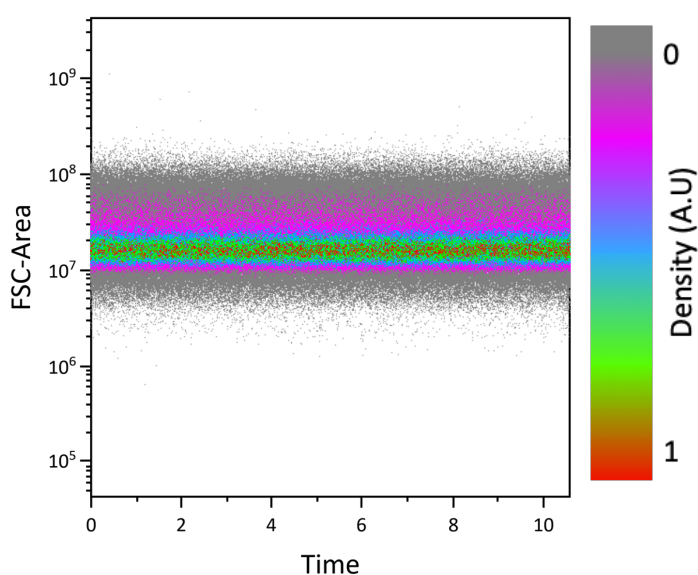
#### 5.2.3.1. Analysis - Flow Cytometry Gate

Flow cytometry uses an analytical method to group a population cells with the same characteristics, as well as to exclude data points and set value limits (to account for background noise); these are known as 'gates' or 'gating'. Gating is conducted on the experimental data using a flow cytometry analysis software, such as Kaluza (Beckman). Four gates were used for the flow cytometry analysis procedure in this work; injection rate; whole cells; single cells; and fluorescence. An example of the methodology for each of these gates is shown below. This gating procedure was conducted on all flow cytometry data in this thesis, however, the gating steps are not shown for every data set.

Cells to be analysed are funnelled through the nozzle within the flow cytometer, creating a single stream of cells which are excited by a laser, one cell at a time. The flow cytometer records the light scatter produced by the cells: side scatter and forward scatter. Forward scatter is the measure of light scattered along the same path as the excitation laser. Side scatter measures the light scatter at 90° to the path of the laser.

### 5.2.3.2. Gate 1 - Gate for Injection Rate

The first gate was created for injection rate, this shows if there is a steady rate of injection of cells into the flow cytometer channel. Figure 5.4 shows an example of a density plot showing; forward scatter (FSC) area against time (s). This is an example of a constant flow rate, as there are no notable breaks in the vertical scatter of data. No gate was needed for this data set.

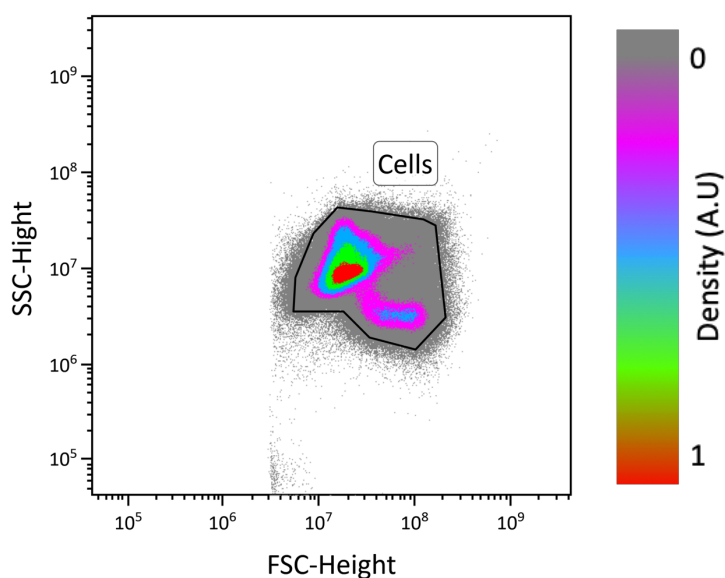


**Figure 5.4 Flow Cytometry Gate for Injection Rate**

Density plot of Forward Scatter (FCS) area against time (s) to show rate of injection.

### 5.2.3.3. Gate 2 - Gate for Whole Cells

A second gate was created to distinguish whole cells from debris and lysed cells. Comparing side scatter (SSC) height and forward scatter (FSC) height will show whole cells. SSC is used to look at granularity and structure, whereas FSC is used to look at the proportional size of the cell. This gate is generally more useful when gating for multiple cell types in a single sample, however it can be applied to remove debris. Figure 5.5 shows an example of a density plot showing; side scatter (SSC) against forward scatter (FCS). “Cells” shows the gate for whole cells, which will be analysed.

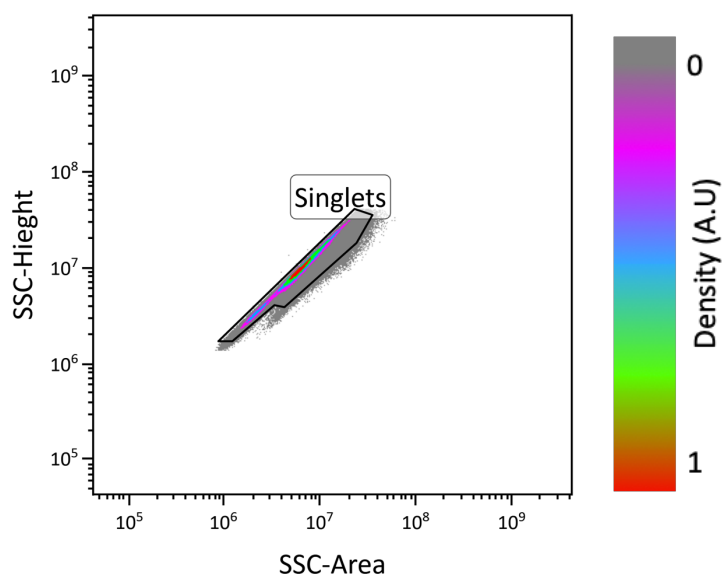


**Figure 5.5 Flow Cytometry Gate for Whole Cells**

Density plot of Side Scatter (SSC) height against Forward Scatter (FCS) height to distinguish whole cells. “Cells” shows gate of whole cells.

#### 5.2.3.4. Gate 3 - Gate for Single Cells.

A third gate is used to determine single cells. Comparing side scatter (SSC) area and side scatter (SSC) height can be used to distinguish single cells from cells which may have clumped together. The theory is that single cells will have almost identical SSC height and area, however, the shape of the cell will affect this. If cells are stuck together, there would be an obvious increase in either SSC height or SSC area, depending on the orientation when passing through the flow cell. Figure 5.6 shows an example of a density plot showing; side scatter (SSC) area against SSC height. “singlets” shows the gate for single cells.



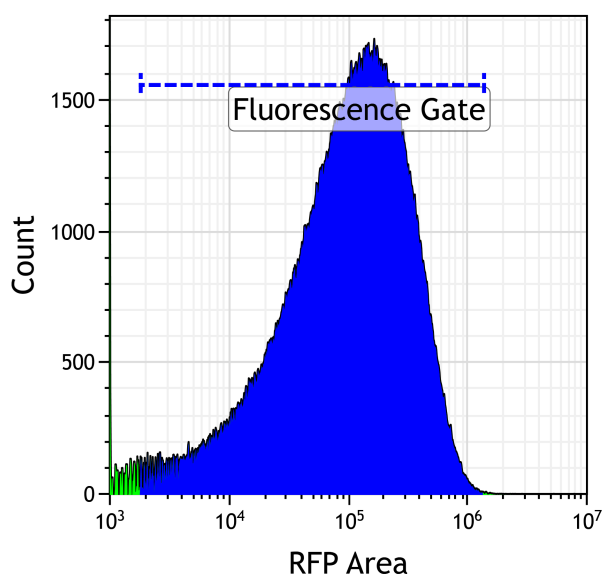
**Figure 5.6 Flow Cytometry Gate for Single Cells**

Density plot of Side Scatter (SSC) height against Side Scatter (SSC) area, to determine single cells. “Singlets” shows gate form single cells. Gated for whole cells from Figure 5.5.



### 5.2.3.5. Gate 4 –Gate for Fluorescence

A final gate was used to distinguish actual fluorescence data from background fluorescence noise. Background noise can be due to a variety of reasons, including: thermionic emission; stray light; and electrical circuits. Noise was removed by comparing RFP (area) against cell count. The auto-gate function within the Kaluza Analysis Software 1.5 (Beckman Coulter) was used for this gate. Figure 5.7 shows an example of a histogram showing RFP area against cell count. “Fluorescence gate” shows gate for true fluorescence data. The large fluctuation in fluorescence in low fluorescent values suggests that the data is not real and is noise.



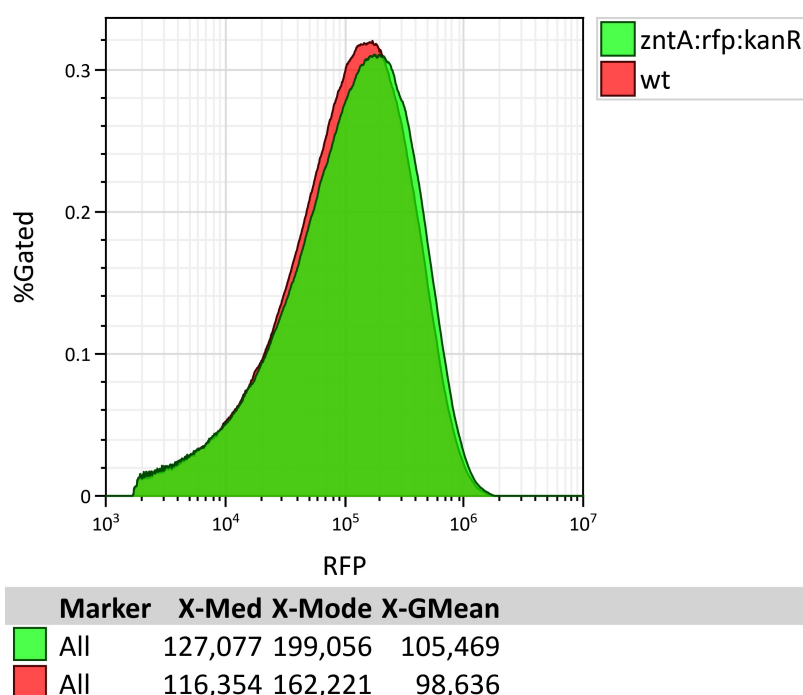
**Figure 5.7 Flow Cytometry Gate for Fluorescence**

Histogram plot of RFP fluorescence area against cell count. Auto gate function used to fluorescence noise. “Fluorescent Gate” shows gate for foreground fluorescence data.

#### 5.2.4. Flow Cytometry of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* in LB

##### 5.2.4.1. Background Fluorescence

An initial flow cytometry experiment was conducted to compare the fluorescence of *E. coli* MG1655 *wt* (henceforth referred to as *wt* or *E. coli wt*) and *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* (henceforth referred to as *zntA:rfp* or *E. coli zntA:rfp*) grown in LB only. *E. coli wt* was used, as in theory this should produce the lowest level of fluorescence, as no fluorescent protein is present in the cell; any fluorescence can be viewed as autofluorescence of the cells and/or media. Figure 5.8 shows a histogram for fluorescence of *E. coli wt* and *zntA:rfp* grown in LB only. *E. coli zntA:rfp* grown in LB should show low fluorescence, but not as low as *wt*, as there will be basal *zntA* expression. Previous data in Chapter 3 shows that even in zinc depleted conditions, there is still basal expression of  $P_{zntA}$ . Heterogenous expression of *zntA* is still predicted in normal zinc concentrations (Takahashi *et al.*, 2015), which could suggest why there is always a basal level of expression of *zntA*. From looking at Figure 5.8 there is little observed difference between the two samples, however *zntA:rfp* has slightly higher fluorescence, and the histogram is shifted to the right compared to the *wt*. The median *zntA:rfp* fluorescence is considerably higher than in the *wt*. This confirms the shift of the histogram to the right. The mode is also much higher in *zntA:rfp* than in the *wt*, meaning the peak fluorescence of *zntA:rfp* is higher than *wt*, which is to be expected.



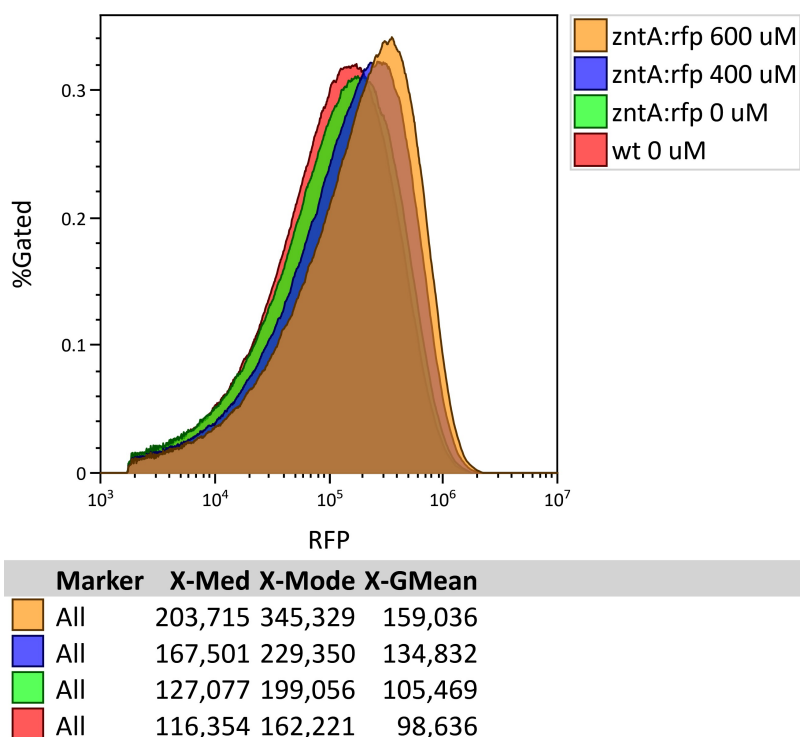
**Figure 5.8 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* MG1655 (wt and *zntA::rfp::kan<sup>R</sup>*) in LB Only**

*E. coli* MG1655 wt and *zntA::rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in LB only. After 16 hr incubation, 100  $\mu$ L of bacterial culture was pelleted (13,000  $\times g$  for 1 min), re-suspended in 1000  $\mu$ L 4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

#### 5.2.4.2. Foreground Fluorescence

To help determine if *zntA* mounts a heterogeneous gene expression in response to zinc, *E. coli zntA::rfp* was grown in LB with the addition of  $\text{ZnSO}_4$  (0, 400, and 600  $\mu\text{M}$ ), in order to look at the distribution of fluorescence (Figure 5.9). 400  $\mu\text{M}$   $\text{ZnSO}_4$  was chosen to add to LB, as previous end point fluorescence data showed that 400  $\mu\text{M}$   $\text{ZnSO}_4$  had the highest increase in fluorescence (Figure 3.8). Takahashi *et al.*, (2015) demonstrated potential heterogeneity of ZntA with the lower addition of 25  $\mu\text{M}$  or 100  $\mu\text{M}$  zinc. However, when this lower addition of  $\text{ZnSO}_4$  was added in the flow cytometry experiment, no significant difference was observed in the flow cytometry fluorescence compared to 0  $\mu\text{M}$   $\text{ZnSO}_4$  (data not shown). Therefore, the added  $\text{ZnSO}_4$  was increased to 400 and 600  $\mu\text{M}$   $\text{ZnSO}_4$  in the hopes to increase gene expression and

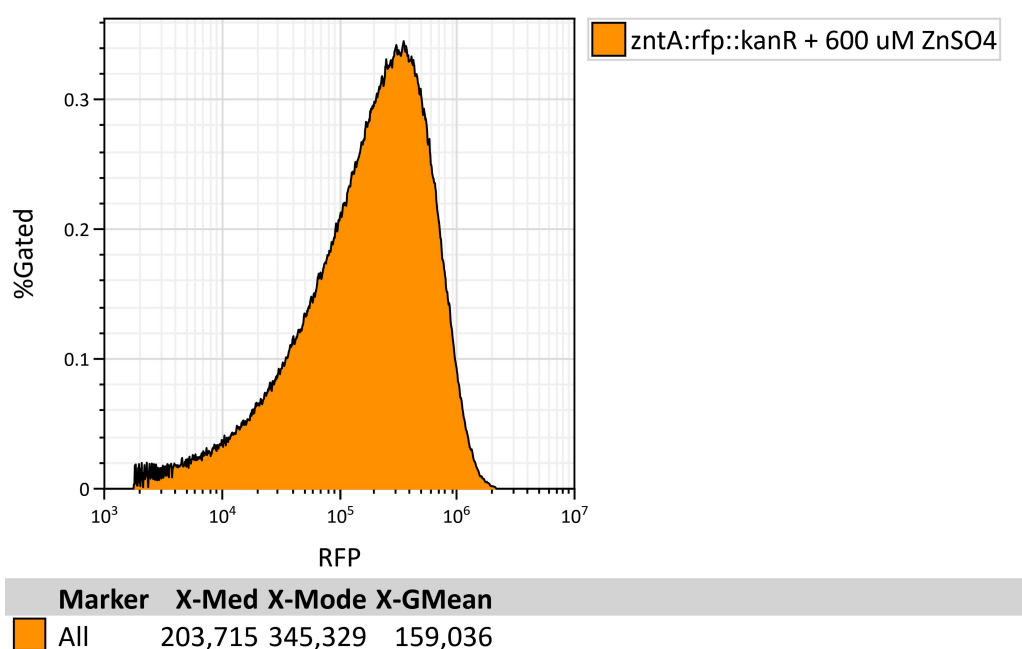
return measurable fluorescence. *E. coli zntA::rfp* grown with 600  $\mu\text{M}$   $\text{ZnSO}_4$  shows the highest fluorescence; this is supported by median, mode, and geometric mean statistics. All stats show significant increase in fluorescence compared to the background, *E. coli wt* (two-tailed, unpaired, t-test). *E. coli zntA::rfp* grown in 400  $\mu\text{M}$   $\text{ZnSO}_4$  shows higher fluorescence (median, mode and geometric mean) compared to *E. coli wt*, but less than *E. coli zntA::rfp* grown in 600  $\mu\text{M}$   $\text{ZnSO}_4$ . This data agrees with the end point fluorescence assays in Chapter 3; that an increase in zinc concentration with increase *zntA* gene expression.



**Figure 5.9 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* MG1655 (*wt* and *zntA::rfp::kan<sup>R</sup>*) in LB with Added  $\text{ZnSO}_4$**

*E. coli* MG1655 *wt* and *zntA::rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in LB (0, 400 or 600  $\mu\text{M}$   $\text{ZnSO}_4$ ). After 16 hr incubation, 100  $\mu\text{L}$  of bacterial culture was pelleted (13,000  $\times g$  for 1 min), re-suspended in 1000  $\mu\text{L}$  4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

However, the aim of this chapter was not to look at promoter activity, but to look at stochastic gene expression to determine if there is heterogeneity of gene expression. For this, it made sense to plot a histogram for just *E. coli zntA:rfp* + 600  $\mu$ M ZnSO<sub>4</sub>. The addition of 600  $\mu$ M rather than 400  $\mu$ M ZnSO<sub>4</sub> was chosen as this gave the greatest increase in fluorescence. Figure 5.10 shows a histogram and distribution of fluorescence for *E. coli zntA:rfp* + 600  $\mu$ M ZnSO<sub>4</sub>. It is obvious that *zntA:rfp* does not show normal Gaussian distribution, it shows a negative skew (sometimes referred to as left skew) in distribution. If comparing the distribution to *E. coli wt* or *zntA:rfp* + 0  $\mu$ M ZnSO<sub>4</sub> (Figure 5.8), they also show a negative skew in distribution, but *E. coli zntA:rfp* + 600  $\mu$ M ZnSO<sub>4</sub> shows a stronger negative skew.

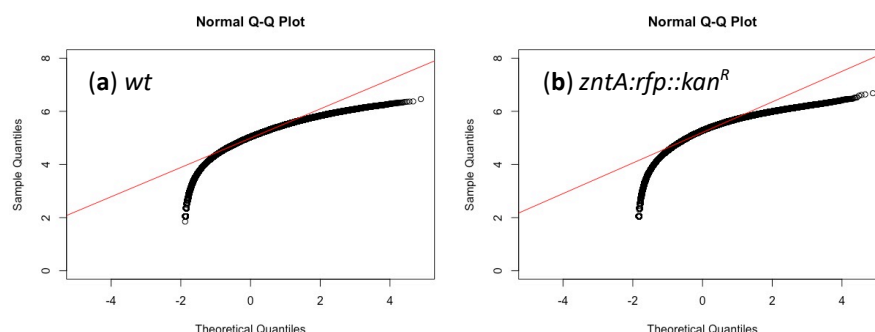


**Figure 5.10 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* in LB + 600  $\mu$ M ZnSO<sub>4</sub>**

*E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in LB + 600  $\mu$ M ZnSO<sub>4</sub>. After 16 hr incubation, 100  $\mu$ L of bacterial culture was pelleted (13,000  $\times$  g for 1 min), re-suspended in 1000  $\mu$ L 4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

A quantile-quantile (Q-Q) plot was created for *E. coli zntA:rfp* grown in LB + 600  $\mu$ M ZnSO<sub>4</sub> and *E. coli wt* grown in LB only (Figure 5.11). The Q-Q plot clearly shows that both samples show a negative skew, meaning that most of the data is distributed in the right, and shows an extended left tail. It could be argued that the Q-Q plot for *zntA:rfp* (Figure 5.11b) shows less alignment to ‘theoretical normal distribution’ at the higher end values compared to *wt*. However, comparing Q-Q plots this way is very subjective and is not quantitative.

From looking at the histogram and Q-Q plot, there is no obvious bimodal distribution, which would have been a strong indicator for heterogeneity. However, the negative skew in distribution, could suggest that the data shows slight heterogeneity; that there is more variety of *zntA* gene expression at low level expression level than at high expression levels



**Figure 5.11 Q-Q plot *E. coli* MG1655 *wt* and *zntA:rfp::kan<sup>R</sup>* in LB**

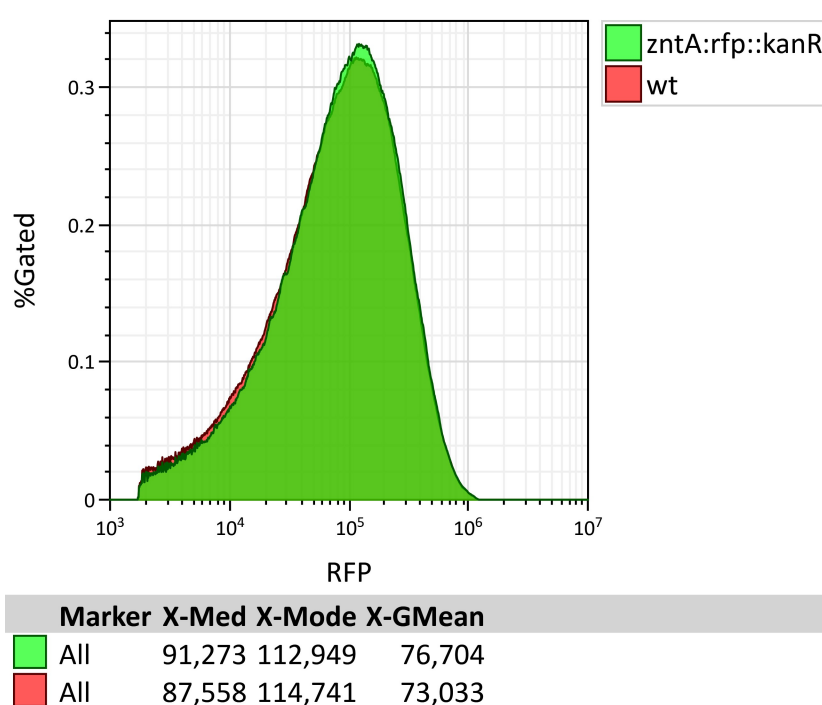
(a) Q-Q plot showing distribution of *E. coli* MG1655 *wt* grown LB only. (b) Q-Q plot showing distribution of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* grown in LB + 600  $\mu$ M ZnSO<sub>4</sub>. Red line shows theoretical normal distribution.

### 5.2.5. Flow Cytometry of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* in NH

#### 5.2.5.1. Background Fluorescence

An initial flow cytometry experiment was conducted to compare the background fluorescence of *E. coli wt* and *E. coli zntA:rfp* grown in zinc depleted Neidhardt's MOPS

minimal media (NH) only. Figure 5.12 shows a histogram of *E. coli wt* and *zntA:rfp* fluorescence in NH only. This plot suggests that the fluorescence distribution between *wt* and *zntA:rfp* is similar; they both show close values in median, mode, and geometric mean. However, *zntA:rfp* does have a higher median and geometric mean than *E. coli wt*, which suggests that *zntA:rfp* shows slightly higher fluorescence than *wt*. Again, this is to be expected; previous plasmid-based end point fluorescence assays for  $P_{zntA}$  show basal expression in NH media (Figure 3.11). Interestingly, when looking at mode, which indicated the peak fluorescence expression, it is higher in *wt* than in *zntA:rfp*.



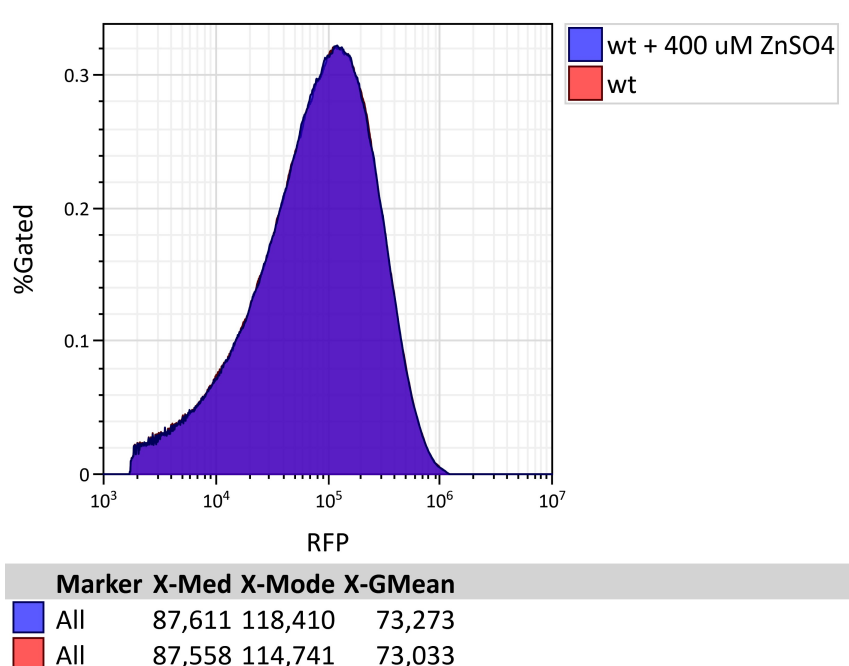
**Figure 5.12 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* MG1655 (*wt* and *zntA:rfp::kan<sup>R</sup>*) in NH Only**

*E. coli* MG1655 *wt* and *zntA:rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in NH only. After 16 hr incubation, 100 µL of bacterial culture was pelleted (13,000 x *g* for 1 min), re-suspended in 1000 µL 4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

To ensure the background fluorescence of *E. coli wt* was auto-fluorescence of the cells and/or media, *E. coli wt* was grown in NH only and NH + 400 µM ZnSO<sub>4</sub> (Figure 5.13).

Figure 5.13 shows an almost identical RFP fluorescence distribution between *wt* +

0  $\mu\text{M}$   $\text{ZnSO}_4$  and wt + 400  $\mu\text{M}$   $\text{ZnSO}_4$ . The geometric mean and median are similar, much closer than wt and *zntA:rfp* in NH only (Figure 5.12). This suggests that the addition of 400  $\mu\text{M}$   $\text{ZnSO}_4$  does not affect the fluorescence of *E. coli* wt, further confirming that the fluorescence observed in *E. coli* wt is purely auto-fluorescence of the cells and/or media and is not influenced by zinc concentrations; either directly or indirectly. Again, *E. coli* wt will be used as background data.



**Figure 5.13 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* MG1655 wt in NH with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$**

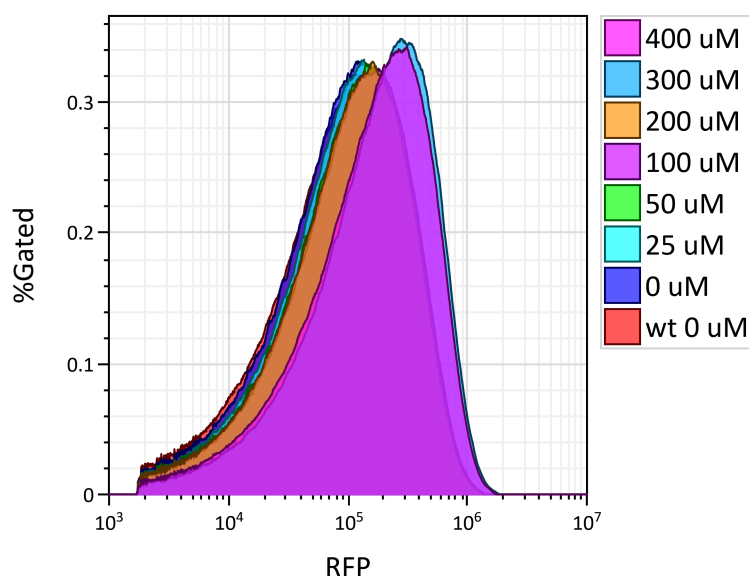
*E. coli* MG1655 wt aerobically incubated at 37°C with shaking (200 RPM) in NH with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$ . After 16 hr incubation, 100  $\mu\text{L}$  of bacterial culture was pelleted (13,000  $\times g$  for 1 min), re-suspended in 1000  $\mu\text{L}$  4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/620 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

#### 5.2.5.2. Foreground Fluorescence

*E. coli zntA:rfp* was grown in increasing concentrations of added  $\text{ZnSO}_4$  (0-400  $\mu\text{M}$   $\text{ZnSO}_4$ ) and compared to *E. coli* wt with no added zinc (Figure 5.14). The mode of each sample shows there is a clear relationship between increased  $\text{ZnSO}_4$  concentrations with increased fluorescence. This can be observed in the Mode, and to a lesser extent



in the GMean (geometric mean). It is hard to see, but as  $\text{ZnSO}_4$  concentration increases, the fluorescent distribution becomes less normally distributed, and shows a greater negative skew in distribution.



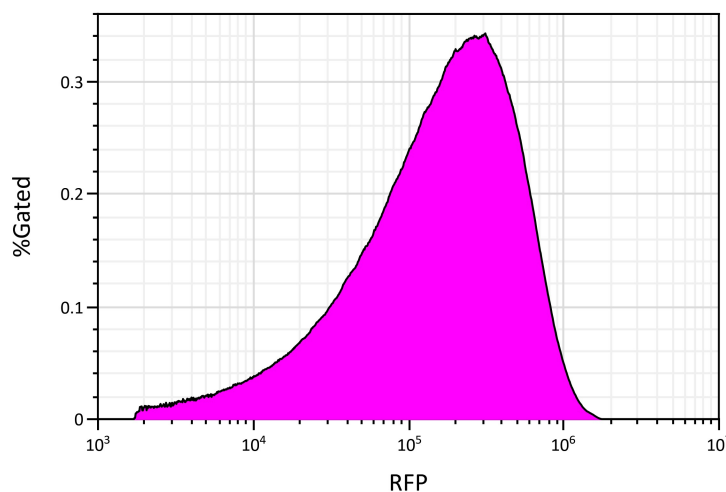
Marker	X-Med	X-Mode	X-GMean
All	175,098	304,471	139,643
All	189,774	285,892	151,405
All	111,952	159,688	93,334
All	115,782	170,065	96,389
All	109,463	145,297	91,256
All	100,889	134,300	84,680
All	91,273	112,949	76,704
All	87,558	114,741	73,033

**Figure 5.14 Histogram (Flow Cytometry Fluorescence Data) of *E. coli* (wt and *zntA::rfp*) grown in NH with increasing  $\text{ZnSO}_4$**

*E. coli* MG1655 wt and *zntA::rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in NH with increasing  $\text{ZnSO}_4$  concentrations. After 16 hr incubation, 100  $\mu\text{L}$  of bacterial culture was pelleted (13,000  $\times g$  for 1 min), re-suspended in 1000  $\mu\text{L}$  4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

Figure 5.15 shows a histogram of the fluorescence distribution of *E. coli zntA::rfp* grown in NH + 400  $\mu\text{M}$ . *E. coli zntA::rfp* grown in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$  shows there is not a normal Gaussian distribution in the fluorescence data and shows a negative skew in the data. This negative skew can be more easily observed as a Q-Q plot (Figure 5.16b).

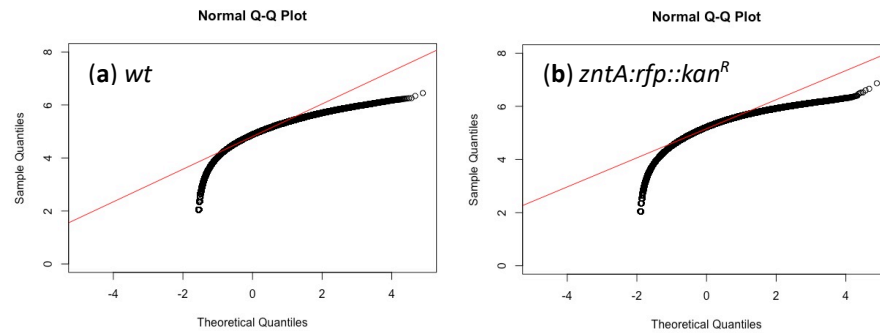
There are very few visible differences between Q-Q plots for *wt* in NH only (Figure 5.16a) and *zntA::rfp* in NH + 400  $\mu$ M ZnSO<sub>4</sub> (Figure 5.16b).



**Figure 5.15 Histogram (Flow Cytometry Fluorescence Data) of *E. coli zntA::rfp* in NH + 400  $\mu$ M ZnSO<sub>4</sub>**

*E. coli* MG1655 *zntA::rfp::kan<sup>R</sup>* aerobically incubated at 37°C with shaking (200 RPM) in NH + 400  $\mu$ M ZnSO<sub>4</sub>. After 16 hr incubation, 100  $\mu$ L of bacterial culture was pelleted (13,000  $\times g$  for 1 min), re-suspended in 1000  $\mu$ L 4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/20 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.

Like the hypothesis of heterogeneous gene expression of *zntA* in LB (5.2.4.2) this negative skew could suggest heterogeneous gene expression of *zntA*. However, the data presented thus far does not draw a justified conclusion. Further analysis is needed to truly understand these results.

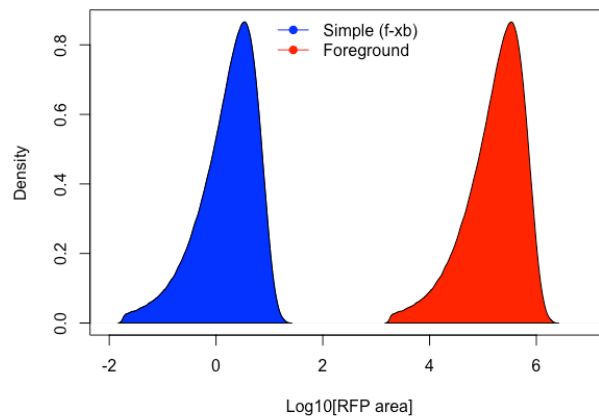


**Figure 5.16 Q-Q plot *E. coli* MG1655 *wt* and *zntA:rfp::kan<sup>R</sup>* in NH**

(a) Q-Q plot showing distribution of *E. coli* MG1655 *wt* grown NH only. (b) Q-Q plot showing distribution of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* grown in NH + 400  $\mu$ M ZnSO<sub>4</sub>. Red line shows theoretical normal distribution.

### 5.2.6. A Bayesian Background Subtraction Approach for Flow Cytometry Data

Flow cytometry results thus far have not confirmed a heterogeneous stochastic gene regulation of *zntA* in response to zinc, only alluded to it. As seen in the above data, the background fluorescence (*E. coli wt*) shows significant fluorescence, and largely overlaps with all foreground fluorescence data of the samples (*E. coli zntA:rfp* +  $x$   $\mu$ M ZnSO<sub>4</sub>). This makes distinguishing foreground data from background fluorescence extremely difficult and problematic. In standard end point fluorescence assays, such as those used throughout Chapter 3 and Chapter 4, simple subtraction of background data from foreground data can be used to account for auto-fluorescence of cells and media. However, as flow cytometry data considers the fluorescence of all individual cells, rather than the population, removing background auto-fluorescence is not as simple as foreground minus background. Figure 5.17 exemplifies this simple equation: subtracting the background mean ( $x_b$ ) from the foreground data ( $f$ ); all this does is negatively shift the data. Another issue with this simple subtraction is that the adjusted data ( $f - x_b$ ) now has negative values, but it is not possible for results to show negative fluorescence.



**Figure 5.17 Illustration of Simple Foreground Minus Background Mean**

Foreground data (f) – background mean (xb). f-xb. Foreground data, *zntA::rfp* in LB + 600  $\mu$ M. Background data, *wt* in LB only.

A similar issue also plagues DNA microarray data; that by simply subtracting the ‘background’ noise/data from the foreground data, it often leads to data being lost or giving false results. Kooperberg *et al.*, (2002) developed a Bayesian background subtraction approach which provides a more accurate estimation of the true foreground. I have further developed this model and altered it to be used on flow cytometry data.

DNA microarray data is like flow cytometry data. DNA microarray data provides a; mean, median and standard deviation, which is calculated from all pixels within a defined region on the array. Flow cytometry data will also provide a; mean, median and standard deviation which is calculated from all cells within a sample.

The Bayesian background subtraction approach by Kooperberg *et al.*, (2002) takes into consideration four points;

(1) The observed foreground ( $X_f$ ) and observed background ( $X_b$ ) may differ from the true foreground ( $\mu_f$ ) and true background ( $\mu_b$ ).

(2) The true data ( $\mu_t$ ) must be equal to or greater than 0 [ $\mu_t \geq 0$ ], and therefore

$$\mu_f = \mu_t + \mu_b$$

(3) If  $X_b > X_f$ , knowing that  $\mu_f \geq \mu_b$ , this cannot be true; which means the estimation of

$\mu_b = X_b$  is not true, the same goes for  $\mu_f = X_f$ . Therefore  $\mu_b < X_b$  and that  $\mu_f > X_f$ .

(4) Therefore  $\mu_t = \mu_f - \mu_b > X_f - X_b$ .

The Bayesian background subtraction approach essentially subtracts the background data from the foreground data by constructing a posterior distribution of the true foreground data (given the observed foreground and background data), knowing that the true foreground cannot be negative given the above four points.

Kooperberg's Bayesian background subtraction equation (Figure 5.18a) was slightly modified to work with flow cytometry data, rather than its intended use in DNA microarray data. The original Bayesian subtraction equation (Figure 5.18a) calculated the point estimation (mean) of the posterior distribution, knowing the mean of the foreground ( $X_f$ ) and the background ( $X_b$ ). However, the R code (Figure 5.18b) was adapted rather than the mean of the foreground ( $X_f$ ) and background ( $X_b$ ), the adapted R code allowed  $X_f$  to be each individual data point ( $f$ ), rather than the mean. This returned a point estimation of the prior distribution for every data point (or cell). The R code for the flow cytometry compatible Bayesian background subtraction, based on (Kooperberg *et al.*, 2002) was written in R as a function, and named 'Ingram' (Figure 5.18b). This R code was based on an R code provide by Dov Stekel, The University of Nottingham.

Before the Ingram function can be applied, the flow cytometry data was imported into R studio and converted to Log [ $\text{Log}_{10}$ ]. The 'sapply' function was used to apply the Ingram function to every data point within a data set [Results = sapply( $f$ , Ingram)].

The Ingram function was applied to data previously seen in this chapter (5.2.4-5.2.5), no new experiments were conducted, and no new raw data will be presented.

**a**

$$p(\mu_t | \sigma_b, \sigma_f, X_b, X_f) = \frac{\phi\left(\frac{X_f - \mu_t - X_b}{\sigma_d}\right) \Phi\left(\frac{(X_f - \mu_t)\sigma_b^2 + X_b\sigma_f^2}{\sigma_f\sigma_b\sigma_d}\right)}{\sigma_d \int_0^\infty \Phi\left(\frac{X_f - v}{\sigma_f}\right) \phi\left(\frac{X_b - v}{\sigma_b}\right) dv},$$

**b**

```

Ingram = function(f) {

  SDS = 4
  sf = sb/10
  sd = sqrt(sf*sf+sb*sb)

  #Range
  low = max(0, as.integer(f-xb-sds*sd))
  hi = as.integer(f-xb+sds*sd+1)
  x = seq(low, hi)
  x

  #kooper
  y = dnorm((f-x-xb)/sd)*pnorm(((f-x)*sb*sb+xb*sf*sf)/(sf*sb*sd))
  y

  #Backsub
  Z = sum(x*y)/sum(y)
}

```

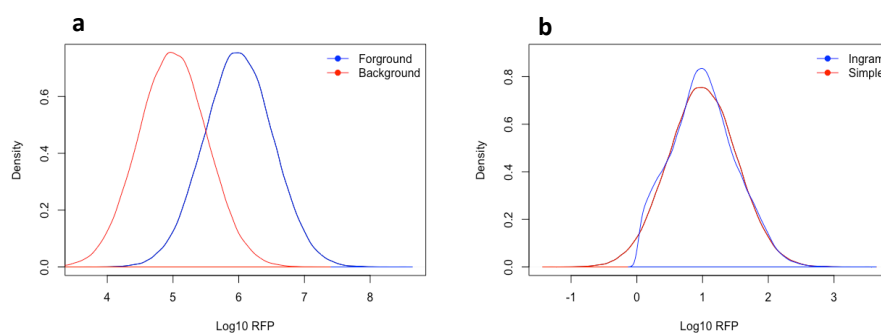
f = every foreground data point  
 xb = mean background  
 sf = standard deviation foreground  
 sb = standard deviation background

### Figure 5.18 Bayesian Background Subtraction

**(a)** Kooperberg *et al.*, (2002) equation for posterior distribution of observed data  
**(b)** R code of 'Ingram' function. Bayesian approach to substitution of background fluorescence from foreground fluorescence of Flow Cytometry data. Based on work by Kooperberg *et al.*, (2002). Code adapted from Dov Stekel.

### 5.2.6.1. Bayesian Background Subtraction: Proof of Principle

Before the Bayesian background subtraction was applied to the flow cytometry data, a simple test was conducted. Two sets of data were randomly generated in R, both with normal distribution. Background and foreground data were created, with foreground having a mean of background mean + 1 ( $X_f = X_b + 1$ ) (Figure 5.19a). Figure 5.19a shows a density plot of the randomly generated foreground and background data, which clearly shows normal distribution and shows an overlapping area of foreground and background data. Figure 5.19b shows two approaches to analyse the true foreground: the 'Simple' model (foreground – background mean [ $f - X_b$ ]); and the Bayesian background subtraction (Ingram). The 'Simple' model shows a normal distribution, but ranges into the negative, which we know not to be true. The Bayesian background subtraction (Ingram) approach of the data shows a non-normal distribution, a higher peak fluorescence, and no data point in the negative range. Figure 5.19 validates the use of the Bayesian model approach and the use of the 'Ingram' function to be used on flow cytometry data.



**Figure 5.19 Bayesian Background Subtraction: Proof of Principle**

(a) Density plot of randomly generated data with normal distribution. (b) Density plot comparison of analysed foreground data. Bayesian background subtraction analysis of foreground data using the 'Ingram' function. 'Simple' analysis of foreground data ( $f - X_b$ ). Key point, 'Ingram' function values are  $\geq 0$ .

### 5.2.7. Bayesian Background Subtraction: True Foreground Fluorescence of *E. coli* MG1655 *zntA::rfp::kan<sup>R</sup>* in LB

The Bayesian background subtraction approach was used to estimate the true foreground fluorescence of *zntA* gene expression in LB with the addition of ZnSO<sub>4</sub> (0, 400, or 600 µM) (Figure 5.20). This used *E. coli* wt in LB only as the background, and *E. coli zntA::rfp* in LB + ZnSO<sub>4</sub> (0, 400 or 600 µM) as the foreground. A density plot, with kernel density estimator (Gaussian) was plotted to visualise a smooth graph, negating the need for 'bins' in histogram plots. Figure 5.20 shows all three conditions and shows a pattern which is highly dissimilar to the raw fluorescence data (Figure 5.9). For ease of visualization and discussion, the data presented in (Figure 5.20a) has been split into two sub-section: low intensity values (<0.35 Log<sub>10</sub>[RFP area], Figure 5.20b); and high intensity values (>0.35 Log<sub>10</sub>[RFP area], Figure 5.20c).

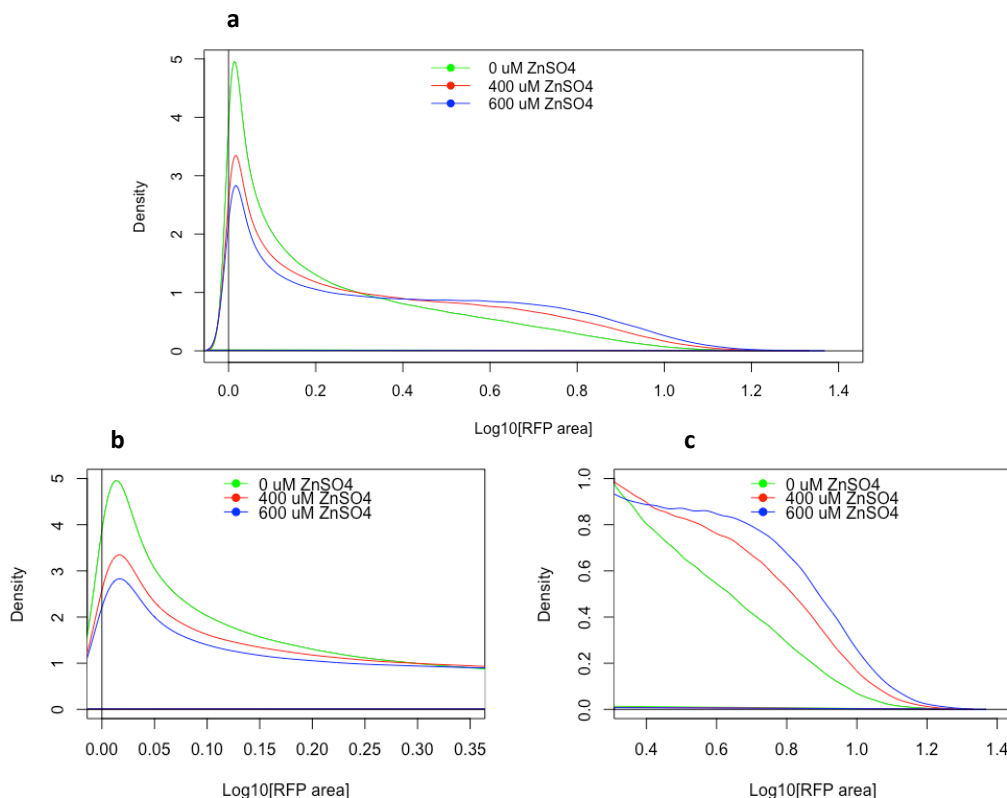
Figure 5.20 density plot shows that in all three conditions show a higher proportion of cells expressing ZntA at low levels (<0.35), compared to proportion of cells expression ZntA at the higher levels (>0.35). This data clearly shows that there is heterogenous gene expression of *zntA*; this data does not show normal Gaussian distribution, and the distribution observed is due to heterogenous gene expression of *zntA*. This data agrees with the hypothesis by Takahashi *et al.*, (2015) that *E. coli* mount a heterogenous gene expression of *zntA*, in both normal zinc concentration, and during zinc stress. This also demonstrates the sub-population which are expressing ZntA at different levels, there are clearly two sub-population: expression at low (<0.35) and high (>0.35) levels.

Looking specifically at expression observed with the addition of 0 µM ZnSO<sub>4</sub>, there is clear distribution of cells expression both at low and high levels. Again this agrees with Takahashi *et al.*, (2015) hypothesis of bet-hedging, that *E. coli*, even under normal zinc



conditions, will express *zntA* at low and high level to prepare to an increasing zinc concentrations. This is also true with the addition of 400  $\mu\text{M}$  and 600  $\mu\text{M}$   $\text{ZnSO}_4$ .

It is clear to see that cells with 0  $\mu\text{M}$   $\text{ZnSO}_4$  show a higher proportion of cells expressing ZntA at the low levels ( $<0.35$ ), compared to 400 or 600  $\mu\text{M}$   $\text{ZnSO}_4$  (Figure 5.20b). In contrast, cells with 0  $\mu\text{M}$   $\text{ZnSO}_4$  show a much lower proportion of cells expressing ZntA at the higher levels ( $>0.35$ ). Further to this, cells with 400  $\mu\text{M}$   $\text{ZnSO}_4$  show a higher proportion of cells at low levels ( $<0.35$ ) and a lower portion of cells at the higher level ( $>0.35$ ) compared to 600  $\mu\text{M}$   $\text{ZnSO}_4$ . This data suggests that with increasing concentrations of  $\text{ZnSO}_4$ , there is a shift in the distribution of the population expressing ZntA at different levels. Increased  $\text{ZnSO}_4$  reduced the proportion of the population expressing ZntA at low levels ( $<0.35$ ) and shift it to high ZntA expression levels ( $>0.35$ ), increasing the proportion of cells expression at the higher expression levels. This mechanism allows *E. coli* to more adequately prepare and survive zinc stress.



**Figure 5.20 Bayesian Background Subtraction: Fluorescence of *E. coli zntA:rfp* Grown in LB Media**

Bayesian background subtraction approach to determine true foreground fluorescence of flow cytometry (RFP area) data of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* grown in LB with added ZnSO<sub>4</sub>. Background data *E. coli* MG1655 *wt* grown in LB only. Density plot using kernel density estimator (Gaussian). (a) All data. (b) Low intensity values (<0.35 Log<sub>10</sub>[RFP area]). (c) High intensity values (>0.35 Log<sub>10</sub>[RFP area]). Raw data from (Figure 5.9).

### 5.2.8. Bayesian Background Subtraction: True Foreground Fluorescence of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* in NH

The Bayesian background subtraction approach was applied to flow cytometry data for *E. coli zntA:rfp* grown in zinc depleted NH media to determine the true foreground fluorescence. A wider variety of zinc concentrations was added to NH compared to LB. *E. coli wt* grown in NH only was used for background fluorescence. Three samples were chosen for Bayesian background subtraction; *E. coli zntA:rfp* grown in NH + 0μM, 25 μM or 400 μM ZnSO<sub>4</sub>. These Three conditions were chosen as they demonstrate low and high zinc concentrations. This is reflected in the raw fluorescence as seen in

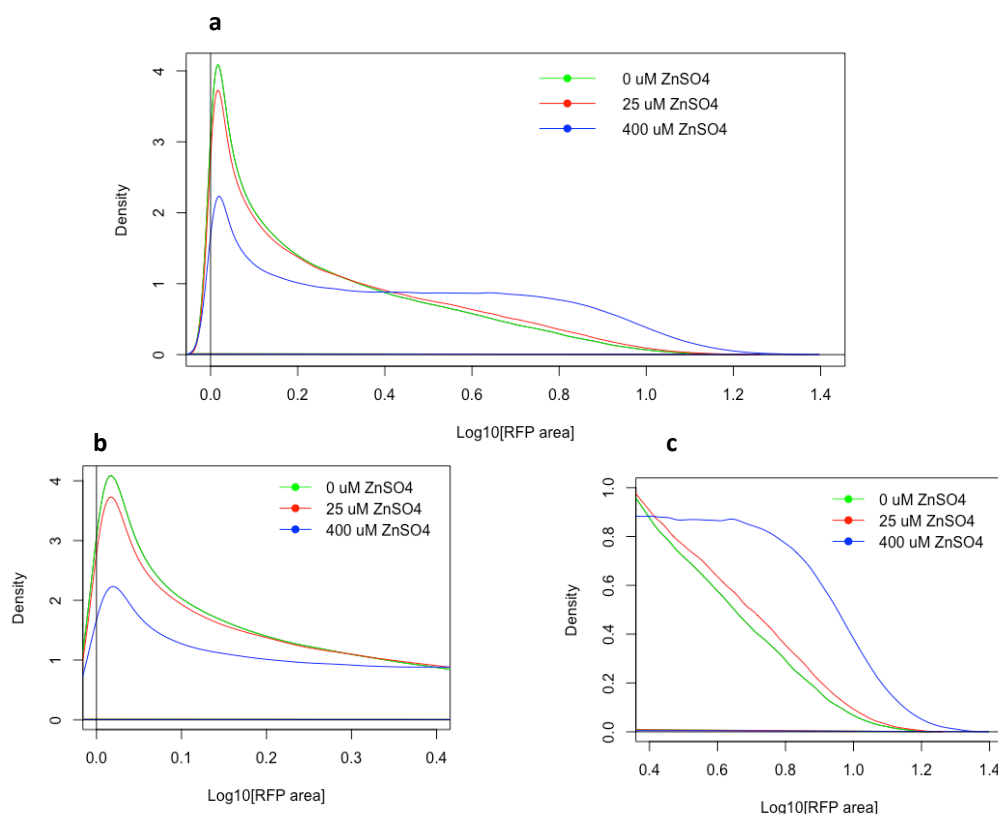
Figure 5.14. The addition of 25  $\mu\text{M}$   $\text{ZnSO}_4$  shows lower raw fluorescence than 400  $\mu\text{M}$   $\text{ZnSO}_4$ , with 0  $\mu\text{M}$  (NH only) shows the lowest raw fluorescence.

Figure 5.21 shows the density plot, with kernel density estimator (Gaussian), of the true foreground fluorescence of *E. coli zntA::rfp* grown in NH media with added  $\text{ZnSO}_4$  (0, 25 or 400  $\mu\text{M}$ ). Similar to data presented in Figure 5.20, the data presented has been split into two sub-subsections: low intensity values ( $<0.4 \text{ Log}_{10}[\text{RFP area}]$ , Figure 5.21**b**); and high intensity values ( $>0.4 \text{ Log}_{10}[\text{RFP area}]$ , Figure 5.21**c**).

Figure 5.21 shows a similar pattern in the distribution of fluorescence as that observed when grown in LB (Figure 5.20); that a higher proportion of cells express ZntA at low levels ( $<0.4$ ), compared to a lower proportion of cells expressing ZntA at higher levels ( $>0.4$ ). Again, this data agrees with the hypothesis by Takahashi *et al.*, (2015); that *E. coli* mount a heterogenous *zntA* gene expression in both normal zinc concentrations and under zinc stress.

The distribution of fluorescence between different zinc concentrations shows a relationship between the distribution and increasing zinc concentration. Higher zinc concentrations show a lower density of cells with low fluorescence values ( $<0.4$ , Figure 5.21**b**). Looking at high fluorescence values ( $>0.4$ , Figure 5.21**c**), higher zinc concentration now shows higher density of high fluorescence values. This concurs with data from Figure 5.20, that there is a shift in fluorescence distribution in relationship to increasing  $\text{ZnSO}_4$ ; increased  $\text{ZnSO}_4$  reduced the proportion of the population expressing *zntA* at low levels ( $<0.4$ ) and shift it to high *zntA* expression levels ( $>0.4$ ), increasing the proportion of cells expression at the higher expression levels.

Looking specifically at *E. coli zntA:rfp* in 400  $\mu\text{M}$   $\text{ZnSO}_4$ , between 0.2-0.8  $\text{Log}_{10}[\text{RFP area}]$ , there is a clear ‘plateau’ of density. This suggests that there is a large variety of gene expression levels that are equally expressed within the population. This ‘plateau’ is not observed with either 0 or 25  $\mu\text{M}$   $\text{ZnSO}_4$ . This would suggest that heterogenous gene expression is more predominant, and equally distributed at higher zinc concentrations. *E. coli zntA:rfp* in NH + 0 and 25  $\mu\text{M}$  show maximum fluorescence around 1.2  $\text{Log}_{10}[\text{RFP area}]$ , whereas the addition of 400  $\mu\text{M}$  shows maximum fluorescence  $>1.2 \text{Log}_{10}[\text{RFP area}]$ , suggesting that the increased zinc concentration also increases maximum *zntA* gene expression.



**Figure 5.21 Bayesian Background Subtraction: Fluorescence of *E. coli zntA:rfp* Grown in NH Media**

Bayesian background subtraction approach to determine true foreground fluorescence of flow cytometry (RFP area) data of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>* grown in NH with added  $\text{ZnSO}_4$ . Background data *E. coli* MG1655 wt grown in NH only. Density plot using kernel density estimator (Gaussian) (a). All data. (b) Low intensity values ( $<0.4 \text{Log}_{10}[\text{RFP area}]$ ). (c) High intensity values ( $>0.4 \text{Log}_{10}[\text{RFP area}]$ ). Raw data from Figure 5.14.

### 5.3. Discussion and Conclusion

#### 5.3.1. Construction of *E. coli* MG1655 *zntA:rfp::kan<sup>R</sup>*

The gene doctoring method (Lee *et al.*, 2009) was used to chromosomally fuse mRFP1 to the C-terminal end of ZntA in *E. coli* MG1655, while introducing *kan<sup>R</sup>* as a selective marker. Colony PCR, patch plating and sequence analysis was used to confirm the presence of the *zntA:rfp::kan<sup>R</sup>* chromosomal fusion, of which all showed positive for insertion. A simple end point fluorescence assay was conducted in *E. coli zntA:rfp* to ensure fluorescence was phenotypically observed as expected, that increasing zinc concentration will increase fluorescence. Figure 5.3 showed the expected phenotype: that increased zinc concentration increased end point fluorescence. In conclusion, the gene doctoring method provided a simple and effective method to chromosomally fuse mRFP1 to the C-terminal end of ZntA.

#### 5.3.2. Raw Flow Cytometry Fluorescence Data

The raw fluorescence data of *E. coli zntA:rfp* grown in LB (Figure 5.9) did not draw any meaningful conclusions around heterogenous gene expression. The raw data is hard to interpret as there are large areas which overlap the background control fluorescence data (*E. coli wt*). The same is true for the raw fluorescence data of *E. coli zntA:rfp* grown in NH media (Figure 5.14), again due to large overlapping area of the control and the sample data, no meaningful concluding can be made.

To ensure the background fluorescence of *E. coli* MG1655 *wt* was not influenced by additional zinc, a flow cytometry assay was conducted with *E. coli wt* grown in NH media with and without 400  $\mu$ M ZnSO<sub>4</sub>. The addition of 400  $\mu$ M ZnSO<sub>4</sub> did not alter the raw fluorescence of *E. coli wt* (Figure 5.13). This helped to validate the use of *E. coli wt* as background data.

All flow cytometry data recorded: both foreground (*E. coli zntA::rfp* +  $x$  ZnSO<sub>4</sub>) and background (*E. coli wt* + 0  $\mu$ M ZnSO<sub>4</sub>), showed a negative skew in the distribution, non-normal distribution was observed. As this negative skew was observed in the background as well as foreground, it is possible that the skew could just be an artefact of the data and cannot be used to draw any conclusions from.

Both flow cytometry data recorded in LB and NH showed similar results and provided data to support the zinc tuneable effects of *zntA* gene expression. Previous *zntA* expression assays, such as (Brocklehurst *et al.*, 1999) and my data in Chapter 3 used end point assays based on promoters cloned into a plasmid. This is the first time, to my knowledge, that any *zntA* expression assay has been conducted using a chromosomally fused fluorescent marker, and not a plasmid-based system. This chapter's aim was not to look at *zntA* gene expression beyond stochastics, but inadvertently has provided new data to support the tuneable *zntA* response to increasing zinc concentrations. Using a chromosomally fused fluorescent marker, rather than a reporter plasmid, provided some invaluable advantages, such as plasmid copy number not affecting results. Further, a paper by Bryant *et al.*, (2014) demonstrated that the position of a gene on the chromosome can have a significant effect on gene expression; up to 300-fold difference in expression was recorded depending on the chromosomal location. Using chromosomal gene fusions for end points assays is doubly advantageous. Plasmid copy number will not impact data, and by recording gene expression from its native chromosomal location will provide more accurate gene expression data of a gene *in vivo*.

### 5.3.3. Bayesian Background Subtraction of Flow Cytometry Fluorescence Data

To further help understand the flow cytometry data, a Bayesian background subtraction approach was used to estimate the true foreground data. This Bayesian background subtraction approach was based on a Bayesian model adapted from DNA microarray analysis (Kooperberg *et al.*, 2002), which is used to more accurately estimate the true foreground data. The Bayesian background subtraction approach is based around the understanding of the relationship between the foreground and background;  $[\mu_t = \mu_f - \mu_b > X_f - X_b]$ . The statistical model by Kooperberg *et al.*, (2002), developed for DNA microarray data, was adapted to work with flow cytometry data, and written as an R code, termed 'Ingram' (Figure 5.18).

A proof of principle test was carried out on a randomly generated data set, with normal distribution (Figure 5.19). This test proved the power of the Bayesian background subtraction on flow cytometry data and showed how it is better than a standard simple  $(f-X_b)$  method.

All Bayesian background subtracted data showed a large peak of fluorescence at low fluorescence values. This is a product of the Bayesian model, knowing  $\mu_t \geq 0$  and cannot be less than zero. The distribution of the foreground gene expression, characterised by gene stochastics, means low fluorescence values overlap with the distribution of background auto-fluorescence.

This foreground distribution, or gene stochastics, is either intrinsic or extrinsic. Intrinsic defines the variation in gene expression due to the biochemical process involved in the central dogma. Extrinsic defines the variation in gene expression due to factors outside of the biochemical process of the central dogma, which impact on gene expression (Swain, Elowitz and Siggia, 2002). Takahashi *et al.*, (2015) demonstrated *zntA*

heterogeneity based on intrinsic models, which would suggest that data obtained in this chapter is likely to be due to intrinsic variation, rather than extrinsic variation.

The Bayesian background subtraction analysis of *E. coli zntA::rfp* in LB with added  $\text{ZnSO}_4$  (0, 400, and 600  $\mu\text{M}$ ) all show evidence to support heterogeneous gene expression (Figure 5.20). The density distribution of fluorescence shows a significant difference to that of the raw data and reflects the true foreground fluorescence. All three conditions analysed (0, 400, and 600  $\mu\text{M}$   $\text{ZnSO}_4$ ) showed a distribution of fluorescence, which is a distribution of *zntA* gene expression. These three conditions show a higher density of low fluorescence values, and lower density of high fluorescence values: evidence of heterogeneous gene expression. Interestingly, as  $\text{ZnSO}_4$  is increased, it shifts the distribution of fluorescence, reducing low values and increasing high values. This data also suggested that in higher zinc concentrations, the distribution of *zntA* expression is more equally expressed, compared to lower zinc concentrations.

The Bayesian background subtraction analysis of *E. coli zntA::rfp* in NH with added  $\text{ZnSO}_4$  (0, 25 and 400  $\mu\text{M}$ ) also support evidence for heterogeneous *zntA* gene expression. The growth condition in NH media was conducted with a larger range of zinc concentrations, this was more feasible in NH media due to its zinc depleted nature. One would assume a 25  $\mu\text{M}$   $\text{ZnSO}_4$  increase in LB would be less noticeable than in NH. Again, like in LB, all three conditions show a higher density of low fluorescence values, and a lower density of higher fluorescence values. The three conditions also show a shift in the distribution of the fluorescence density, increasing zinc concentration reduces low values and increases high values.

Looking purely at the Bayesian background subtraction data one cannot tell the reason for the change in distribution, it could be due to two reasons: firstly, that cells expressing *zntA* at low levels have died due to toxic zinc concentration, which shifts



the distribution to cells expression *zntA* at higher levels; this was hypothesised by Takahashi *et al.*, (2015). Or secondly, that no cells have died, and the shift in distribution is purely due to a higher proportion of the population of cells expressing *zntA* at higher levels. However, if one corresponds this data to the ZnSO<sub>4</sub> toxicity test from earlier in this thesis (Figure 3.3), this can help the understanding of the shift. The ZnSO<sub>4</sub> toxicity test showed no significant difference in either absorbance<sub>600</sub> or in the ratio of live:dead cells between 0-600 µM ZnSO<sub>4</sub>; I hypothesise that the shift in distribution is purely due to a higher proportion of the cell's expressing *zntA* at higher levels. However more experimental data is needed to prove this definitively.

#### 5.3.4. Concluding Remarks

The Bayesian background subtraction approach for true foreground fluorescence of flow cytometry data has provided new, invaluable insight into understanding true flow cytometry fluorescence data. The data presented in this chapter, particularly the Bayesian background subtraction data, showed experimental evidence which is consistent with the hypothesis by Takahashi *et al.*, (2015): that *zntA* gene expression is heterogeneous and thus indicates bet-hedging by *E. coli*.

This data lead to my hypothesis: “*zntA* heterogeneity is observed in both normal and increased zinc concentration. In higher zinc concentration *zntA* expression is shifted; high *zntA* expression levels are increased while low *zntA* expression levels are decreased”.

#### 5.3.5. Future Work

The data in this chapter gives good evidence for heterogeneous *zntA* gene expression. To further help the understanding of heterogeneous and homogeneous expression with the Bayesian background subtraction, it would be useful to introduce more

controls. Two controls would be useful: a known gene showing homogenous gene expression; and a known gene showing heterogeneous gene expression. This would allow one to compare the distribution of the control to *zntA*, to truly understand if heterogeneous expression is observed.

The use of microfluidics, in combination with *E. coli zntA::rfp* could be an invaluable tool to examine stochastic gene expression of *zntA*. Microfluidics allows for a few key advantages over flow cytometry. Flow cytometry only records fluorescence intensity of individual cells within a population. However, a major drawback of this is that often cells are fixed or are analysed at a single time point. Flow cytometry data will only give a snapshot of stochasticity at a chosen time point. Using microfluidics can improve on this, as it allows one to record data on a single cell, over time. Microfluidics, with its combination of microscopy, will allow one to understand how the cell cycle is affected under zinc stress, as well as how *zntA* gene regulation is altered. Microfluidics will provide results which flow cytometry simply cannot. Essentially microfluidics will track gene expression in lineages of individual cells, linking gene expression (via *zntA::rfp*) to cell fate, division rate, growth rate, cell length and cell death. This allows one to have a unique opportunity to test and measure stochasticity within a bacterial population.

As mentioned above, flow cytometry only gives a snapshot of gene stochastics. It would be useful to conduct the same flow cytometry experiments in this chapter, but rather than just 16 hr growth time point, conduct analysis on time points <16 hr.

To be able to further understand the reason for the shift in *zntA* expression distribution, it would be useful to measure the fluorescence of cells, and directly compare it to cells viability (dead or alive). This could be accomplished by simply staining the cells, prior to flow cytometry, with the LIVE/DEAD BacLight viability assay kits (Thermo Fisher, L34856), and recording the fluorescence of STYO9 (live), and propidium iodide

(dead) dyes. Unfortunately, the propidium iodide fluorescence would interfere with mRFP1 fluorescence, as both are red fluorescent. Either an alternative fluorescent protein for chromosomal fusion could be used or, conduct the LIVE/DEAD assay with just SYTO9 and record live cells. This would give an indication of which cells are alive and what level of *zntA* expression they have.

## Chapter 6

# Determining the Phenotype of the Novel Zur Regulated Operon, C1265-7

### 6.1. Introduction

A study by Clayton, (2012) described a novel Zur regulated operon, *c1265-7* in uropathogenic *E. coli* strain CFT073, which is not found in *E. coli* K-12 strains. Homology to *c1265-7* was only found in 11 *E. coli* strains, of which most are human pathogens, and only found chromosomally in UPEC strains. This operon showed a significant increase (60x) in transcripts of *c1265-7* mRNA when expressed in *E. coli* CFT073  $\Delta$ *zur* compared to *E. coli* CFT073 wt. Further, Patser analysis of potential Zur binding sites in *E. coli* CFT073 returned a high score (19.90) for the *c1265-7* promoter ( $P_{c1265}$ ). Interestingly the  $P_{c1265}$  Patser score was higher than the known Zur binding of  $P_{znuA}$  (Patser score 18.91) (Clayton, 2012). Clayton, (2012) concluded that C1265 showed homology to the TonB-dependent transport, ShuA, of *Shigella dysenteriae*, and that C1267 showed homology to a COG0523 protein, both of which can be associated with metal import systems. At the time, Clayton, (2012) found no homologous proteins to C1266.

*E. coli* CFT073 is a Uropathogenic *E. coli* (UPEC), which was isolated from a blood sample from a female patient who was suffering from acute pyelonephritis (Mobley *et al.*, 1990). UPEC as the name suggests, are associated with urinary tract infections. The FimH adhesion protein, found at the tip of type-1 pili, aids in attachment to the uroepithelium (Tchesnokova *et al.*, 2011). Here, UPEC can invade human epithelial cells causing apoptosis.

Promoter activity data in this thesis, and transcriptomic data from Clayton, (2012) suggested that *c1265-7* gene expression is regulated by Zur. To my knowledge, all Zur regulated genes are associated with acquisition of zinc, either by; a zinc import system, or alternative zinc free paralogue protein. Knowing all this, it would be right to assume

that C1276-7 is involved in zinc acquisition in some form. However, the known route of zinc acquisition is not known.

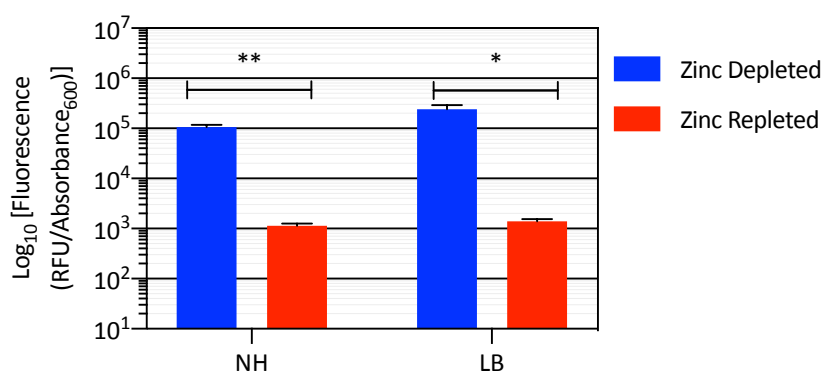
#### 6.1.1. Aim and Objectives

C1265-7 is associated with zinc acquisition when in a zinc-depleted state; therefore, the main aim of this chapter is to determine what is the function of C1265-7 during a zinc depleted state.

## 6.2. Results

### 6.2.1. Promoter Activity

Recapping the data presented in Chapter 4,  $P_{c1265}$  showed a significant increase in promoter activity when expressed in a zinc depleted growth media compared to a zinc repleted media. Figure 6.1 shows a summary of the promoter activity data obtained through fluorescence assays in Chapter 4. There is a significant difference in the promoter activity (fluorescence) of  $P_{c1265}$  between zinc depleted and zinc replete conditions, which is true for both NH and LB media. In both media there is an increase in promoter activity when the growth media is zinc depleted. This data suggests that the regulation of *c1265*, and thus the *c1265-7* operon, is upregulated in response to reduced zinc concentrations. Therefore, one can hypothesise that the C1265-7 proteins should play a role in zinc acquisition, or aid survival in zinc depleted conditions.



**Figure 6.1 End Point Fluorescent Assay of  $P_{c1265}$  in LB and NH (Summary)**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $NH_{kan50}$  or  $LB_{kan50}$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Zinc depleted ( $LB + 20 \mu M$  TPEN /  $NH$  only), Zinc repleted ( $LB$  only /  $NH + 400 \mu M$   $ZnSO_4$ ). Error bars indicated S.D. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test). Data summary from Chapter 4.

### 6.2.2. Bioinformatics

Bioinformatic analysis was used to search for homologous proteins to C1265-7. NCBI protein BLAST (Basic Local Alignment Search Tool) was used for amino acid sequence comparison (Altschul *et al.*, 1997). The Blastp (Protein-Protein BLAST) algorithm was used for protein BLAST. Alongside BLAST, Swiss-Model was used to search for protein and structure homology. Swiss-Model (Waterhouse *et al.*, 2018) is a “fully automated protein structure homology-modelling server”. As part of this, Swiss-Model searches for template using BLAST and HHblits. HHblits is a sequence based protein function and structure prediction tool which uses hidden Markov models (HMM-HMM) (Zimmermann *et al.*, 2018). Swiss-model templates are proteins, which have their structure defined, through experimental data, such as X-ray crystallography.

It is worth mentioning that bioinformatic analysis was conducted on C1265-7 by Clayton, (2012), however I have chosen to conduct a more recent search. In the eight years since Clayton’s thesis submission there is the potential that new genes and proteins since have been discovered and described (and which have subsequently been added to online databases which can be searched against).

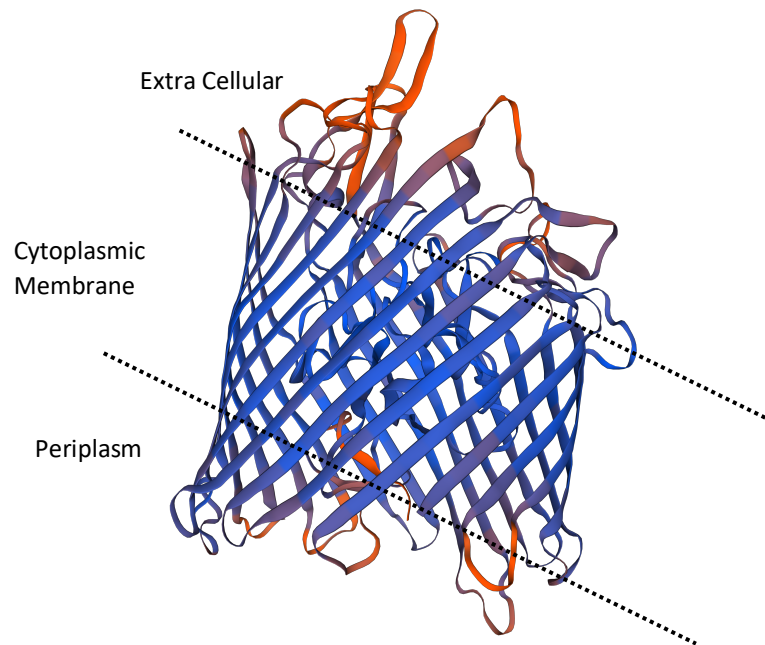
#### 6.2.2.1. Bioinformatic Analysis of C1265

NCBI protein BLAST search was used to search for homologous proteins to C1265. The top results showed 100% homology of a domain of C1265 to a TonB-dependent receptor protein. Results returned with a 100% identity, a score of 1352/1352 and an E-value of 0.0. Specifically the TonB-receptor protein has been identified as a TonB-dependent heme/haemoglobin receptor family protein [NCBI accession WP\_001445605.1] (Buchanan *et al.*, 1999). Many of the other top results were associated with TonB-dependent receptors. However, when conducting a deeper



bioinformatic search into an annotated CFT073 genome [accession AE014075], I could not find any homology or partial domain homology of C1265 to any other annotated gene, besides 'C1265'. This could suggest that the CFT073 annotated genome used automatic gene annotation, and thus the homology gene to C1265 had not been experimentally proven.

A second protein bioinformatic search was conducted which searched for homologous proteins using Swiss-Model. C1265 showed 41.6% complete sequence identity and 0.74 GMQE (Global Model Quality Estimation) to the *Shigella dysenteriae* outer membrane heme receptor, ShuA [accession 3fhh.1.A]. GMQE is a quality estimator, between 0–1, which indicates the global (entire structure) expected accuracy of the built model. Clayton, (2012) found the same homology of C1265 to ShuA. Interestingly, when looking at the predicted model structure of C1265 (Figure 6.2), based on ShuA, one can see that the region with high QMEAN (Qualitative Model Energy Analysis) values are in the  $\beta$ -barrel, and low QMEAN regions are located in the extracellular domain. QMEAN is a composite estimator, which indicates the degree of nativeness, and provides both a global (structure) and local (residue) value (Benkert, Biasini and Schwede, 2011). It has been shown, in similar TonB-dependent receptor proteins, that the two loops on the extracellular domain are associated with ligand binding (Buchanan *et al.*, 1999). The C1265 loop is predicted from the Swiss-Model but shows a low local QMEAN score. This could suggest that C1265 has a similar function to ShuA, however the ligand C1265 interacts with may be different from ShuA.



**Figure 6.2 Swiss-Model Structure of C1265**

Model structure of C1265 based on the ShuA (*S. dysenteriae*) crystal structure. Swiss-Model was used model this crystal structure. **Blue** = high quality local QMEAN, **Red** = low quality local QMEAN. Global QMEAN = -3.0. QMEAN “Qualitative Model Energy Analysis”.

C1265 also showed homology to an outer membrane protein, ZnuD, from *Neisseria meningitidis* [accession 4rdr.1.A]. However, this only showed 16.26% sequence identity, with a 0.54 GMQE score. ZnuD is an outer-membrane protein of *N. meningitidis*, which has been shown to facilitate translocation of zinc from outside of the cell to the periplasm. ZnuD is also a TonB-dependent receptor (Calmettes *et al.*, 2015). When comparing amino acid sequences and looking at which residues  $\text{Zn}^{2+}$  interacts with in ZnuD, there is little homology to C1265. C1265 residues; A115, S117, D159, D443, and H446 show residues with functional similarity to the aligned residues in ZnuD, only H446 shows a conserved residue between C1265 and ZnuD (Figure 6.3).

Protein	Sequence	Position
znuD	-----MAHHHHHHLVPRGSHETE--QSDLETVSVVGKSRPRATSGLLHTSTASDKIIS	52
c1265	MYMNVIRTVICTLLIILPVGLQAATSHSSVMVKDTTITATGNQ---NTVFETPSMVSVVTN	57
	: : * * : : . * : * : * : * . . : * : . : .	
znuD	GDTLRQKAVNLGDLALDGVPGIHASQYGGGASA-PVIRGQTGRRIKVLNHHGETGDMADFS	111
c1265	DTPWSQNAVTSAGMLKGVAGLSQTGAGRTNGQTFNLRGYDKS--GVLVLVDGVRQLSDM	115
	. * : * . . * . * * : : * . . : * * . . : * : *	
ZnuD	PDHAIMVDTALSQQVEILRGFVPTLLYSSGNVAGLVADVADGKIPKMPENGVSSELGLRLS	171
C1265	KSGSTYLDPALVKRIEVVRGPNSSLYSGGLGGVVDFTADAAFLPPGETNGLS---LW	172
	. . : * * * : : * : * * : : * : * . . : * . . * *	
ZnuD	SGNLEKLTSGGINIGLGKNFV-----LHTEGLYRKSGDYAVPRYRNLRKLP	217
C1265	----GNIASGDHSTSGSLTWFGKTGKTDALLSVIMRKGNINYQSDGEHAPNKEK-----	222
	: : * * . . * * . : . : * : * : * * : :	
ZnuD	DSHADSQTGSIGLSWVGKEGFIGVAYSDRRDQYGLPAHSHEYDD-----CHA-DIIWQK	270
C1265	-PAALFAKGSVGITDSNKAGASLRLYRNNTTEPGNSTQTHGDSGLRDRKTVQNDVQFWYQ	281
	* . * : * : . : * * : * : * : * . . : * : *	
ZnuD	-SLINKRYLQLYPHLLTEEDIDYDNPLGSCGFHDDDNAHAHSTHSGRPWIDLNRKRYEL-R	328
C1265	YAPVDNSLINVKSTL-YLSDITIKTN-----GH---NKTAEWRNNRTSGVNVVN	326
	: : : : : : * . * * . . . * : . : * : * . : : .	
ZnuD	AEWKQFPFGFEALRVHLNLRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWG	388
C1265	RSHTLIFPGAHQ-----LSYG	342
	. . * * * . * : *	
ZnuD	VQYLQOKSSALSASEAVKQPMLLDNKVQHYSFFGVEQANWDNFTL--EGGVRVEKQKAS	446
C1265	AEYYRQQKPEGSAT-----LYPEGNIDFTSLYQDEMTEMKSYPVNIIVGSR-----	389
	. : * : * : . . : : : : : * : : . . : : * *	
ZnuD	IQYDKALIDRENYNHPPLDGAHRQTARS--FALSGNWFYTPQHKLSLTASHQERLPST	504
C1265	--Y----DRYKSFNPRAGELKAERLSPRAAISVSPDWMMLMYS-----ISSAFRAPTM	437
	* * : : * : * * * : * . : * : * * * :	
ZnuD	QELYAHGKHVAT---NTFEVGNKHLNKRSSNNIELALGY-----EGDRWQYNLALYRN	554
C1265	AEMYRQDVVFYRKGKPNYVWPNLNLKPENNITREIGAGIQLDGLLTDNDRQLKGGYFGT	497
	* : * . . * . . : * * : * : * . . * : * . : * * * : . : .	
ZnuD	RFGNYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSGADFYGAEG-----	597
C1265	DARNYIATR-----VDMKMRMSYSYNVSRARIWGWDMQGNYQSDYVDWMLSYNR	546
	* * * : : * . . : * * * * : * : *	
ZnuD	-----EIYFKPTPRYRIGVSGDYVRGLKNLPSLP---G--REDAYGN	635
C1265	TESMDASSREWLGSGNPDTLISDISIPVGHARGVYAGWRAELASATHVKKGDPHQAGYTI	606
	: : * * * . * : * * : * . : *	
ZnuD	RPFIAQDDQNAPRVPAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLN	695
C1265	HSFS-----LSYKPVSVKGFASVTLT-----NAFN--KLAM-----NGKG-----	640
	: * * : * . . : * . . * * * . *	
ZnuD	LGANYRRNTRYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKF	748
C1265	VPLSGRTVSLYTRYQW-----	656
	: . * : * : * : *	

**Figure 6.3 Amino Acid Sequence Alignment of C1265 and ZnuD**

Amino acid sequence alignment of C1265 and ZnuD. **Bold underline** indicates ZnuD residues which interact with  $\text{Zn}^{2+}$ . **Green** highlighted indicates aligned C1265 residues which potentially interact with  $\text{Zn}^{2+}$ . '\*' fully conserved residue. ':' conservation between groups of strongly similar properties, scoring > 0.5 in the Gonnet PAM 250 matrix. '.' conservation between groups of weakly similar properties, scoring  $\leq 0.5$  in the Gonnet PAM 250 matrix.

### 6.2.2.2. Bioinformatic Analysis of C1266

Bioinformatic analysis of C1266 was less fruitful than that of C1265. Protein blast returned a domain which showed 100% identity to a DUF4198, a domain of unknown function of from *E. coli*.

Using Swiss-Model to search for protein structures with similarities, results showed a protein with 20.48% sequence identity and 0.13 GMQE score to a cell wall surface anchor family protein, RrgC of *Streptococcus pneumoniae* [accession 4oq1.1.A]. Interestingly, RrgC folds into three distinct domains (D1, D2, and D3). When comparing sequence homology from C1266 to RrgC, C1266 only shows homology to the D2 domain, not D1 or D3. It has been suggested the domain D2 of RrgC is a IgG-like domain, which is known to be involved in bacterial adhesin, and potentially pathogenicity (Shaik *et al.*, 2014).

### 6.2.2.3. Bioinformatic Analysis of C1267

A Swiss-Model search returned high similarity of C1267 to a GTP-binding protein YjiA, of *E. coli* K-12. With a 22.22% sequence identity and a 0.47 GMQE score [accession 4ixn.1.A]. YjiA is a member of the metal homeostasis COG0523 family of GTPases and was the first COG0523 member to have a high-resolution (2.05 Å) structure analysed (Sydor *et al.*, 2013). YjiA has been shown to interact with  $\text{Zn}^{2+}$ . Figure 6.4 shows the amino acid sequence alignment of C1267 and YjiA, (bold underline indicates the residues which  $\text{Zn}^{2+}$  interacts with in YjiA). There is little homology between C1267 and YjiA at the amino acid residues which binds to  $\text{Zn}^{2+}$  in YjiA. D44 and D83 of C1267 shows similar functional properties to the aligned  $\text{Zn}^{2+}$  binding residue of YjiA, and H118 shows a conserved residue between C1267 and the  $\text{Zn}^{2+}$  binding residue of YjiA. The yellow highlighted sequence shows the identical sequence, which is the suggested

GTP-binding domain of YjiA (Sydor *et al.*, 2013). A key domain of COG0523 is a putative CxCC in the amino acid sequence, it is clearly present in YjiA, but is not found in C1267 (Haas *et al.*, 2009). Additional protein BLAST suggests that C1267 may be a GTP-binding protein, which would agree with the results from Swiss-Model.

C1267	MNRPVIIIN <b>GFLGAGKTTL</b> LKNLLTQANRNHLAVSVIVNDMSLDVDGVLIANTEI--V	58
Yjia	MNPiAvTLTL <b>GFLGAGKTTL</b> RHILNEQHGYK--IAVIENEFG <b>EVS</b> VDDQLIGDRATQIK	58
	** * * :.:*****:::.*: : : :*:~::~:..**.	
C1267	NIANNFVSITSDSISSQSGIKTV <b>K</b> AIKANMLR----HYLPDVIFVETSGGSHPLPLIK	113
Yjia	TLTNGCI <b>C</b> CSRS-----NE <b>E</b> DALLDLDLDNLDKGNIQFDRLVIECTGMADPGPIIQ	109
	::*. : . * : ::*: ::* : * :.:* :* :.* **:	
C1267	YLQT <b>E</b> RRLQKS----LLSLVDVTMLNDDYNGGKALIPAFQENLFHGKRRIENLLAEQIM	169
Yjia	TFFS <b>H</b> EVLRCQRYLGDVIALVDVAHVDEQMNQ-----F-----TIAQSQVG	150
	: :** * : :****:* ::: * *	
C1267	FCSTLLLTKKDRLSFDIVTDVAKAIIHPLNPYNVIVASWGNLKLAELLTLDPDYDFNRVGL	229
Yjia	YADRILLTKTDVAGE-- <b>A</b> E <b>K</b> <b>H</b> ERLARINARAPVYTV <b>H</b> GDIDLGLLFNT-----NGF	201
	::.:****.* . .:~::~: : * . * *:~::~:..**.	
C1267	LIRELEDITITVEEKISAQNGEIIISRVIKDDRPFHP---QRLWETYHYFMGMGIYRSKGF	286
Yjia	MLEENVSTKPRFHFIADKQNDISSIVVELYPVDISEVSRVMENLLESADKLRLRYKGM	261
	::.* : . . *:~::~:* * *:~::~* *. . *:~::~* . : * **:	
C1267	FWLPGRPDMALLWNQAAGSINLEFISYWNSGVLADPDNHLTHEERSVLQKKVNKTMRFRG	346
Yjia	LWIDGENRLLFQG-VQRLYSADWDREPWGDEK---PHSTMVFIGIQLP <del>EEEE</del> IRA---AFA	314
	*: *.*: *:~::~. . :~::~* . . *~::~. . :~::~. . :	
C1267	DRRCHLTIIGRSDEVHDFTSALINCFLSEEEIVWWQSGGVFADPWPVNISR LN	399
Yjia	GLRK-----	318

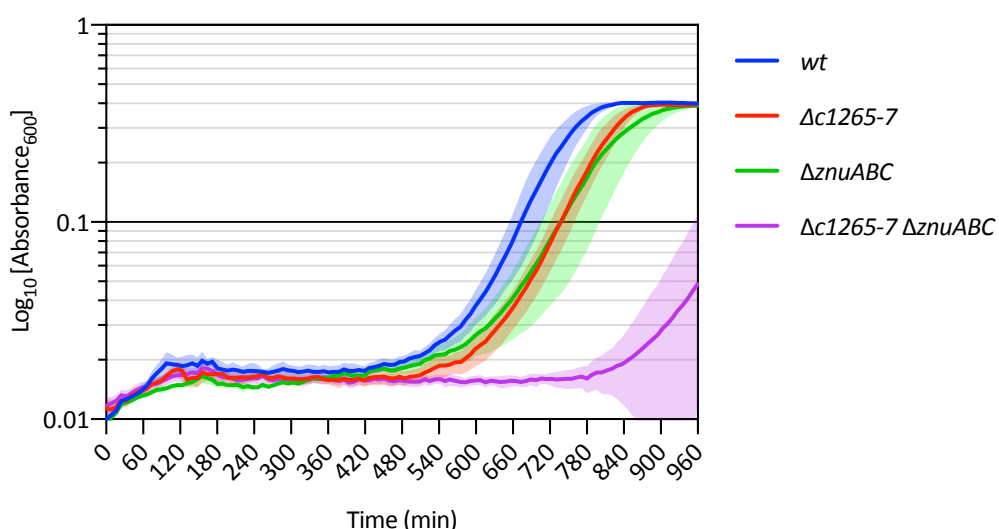
**Figure 6.4 Amino Acid Sequence Alignment of C1267 and YjiA**

Amino acid sequence alignment of C1267 and YjiA. **Bold Underline** indicated a Zn<sup>2+</sup> binding site of YjiA. **Yellow** highlighted indicated a GTP-binding domain. **Green** highlighted indicated aligned C1265 residues which potentially interact with Zn<sup>2+</sup>. '\*' fully conserved residue. ':' conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix.

### 6.2.3. Microtiter Plate Growth Assay

To try and determine if C1265-7 has an essential, and phenotypically detectable role during growth in zinc depleted conditions, a simple growth curve experiment was conducted. An initial growth curve was conducted in a 96-well microtiter plate, in zinc depleted Neidhardt's MOPS minimal media (NH), with CFT073 *wt*,  $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$  (Figure 6.5). When grown in NH media, the double knockout (KO) strain CFT073  $\Delta c1265-7\Delta znuABC$  strain shows significantly reduced growth; it

takes longer to reach exponential phase. There are large error bars in the double KO strain, suggesting that not all the biological repeats started exponential growth at the same time. When looking at the *wt* and the single KO strains, it is hard to determine if there is any difference. The graph suggests that the *wt* has a slight advantage over the single KO strains; that is, it reaches exponential phase quicker. This data suggests that a single KO of either  $\Delta c1265-7$  or  $\Delta znuABC$  does not cause a noticeable difference in growth, however a double KO of both genes causes a reduced growth efficiency. This data could suggest that in the absence of *znuABC*, *c1265-7* acts as an alternative zinc import system, and vice versa. Further strengthening the argument that *c1265-7* acts as a zinc import system.



**Figure 6.5 Microtiter Growth Curve Assay of *E. coli* CFT073 (*wt* and KO strains) in NH Media**

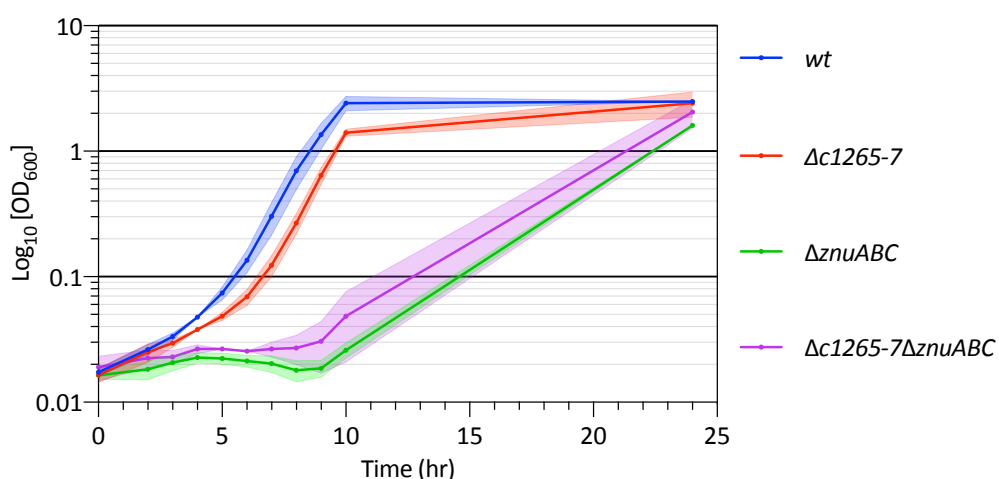
*E. coli* CFT073 (*wt*,  $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$ ) were grown with aerobic incubation at 37°C, with shaking (200 RPM), in NH, with five biological repeats. Absorbance<sub>600</sub> was recorded every 10 min for 16 hr in a TECAN GENios Pro plate reader. Shaded areas indicate S.D.

#### 6.2.4. Shake Flask Growth Curve Assay

Conducting a growth curve assay in a microtiter plate has its disadvantages, such as; small volume which tends to dry out over time, and small pathlength for absorbance reading. Microtiter plate readers do not record OD<sub>600</sub>, but rather Absorbance<sub>600</sub>. This

is because the small sample volume used in microtiter plates does not allow for the required 1 cm pathlength.

A standard shake flask growth curve assay was conducted with *E. coli* CFT073 *wt* and KO strains ( $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$ ), grown in NH media (Figure 6.6). *E. coli* CFT073  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$  shows reduced growth compared to *wt* and  $\Delta c1265-7$ .  $\Delta znuABC$  and  $\Delta c1265-7\Delta znuAB$  take longer to reach exponential phase, which starts at around 9-10 hr incubation, however this is hard to tell as optical density readings were not recorded between 11-23 hr. Whereas *wt* and  $\Delta c1265-7$  starts exponential phase at 4-5 hr. *E. coli* CFT073 *wt* has slightly better growth than  $\Delta c1265-7$ . If one looks at the rate of growth ( $OD_{600} \text{ hr}^{-1}$ ), *E. coli* CFT073 *wt* has a significantly higher rate than *E. coli* CFT073  $\Delta c1265-7$  (Figure 6.7). This suggests that,  $\Delta c1265-7$  does not cause a large measurable reduction in growth, but a small reduction, meaning that *c1265-7* may play a minor role in zinc acquisition in a zinc depleted state.

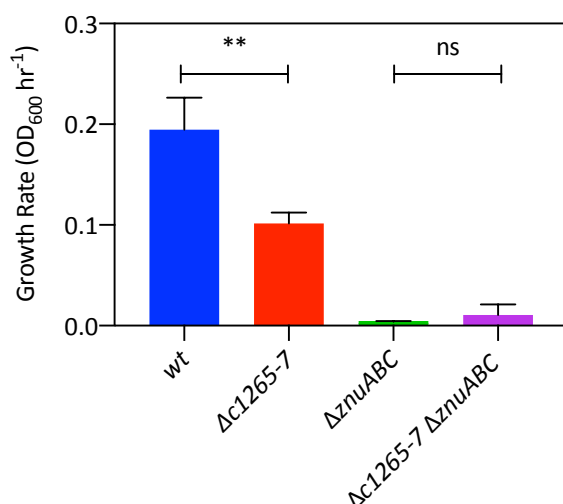


**Figure 6.6 Shake Flask Growth Curve Assay of *E. coli* CFT073 in NH**

*E. coli* CFT073 (*wt*,  $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$ ) were grown aerobically at 37°C, with shaking (200 RPM) in 35 mL NH media, with three biological repeats.  $OD_{600}$  was recorded every hour for 10 hr, and a final reading at 24 hr Incubation. Shaded areas indicate S.D.

Interestingly, after 24 hr incubation, the *wt* and the three KO strains all have similar final  $OD_{600}$  values. A one-way ANOVA was conducted and showed there is no significant

variation between samples in the final 24 hr reading (data not shown). This suggests that the deletion in the CFT073  $\Delta znuABC$  and  $\Delta c1265-7\Delta znuABC$  strains reduced the growth rate but did not reduce their ability to grow to the same final cellular concentration as *E. coli* CFT073 *wt*. This could further suggest that other less effective import systems acquire zinc in the absence of  $\Delta znuABC$ .



**Figure 6.7 Shake Flask Growth Rate of *E. coli* CFT073 (*wt* and KO) in NH**

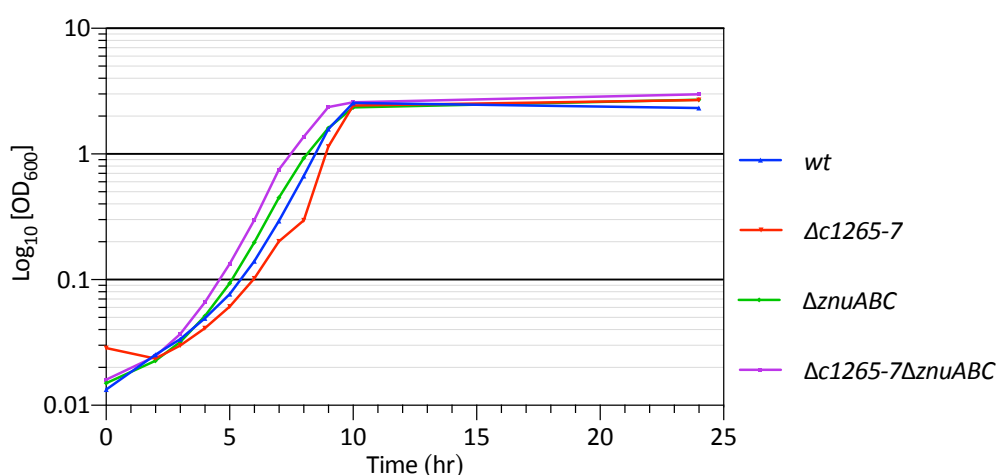
Growth Rate (OD<sub>600</sub> hr<sup>-1</sup>) of *E. coli* CFT073 *wt*,  $\Delta znuABC$ ,  $\Delta c1265-7$ ,  $\Delta c1265-7 \Delta znuABC$ . Grown aerobically at 37°C with shaking (200 RPM) Error bars indicate S.D. \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test). Data from Figure 6.6

There is little variation between  $\Delta znuABC$  and  $\Delta c1265-7\Delta znuABC$ , however there is a difference between  $\Delta c1265-7$  and the aforementioned deletion strains. Comparing growth rates of  $\Delta c1265-7$  and  $\Delta c1265-7\Delta znuABC$ , there is a significant difference in growth rate (two-tailed, unpaired, t-test).

*E. coli* CFT073 *wt* and KO ( $\Delta znuABC$ ,  $\Delta c1265-7$ , and  $\Delta c1265-7 \Delta znuABC$ ) strains were grown in zinc replete NH media (NH + 50  $\mu$ M ZnSO<sub>4</sub>). This experiment was conducted to show that in zinc replete conditions, the KO strains do not have a secondary burden on growth. Figure 6.8 shows *E. coli* CFT073 *wt* and KO strains grown in NH + 50  $\mu$ M ZnSO<sub>4</sub>. It is important to note that no standard deviation was plotted, this was done simply because if SD was plotted the graph becomes overcrowded and obscures data,



however statistical tests were conducted using all biological repeats. The graph shows all four strains show similar growth curves. All strains have a similar final OD<sub>600</sub> at 24 hr and show similar growth rates in exponential phase. A one-way ANOVA was conducted on growth rate during exponential growth, and results show no significant variation between the growth rate of all four strains (data not shown). This shows that the KO strains do not cause any secondary, or unknown side effects on growth when grown in a normal or zinc replete media. The alterations in growth observed in Figure 6.6 are due to the zinc depleted conditions.



**Figure 6.8 Shake Flask Growth Curve Assay of *E. coli* CFT073 (*wt* and KO) in NH + 50  $\mu$ M ZnSO<sub>4</sub>**

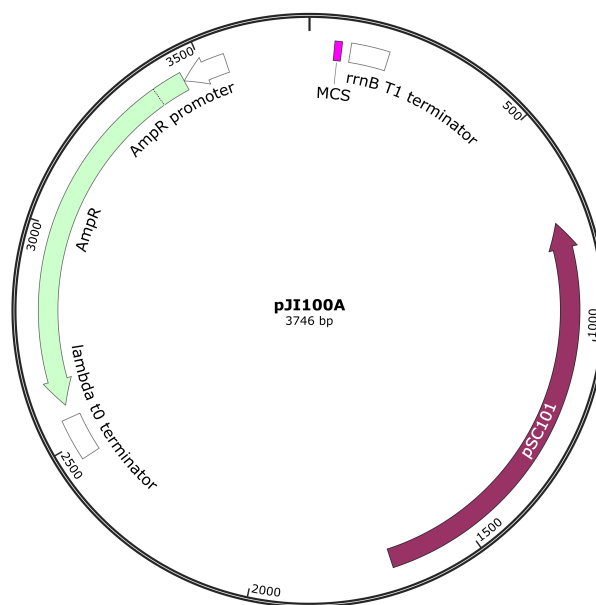
*E. coli* CFT073 (*wt*,  $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$ ) were grown aerobically at 37°C, with shaking (200 RPM) in 35 mL NH + 50  $\mu$ M ZnSO<sub>4</sub>, with three biological repeats. OD<sub>600</sub> was recorded every hour for 10 hr, and a final reading at 24 hr Incubation.

#### 6.2.5. Construction and validation of pJI100A and pJI102: C1265-7 Expression Plasmid

Growth curve assays in *E. coli* CFT073 *wt* and KO strains did not give a definitive answer on the phenotype of *c1265-7*. To help define the phenotype of *c1265-7*, a low copy number expression plasmid was constructed. This expression plasmid was created for two main reasons: to allow recombinant gene expression in a lab strain of *E. coli*

(MG1655) where *c1265-7* is not present; and to increase the copy number of *c1265-7* per cell.

Firstly, the empty vector plasmid was constructed, pJI100. pJI100 was based on pLUX (Burton *et al.*, 2010). NEB Q5 polymerase was used to amplify the *pSC101*, *rrnB*, and *kan<sup>R</sup>* elements of pLUX, whilst incorporating a BamHI site at the 3' and 5' end, which was then used to circularise the PCR product. An ampicillin resistant alternative plasmid was created. pJI100 was firstly linearized up and down stream of *kan<sup>R</sup>*, *amp<sup>R</sup>* was linearized from pDOC-G, and the NEBuilder HiFi DNA Assembly kit was used to clone these products together, creating pJI100A (Figure 6.9).

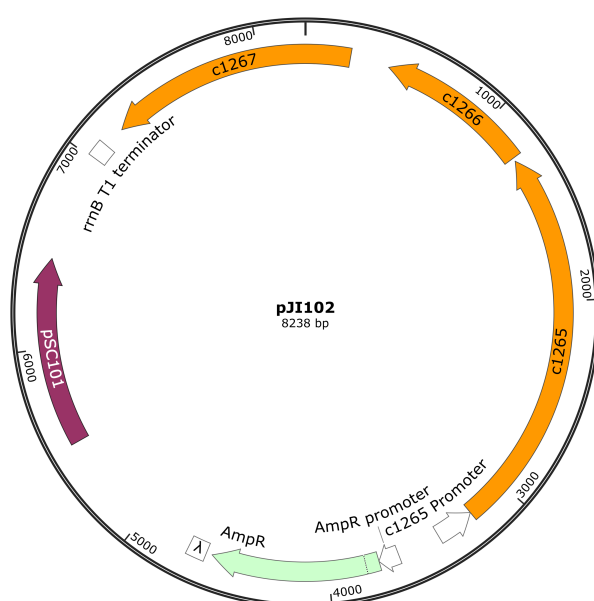


**Figure 6.9 Plasmid Map of pJI100A**

Plasmid map of pJI100A (empty Vector) showing relevant genes.

pJI102 (Figure 6.10), a *c1265-7* expression plasmid was created by cloning *c1265-7* into pJI100A. *c1265-7*, along with 200 bp upstream of *c1265* was amplified from *E. coli* CFT073, pJI100A was linearized at the MCS site, and NEB HiFi DNA Assembly kit was used to clone these two products together.

To ensure correct construction, primers *rrnB\_seq\_R* and *seq\_204\_F* were used to sequence pJI102. ClustalW was used to align the returned sequence data with the expected sequence data; the results showed correct insertion (data not shown). The sequence data was further analysed by visualising the Chromatogram data, to ensure high quality sequenced data. Chromatogram data showed high quality sequence data (Data not shown).



**Figure 6.10 Plasmid Map of pJI102**

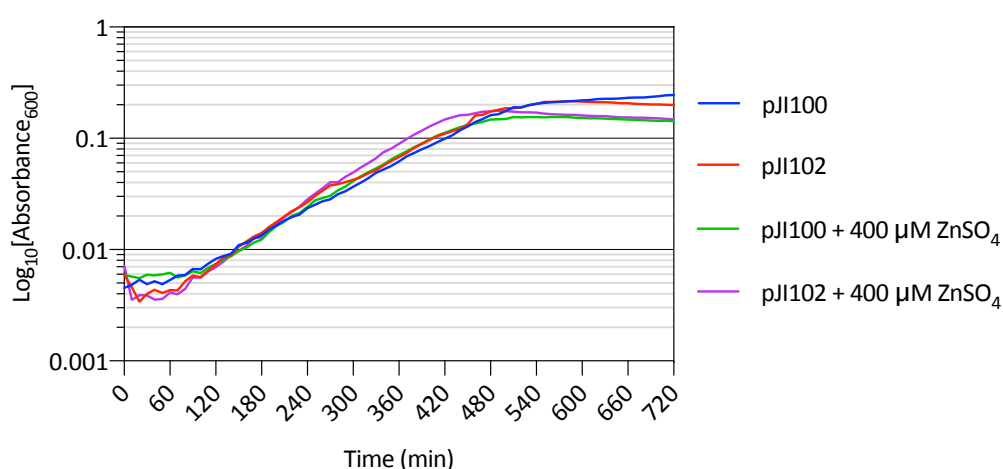
Plasmid map of pJI102, *c1265-7* expression plasmid, showing relevant genes.

#### 6.2.6. Growth Assays with the *c1265-7* Expression Vector (pJI102)

Standard growth curve assays, in both microtiter plates (6.2.3) and shake flasks (6.2.4) suggests that *c1265-7* may not be vital in survival in a zinc depleted conditions, unless the major zinc import system (*znuABC*) has been deleted. Results suggest that a double knockout of  $\Delta c1265-7\Delta znuABC$  may grow less well in zinc depleted conditions than a single KO strain, but results are not conclusive and the microtiter and shake flask growth curve assays suggest different conclusions.

pJI102, a *c1265-7* expression plasmid was created, allowing recombinant gene expression in *E. coli* MG1655. Recombinant expression in *E. coli* MG1655 *wt*,  $\Delta znuA$  and  $\Delta znuCB$  strains may allow us to further understand the phenotype of *c1265-7*.

Initially, plasmids pJI100A and pJI102 were transformed and expressed in *E. coli* MG1655 in NH with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$  (Figure 6.11). Results show that in NH only media, there is no variation in growth between strains carrying the two plasmids; the empty vector (pJI100A), or the *c1265-7* expression vector (pJI102). Further, there is no variation with NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ . This suggests that in *E. coli* MG1655 *wt*, the presence of *c1265-7* confers no advantages or disadvantages to growth. Data also suggests that in *E. coli* MG1655 *wt*, the presence or absence of *c1265-7* in either zinc depleted, or zinc replete conditions does not provide any advantages in growth.



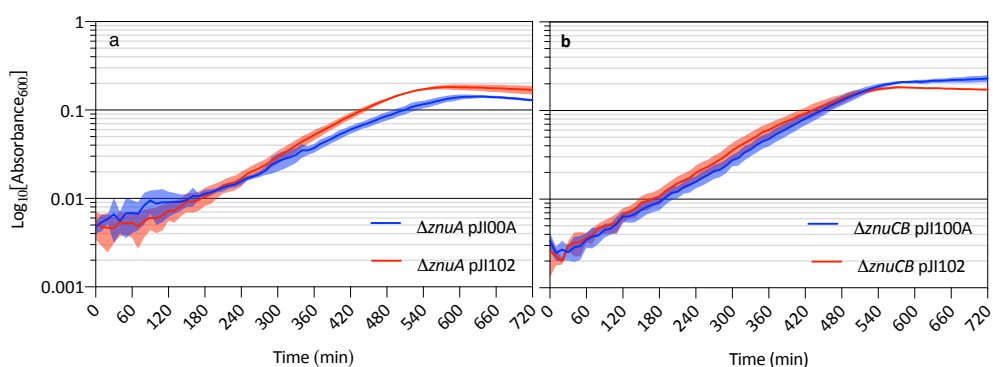
**Figure 6.11 Microtiter Growth Assay of *E. coli* MG1655 *wt* (pJI100A or pJI102) in NH**

Plasmid constructs (pJI100A, pJI102) transformed and expressed in *E. coli* MG1655 *wt* with aerobic incubation at 37°C, with shaking (200 RPM) in  $\text{NH}_{\text{amp100}}$  with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> recorded every 10 min for 12 hr in a TECAN GENios Pro Plate reader. pJI100A empty vector. pJI102, *c1265-7* expression vector.

As the *c1265-7* expression plasmid does not seem to confer any advantages or disadvantages when expressed in *E. coli* MG1655 *wt*, deletion of the *znuABC* zinc import system was the next logical step. Data in Figure 6.6 suggested that  $\Delta znuABC$

plays a major disadvantage to growth in *E. coli* CFT073 when grown in zinc depleted media. To further understand the role of *c1265-7* and determine if it replaces the zinc import function of *znuABC* in a  $\Delta znuABC$  strain, pJI102 was expressed in *E. coli* MG1655  $\Delta znuA$  and  $\Delta znuCB$  strains. By knocking out genes separately within the *znuABC* operon it may help to further understand the role of *c1265-7* in zinc acquisition.

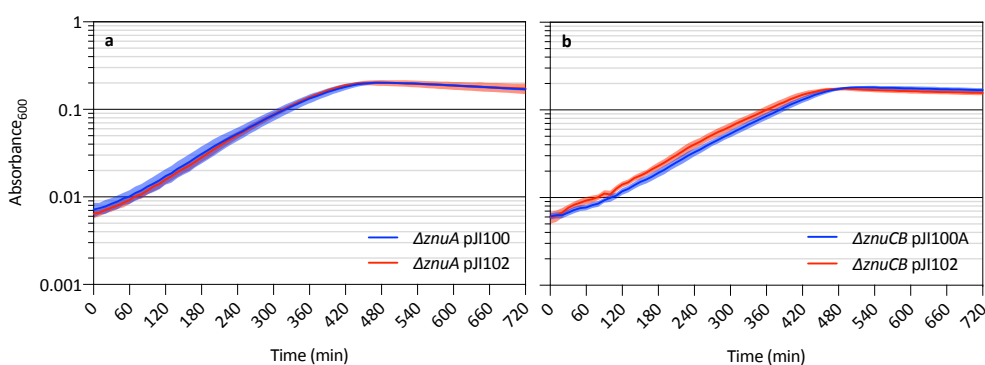
*E. coli* MG1655  $\Delta znuA$  and  $\Delta znuCB$  were transformed with both pJI100A and pJI102 and grown in NH media (Figure 6.12). Results in Figure 6.12a suggest that *E. coli* MG1655  $\Delta znuA$  is disadvantaged when grown without *c1265-7* (pJI102) compared to the empty vector (pJI100A). When comparing the growth rate of *E. coli* MG1655  $\Delta znuA$  transformed with either pJI100A or pJI102, there is a significant higher growth rate with pJI102 than with pJI100A (two-tailed, unpaired, t-test). In comparison, the growth rate of *E. coli* MG1655  $\Delta znuCB$  transformed with pJI100A or pJI102, there is no significant difference in growth rate (two-tailed, unpaired, t-test) (Figure 6.12b).



**Figure 6.12 Microtiter Growth Assays of *E. coli* MG1655  $\Delta znuA$  &  $\Delta znuCB$  (pJI100A or pJI102) in NH**

Plasmid constructs transformed and expressed in *E. coli* MG1655  $\Delta znuA$  or  $\Delta znuCB$  with aerobic incubation at 37°C, with shaking (200 RPM) in  $NH_{amp100}$ , with three biological repeats. Absorbance<sub>600</sub> recorded every 10 min for 12 hr in a TECAN GENios Pro Plate reader. pJI100A empty vector. pJI102, *c1265-7* expression vector. Shaded areas indicate S.D. (a)  $\Delta znuA$ . (b)  $\Delta znuCB$ .

However, the variation in growth between *E. coli* MG1655  $\Delta znuA$  with either pJI100A and pJI102 is very small and does not necessarily suggest a phenotype of *c1265-7*. The experiment in Figure 6.12 was re-run with increased biological repeats (five repeats) with the aim to reduce error and to draw a more statistically robust conclusion from the data (Figure 6.13). From initial comparison of the three biological repeat data (Figure 6.12) to the five biological repeat data (Figure 6.13), the five biological repeat data set is more consistent; the error bars are smaller, especially between 0-180 min, as there is less variation between biological repeats. On examination of *E. coli* MG1655  $\Delta znuA$  (Figure 6.13a), there is little to no variation between *E. coli* MG1655  $\Delta znuA$  transformed with either pJI100A or pJI102, which suggests that in the absence of *znuA*, *c1265-7* does not provide any advantages. This is the same for *E. coli* MG1655  $\Delta znuCB$  transformed with either pJI100A or pJI102 (Figure 6.13b). There is very little variation that one cannot significantly say there is any variation.



**Figure 6.13 Microtiter Growth Assay of *E. coli* MG1655  $\Delta znuA$  &  $\Delta znuCB$  (pJI100A, pJI102) in NH, Five Biological Repeats**

Plasmid constructs transformed and expressed in *E. coli* MG1655  $\Delta znuA$  or  $\Delta znuCB$  with aerobic incubation at 37°C, with shaking (200 RPM) in  $NH_{Amp^{100}}$ , with five biological repeats. Absorbance<sub>600</sub> recorded every 10 min for 12 hr in a TECAN GENios Pro Plate reader. pJI100A empty vector. pJI102, *c1265-7* expression vector. Shaded areas indicate S.D. (a)  $\Delta znuA$ . (b)  $\Delta znuCB$ .

From the data collected in Section 6.2.6, no conclusion of the phenotype of *c1265-7* can be drawn. The hypothesis that *c1265-7* may be vital for survival in zinc depleted conditions has not been proven. Neither results from Section 6.2.4 helped conclude a phenotype for *c1265-7*. It is possible that the variation between samples are too small for a standard growth curve assay to record, or that advantages to cell growth under zinc depleted conditions are found in conditions that were not tested in these experiments.

#### 6.2.7. BacTiter-Glo Cell Viability Assay

To improve the sensitivity of a standard growth curve, the BacTiter-Glo™ Microbial Cell Viability Assay kit was used (Promega, G8230). The BacTiter-Glo kit measures cell viability by quantifying the presence of ATP. Promega claim that the BacTiter-Glo kit is 1000x more sensitive than a standard absorbance-based assay. This improved sensitivity to cell count may help with the understanding of the phenotype of C1265-7.

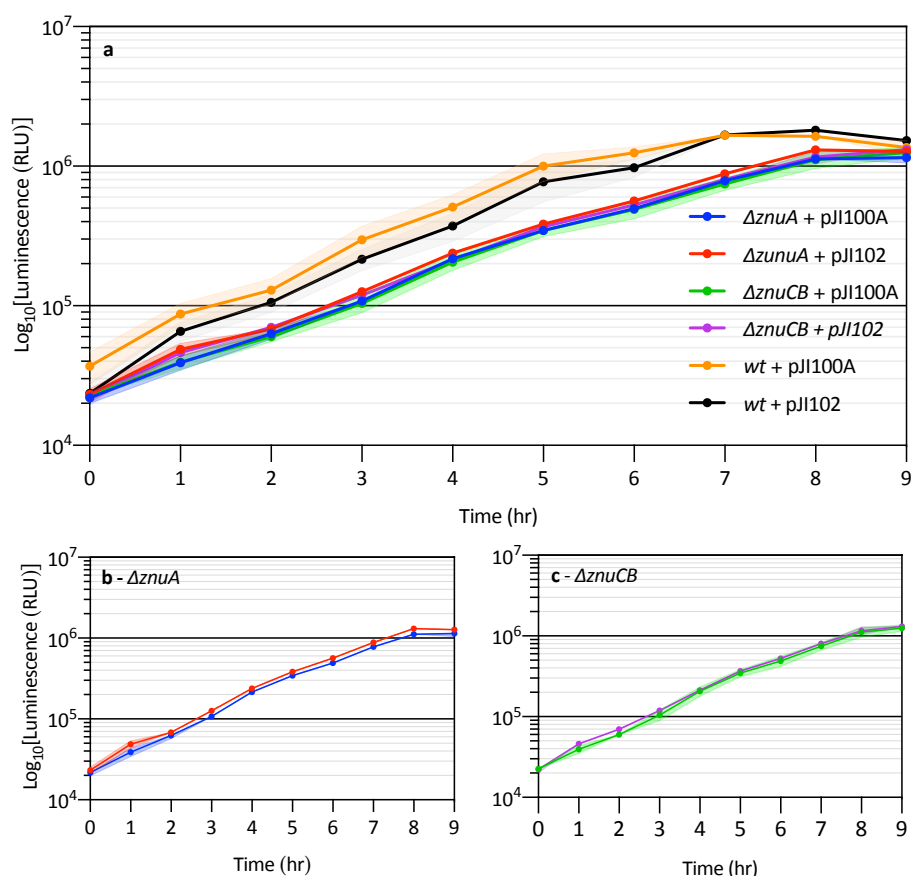
*E. coli* MG1655 *wt*,  $\Delta znuA$ , and  $\Delta znuCB$  were transformed with both pJI100A and pJI102, grown in NH media, with BacTiter-Glo readings taken hourly for 9 hr (Figure 6.14). Initially, from comparing all strains and transformants (Figure 6.14a) *E. coli* MG1655 *wt* grew better than either of the KO strains. Growth rates were determined for the *wt* strains, and a two-tailed unpaired t-test was conducted. A t-test concluded that there is no significant difference in growth rate between *E. coli* MG1655 *wt* + pJI100A and *E. coli* MG1655 *wt* + pJI102 (Figure 6.15). This finding confirms that the presence of *c1265-7* in *E. coli* MG1655 does not cause a burden growth, compared to *E. coli* MG1655 *wt* not containing *c1265-7*.

Looking at the effect of *c1265-7* on  $\Delta znuA$  strains (Figure 6.14b) there looks to be improved growth in *E. coli* MG1655  $\Delta znuA$  with pJI102 compared with pJI100A.

Comparing growth rates (RLU hr<sup>-1</sup>) of this data shows a significant difference in growth rate of *E. coli* MG1655  $\Delta znuA$  transformed with pJI100A, compared to pJI102, that shows increased growth rate. This data suggests that in the absence of *znuA*, the presence of *c1265-7* shows slightly advantageous growth. This could further suggest that one of the *c1265-7* operon genes may play a similar role to ZnuA in zinc import.

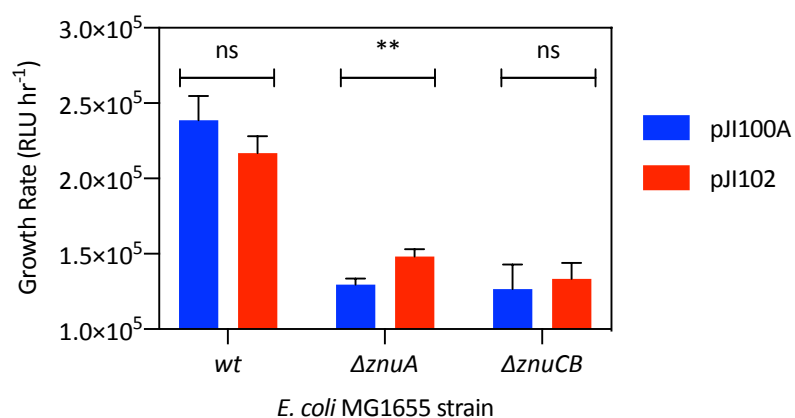
When comparing *E. coli* MG1655  $\Delta znuCB$  transformed with pJI100A and pJI102 (Figure 6.14c) there seems to be less variation between the two samples and overlapping error bars. When comparing the growth rate of *E. coli* MG1655  $\Delta znuCB$  transformed with either pJI100A or pJI102, there is no significant difference in growth rate (Figure 6.15)





**Figure 6.14** BacTiter-Glo Assay of *E. coli* MG1655 *wt*,  $\Delta znuA$  and  $\Delta znuCB$  (pJI100A, pJI102) in NH

*E. coli* MG1655 (*wt*,  $\Delta znuA$ , and  $\Delta znuCB$ ) transformed with either (pJI100A or pJI102) were grown aerobically at 37°C, with shaking (200 RPM) in 25 mL  $\text{NH}_{\text{amp}100}$  media, with three biological repeats. BacTiter-Glo readings were taken hourly, according to manufactures instructions. Luminescence (integration 250 ms) was recorded on a TECAN GENious Pro. Shaded areas indicate S.D. (a) All data. (b) *E. coli* MG1655  $\Delta znuA$ . (c) *E. coli* MG1655  $\Delta znuCB$ .



**Figure 6.15 Growth Rate of *E. coli* MG1655 wt,  $\Delta znuA$  and  $\Delta znuCB$  (pJI100A or pJI102)**

Growth Rate (RLU hr<sup>-1</sup>) of *E. coli* MG1655 wt,  $\Delta znuA$ ,  $\Delta znuCB$  transformed with either pJI100A or pJI102. Growth determined by BacTiter-Glo. Error bars indicate S.D. ns  $p \geq 0.05$ , \*\*  $p \leq 0.01$  (two-tailed, unpaired, t-test) comparison of pJI100A to pJI102. Growth rate determined from data in Figure 6.14.

### 6.3. Discussion and Conclusion

This chapter aimed to determine the phenotype of the novel *c1265-7* gene operon from *E. coli* CFT073. Previous data shows that *c1265-7* is significantly upregulated during zinc depleted conditions, which suggests that C1265-7 may play a role in zinc acquisition. Further to this, transcriptomic and bioinformatic data previously conducted by Clayton, (2012) suggested that *c1265-7* is regulated by Zur. All known Zur regulated genes play a role in either zinc acquisition, or are alternative paralogue proteins which do not require zinc, thus releasing zinc in the cytoplasm during times of zinc shortage

#### 6.3.1. Bioinformatics

Bioinformatics were used to look for homologous proteins from online databases. Analysis of C1265 showed strong sequence and structural homology to the outer heme receptor, ShuA, of *S. dysenteriae*, a TonB-dependent receptor (Cobessi, Meksem and Brillet, 2010). Further analyses using BLASTp showed homology to many uncharacterised TonB-dependent receptors. When comparing the structure of ShuA and C1265 (Swiss-Model predicted), there is an interesting pattern (Figure 6.2). The  $\beta$ -barrel structure of the predicated model of C1265 shows a high QMEAN value to ShuA, whereas the extra cellular domains of C1265 show a low QMEAN score. This suggests that the  $\beta$ -barrel of C1265 shows a similar modelled structure to that of ShuA, and one could assume that the structure and function would be the same. However, the QMEAN value of the extra cellular domain, the loops, is low. Knowing that the loops of the extra cellular domain are responsible for ligand binding (Buchanan *et al.*, 1999), I hypothesise that C1265 has the same function as ShuA (TonB-dependent receptor), but C1265 binds to a different ligand, potentially  $\text{Zn}^{2+}$ . To further support this argument, C1265 shows homology to ZnuD, a TonB-dependent receptor which

translocates zinc in *N. meningitidis* (Calmettes *et al.*, 2015). C1265 shows moderate sequence homology to the zinc binding residues of ZnuD, with five zinc binding residues which align to ZnuD (Figure 6.3). Moreover, *znuD* has been shown to be regulated by Zur (Pawlik *et al.*, 2012).

Bioinformatic analysis of C1266 was less useful in determining the phenotype. C1266 showed homology to DUF1498. However, it did show ~20% identity to RrgC, which is a IgG-like domain which is involved in host attachment in *S. pneumoniae* (Shaik *et al.*, 2014).

Bioinformatic analysis of C1267 showed 22.22% sequence homology to YjiP, a COG0523 GTPase which has been shown to interact with  $\text{Zn}^{2+}$ . There is little sequence homology in YjiA residues which interact with  $\text{Zn}^{2+}$ . Two residues show amino acids with similar function, and one fully conserved residue. COG0523 is subfamily of the G3E family of P-loop GTPases (G3E family). A function of a G3E family are metallochaperones, or for the storage and delivery of metals. C1267 does not contain the conserved putative metal binding domain (CxCC) found in all COG0523 protein (Haas *et al.*, 2009). It has been shown that some COG0523 genes are regulated by Zur, and are often associated with pathogenic bacteria (Gaballa and Helmann, 1998; Maciąg *et al.*, 2007; Haas *et al.*, 2009).

From looking at the homologous proteins to C1265, it can be strongly argued that C1265 is a TonB-dependent receptor, which is highly likely to be involved with the import of  $\text{Zn}^{2+}$ . However, there is not enough bioinformatic data to strongly argue the role of C1266 or C1267. Bioinformatics does allude to C1267 being a COG0523 protein.

### 6.3.2. Growth Curve Experiments

Standard growth curve experiments, both in shake flasks and microtiter plates did not help define the phenotype of C1265-7. When conducted in a microtiter plate, in zinc depleted media (NH), results suggested that in the *E. coli* CFT073 dual KO strain  $\Delta c1265-7\Delta znuABC$  there was a reduced growth rate compared to *wt* (Figure 6.5). However, when conducted in a shake flask in zinc depleted media, results indicated that both  $\Delta znuABC$  and  $\Delta c1265-7\Delta znuABC$  reduced the growth rate compared to *wt* (Figure 6.6). As  $\Delta c1265-7$  did not cause a reduction in growth rate compared to *wt*, this could suggest that ZnuABC plays a more vital role in survival in zinc depleted media than C1265-7. A previous study by Sabri, Houle and Dozois, (2009) showed that ZnuABC was the predominant zinc import system of *E. coli* CFT073, and that deletion of *znuABC* will reduce its pathogenicity in urinary tract infections.

Growth curve experiments using the C1265-7 expression plasmid (pJI102) in *E. coli* MG1655 also did not help define a phenotype. The growth curve assay in a microtiter plate (Figure 6.12) showed a significant increase in growth rate of *E. coli*  $\Delta znuA$  when grown with the C1265-7 expression plasmid (pJI102) compared to the empty vector (pJI100A), however this variation was very little and may not actually suggest phenotype. This experiment was repeated with increased biological repeats, from three to five biological repeats (Figure 6.13). Results with increased biological repeats reduced variation between samples and showed no variation in growth rate between the pJI102 or pJI100A in either *E. coli* MG1655  $\Delta znuA$  or *E. coli* MG1655  $\Delta znuCB$ .

Finally, a shake flask growth curve experiment was conducted which used BacTiter-Glo, an ATP assay, as an improved method to measure cell viability, compared to optical density measurement. BacTiter-Glo measures ATP, and as such only measures live cells. Results showed that there is a significant increase in growth rate

of *E. coli* MG1655  $\Delta znuA$  with the C1265-7 expression plasmid (pJI102) compared to *E. coli* MG1655  $\Delta znuA$  with the empty vector (pJI100A) (Figure 6.14). There was no significant difference in growth rates with *E. coli* MG1655  $\Delta znuCB$  with pJI102 or pJI100A (Figure 6.15).

Growth curve results suggest that in the absence of ZnuA, a periplasmic binding protein, C1265-7 may replace its function. Clayton, (2012) proposed that C1266 may be a zinc chaperone found in the periplasm, transporting zinc to C1267, which could link to the growth curve experiment data. In the absence of ZnuA, C1266 may 'deliver' zinc to ZnuCB, which transports it across the inner membrane to the cytoplasm. If the presence of C1265-7 acted as a completely independent zinc import system, then  $\Delta znuA$  and  $\Delta znuCB$  should in theory have the same growth rate, as all three genes, *znuABC*, are needed for zinc import. But as an improved growth rate is only seen in  $\Delta znuA$  with C1265-7 expression, this supports the idea that C1266 may replace the functional role of ZnuA, but C1265-7 does not replace the functional role of ZnuCB. However, the significant differences in the growth rate are very low. Therefore, the evidence would not be strong enough to support any theory for the function of C1265-7.

### 6.3.3. Concluding Remarks

With the strong promoter homology of  $P_{c1265}$  to the Zur binding consensus, and the significant increase in promoter activity in zinc depleted condition, it is clear that *c1265-7* is under the regulation of Zur.

Bioinformatics gives support that C1265 is a TonB-dependent receptor, which is likely to translocate  $Zn^{2+}$ . C1266 may play a role in host adhesion or zinc chaperone, and that C1267 may be a COG0523, a metallochaperone protein.

The growth curve experiments and BacTiter-Glo assays did not definitively show any phenotype of C125-7, or indeed help to answer what their function is. Following on from Clayton, (2012) it could be argued that C1266 acts as a paralogue for ZnuA in its absence. To be able to see the true function of C1265-7 it may be essential for an infection model to be used. Host nutritional immunity is a mechanism in which the host immune system sequesters metal from a pathogen. The idea is simple, transition metals such as  $\text{Fe}^{2+}$  and  $\text{Zn}^{2+}$  are essential for life, limiting these from the pathogen inhibits bacterial growth. An example of this is a vertebrate calprotectin, S100, which sequesters both iron and potentially zinc (Kehl-Fie and Skaar, 2010). It has been shown that a KO strain of *E. coli* CFT07 ( $\Delta\text{znuABC}$  and  $\Delta\text{zupT}$ ) significantly reduced bacterial count in both kidney and bladder infection models, showing their functionality in pathogenicity (Sabri, Houle and Dozois, 2009). Knowing that C1265-7 was isolated from pathogenic *E. coli*, and the C1266 shows minor homology to D2 of the RrgC, IgG-like domain, which is involved in host bacterial adhesion, it is possible that the phenotype of C1265-7 will only be observed in an infection model. All this knowledge may suggest that C1265-7 has a role in pathogenicity, or to increase survival in a host, and a phenotype may only be observed during pathogenicity (e.g. an infection model). But unfortunately, the data collected in this chapter has not definitively defined the function of C1265-7.

#### 6.3.4. Future Work

All previous data suggests that C1265-7 is under the regulation of Zur, both end point fluorescence assays, and bioinformatics analysis predicting the Zur binding site. To definitively show that *c1265-7* is under the regulation of Zur, a DNA foot-printing assay should be conducted; as was used by Outten *et al.*, (2001) to determine the Zur binding site of *c1265-7*.

To understand if the phenotype of C1265-7 is more prevalent during infection and pathogenicity, it would be useful to conduct an infection model assay; *E. coli* CFT073 *wt*,  $\Delta c1265-7$ ,  $\Delta znuABC$ , and  $\Delta c1265-7\Delta znuABC$  grown in a kidney and bladder infection model, such as CBA/J mice as described by (Sabri, Houle and Dozois, 2009). By observing growth and bacterial count, one should be able to see if the presence of C1265-7 aids in pathogenicity.

Finally, crystallography of C1265-7 would provide invaluable information regarding the structure and function of these genes.



## Chapter 7

# General Discussion, Conclusion, and Future Work

### 7.1. Discussion

The main aim of this thesis was to develop a novel inducible promoter system based on the highly controlled zinc homeostasis system found in *Escherichia coli*. Previous studies indicated that the gene regulation of these zinc homeostasis genes could be adapted for a novel inducible promoter system. Particularly, the ZntR regulation of  $P_{zntA}$  was of interest, due to the evidence that it was a tuneable promoter. As well, Zur regulated zinc homeostasis gene promoters could also be used as a novel inducible promoter system. This thesis used a combination of fluorescence assays, luminescence assays, RT-qPCR analysis, and flow cytometry to evaluate the potential of a novel inducible system based on either ZntR or Zur regulation. Seven promoters were chosen for analysis: ZntR regulated  $P_{zntA}$ ; Zur regulated  $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$  and  $P_{c1265}$ . As well, several site directed mutated  $P_{zntA}$  promoters were created and analysed.

The understanding of ZntR regulation of  $P_{zntA}$  gained in this thesis suggests that  $P_{zntA}$  is a highly zinc tuneable promoter. However, little is known about the stochastic gene regulation  $P_{zntA}$ . This led to the development of studies on individual cells within a population, which tested the hypothesis by Takahashi *et al.*, (2015): that *E. coli* mounts a heterogeneous gene regulation of *zntA*.

The promoter of the novel C1265-7 gene operon ( $P_{c1265}$ ), only found chromosomally in UPEC strains, showed a highly significant increase in promoter activity in low zinc environments. However, the function of C1265-7 had not yet been determined. This led to the development of Chapter 6, which aimed to further understand the role of C1265-7 during zinc depleted conditions.

### 7.1.1. Zinc Depleted Neidhardt's MOPS Minimal Media (NH)

Zinc depleted Neidhardt's MOPS minimal media (NH) was chosen as the growth medium to conduct zinc depleted assays in. Neidhardt's MOPS minimal media was developed by Neidhardt, Bloch and Smith, (1974) as chemically defined media for use as an alternative to tradition media (e.g. LB) for the growth of *Enterobacteriaceae*. Clayton, (2012) adapted the original formula of Neidhardt's MOPS minimal media to significantly reduce the zinc concentration. This adapted zinc depleted Neidhardt's MOPS minimal media (NH) formulated by Clayton, (2012) was the basis in which NH was made in this thesis. ICP-MS analysis conducted on NH media determined the zinc concentration in the media to be 115 nM. Compared to the zinc concentration of LB (12.2  $\mu$ M), NH showed a significant reduction in zinc concentration (Table 3.1).

Another, easier, approach to reduce the available zinc in the media is with the addition of TPEN [N,N,N',N'-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine], a membrane permeable chelator. TPEN is a zinc chelator, which reduces the zinc available by binding to  $Zn^{2+}$  ions. TPEN is an often used as zinc chelator for studies involving zinc homeostasis in *E. coli* and other bacteria (Outten and O'Halloran, 2001; Stork *et al.*, 2010; Clayton, 2012; Latorre *et al.*, 2015; Goethe *et al.*, 2020). Often a concentration of 20-25  $\mu$ M TPEN is added to media when working with *E. coli* (Outten and O'Halloran, 2001; Clayton, 2012). A simple TEPN toxicity test was conducted in *E. coli* which concluded that up to 40  $\mu$ M TPEN does not cause any side effects on *E. coli* growth (Figure 3.4), and therefore the addition of 20  $\mu$ M TPEN to LB, used through this study, is safely below the toxic level.

### 7.1.2. A Novel $P_{zntA}$ Zinc Inducible Promoter

Previous studies, such as (Brocklehurst *et al.*, 1999; Outten and O'Halloran, 2001; Takahashi *et al.*, 2015) have shown evidence that the promoter activity of  $P_{zntA}$  increases with increased zinc concentration. However, no published work has examined whether the tuneable promoter of  $P_{zntA}$  could be used to develop a novel inducible promoter system.

End point fluorescence assays, based on the pJI301  $P_{zntA}$  reporter plasmid, demonstrated the strong correlation between increased zinc concentration ( $\text{ZnSO}_4$ ) and the recorded fluorescence (a measurable proxy for promoter gene expression) (Figure 3.9 and Figure 3.12). Regression analysis of the relationship between  $\text{ZnSO}_4$  concentration and fluorescence returned an  $R^2$  value of 0.99 for both end point assays in LB and NH. This agrees with the data presented by (Brocklehurst *et al.*, 1999), however Brocklehurst *et al.*, (1999) showed an induction of  $P_{zntA}$  with up to 1.2 mM added zinc, whereas the zinc toxicity test in this thesis suggested added zinc ( $\text{ZnSO}_4$ )  $\geq 600 \mu\text{M}$  causes a toxic effect on *E. coli* (Figure 3.3), and therefore end point assays in this thesis did not use  $\geq 600 \mu\text{M}$  of added zinc. This data in combination with the flow cytometry data showing the distribution of fluorescence of a population with increasing  $\text{ZnSO}_4$  concentration (Figure 5.20 and Figure 5.21) suggests the mechanisms behind this correlation. As zinc concentration is increased, the distribution of *zntA* gene expression is shifted in the population, reducing low gene expression and increasing high gene expression, which on average increases the overall end point fluorescence. This flow cytometry data experimentally demonstrated heterogeneous gene expression of *zntA*, which is consistent with an observed population basal expression of *zntA*, even during zinc depleted conditions.

This is the first time, to my knowledge, that temporal luminescence assays have been conducted on  $P_{zntA}$  with the variety of zinc concentrations tested in this thesis, and in zinc depleted NH media. Takahashi *et al.*, (2015) demonstrated temporal luminescence of  $P_{zntA}$  in LB with only the addition of 0, 12.5 and 100  $\mu\text{M}$   $\text{ZnSO}_4$ , and only up to 60 min. Temporal luminescence assay data in both LB and NH media demonstrated that with increased  $\text{ZnSO}_4$ , not only does the maximum level of luminescence increase, but the rate of expression also increases (Figure 3.23 and Figure 3.28). Further, the rate of expression of induced  $P_{zntA}$  (400  $\mu\text{M}$   $\text{ZnSO}_4$ ) in NH is significantly higher than the induced control  $P_{trc}$  (100  $\mu\text{M}$  IPTG), and induced  $P_{zntA}$  and  $P_{trc}$  showed no significant difference in rate of expression when grown in LB (Figure 3.27 and Figure 3.31). These temporal luminescence assays showed that induced  $P_{zntA}$  demonstrates similar desirable characteristics to the IPTG inducible  $P_{trc}$ .

RT-qPCR was employed to directly quantify *zntA* mRNA as a measure of *zntA* gene expression. The RT-qPCR data concluded two points; firstly that the correlation of zinc concentration and gene regulation is contributed to by transcription gene regulation (Figure 3.20); secondly, the mRFP1 fluorescent protein (used throughout this thesis in end point gene reporter assays) showed comparative results to the RT-qPCR data (Figure 3.22). This validated the use of mRFP1 as a reporter element to record transcriptional gene regulation of *zntA*.

Increasing the spacer between the 22 bp (11-11) palindromic binding site of ZntR by either 1 or 2 bp [ $P_{zntA(-25^{24}insT)}$  and  $P_{zntA(-25^{24}insTG)}$ ], it completely removed ZntR regulation of  $P_{zntA}$  and expression from  $P_{zntA}$  occurred at a level comparable to that observed in the *wt*  $P_{zntA}$  in zinc depleted media (Figure 3.15). These mutations increased the spacer between the -10 and -35 elements to 21/22 bp, which is not optimal ( $17 \pm 1$ ) (Harley and Reynolds, 1987), reducing successful  $\sigma^{70}$  interactions with the -10 and -35 element.

A study by Parkhill and Brown, (1990) increased the spacer region of  $P_{mer}$ , which is regulated by MerR (a member of the same family of regulators as ZntR). The  $P_{mer}$  spacer was increased by either 1 or 2 bp, and an end point assay concluded that the increased spacer removed MerR induction; like that observed with the  $P_{zntA}(-25^{+24}insT)$  and  $P_{zntA}(-25^{+24}insTG)$ . The data presented in this thesis, in combination with the data by Brocklehurst *et al.*, (1999) on decreasing the spacer region shows the importance of the spacer region in the ZntR regulation of  $P_{zntA}$ . This 20 bp spacer region is essential for correct ZntR regulation of  $P_{zntA}$ ; a mutation which either increases or decreases the spacer by a single base pair will significantly alter the highly controlled ZntR regulation of  $P_{zntA}$ . This combination of data on increasing and decreasing the spacer in  $P_{zntA}$  agrees with data present by Parkhill and Brown, (1990) on  $P_{mer}$ . Parkhill and Brown, (1990) demonstrated that by increasing the spacer in  $P_{mer}$  by 1 or 2 bp it prevented MerR induction and decreasing the spacer by 1 or 2 bp showed constant unregulated expression, true for ZntR regulation of  $P_{zntA}$ . Both MerR and ZntR are members of the same family of regulators (MerR) it shows the importance of this non-optimal 20 bp spacer, and its key role in regulation.

Bryant *et al.*, (2014) demonstrated the importance of chromosomal gene location with regards to gene regulation, showing that a 300-fold difference in gene regulation can be observed depending on chromosomal location. By conducting a promoter activity assay in the native *zntA* chromosomal location provided a more accurate view on *zntA* regulation. End point assays using *E. coli* MG1655 *zntA::rfp::kan<sup>R</sup>* showed the same relationship between zinc and promoter activity that was observed in plasmid-based end point fluorescence assays in this thesis; a significant increase in promoter activity with the addition of 400  $\mu$ M  $ZnSO_4$ , and a significant reduction with the addition of 20  $\mu$ M TPEN compared to LB only (Figure 5.3). Firstly, this shows that the plasmid-based assays used in this thesis are a reliable way of measuring *zntA* regulation, and that the

relationship between zinc and gene regulation is not purely a consequence of plasmid-based assay; it is observed natively in *E. coli*.

All the data collected in this thesis gives promise for the ZntR regulated  $P_{zntA}$  to be used as a novel, tuneable, zinc inducible promoter.  $P_{zntA}$  shows comparable, if not better fundamentals of an inducible promoter system than the IPTG inducible  $P_{trc}$ .

### 7.1.3. Zur Regulated Promoters for Novel Inducible Promoters

Six Zur regulated promoters were analysed for their potential as a novel Zur regulated promoter ( $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$ ,  $P_{c1265}$ ). Results indicated that the most useful and most practical method to utilise Zur regulated promoters as an inducible system is to use TPEN as an inducer; TPEN chelates zinc effectively reducing free zinc concentration, which upregulates expression from Zur repressed promoters. All six Zur regulated promoters showed a significant increase in promoter activity when induced with 20  $\mu$ M TPEN compared to LB only (Figure 4.4); however,  $P_{ykgM}$  and  $P_{c1265}$  show the most promise to be used as a highly inducible TPEN promoter.  $P_{ykgM}$  showed a 270-fold increase in promoter activity when induced with TPEN, and  $P_{c1265}$  showed 174-fold increase; this compared to the 63-fold increase of the IPTG inducible  $P_{trc}$  is substantially larger (Table 4.2).

This data agrees with the transcriptional microarray data by Sigdel, Easton and Crowder, (2006) and Hensley *et al.*, (2012) which found *ykgM* showed the highest fold increase when comparing growth in zinc repleted media to zinc depleted media. Microarray data conducted by Clayton, (2012) showed that in a zinc depleted state compared to a normal state, the strongest upregulated gene of *E. coli* CFT073 was *ykgM*, closely followed by *c1265*. However, in contrast Sigdel, Easton and Crowder, (2006) found *zinT* (*yodA*) to be the second highest regulated in *E. coli* K-12 when

induced with TPEN. Additionally, (Clayton, 2012) found *zinT* to be in the top five upregulated gene (zinc depleted to zinc replete) in both *E. coli* K-12 and *E. coli* CFT073. Results in this thesis found *zinT* to be one of the least upregulated gene in a zinc depleted state. This disparity could be due to the method of analysis; oligonucleotide transcriptional microarray and plasmid-based end point fluorescence assays. The promoter sequence of  $P_{zinT}$  for plasmid-based assay in this thesis cloned a 92 bp sequence upstream of the +1 start site. A study by Puškárová *et al.*, (2002), which used a 1200 bp promoter sequence for plasmid-based assays, demonstrated the involvement of Fur and SoxS in the regulation of *zinT* (*yodA*), however the binding location of these transcription factors was not established. Moreover, Puškárová *et al.*, (2002) also demonstrated that ppGpp is needed for *zinT* induction, but induction cannot be achieved solely by ppGpp. Considering that ZinT was originally discovered as a cadmium-induced protein (Ferianc, Farewell and Nyström, 1998), and later shown its involvement in zinc acquisition (Kershaw, Brown and Hobman, 2007) and has been shown to be regulated by SoxS, Fur and Zur (Puškárová *et al.*, 2002; Panina, Mironov and Gelfand, 2003). It is clear a lot of transcription factors are involved in the regulation of *zinT*, and it is possible the shorter 92 bp promoter sequence used for cloning in the reporter plasmids for this thesis missed vital TF binding locations, and thus lead to considerably different data to that previously reports on *zinT* regulation. Microarray data measured hybridisation of cDNA, transcribed from the chromosome; this would show gene transcription involving any and all transcription factors. It is also possible that the zinc concentration range used in the end point fluorescence assays in this thesis were not in the range to cause a measurable variation in  $P_{zinT}$  regulation; it is possible that the 115 nM zinc concentration in NH media (as determine by ICP-MS) is high enough that it causes maximum repression of  $P_{zinT}$ .



Temporal luminescence data of  $P_{c1265}$  showed the highest rate of expression and the highest maximum luminescence, followed by  $P_{ykgM}$ . All six promoters showed a higher rate of expression when induced with 20  $\mu$ M TPEN compared to LB only, which showed stabilised expression level throughout the 16 hr assay (Figure 4.9).

This high rate of expression, and stabilized expression level over 16 hr gives great potential for Zur regulated promoters to be used as novel TPEN inducible promoters. The six Zur regulated promoters tested all showed different levels of expression when induction with 20  $\mu$ M TPEN, which could be advantageous; one could induce different expression levels of genes within an operon by using different Zur regulated promoters for each gene, however the basal level of expression is also different for all Zur regulated promoters. If one is merely looking for a highly inducible promoter,  $P_{c1265}$  and  $P_{ykgM}$  are the best option, due to their low basal expression, and significantly higher induced expression. Finally, the induction with 20  $\mu$ M TPEN, as used in this thesis, does not show any toxic side effects on *E. coli* growth.

#### 7.1.4. Heterogeneous *zntA* Gene Expression

mRFP1 was fused chromosomally to the 3' end of *zntA* (*E. coli* MG1655 *zntA::rfp::kan<sup>R</sup>*). An initial end point fluorescence assay of the chromosomal fusion showed a significant increase in promoter activity with the addition of 400  $\mu$ M ZnSO<sub>4</sub>, and a significant decrease with the addition of 20  $\mu$ M TPEN, compared to LB only. This data agrees with the plasmid based assays in this thesis, as well as plasmid based assays on  $P_{zntA}$  promoter activity in response to zinc by Brocklehurst *et al.*, (1999), and Outten, Outten and O'Halloran, (1999).

Flow cytometry analysis of *E. coli zntA::rfp* suggested heterogenous gene expression of *zntA*. Firstly, when grown in a medium (both LB and NH) with increasing concentrations

of added  $\text{ZnSO}_4$ , overall fluorescence of the population increased in line with increasing  $\text{ZnSO}_4$  (Figure 5.9 and Figure 5.14). This justifies the correlation observed between increasing  $\text{ZnSO}_4$  concentration and end point fluorescence of  $P_{zntA}$  from plasmid-based assays (Figure 3.8 and Figure 3.11). This correlation is not just an observed consequence of plasmid-based assays, but also observed in the native chromosomal location. It has been shown that the chromosomal gene location in *E. coli* K-12 can alter gene regulation by up to 300-fold (Bryant *et al.*, 2014) This shows that even chromosomally there is a strong correlation between  $\text{ZnSO}_4$  and *zntA* gene regulation.

Population fluorescence of *E. coli* MG1655 *zntA::rfp* grown in LB only or LB + 600  $\mu\text{M}$   $\text{ZnSO}_4$  suggests *zntA* heterogenous gene expression (Figure 5.9). The density plot shows a negative skew in both conditions, with an extended left tail; this is supported by Q-Q plots (Figure 5.11). This negatively skewed data suggests heterogenous gene expression, with a higher proportion of the population expressing *zntA* at high level, and a lower proportion of the population expressing *zntA* at low level. This can also be observed with *E. coli zntA::rfp* in NH and NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$  (Figure 5.14 and Figure 5.16). These density plots show that with increased  $\text{ZnSO}_4$  concentrations, the negative skew is increased, suggesting that there is a higher distribution of *zntA* heterogenous gene expression with higher zinc concentrations. However, a major issue with this data is the large overlapping background population fluorescence (*E. coli wt* in media only), a measure of autofluorescence of cells, which makes distinguishing foreground and background data problematic.

#### 7.1.5. Bayesian Background Subtraction Approach to Flow Cytometry Data

A Bayesian background subtraction approach was used to further analyse the flow cytometry data; this approach subtracts the background data from the foreground data by constructing a posterior distribution of the true foreground data (given the

observed foreground and background data), knowing that the true foreground cannot be negative. This approach returns a more accurate estimation of the true foreground. This Bayesian subtraction approach was a development from the Bayesian analysis used with microarray data (Kooperberg *et al.*, 2002). To my knowledge, a Bayesian approach to estimate the true foreground fluorescence of flow cytometry has not previously been used. A study by Galbusera *et al.*, (2019) developed an R code based on Gaussian and uniform distribution. This R code estimates the fluorescence mean and variation. This is another approach to understand the true foreground fluorescence from flow cytometry data. This approach would not be useful for stochastics, as it does not estimate the distribution for true fluorescence, but rather the mean estimation of the population.

An initial proof of principle test was conducted on a randomly generated set of data and compared to a simple background subtraction approach ( $f - X_b$ ) (Figure 5.19). This proof of principle test showed the power of the Bayesian background subtraction approach, showing improvement over a simple subtraction approach. The Bayesian approach returned values all greater than zero and showed a more accurate estimation of the true foreground data.

The Bayesian background subtraction analysis of *E. coli zntA::rfp* in LB (0, 400 or 600  $\mu\text{M}$   $\text{ZnSO}_4$ ) returned a distribution of fluorescence dissimilar to the raw flow cytometry fluorescence (Figure 5.20). All three conditions (0, 400 and 600  $\mu\text{M}$   $\text{ZnSO}_4$ ) showed evidence of heterogeneous *zntA* gene expression; the distribution of fluorescence showed a large proportion of cells expressing *zntA* at low levels, and a smaller proportion of cells expression *zntA* at high levels. Moreover, as  $\text{ZnSO}_4$  concentration increased, it shifted the distribution of fluorescence, reducing low level and increasing high level expression of *zntA*. This shift in the distribution in fluorescence of the

population can also be observed when *E. coli zntA::rfp* was grown in NH (0, 25, and 400  $\mu$ M ZnSO<sub>4</sub>) (Figure 5.21).

The distribution of fluorescence of *E. coli zntA::rfp* in both NH and LB (in all conditions) showed a similar pattern of distribution to the stochastic simulation of ZntA protein abundance model by Takahashi *et al.*, (2015) [Figure 8.c]. A large distribution in low protein abundance, and small distribution in higher protein abundance. This contrasted with the ZnuABC protein abundance model, which showed a distribution more aligned with normal Gaussian distribution (Takahashi *et al.*, (2015) Figure 8.b). The Bayesian background subtract data of experimental flow cytometry data is reinforced by the stochastic simulation of the ZntA protein abundance model, which thus further validated the hypothesis by Takahashi *et al.*, (2015): in both normal and zinc stress, *E. coli* mount *zntA* heterogeneous gene expression.

#### 7.1.6. *In silico* Approach to Determine the C1265-7 Phenotype

The novel Zur regulated C1265-7 operon was discovered by Clayton, (2012) through transcriptomic microarray analysis of Zur regulated genes in UPEC *E. coli* CFT073, and bioinformatics showed its chromosomal presence only in UPEC strains. Bioinformatic analysis by Clayton, (2012) did not find major homologues to C1265-7, and thus the phenotype and function of C1265-7 not yet determined. Bioinformatic analysis of C1265-7 in this thesis provided more data on the function but has not yet fully determine the true function of C1265-7.

C1265 showed 42% amino acid sequence similarity to the ShuA (TonB dependent receptor) of *S. dysenteriae*, this homologous protein was also found by Clayton, (2012). More interestingly, Swiss-Model was used to predict the structure of C1265 (Figure 6.2); the model showed high QMEAN in the periplasm bound  $\beta$ -barrel domain, and low

QMEAN score in the extracellular domain, particularly in the loops. These loops have been shown to be associated with ligand binding in other TonB proteins (Buchanan *et al.*, 1999); this suggests the function of C1265 is a TonB-dependent receptor, which binds to an alternative ligand than ShuA, potentially zinc. To further argue a role in zinc acquisition, C1265 showed low similarity to ZnuD (a TonB-dependent receptor) of *N. meningitidis* (Calmettes *et al.*, 2015); which showed moderate homology to the zinc binding residues of ZnuD. Interestingly, ZnuD has been misannotated as a TonB-dependent haem transporter (Kumar, Sannigrahi and Tzeng, 2012), partially due to its homology to ShuA, but has been shown to be involved with zinc acquisition under the regulation of Zur (Stork *et al.*, 2010). Moreover, ZnuD has been shown as a vital component for survival against host nutritional immunity in *N. meningitidis* (Stork *et al.*, 2013), knowing that C1265 is only found chromosomally in UPEC strains it is sensible to assume that C1265 may play a role in survival during nutritional immunity. I hypothesise that the role of C1265 is like that of ShuA, a TonB-dependent receptor, but binds to zinc, which aids in zinc acquisition during zinc depleted conditions such as nutritional immunity.

C1266 showed a low 20% amino acid sequence homology to RrgC, an IgG-like domain which is involved in host attachment and thus virulence of *S. pneumoniae* (Shaik *et al.*, 2014). RrgC folds into three domains (D1, D2 and D3), interestingly when comparing the amino acid sequence homology, homology is only found in the D2 domain, the domain involved in bacterial adhesion. This suggests that C1266 may play a role in virulence and bacterial adhesion; knowing that C1265-7 is only found chromosomally in UPEC strain (Clayton, 2012), it is further possible to hypothesise that C1266 attachment is specific to the urinary tract.

Finally, C1267 showed low amino acid sequence identity (22.2%) to YjiA, a GTP-binding protein (COG0523 member) of *E. coli* K-12 (Figure 6.4). YjiA has been shown to interact with  $\text{Zn}^{2+}$  (Sydor *et al.*, 2013); C1267 showed functional similar residues to YjiA which interact with  $\text{Zn}^{2+}$  (Figure 6.4). Furthermore, C1267 showed 100% identity to the GTP-binding domain of YjiA. Unfortunately, C1267 does not show homology to the conserved putative metal binding domain found in all COG0523 member. Further evidence that C1267 may be a COG0523 member is that COG0523 proteins have been shown to be regulated by Zur (Gaballa and Helmann, 1998; Maciąg *et al.*, 2007; Haas *et al.*, 2009)

*In silico* analysis of C1265-7 suggests that it plays a role in zinc acquisition, and potentially a role in virulence. This gene operon may play a vital role in zinc acquisition when infecting a host, specifically when a host employs nutritional immunity as a defence mechanism. In the case of Enteroaggregative *E. coli* (EAEC), host nutritional immunity, which reduces available zinc, leads to increased stress of EAEC and increases virulence factors, worsening diarrhoea in the host (Bolick *et al.*, 2015).

#### 7.1.7. *In vitro* Approach to Determine the C1265-7 Phenotype

As previously mentioned,  $P_{c1265}$  is highly induced with added TPEN, or when grown in a zinc depleted media. Previous data suggested that  $P_{znuC}$ , regulated by Zur shows tuneable regulation in response to zinc (Outten and O'Halloran, 2001; Takahashi *et al.*, 2015), however as  $P_{c1265}$  is a novel Zur regulated gene, the zinc response has not been evaluated, but one could presume that a zinc response regulation may play a role in  $P_{c1265}$  regulation. Data in this thesis shows that  $P_{c1265}$  does not show any tuneable regulation between zinc concentrations of 3.240  $\mu\text{M}$  to 400  $\mu\text{M}$  in NH media (Figure 4.11). Moreover, repression of  $P_{c1265}$  can be observed when grown in NH, which has a zinc concentration of 115 nM. It is possible that a zinc concentration between 115 nM

to 3.24  $\mu\text{M}$  or less than 115 nM zinc is needed to see a tuneable regulation, but data presented in this thesis suggests  $P_{c1265}$  may not be tuneable in response to zinc concentration.

Interestingly, data by Outten and O'Halloran, (2001) showed an overlap of zinc concentration which demonstrated tuneable regulation in both Zur and ZntR: as ZntR regulation increases, Zur regulation decreases (Outten and O'Halloran, 2001) [Fig. 4]. Data in this thesis demonstrated a correlation between ZntR regulation of  $P_{zntA}$  with zinc concentration in the range of 115 nM to 400  $\mu\text{M}$  zinc. Therefore, one would expect to see a  $P_{c1265}$  Zur regulated response to zinc concentrations within the same tested range. It is possible that the promoter sequence cloned into the reporter plasmid is too short and removed potential involvement of additional and unknown transcription factors. There is a need for more fundamental understanding of the  $P_{c1265}$  promoter.

Standard growth curve assays of *E. coli* CFT073 (*wt*,  $\Delta c1265$ -7,  $\Delta znuABC$ , and  $\Delta c1265$ -7 $\Delta znuABC$ ) suggested that in the absence of C1265-7 bacterial growth was slightly disadvantaged, with *E. coli* CFT073 *wt* showing a significantly higher growth rate than *E. coli* CFT073  $\Delta c1265$ -7 (Figure 6.7). This growth curve also showed the importance of *znuABC* in *E. coli* CFT073 growth. A study by Gunasekera, Herre and Crowder, (2009) demonstrated the importance of ZnuABC in *E. coli* CFT073 growth; that *E. coli*  $\Delta znuB$  did not grow in minimal media unless supplemented with excess zinc. Further microtiter growth curve assays did not further help understand the phenotype of C1265-7, partially due to the issue with conducting growth curve assays in a microtiter plate.

To overcome the issues with both standard and microtiter growth assays, *c1265*-7 was cloned into an expression plasmid (pJI102), and along with an empty vector (pJI100A) were transformed into *E. coli* MG1655 (*wt*,  $\Delta znuA$ ,  $\Delta znuCB$ ) and a growth assay

conducted using a BacTiter-Glo cell viability assay (Figure 6.14). Firstly, this assay concluded that deletion of either  $\Delta znuA$  or  $\Delta znuCB$  reduces the growth rate of *E. coli* MG1655 when transformed with the empty vector. This agrees with Gunasekera, Herre and Crowder, (2009) who showed *E. coli* K-12  $\Delta znuB$  showed a reduced growth rate compared to the *wt* strain. Interestingly Gunasekera, Herre and Crowder, (2009) showed that the deletion of  $\Delta znuB$  reduced the growth rate of *E. coli* K-12, but completely stunted growth of *E. coli* CFT073 when grown in minimal media; suggesting the more vital role of ZnuABC in *E. coli* CFT073. When *E. coli* MG1655  $\Delta znuA$  was transformed with the *c1265-7* expression plasmid, and grown in NH, the growth rate was significantly higher than when transformed with the empty vector. The same is not true for *E. coli* MG1655  $\Delta znuCB$ ; the growth rate with or without the *c1265-7* expression plasmid was similar. This suggests that in the absence of ZnuA, C1265-7 may play a replacement role for ZnuA in zinc acquisition. However, on the contrary, even though there is a statistically significant difference in the growth rate, the difference observed would not be strong enough to justify a phenotype of C1265-7.

Unfortunately, the *in vitro* experimental results to determine the phenotype of C1265-7 have not helped to further understand the role of C1265-7 in zinc acquisition. An infection model may be needed to see the true phenotype of C1265-7. Sabri, Houle and Dozois, (2009) demonstrated that deletions of  $\Delta znuABC$  and  $\Delta zupT$  (zinc import genes) in *E. coli* CFT073 significantly reduced the bacterial count in both kidney and bladder infection model. A true phenotype of C1265-7 may only be observed in an infection model. This is further backed by the *in silico* analysis, which suggested C1265 is a zinc TonB-dependent receptor which may play an important role during zinc nutritional immunity, and the C1266 homology to RrgC, specifically the homology in the adhesion domain (D2).



## 7.2. Conclusion

Data presented in Chapter 3 showed the potential for  $P_{zntA}$  to be used as a novel, tuneable, zinc inducible promoter system.  $P_{zntA}$  showed higher levels of induced expression, with a correlation between promoter activity and zinc concentrations, as well as the comparison to  $P_{trc}$  which showed  $P_{zntA}$  demonstrated desirable characteristics of an inducible promoter. The potential of  $P_{zntA}$  is only let down by its undesired higher basal level of expression, which is partially due to the heterogeneity of *zntA* gene expression.

Six Zur regulated promoter ( $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$ ,  $P_{c1265}$ ) showed promise to be used as an “off/on” TPEN inducible promoter system. These six promoters all showed a significant variation in promoter activity with 20  $\mu$ M TPEN induction:  $P_{c1265} > P_{ykgM} > P_{znuA} > P_{znuCB} > P_{pliG} > P_{zinT}$ .  $P_{ykgM}$  and  $P_{c1265}$  indicated the most promise, showing more desirable promoter characteristics than  $P_{trc}$ : lower basal expression, and higher induced expression.

A Bayesian background subtracted approach to flow cytometry data provided strong experimental evidence that *zntA* gene expression is heterogeneous. This is further supported by the modelled data by Takahashi *et al.*, (2015). The experimental data also showed a relationship between zinc concentration and *zntA* expression: in higher zinc concentrations *zntA* expression is shifted; high *zntA* expression is increased while low *zntA* expression is decreased. The Bayesian background subtraction approach provided a new and powerful approach to estimate the true foreground from flow cytometry data.

*In vitro* analysis of *c1265-7* further proved Zur regulation of the operon but did not suggest any tuneable promoter activity between 3.240  $\mu$ M to 400  $\mu$ M zinc. Growth

curve assays did not provide any strong evidence for the observed phenotype or role of C1265-7 in *E. coli* CFT073. *In silico* analysis showed strong evidence that C1265 is a TonB-dependent receptor which is likely to translocate zinc. C1266 plays a role in host adhesion or a zinc chaperone, and C1267 may be a COG0523 (a metallochaperone protein). It is likely that an observed phenotype of C1265-7 can only be observed in an infection model.

The zinc homeostasis system of *E. coli* is extremely complex, the data in this thesis has enhanced our knowledge of ZntR and Zur regulation, experimentally proved *zntA* heterogenous gene regulation, and utilized the zinc homeostasis system for biotechnological gain.

### 7.3. Future Direction of Work

To advance the development of a ZntR and Zur based inducible promoter system, 'real world' experiments should be conducted to gain further insight in the application of these promoters in biotechnology settings. All experiments in this thesis were conducted using reporter elements, in data driven assays. This is more applicable to how an inducible promoter would be used in a research background. However, it is not very indicative of industrial biotechnological scenario. A real-world application using these promoter systems, in combination with a biotechnological recombinant gene should be used and grown in industrial or biotechnological environments, such as turbidostat for continuous product production.

Bayesian background subtraction of flow cytometry data gave strong evidence that *zntA* gene expression is heterogenous. To further understand the heterogeneity of *zntA*, and the role of Bayesian background subtraction approach to data analysis, flow cytometry should be conducted with a known heterogenous and homogenous fluorescently labelled gene. This would first help to understand the role of the Bayesian approach to the data and provide a direct comparison of gene stochastics. Flow cytometry only shows a snap shot of gene stochastics, and in the data presented in this thesis, flow cytometry was conducted after 16 hr of growth. Conducting flow cytometry analysis to look at *zntA* stochastics during all growth phases would help to develop our knowledge of how *zntA* regulation adapts over time. A further improvement on this would be to use microfluidics to measure *zntA* regulation over time and cell lineages.

To further understand the novel C1265-7 from UPEC strains, it is essential for an infection models to be carried out. Knowing that C1266 shows potential to be involved with host attachment, C1265 show potential as a zinc TonB-dependent receptor, and

that host immune system employs nutritional immunity as a defence mechanism, it is only logical to assume that a phenotype may only be seen during infection.

## Chapter 8

# References

- Abrantes, M. C., de Fátima Lopes, M. and Kok, J. (2011) 'Impact of manganese, copper and zinc ions on the transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583', *PLoS ONE*.
- Abrantes, M. C., Kok, J. and Lopes, M. de F. S. (2014) 'Enterococcus faecalis zinc-responsive proteins mediate bacterial defence against zinc overload, lysozyme and oxidative stress', *Microbiology (United Kingdom)*.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) 'Gapped BLAST and PSI-BLAST: A new generation of protein database search programs', *Nucleic Acids Research*, 25(15), pp. 3389–3402.
- Ammendola, S., Pasquali, P., Pistoia, C., Petrucci, P., Petrarca, P., Rotilio, G. and Battistoni, A. (2007) 'High-affinity Zn<sup>2+</sup> uptake system ZnuABC is required for bacterial zinc homeostasis in intracellular environments and contributes to the virulence of *Salmonella enterica*', *Infection and Immunity*.
- Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2006) 'Zinc through the three domains of life', *Journal of Proteome Research*, 5(11), pp. 3173–3178.
- Angerer, A. and Braun, V. (1998) 'Iron regulates transcription of the *Escherichia coli* ferric citrate transport genes directly and through the transcription initiation proteins', *Archives of Microbiology*, 169, pp. 483–490.
- Ansari, A., Cheal, M. and O'Halloran, T. V. (1992) 'Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR', *Letters to Nature*, 355(January), pp. 87–89.
- Ausubel, F. M., Brent, R., Kingston, R. E., et al. (2003) *Current Protocols in Molecular Biology*.
- Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S. and Redwan, E. M. (2014) 'Cell factories for insulin production', *Microbial Cell Factories*, 13(1).
- Barne, K. A., Bown, J. A., Busby, S. J. W. and Minchin, S. D. (1997) 'Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the "extended-10" motif at promoters.', *The EMBO journal*, 16(13), pp. 4034–4040.
- Beard, S. J., Hashim, R., Wu, G., et al. (2000) 'Evidence for the transport of zinc(II) ions via the Pit inorganic phosphate transport system in *Escherichia coli*', *FEMS Microbiology Letters*, 184(2), pp. 231–235.
- Benkert, P., Biasini, M. and Schwede, T. (2011) 'Toward the estimation of the absolute quality of individual protein structure models', *Bioinformatics*, 27, pp. 343–350.
- Bertani, G. (1951) 'Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*.', *Journal of bacteriology*, 62(3), pp. 293–300.
- Bervoets, I. and Charlier, D. (2019) 'Diversity, versatility and complexity of bacterial gene regulation mechanisms: Opportunities and drawbacks for applications in synthetic biology', *FEMS Microbiology Reviews*, 43, pp. 304–339.
- Bingle, L. E. H., Constantinidou, C., Shaw, R. K., et al. (2014) 'Microarray analysis of the Ler regulon in enteropathogenic and enterohaemorrhagic *Escherichia coli* strains', *PLoS ONE*.
- Bio-Rad (2020) *Chelex 100 Resin*. Available at: <https://www.bio-rad.com/en-uk/product/chelex-100-resin?ID=6448ab3e-b96a-4162-9124-7b7d2330288e> (Accessed: 25 March 2020).
- Blaby-Haas, C. E., Flood, J. A., Crécy-Lagard, V. De and Zamble, D. B. (2012) 'YeiR: A metal-binding GTPase from *Escherichia coli* involved in metal homeostasis', in *Metallomics*.
- Blattner, F. R., Plunkett, G., Bloch, C. A., et al. (1997) 'The complete genome sequence of *Escherichia coli* K-12', *Science*, pp. 1453–1462.
- de Boer, H. A., Comstock, L. J. and Vasser, M. (1983) 'The tac promoter: a functional hybrid derived from the trp and lac promoters.', *Proceedings of the National Academy of Sciences of*

*the United States of America*, 80(1), pp. 21–25.

Bolick, D. T., Kolling, G. L., Moore, J. H., *et al.* (2015) 'Zinc deficiency alters host response and pathogen virulence in a mouse model of enteroaggregative escherichia coli-induced diarrhea', *Gut Microbes*, 5(5), pp. 618–627.

Branen, J. K. and Davidson, P. M. (2004) 'Enhancement of nisin, lysozyme, and monolaurin antimicrobial activities by ethylenediaminetetraacetic acid and lactoferrin', *International Journal of Food Microbiology*.

Brinkmann, H., Göker, M., Koblížek, M., Wagner-Döbler, I. and Petersen, J. (2018) 'Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae', *ISME Journal*, 12(8), pp. 1994–2010.

Brocklehurst, K. R., Hobman, J. L., Lawley, B., Blank, L., Marshall, S. J., Brown, N. L. and Morby, A. P. (1999) 'ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli', *Molecular Microbiology*, 31(3), pp. 893–902.

Brodl, E., Winkler, A. and Macheroux, P. (2018) 'Molecular Mechanisms of Bacterial Bioluminescence', *Computational and Structural Biotechnology Journal*, 16, pp. 551–564.

Brosiuss, J., Erfle, M. and Storella, J. (1985) 'Spacing of the -10 and -35 Regions in the tac Promoter', *The Journal of biological chemistry*, 260(6), pp. 3539–3541.

Brown, N. L., Stoyanov, J. V., Kidd, S. P. and Hobman, J. L. (2003) 'The MerR family of transcriptional regulators', *FEMS Microbiology Reviews*, 27, pp. 145–163.

Browning, D. F. and Busby, S. J. W. (2004) 'The regulation of bacterial transcription initiation', *Nature Reviews Microbiology*, 2(1), pp. 57–65.

Browning, D. F. and Busby, S. J. W. (2016) 'Local and global regulation of transcription initiation in bacteria', *Nature Reviews Microbiology*. Nature Publishing Group, 14(10), pp. 638–650.

Bryant, J. A., Sellars, L. E., Busby, S. J. W. and Lee, D. J. (2014) 'Chromosome position effects on gene expression in Escherichia coli K-12', *Nucleic Acids Research*, 42(18), pp. 11383–92.

Buchanan, S. K., Smith, B. S., Venkatramani, L., *et al.* (1999) 'Crystal structure of the outer membrane active transporter FepA from Escherichia coli', *Nature Structural Biology*, 6(1), pp. 56–63.

Burgess, R. R. (1969) 'Separation and characterization of the subunits of ribonucleic acid polymerase.', *Journal of Biological Chemistry*, 244(22), pp. 6168–76.

Burr, T., Mitchell, J., Kolb, A., Minchin, S. and Busby, S. (2000) 'DNA sequence elements located immediately upstream of the -10 hexamer in Escherichia coli promoters: a systematic study.', *Nucleic acids research*, 28(9), pp. 1864–1870.

Burton, N. A., Johnson, M. D., Antczak, P., Robinson, A. and Lund, P. A. (2010) 'Novel Aspects of the Acid Response Network of E. coli K-12 Are Revealed by a Study of Transcriptional Dynamics', *Journal of Molecular Biology*. Elsevier Ltd, 401(5), pp. 726–742.

Busby, S. and Ebright, R. H. (1999) 'Transcription activation by catabolite activator protein (CAP)', *Journal of Molecular Biology*, 293(2), pp. 199–213.

Bustin, S. A., Benes, V., Garson, J. A., *et al.* (2009) 'The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments', *Clinical Chemistry*, 55(4), pp. 611–622.

Butof, L., Schmidt-vogler, C., Herzberg, M. and Grolle, C. (2017) 'The components of the unique Zur regulon of Cupriavidus metallidurans mediate cytoplasmic zinc handling', *JOURNAL OF BACTERIOLOGY*, 99(21).

Calmettes, C., Ing, C., Buckwalter, C. M., *et al.* (2015) 'The molecular mechanism of Zinc acquisition by the neisserial outer-membrane transporter ZnuD', *Nature Communications*. Nature Publishing Group, 6(1), pp. 1–11.

Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A. and Tsien,

- R. Y. (2002) 'A monomeric red fluorescent protein', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 99(12), pp. 7877–7882.
- Campoy, S., Jara, M., Busquets, N., Pérez De Rozas, A. M., Badiola, I. and Barbé, J. (2002) 'Role of the high-affinity zinc uptake znuABC system in Salmonella enterica serovar Typhimurium virulence', *Infection and Immunity*.
- Casali, N. and Preston, A. (2003) *Methods in Molecular Biology: E. coli Plasmid Vectors, Methods and Applications, E. coli Plasmid Vectors*.
- Chandra, B. R., Yogavel, M. and Sharma, A. (2007) 'Structural Analysis of ABC-family Periplasmic Zinc Binding Protein Provides New Insights Into Mechanism of Ligand Uptake and Release', *Journal of Molecular Biology*, 367(4), pp. 970–982.
- Chang, C. C., Liao, W. F. and Huang, P. C. (1998) 'Cysteine contributions to metal binding preference for Zn/Cd in the  $\beta$ -domain of metallothionein', *Protein Engineering*, 11(1), pp. 41–46.
- Changela, A., Chen, K., Xue, Y., Holschen, J., Outten, C. E., O'Halloran, T. V. and Mondragón, A. (2003) 'Molecular Basis of Metal-Ion Selectivity and Zeptomolar Sensitivity by CueR', *Science*, 301(5638), pp. 1383–1387.
- Chaoprasid, P., Dokpikul, T., Johnrod, J., Sirirakphaisarn, S., Nookabkaew, S., Sukchawalit, R. and Mongkolsuk, S. (2016) 'Agrobacterium tumefaciens Zur regulates the high-affinity zinc uptake system TroCBA and the putative metal chaperone YciC, along with ZinT and ZnuABC, for survival under zinc-limiting conditions', *Applied and Environmental Microbiology*.
- Cho, Y.-E., Lomeda, R.-A. R., Ryu, S.-H., Lee, J.-H., Beattie, J. H. and Kwun, I.-S. (2007) 'Cellular Zn depletion by metal ion chelators (TPEN, DTPA and chelex resin) and its application to osteoblastic MC3T3-E1 cells', *Nutrition Research and Practice*. Korean Society of Community Nutrition and the Korean Nutrition Society, 1(1), p. 29.
- Clayton, S. R. (2012) *The Zur (Zinc Uptake Regulator) Regulon of Pathogenic and Non-Pathogenic Escherichia coli*. Univeristy of Nottingham.
- Cobessi, D., Meksem, A. and Brillet, K. (2010) 'Structure of the heme/hemoglobin outer membrane receptor ShuA from shigella dysenteriae: heme binding by an induced fit mechanism', *Proteins: Structure, Function and Bioinformatics*, 78(2), pp. 286–294.
- Colaço, H. G., Santo, P. E., Matias, P. M., Bandejas, T. M. and Vicente, J. B. (2016) 'Roles of Escherichia coli ZinT in cobalt, mercury and cadmium resistance and structural insights into the metal binding mechanism', *Metallomics*. Royal Society of Chemistry, 8(3), pp. 327–336.
- Coleman, J. E. (1998) 'Zinc enzymes.', *Current opinion in chemical biology*, 2(2), pp. 222–234.
- Corbin, B. D., Seeley, E. H., Raab, A., et al. (2008) 'Metal chelation and inhibition of bacterial growth in tissue abscesses', *Science*.
- Cournac, A. and Plumbridge, J. (2013) 'DNA Looping in Prokaryotes: Experimental and theoretical approaches', *Journal of Bacteriology*. American Society for Microbiology Journals, 195(6), pp. 1109–1119.
- Crick, F. H. C. (1970) 'Central Dogma of Molecular Biology', *Nature*, 227(5258), pp. 561–563.
- Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E. (2004) 'WebLogo: A sequence logo generator', *Genome Research*.
- Dalet, K., Guin, E., Cenatiempo, Y., Cossart, P. and Héchard, Y. (1999) 'Characterisation of a new operon encoding a Zur-like protein and an associated ABC zinc permease in Listeria monocytogenes', *FEMS Microbiology Letters*, 174(1), pp. 111–116.
- Datsenko, K. A. and Wanner, B. L. (2000) 'One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products', *Proceedings of the National Academy of Sciences of the United States of America*.
- Datsenko, K. A., Wanner, B. L. and Beckwith, J. (2002) 'One-step inactivation of chromosomal



- genes in *Escherichia coli* K-12 using PCR products', *Proceedings of the National Academy of Sciences*, 97(12), pp. 6640–6645.
- David, R. (2011) 'Antimicrobials: Why zinc is bad for bacteria', *Nature Reviews Microbiology*, 10(1), pp. 4–4.
- Davis, L. M., Kakuda, T. and DiRita, V. J. (2009) 'A *Campylobacter jejuni* znuA orthologue is essential for growth in low-zinc environments and chick colonization', *Journal of Bacteriology*.
- Demple, B. (1996) 'Redox signaling and gene control in the *Escherichia coli* soxRS oxidative stress regulon - A review', *Gene*, 179(1), pp. 53–57.
- Ebright, R. H. and Busby, S. (1995) 'The *Escherichia coli* RNA polymerase alpha subunit: structure and function.', *Current opinion in genetics & development*, 5(2), pp. 197–203.
- Egler, M., Grosse, C., Grass, G. and Nies, D. H. (2005) 'Role of the Extracytoplasmic Function Protein Family Sigma Factor RpoE in Metal Resistance of *Escherichia coli*  $\epsilon^+$ ', *Journal of Bacteriology*, 187(7), pp. 2297–2307.
- Erickson, J. W. and Gross, C. A. (1989) 'Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression.', *Genes & development*. Cold Spring Harbor Laboratory Press, 3(9), pp. 1462–1471.
- Ewers, C., Li, G., Wilking, H., *et al.* (2007) 'Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: How closely related are they?', *International Journal of Medical Microbiology*.
- Ferianc, P., Farewell, A. and Nyström, T. (1998) 'The cadmium-stress stimulon of *Escherichia coli* K-12', *Microbiology*, 144(4), pp. 1045–1050.
- Fillat, M. F. (2014) 'The fur (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators', *Archives of Biochemistry and Biophysics*, 546, pp. 41–52.
- Fritz, G., Walker, N. and Gerland, U. (2019) 'Heterogeneous Timing of Gene Induction as a Regulation Strategy', *Journal of Molecular Biology*, 431(23), pp. 4760–4774.
- Gaballa, A. and Helmann, J. D. (1998) 'Identification of a zinc-specific metalloregulatory protein, Zur, controlling zinc transport operons in *Bacillus subtilis*', *Journal of Bacteriology*, 180(22), pp. 5815–21.
- Gabbianelli, R., Scotti, R., Ammendola, S., Petrarca, P., Nicolini, L. and Battistoni, A. (2011) 'Role of ZnuABC and ZinT in *Escherichia coli* O157:H7 zinc acquisition and interaction with epithelial cells.', *BMC microbiology*, 11(1), p. 36.
- Gabriel, S. E. and Helmann, J. D. (2009) 'Contributions of Zur-controlled ribosomal proteins to growth under zinc starvation conditions', *Journal of Bacteriology*.
- Galbusera, L., Bellement-Theroue, G., Urchueguia, A., Julou, T. and Nimwegen, E. van (2019) 'Using fluorescence flow cytometry data for single-cell gene expression analysis in bacteria', *bioRxiv*, pp. 1–31.
- Gentry, D. R. and Burgess, R. R. (1993) 'Cross-Linking of *Escherichia coli* RNA Polymerase Subunits: Identification of  $\beta'$  as the Binding Site of  $\omega$ ', *Biochemistry*, 32(41), pp. 11224–11227.
- Ghosh, P., Ishihama, A. and Chatterji, D. (2001) 'Escherichia coli RNA polymerase subunit  $\omega$  and its N-terminal domain bind full-length  $\beta'$  to facilitate incorporation into the  $\alpha 2\beta$  subassembly', *European Journal of Biochemistry*, 268(17), pp. 4621–4627.
- Ghosh, P., Ramakrishnan, C. and Chatterji, D. (2003) 'Inter-subunit recognition and manifestation of segmental mobility in *Escherichia coli* RNA polymerase: A case study with  $\omega$ - $\beta'$  interaction', *Biophysical Chemistry*, 103(3), pp. 223–237.
- Gilston, B. A., Wang, S., Marcus, M. D., Canalizo-Hernández, M. A., Swindell, E. P., Xue, Y., Mondragón, A. and O'Halloran, T. V. (2014) 'Structural and Mechanistic Basis of Zinc Regulation Across the *E. coli* Zur Regulon', *PLoS Biology*, 12(11).
- Goethe, E., Laarmann, K., Lührs, J., Jarek, M., Meens, J., Lewin, A. and Goethe, R. (2020) 'Critical

- Role of Zur and SmtB in Zinc Homeostasis of *Mycobacterium smegmatis* ', *mSystems*, 5(2), pp. 1–22.
- Graham, A., Hunt, S., Stokes, S., Bramall, N., Bunch, J., Cox, A., McLeod, C. and Poole, R. (2009) 'Severe zinc depletion of *Escherichia coli*: Roles for high affinity zinc binding by ZinT, zinc transport and zinc-independent proteins', *Journal of Biological Chemistry*, 284(27), pp. 18377–18389.
- Grass, G., Fan, B., Rosen, B. P., Franke, S., Nies, D. H. and Rensing, C. (2001) 'ZitB (YbgR), a Member of the Cation Diffusion Facilitator Family, Is an Additional Zinc Transporter in *Escherichia coli*', *Journal of bacteriology*, 183(15), pp. 4664–4667.
- Grass, G., Franke, S., Taudte, N., Nies, D. H., Kucharski, L. M., Maguire, M. E. and Rensing, C. (2005) 'The metal permease ZupT from *Escherichia coli* is a transporter with a broad substrate spectrum.', *Journal of bacteriology*, 187(5), pp. 1604–11.
- Grass, G., Otto, M., Fricke, B., Haney, C. J., Rensing, C., Nies, D. H. and Munkelt, D. (2005) 'FieF (YiiP) from *Escherichia coli* mediates decreased cellular accumulation of iron and relieves iron stress', *Archives of Microbiology*, 183(1), pp. 9–18.
- Grass, G., Wong, M. D., Rosen, B. P., Smith, R. L. and Rensing, C. (2002) 'Zupt is a Zn(II) uptake system in *Escherichia coli*', *Journal of Bacteriology*, 184(3), pp. 864–866.
- Grote, J., Krysciak, D. and Streit, W. R. (2015) 'Phenotypic Heterogeneity, a Phenomenon That May Explain Why Quorum Sensing Does Not Always Result in Truly Homogenous Cell Behavior', *Applied and Environmental Microbiology*, 81(16), pp. 5280–5289.
- GSL Biotech (2019) 'SnapGene softwear'.
- Gunasekera, T. S., Herre, A. H. and Crowder, M. W. (2009) 'Absence of ZnuABC-mediated zinc uptake affects virulence-associated phenotypes of uropathogenic *Escherichia coli* CFT073 under Zn(II)-depleted conditions', *FEMS Microbiology Letters*, 300(1), pp. 36–41.
- Haas, C. E., Rodionov, D. A., Kropat, J., Malasarn, D., Merchant, S. S. and de Crécy-Lagard, V. (2009) 'A subset of the diverse COG0523 family of putative metal chaperones is linked to zinc homeostasis in all kingdoms of life', *BMC Genomics*. BioMed Central, 10(470).
- Haemig, H. A. H. and Brooker, R. J. (2004) 'Importance of conserved acidic residues in MntH, the Nramp homolog of *Escherichia coli*', *Journal of Membrane Biology*. Springer, 201(2), pp. 97–107.
- Han, Y. and Zhang, F. (2020) 'Heterogeneity coordinates bacterial multi-gene expression in single cells', *PLoS Computational Biology*, 16(1), pp. 1–17.
- Hannig, G. and Makrides, S. C. (1998) 'Strategies for optimizing heterologous protein expression in *Escherichia coli*', *Trends in Biotechnology*, 16(2), pp. 54–60.
- Harden, T. T., Herlambang, K. S., Chamberlain, M., *et al.* (2020) 'Alternative transcription cycle for bacterial RNA polymerase', *Nature Communications*. Springer US, 11(1), pp. 1–11.
- Harley, C. B. and Reynolds, R. P. (1987) 'Analysis of *E. coli* promoter sequences.', *Nucleic Acids Research*, 15(5), pp. 2343–61.
- Harris, R. M., Webb, D. C., Howitt, S. M. and Cox, G. B. (2001) 'Characterization of PitA and PitB from *Escherichia coli*', *Journal of Bacteriology*, 183(17), pp. 5008–5014.
- Hawley, D. K. and McClure, W. R. (1983) 'Compilation and analysis of *Escherichia coli* promoter DNA sequences.', *Nucleic acids research*, 11(8), pp. 2237–55.
- Hayashi, K., Morooka, N., Yamamoto, Y., *et al.* (2006) 'Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110', *Molecular Systems Biology*.
- Hayashi, T., Makino, K., Ohnishi, M., *et al.* (2001) 'Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12', *DNA Research*.
- van Helden, J. (2003) 'Regulatory Sequence Analysis Tools', *Nucleic Acids Research*.

- Helmann, J. D. (1991) 'Alternative sigma factors and the regulation of flagellar gene expression', *Molecular Microbiology*, 5(12), pp. 2875–2882.
- Hemm, M. R., Paul, B. J., Miranda-Ríos, J., Zhang, A., Soltanzad, N. and Storz, G. (2010) 'Small stress response proteins in Escherichia coli: Proteins missed by classical proteomic studies', *Journal of Bacteriology*, 192(1), pp. 46–58.
- Hensley, M. P., Gunasekera, T. S., Easton, J. A., *et al.* (2012) 'Characterization of Zn(II)-responsive ribosomal proteins YkgM and L31 in E. coli', *Journal of Inorganic Biochemistry*, 111, pp. 164–172.
- Hensley, M. P., Tierney, D. L. and Crowder, M. W. (2011) 'Zn(II) binding to Escherichia coli 70S ribosomes', *Biochemistry. NIH Public Access*, 50(46), pp. 9937–9.
- Herring, C. D., Glasner, J. D. and Blattner, F. R. (2003) 'Gene replacement without selection: regulated suppression of amber mutations in Escherichia coli', *Gene*. Elsevier, 311, pp. 153–163.
- Hertz, G. Z. and Stormo, G. D. (1999) 'Identifying DNA and protein patterns with statistically significant alignments of multiple sequences', in *Bioinformatics*.
- Hobman, J. L., Patel, M. D., Hidalgo-Arroyo, G. A., Cariss, S. J. L., Avison, M. B., Penn, C. W. and Constantinidou, C. (2007) 'Comparative genomic hybridization detects secondary chromosomal deletions in Escherichia coli K-12 MG1655 mutants and highlights instability in the flhDC region', *Journal of Bacteriology*.
- Holland, P. M., Abramson, R. D., Watson, R. and Gelfand, D. H. (1991) 'Detection of specific polymerase chain reaction product by utilizing the 5' → 3' exonuclease activity of Thermus aquaticus DNA polymerase', *Proceedings of the National Academy of Sciences of the United States of America*, 88(16), pp. 7276–7280.
- Hood, M. I., Mortensen, B. L., Moore, J. L., *et al.* (2012) 'Identification of an Acinetobacter baumannii Zinc Acquisition System that Facilitates Resistance to Calprotectin-mediated Zinc Sequestration', *PLoS Pathogens*, 8(12), pp. 20–24.
- Huang, D. B., Mohanty, A., DuPont, H. L., Okhuysen, P. C. and Chiang, T. (2006) 'A review of an emerging enteric pathogen: Enteraggregative Escherichia coli', *Journal of Medical Microbiology*, 55(10), pp. 1303–1311.
- Huang, D. L., Tang, D. J., Liao, Q., *et al.* (2008) 'The Zur of Xanthomonas campestris functions as a repressor and an activator of putative zinc homeostasis genes via recognizing two distinct sequences within its target promoters', *Nucleic Acids Research*.
- Iqbal, M., Doherty, N., Page, A. M. L. L., *et al.* (2017) 'Reconstructing promoter activity from Lux bioluminescent reporters', *PLoS Computational Biology*, 13(9).
- Irwin, D. M. (2014) 'Evolution of the vertebrate goose-type lysozyme gene family', *BMC Evolutionary Biology*.
- Irwin, D. M., Biegel, J. M. and Stewart, C. B. (2011) 'Evolution of the mammalian lysozyme gene family', *BMC Evolutionary Biology*.
- Irwin, D. M. and Gong, Z. (2003) 'Molecular evolution of vertebrate goose-type lysozyme genes', *Journal of Molecular Evolution*.
- Ishikawa, S., Ogura, Y., Yoshimura, M., *et al.* (2007) 'Distribution of stable DnaA-binding sites on the bacillus subtilis genome detected using a modified ChIP-chip method', *DNA Research*.
- Ito, K., Iwakura, Y. and Tshihama, A. (1975) 'Biosynthesis of RNA polymerase in Escherichia coli. III. Identification of intermediates in the assembly of RNA polymerase', *Journal of Molecular Biology*, 96(2), pp. 257–271.
- Jackson, R. J., Binet, M. R. B., Lee, L. J., Ma, R., Graham, A. I., McLeod, C. W. and Poole, R. K. (2008) 'Expression of the PitA phosphate/metal transporter of Escherichia coli is responsive to zinc and inorganic phosphate levels', *FEMS Microbiology Letters*. Oxford Academic, 289(2), pp.

219–224.

Johnson, F. H., Shimomura, O., Saiga, Y., Gershman, L. C., Reynolds, G. T. and Waters, J. R. (1962) 'Quantum efficiency of Cypridina luminescence, with a note on that of Aequorea', *Journal of Cellular and Comparative Physiology*, 60(1), pp. 85–103.

Johnson, T. J., Kariyawasam, S., Wannemuehler, Y., *et al.* (2007) 'The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes', *Journal of Bacteriology*, 189(8), pp. 3228–3236.

Jordan, M. R., Wang, J., Weiss, A., Skaar, E. P., Capdevila, D. A. and Giedroc, D. P. (2019) 'Mechanistic Insights into the Metal-Dependent Activation of ZnII-Dependent Metallochaperones', *Inorganic Chemistry*.

Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T. and Katayama, T. (2014) 'Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation', *Nucleic Acids Research*, 42(21), pp. 13134–13149.

Kehl-Fie, T. E. and Skaar, E. P. (2010) 'Nutritional immunity beyond iron: a role for manganese and zinc', *Current Opinion in Chemical Biology*. Elsevier Current Trends, 14(2), pp. 218–224.

Kershaw, C. J., Brown, N. L. and Hobman, J. L. (2007) 'Zinc dependence of *zntT* (*yodA*) mutants and binding of zinc, cadmium and mercury by *ZinT*', *Biochemical and Biophysical Research Communications*, 364(1), pp. 66–71.

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., *et al.* (2017) 'The EcoCyc database: Reflecting new knowledge about *Escherichia coli* K-12', *Nucleic Acids Research*, 45(D1), pp. D543–D550.

Khan, S., Brocklehurst, K. R., Jones, G. W. and Morby, A. P. (2002) 'The functional analysis of directed amino-acid alterations in *ZntR* from *Escherichia coli*', *Biochemical and Biophysical Research Communications*, 299(3), pp. 438–445.

Khemthong, S., Nuonming, P., Nookabkaewb, S., Sukchawalit, R. and Mongkolsuk, S. (2019) 'The *Agrobacterium tumefaciens* *atu3184* gene, a member of the COG0523 family of GTPases, is regulated by the transcriptional repressor *Zur*', *Microbiological Research*.

Klug, A. and Rhodes, D. (1987) 'Zinc fingers: A novel protein fold for nucleic acid recognition', *Cold Spring Harbor Symposia on Quantitative Biology*.

Knapp, G., Müller, K., Strunz, M. and Wegscheider, W. (1987) 'Automation in element pre-concentration with chelating ion exchangers: Plenary lecture', *Journal of Analytical Atomic Spectrometry*. The Royal Society of Chemistry, 2(6), pp. 611–614.

Kolb, A. (1993) 'Transcriptional Regulation by cAMP and its Receptor Protein', *Annual Review of Biochemistry*, 62, pp. 749–95.

Kooperberg, C., Fazio, T. G., Delrow, J. J. and Tsukiyama, T. (2002) 'Improved background correction for spotted DNA microarrays', *Journal of Computational Biology*, 9(1), pp. 55–66.

Kotra, L. P., Haddad, J. and Mobashery, S. (2000) 'Aminoglycosides: Perspectives on mechanisms of action and resistance and strategies to counter resistance', *Antimicrobial Agents and Chemotherapy*. American Society for Microbiology (ASM), pp. 3249–3256.

Kumar, P., Sannigrahi, S. and Tzeng, Y. L. (2012) 'The *Neisseria meningitidis* *ZnuD* Zinc Receptor Contributes to Interactions with Epithelial Cells and Supports Heme Utilization when Expressed in *Escherichia coli*', *Infection and Immunity*, 82(2), pp. 657–667.

Kyoto University Bioinformatics Center (2019) *Multiple Sequence Alignment by CLUSTALW*. Available at: <http://www.genome.jp/tools-bin/clustalw>.

Landini, P., Egli, T., Wolf, J. and Lacour, S. (2014) 'sigmaS, a major player in the response to environmental stresses in *Escherichia coli*: role, regulation and mechanisms of promoter recognition', *Environmental Microbiology Reports*, 6(1), pp. 1–13.

De Las Penas, A., Connolly, L. and Gross, C. A. (1997) ' $\sigma(E)$  is an essential sigma factor in *Escherichia coli*', *Journal of Bacteriology*, 179(21), pp. 6862–6864.

- Latorre, M., Low, M., Gárate, E., Reyes-Jara, A., Murray, B. E., Cambiazo, V. and González, M. (2015) 'Interplay between copper and zinc homeostasis through the transcriptional regulator Zur in *Enterococcus faecalis*', *Metallomics*, 7(7), pp. 1137–1145.
- Laupland, K. B. (2013) 'Incidence of bloodstream infection: A review of population-based studies', *Clinical Microbiology and Infection*.
- Lee, D. J., Bingle, L. E. H., Heurlier, K., Pallen, M. J., Penn, C. W., Busby, S. J. W. and Hobman, J. L. (2009) 'Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains.', *BMC microbiology*, 9(252).
- Lee, T. S., Krupa, R. A., Zhang, F., Hajimorad, M., Holtz, W. J., Prasad, N., Lee, S. K. and Keasling, J. D. (2011) 'BglBrick vectors and datasheets: A synthetic biology platform for gene expression', *Journal of Biological Engineering*. BioMed Central, 5(12).
- Lemire, J. A., Harrison, J. J. and Turner, R. J. (2013) 'Antimicrobial activity of metals: mechanisms, molecular targets and applications', *Nature Reviews Microbiology*, 11(6), pp. 371–384.
- Leonhartsberger, S., Huber, A., Lottspeich, F. and Bo, A. (2001) 'The hydHG genes from *Escherichia coli* code for a zinc and lead.pdf', 4, pp. 93–105.
- Lim, J., Lee, K. M., Kim, So Hyun, Kim, Y., Kim, Sae Hun, Park, W. and Park, S. (2011) 'YkgM and ZinT proteins are required for maintaining intracellular zinc concentration and producing curli in enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 under zinc deficient conditions', *International Journal of Food Microbiology*. Elsevier B.V., 149(2), pp. 159–170.
- Liu, J. Z., Jellbauer, S., Poe, A. J., *et al.* (2012) 'Zinc sequestration by the neutrophil protein calprotectin enhances salmonella growth in the inflamed gut', *Cell Host and Microbe*.
- Lloyd, A. L., Rasko, D. A. and Mobley, H. L. T. (2007) 'Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli*', *Journal of Bacteriology*.
- Lobo, S. A. L., Brindley, A. A., Romão, C. V., Leech, H. K., Warren, M. J. and Saraiva, L. M. (2008) 'Two distinct roles for two functional cobaltochelates (CbiK) in *Desulfovibrio vulgaris* Hildenborough', *Biochemistry*.
- Locher, K. P. (2009) 'Structure and mechanism of ATP-binding cassette transporters', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1514), pp. 239–245.
- Lu, M. and Fu, D. (2007) 'Structure of the zinc transporter YiiP', *Science*.
- Lund, P. A. and Brown, N. L. (1989) 'Regulation of transcription in *Escherichia coli* from the mer and merR promoters in the transposon Tn501', *Journal of Molecular Biology*. Academic Press, 205(2), pp. 343–353.
- Maciąg, A., Dainese, E., Rodriguez, G. M., *et al.* (2007) 'Global analysis of the *Mycobacterium tuberculosis* Zur (FurB) regulon', *Journal of Bacteriology*, 189(3), pp. 730–740.
- Madigan, M., Martinko, J., Stahl, D. and Clark, D. (2012) *Brock; Biology of Microorganisms*. 13th edn. San Francisco: Pearson.
- Maeda, H; Fujita, N; Ishihama, A. (2005) 'Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase', *Nucleic acids research*, 28(18), pp. 3497–503.
- Makui, H., Roig, E., Cole, S. T., Helmann, J. D., Gros, P. and Cellier, M. F. M. (2000) 'Identification of the *Escherichia coli* K-12 Nramp orthologue (MntH) as a selective divalent metal ion transporter', *Molecular Microbiology*, 35(5), pp. 1065–1078.
- Manen, D. and Caro, L. (1991) 'The replication of plasmid pSC101', *Molecular Microbiology*, 5(2), pp. 233–237.
- Martínez-Antonio, A. and Collado-Vides, J. (2003) 'Identifying global regulators in transcriptional regulatory networks in bacteria', *Current Opinion in Microbiology*, pp. 482–489.
- Mathew, R. and Chatterji, D. (2006) 'The evolving story of the omega subunit of bacterial RNA

- polymerase', *Trends in Microbiology*, 14(10), pp. 450–455.
- Mavridou, D. A. I., Gonzalez, D., Clements, A. and Foster, K. R. (2016) 'The pUltra plasmid series: A robust and flexible tool for fluorescent labeling of Enterobacteria', *Plasmid*, 87–88, pp. 65–71.
- Mazzon, R. R., Braz, V. S., da Silva Neto, J. F. and do Valle Marques, M. (2014) 'Analysis of the *Caulobacter crescentus* Zur regulon reveals novel insights in zinc acquisition by TonB-dependent outer membrane proteins.', *BMC genomics*, 15, p. 734.
- Mcdevitt, C. A., Ogunniyi, A. D., Valkov, E., Lawrence, M. C., Kobe, B., Mcewan, A. G. and Paton, J. C. (2011) 'A Molecular Mechanism for Bacterial Susceptibility to Zinc', *PLoS Pathog*, 7(11).
- Meighen, E. (1991) 'Molecular Biology of Bacterial Bioluminescence', *Annual review of microbiology*, 55(1), pp. 123–142.
- Merriam-Webster (2020) *Biotechnology*. Available at: <https://www.merriam-webster.com/dictionary/biotechnology>.
- Merrick, M. J. (1993) 'In a class of its own — the RNA polymerase sigma factor  $\sigma_{54}$  ( $\sigma_N$ )', *Molecular Microbiology*, 10(5), pp. 903–9.
- Mitchell, J. E., Oshima, T., Piper, S. E., *et al.* (2007) 'The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70', *Journal of Bacteriology*, 189(9), pp. 3489–3495.
- Mobley, H. L., Green, D. M., Trifillis, A. L., Johnson, D. E., Chippendale, G. R., Lockatell, C. V., Jones, B. D. and Warren, J. W. (1990) 'Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: Role of hemolysin in some strains', *Infection and Immunity*, 58(5), pp. 1281–1289.
- Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H. and Yura, T. (1999) 'Translational induction of heat shock transcription factor  $\sigma_{32}$ : Evidence for a built-in RNA thermosensor', *Genes and Development*. Cold Spring Harbor Laboratory Press, 13(6), pp. 655–665.
- Müller, K., Matzanke, B. F., Schünemann, V., Trautwein, A. X. and Hantke, K. (1998) 'FhuF, an iron-regulated protein of *Escherichia coli* with a new type of [2Fe-2S] center', *European Journal of Biochemistry*, 258(3), pp. 1001–1008.
- Murakami, K. S. and Darst, S. A. (2003) 'Bacterial RNA polymerases: The whole story', *Current Opinion in Structural Biology*, 13(1), pp. 31–39.
- Natori, Y., Nanamiya, H., Akanuma, G., Kosono, S., Kudo, T., Ochi, K. and Kawamura, F. (2006) 'A fail-safe system for the ribosome under zinc-limiting conditions in *Bacillus subtilis*', *Molecular Microbiology*, 63(1), pp. 294–307.
- NEB (2018) *NEBase Changer*. Available at: <http://nebasechanger.neb.com> (Accessed: 25 July 2019).
- NEB (2019) *NEBuilder Assembly Tool*. Available at: <https://nebuilder.neb.com/#/> (Accessed: 25 July 2019).
- Neidhardt, F. C., Bloch, P. L. and Smith, D. F. (1974) 'Culture medium for enterobacteria', *Journal of Bacteriology*, 119(3), pp. 736–747.
- Nielubowicz, G. R., Smith, S. N. and Mobley, H. L. T. (2010) 'Zinc uptake contributes to motility and provides a competitive advantage to *Proteus mirabilis* during experimental urinary tract infection', *Infection and Immunity*.
- Niu, W., Kim, Y., Tau, G., Heyduk, T. and Ebright, R. H. (1996) 'Transcription activation at class II CAP-dependent promoters: Two interactions between CAP and RNA polymerase', *Cell*, 87(6), pp. 1123–1134.
- Noinaj, N., Guillier, M., Barnard, T. J. and Buchanan, S. K. (2010) 'TonB-Dependent Transporters: Regulation, Structure, and Function', *Annual Review of Microbiology*.

- Novichkov, P. S., Laikova, O. N., Novichkova, E. S., Gelfand, M. S., Arkin, A. P., Dubchak, I. and Rodionov, D. A. (2009) 'RegPrecise: A database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes', *Nucleic Acids Research*.
- Outten, C. E. and O'Halloran, T. V (2001) 'Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis.', *Science*, 292, pp. 2488–2492.
- Outten, C. E., Outten, F. W. and O'Halloran, T. V (1999) 'DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in *Escherichia coli*', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 274(53), pp. 37517–37524.
- Outten, C. E., Tobin, D. A., Penner-Hahn, J. E. and O'Halloran, T. V. (2001) 'Characterization of the metal receptor sites in *Escherichia coli* Zur, an ultrasensitive zinc(II) metalloregulatory protein', *Biochemistry*, 40(35), pp. 10417–10423.
- Palmgren, M. G. and Nissen, P. (2011) 'P-Type ATPases', *Annual Review of Biophysics*, 40(1), pp. 243–266.
- Panina, E. M., Mironov, A. A. and Gelfand, M. S. (2003) 'Comparative genomics of bacterial zinc regulons: Enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins', *PNAS*. National Academy of Sciences, 100(17), pp. 9912–9917.
- Parkhill, J. and Brown, N. L. (1990) 'Site-specific insertion and deletion mutants in the mer promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR', *Nucleic Acids Research*, 18(17), pp. 5157–5162.
- Patzer, S. I. and Hantke, K. (1998) 'The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*', *Molecular Microbiology*, 28(6), pp. 1199–1210.
- Patzer, S. I. and Hantke, K. (2000) 'The zinc-responsive regulator Zur and its control of the znu gene cluster encoding the ZnuABC zinc uptake system in *Escherichia coli*', *Journal of Biological Chemistry*, 275(32), pp. 24321–24332.
- Pawlik, M., Christin, H., Hubert, K., Joseph, B., Claus, H., Schoen, C. and Vogel, U. (2012) 'The zinc-responsive regulon of *Neisseria meningitidis* comprises 17 genes under control of a Zur element', *Journal of Bacteriology*, 194(23), pp. 6594–6606.
- Peng, S., Stephan, R., Hummerjohann, J. and Tasara, T. (2014) 'Evaluation of three reference genes of *Escherichia coli* for mRNA expression level normalization in view of salt and organic acid stress exposure in food', *FEMS Microbiology Letters*, 355(1), pp. 78–82.
- Perez-Rueda, E. (2000) 'The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12', *Nucleic Acids Research*, 28(8), pp. 1838–1847.
- Petit-Härtlein, I., Rome, K., de Rosny, E., Molton, F., Duboc, C., Gueguen, E., Rodrigue, A. and Covès, J. (2015) 'Biophysical and physiological characterization of ZraP from *Escherichia coli*, the periplasmic accessory protein of the atypical ZraSR two-component system.', *The Biochemical Journal*, 472(2), pp. 205–16.
- Petrarca, P., Ammendola, S., Pasquali, P. and Battistoni, A. (2010) 'The zur-regulated ZinT protein is an auxiliary component of the high-affinity ZnuABC zinc transporter that facilitates metal recruitment during severe zinc shortage', *Journal of Bacteriology*, 192(6), pp. 1553–1564.
- Pohlmann, A., Fricke, W. F., Reinecke, F., *et al.* (2006) 'Genome sequence of the bioplastic-producing "Knallgas" bacterium *Ralstonia eutropha* H16', *Nature Biotechnology*, 24(10), pp. 1257–1262.
- Pruteanu, M., Neher, S. B. and Baker, T. A. (2007) 'Ligand-controlled proteolysis of the *Escherichia coli* transcriptional regulator ZntR', *Journal of Bacteriology*. American Society for Microbiology (ASM), 189(8), pp. 3017–3025.
- Puškárová, A., Ferianc, P., Kormanec, J., Homerová, D., Farewell, A. and Nyström, T. (2002) 'Regulation of yodA encoding a novel cadmium-induced protein in *Escherichia coli*', *Microbiology*, 148(12), pp. 3801–3811.

- Rahman, M., Patching, S. G., Ismat, F., Henderson, P. J. F., Herbert, R. B., Baldwin, S. A. and McPherson, M. J. (2008) 'Probing metal ion substrate-binding to the *E. coli* ZitB exporter in native membranes by solid state NMR', *Molecular Membrane Biology*, 25(8), pp. 683–690.
- Raj, A. and van Oudenaarden, A. (2008) 'Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences', *Cell*, 135(2), pp. 216–226.
- Rauhut, R. and Klug, G. (1999) 'mRNA degradation in bacteria', *FEMS Microbiology Reviews*, 23(3), pp. 353–370.
- Rensing, C., Mitra, B. and Rosen, B. P. (1997) 'The *zntA* gene of *Escherichia coli* encodes a Zn(II)-translocating P-type ATPase', *Proceedings of the National Academy of Sciences of the United States of America*, 94(26), pp. 14326–14331.
- Reznikoff, W. S. (1992) 'The lactose operon-controlling elements: a complex paradigm', *Molecular Microbiology*, 6(17), pp. 2419–2422.
- Richards, J., Sundermeier, T., Svetlanov, A. and Karzai, A. W. (2008) 'Quality control of bacterial mRNA decoding and decay', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1779(9), pp. 574–582.
- Roberts, J. W., Shankar, S. and Filter, J. J. (2008) 'RNA Polymerase Elongation Factors', *Annual Review of Microbiology*, 62(1), pp. 211–233.
- Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R. C. and San Millán, Á. (2021) 'Beyond horizontal gene transfer: the role of plasmids in bacterial evolution', *Nature Reviews Microbiology*.
- Rosano, G. L. and Ceccarelli, E. A. (2014) 'Recombinant protein expression in *Escherichia coli*: Advances and challenges', *Frontiers in Microbiology*, 5(171).
- Sabri, M., Houle, S. and Dozois, C. M. (2009) 'Roles of the extraintestinal pathogenic *Escherichia coli* ZnuACB and ZupT Zinc Transporters during Urinary Tract Infection', *Infection and Immunity*, 77(3), pp. 1155–1164.
- Schauer, K., Gouget, B., Carrière, M., Labigne, A. and De Reuse, H. (2007) 'Novel nickel transport mechanism across the bacterial outer membrane energized by the TonB/ExbB/ExbD machinery', *Molecular Microbiology*.
- ScienceDirect (2008) 'Kanamycin', *Tuberculosis*, pp. 117–118.
- Sevcenco, A. M., Pinkse, M. W. H., Wolterbeek, H. T., Verhaert, P. D. E. M., Hagen, W. R. and Hagedoorn, P. L. (2011) 'Exploring the microbial metalloproteome using MIRAGE', *Metallomics*.
- Shaik, M. M., Maccagni, A., Tourcier, G., Di Guilmi, A. M. and Dessen, A. (2014) 'Structural basis of pilus anchoring by the ancillary pilin RrgC of *Streptococcus pneumoniae*', *Journal of Biological Chemistry*, 289(24), pp. 16988–16997.
- Shin, C. S., Hong, M. S., Bae, C. S. and Lee, J. (1997) 'Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]', *Biotechnology Progress*, 13(3), pp. 249–257.
- Shin, J. H., Oh, S. Y., Kim, S. J. and Roe, J. H. (2007) 'The zinc-responsive regulator zur controls a zinc uptake system and some ribosomal proteins in *Streptomyces coelicolor* A3(2)', *Journal of Bacteriology*, 189(11), pp. 4070–4077.
- Sigdel, T. K., Easton, J. A. and Crowder, M. W. (2006) 'Transcriptional response of *Escherichia coli* to TPEN', *Journal of Bacteriology*, 188(18), pp. 6709–6713.
- Singh, S. S., Typas, A., Hengge, R. and Grainger, D. C. (2011) '*Escherichia coli* Sigma70 senses sequence and conformation of the promoter spacer region', *Nucleic Acids Research*, 39(12), pp. 5109–5118.
- Sørensen, H. P. and Mortensen, K. K. (2005) 'Advanced genetic strategies for recombinant protein expression in *Escherichia coli*', *Journal of Biotechnology*, 115, pp. 113–128.
- Stock, A. M., Robinson, V. L. and Goudreau, P. N. (2000) 'Two-Component Signal Transduction',



*Annu. Rev. Biochem.*, 69, pp. 183–215.

Stoebel, D. M., Dean, A. M. and Dykhuizen, D. E. (2008) 'The cost of expression of Escherichia coli lac operon proteins is in the process, not in the products', *Genetics*, 178(3), pp. 1653–1660.

Stork, M., Bos, M. P., Jongerius, I., de Kok, N., Schilders, I., Weynants, V. E., Poolman, J. T. and Tommassen, J. (2010) 'An outer membrane receptor of Neisseria meningitidis involved in zinc acquisition with vaccine potential', *PLoS Pathogens*, 6(7), pp. 1–10.

Stork, M., Grijpstra, J., Bos, M. P., Mañas Torres, C., Devos, N., Poolman, J. T., Chazin, W. J. and Tommassen, J. (2013) 'Zinc Piracy as a Mechanism of Neisseria meningitidis for Evasion of Nutritional Immunity', *PLoS Pathogens*, 9(10).

Swain, P. S., Elowitz, M. B. and Siggia, E. D. (2002) 'Intrinsic and extrinsic contributions to stochasticity in gene expression', *Proceedings of the National Academy of Sciences of the United States of America*, 99(20), pp. 12795–12800.

Sydor, A. M., Jost, M., Ryan, K. S., Turo, K. E., Douglas, C. D., Drennan, C. L. and Zamble, D. B. (2013) 'Metal binding properties of escherichia coli YjiA, a member of the metal homeostasis-associated COG0523 family of GTPases', *Biochemistry*, 52(10), pp. 1788–1801.

Taboada, B., Verde, C. and Merino, E. (2010) 'High accuracy operon prediction method based on STRING database scores', *Nucleic Acids Research*.

Takahashi, H., Oshima, T., Hobman, J. L., *et al.* (2015) 'The dynamic balance of import and export of zinc in Escherichia coli suggests a heterogeneous population response to stress', *Journal of the Royal Society Interface*, 12(106).

Tanaka, K., Kusano, S., Fujita, N., Ishihama, A. and Takahashi, H. (1995) 'Promoter determinants for Escherichia coli RNA polymerase holoenzyme containing  $\sigma_{38}$  (the rpoS gene product)', *Nucleic Acids Research*, 23(5), pp. 827–834.

Taudte, N. and Grass, G. (2010) 'Point mutations change specificity and kinetics of metal uptake by ZupT from Escherichia coli', *BioMetals*.

Tchesnokova, V., Aprikian, P., Kisiela, D., Gowey, S., Korotkova, N., Thomas, W. and Sokurenko, E. (2011) 'Type 1 fimbrial adhesin FimH Elicits an immune response that enhances cell adhesion of Escherichia coli', *Infection and Immunity*, 79(10), pp. 3895–3904.

Ueta, M., Wada, C. and Wada, A. (2020) 'YkgM and YkgO maintain translation by replacing their paralogs, zinc-binding ribosomal proteins L31 and L36, with identical activities', *Genes to Cells*.

Ullmann, A. (2009) 'Escherichia coli Lactose Operon', *Encyclopedia of Life Sciences*. Chichester, UK: John Wiley & Sons, Ltd.

Vaara, M. (1992) 'Agents that increase the permeability of the outer membrane', *Microbiological Reviews*.

Valentin-Hansen, P., Sogaard-Andersen, L. and Pedersen, H. (1996) 'A flexible partnership: the CytR anti-activator and the cAMP-CRP activator protein, comrades in transcription control.', *Molecular microbiology*, 20(3), pp. 461–6.

Vanderkelen, L., Van Herreweghe, J. M., Callewaert, L. and Michiels, C. W. (2011) 'Goose-type lysozyme inhibitor (PliG) enhances survival of Escherichia coli in goose egg albumen', *Applied and Environmental Microbiology*.

Vanderkelen, L., Van Herreweghe, J. M., Vanoirbeek, K. G. A., *et al.* (2011) 'Identification of a bacterial inhibitor against g-type lysozyme', *Cellular and Molecular Life Sciences*, 68(6), pp. 1053–1064.

Vassilyev, D. G., Sekine, S.-I., Laptenko, O., Lee, J., Vassilyeva, M. N., Borukhov, S. and Yokoyama, S. (2002) 'Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution.', *Nature*, 417, pp. 712–719.

Velasco, E., Wang, S., Sanet, M., *et al.* (2018) 'A new role for Zinc limitation in bacterial pathogenicity: Modulation of  $\alpha$ -hemolysin from uropathogenic Escherichia coli', *Scientific*

*Reports.*

- Walker, J. and Rapley, R. (eds) (2000) *Molecular Biology and Biotechnology*. 4th edn. Cambridge: The Royal Society of Chemistry.
- Wang, D., Hosteen, O., Fierke, C. A. and Da Wanga, Olijahwon Hosteenb, and C. A. F. (2012) 'ZntR-mediated transcription of zntA responds to nanomolar intracellular free zinc', *J Inorg Biochem*, 111(1), pp. 173–181.
- Wang, D., Hurst, T. K., Thompson, R. B. and Fierke, C. A. (2011) 'Genetically encoded ratiometric biosensors to measure intracellular exchangeable zinc in Escherichia coli', *Journal of Biomedical Optics*, 16(8), p. 087011.
- Wang, K., Sitsel, O., Meloni, G., et al. (2014) 'Structure and mechanism of Zn<sup>2+</sup>-transporting P-type ATPases', *Nature*, 514(7253), pp. 518–522.
- Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A. and Jun, S. (2010) 'Robust growth of escherichia coli', *Current Biology*. Cell Press, 20(12), pp. 1099–1103.
- Waterhouse, A., Bertoni, M., Bienert, S., et al. (2018) 'SWISS-MODEL: Homology modelling of protein structures and complexes', *Nucleic Acids Research*, 46, pp. 296–303.
- Webb, E. C. (1992) *Recommendations of the nomenclature committee of the international union of biochemistry and molecular biology on the nomenclature and classification of enzymes*, Academic Press. San Diego, California.
- Wei, Y. and Fu, D. (2005) 'Selective metal binding to a membrane-embedded aspartate in the Escherichia coli metal transporter YiiP (FieF)', *Journal of Biological Chemistry*, 280(40), pp. 33716–33724.
- Welch, R. A., Burland, V., Plunkett, G., et al. (2002) 'Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*, 99(26), pp. 17020–17024.
- Wigneshweraraj, S., Bose, D., Burrows, P. C., et al. (2008) 'Modus operandi of the bacterial RNA polymerase containing the  $\sigma$ 54 promoter-specificity factor', *Molecular Microbiology*. Blackwell Publishing Ltd, 68(3), pp. 538–546.
- Xu, Z., Wang, P., Wang, H., Hang Yu, Z., Ho Yu Au-Yeung, X., Hirayama, T., Sun, H. and Yan, A. (2019) 'Zinc excess increases cellular demand for iron and decreases tolerance to copper in Escherichia coli Downloaded from', *J. Biol. Chem*, 294(45), pp. 16978–16991.
- Yatsunyk, L. A., Easton, J. A., Kim, L. R., et al. (2008) 'Structure and metal binding properties of ZnuA, a periplasmic zinc transporter from Escherichia coli.', *Journal of biological inorganic chemistry*, 13(2), pp. 271–288.
- Yu, N. Y., Laird, M. R., Spencer, C. and Brinkman, F. S. L. (2011) 'PSORTdb-an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea', *Nucleic Acids Research*.
- Zaitoun, M. A. and Lin, C. T. (1997) 'Chelating behavior between metal ions and EDTA in sol - gel matrix', *Journal of Physical Chemistry B*, 101(10), pp. 1857–1860.
- Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. and Darst, S. A. (1999) 'Crystal structure of thermus aquaticus core RNA polymerase at 3.3 Å resolution', *Cell*, 98(6), pp. 811–24.
- Zheng, D., Constantinidou, C., Hobman, J. L. and Minchin, S. D. (2004) 'Identification of the CRP regulon using in vitro and in vivo transcriptional profiling', *Nucleic Acids Research*, 32(19), pp. 5874–5893.
- Zhou, K., Zhou, L., Lim, Q., Zou, R., Stephanopoulos, G. and Too, H. P. (2011) 'Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR', *BMC Molecular Biology*, 12(18).
- Zhou, Y., Kolb, A., Busby, S. J. W. and Wang, Y. P. (2014) 'Spacing requirements for Class I

transcription activation in bacteria are set by promoter elements', *Nucleic Acids Research*, 42(14), pp. 9209–9216.

Zimmermann, L., Stephens, A., Nam, S. Z., *et al.* (2018) 'A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core', *Journal of Molecular Biology*, 430(15).

## Chapter 9

## Appendix

## 9.1. Supplementary Sequences

### 9.1.1. pJ100 Sequence

AGGCCCTTTTCGTCTTCACCTCGAGCCCGGGGGATCCGCGGCCGCAACTAGAGGCATCAAATAAAACGAAA  
GGCTCAGTCGAAAGACTGGGCCCTTTCGTCTTTATCTGTTGTTTGTCCGGTGAACGCTCTCCTGAGTAGGACA  
AATCCGCCGCCCTAGACCTAGCCTTAATTAAGGCTAGGGTACGGGTTTTGCTGCCCCGAAACGGGCTGTT  
CTGGTGTGCTAGTTTGTATCAGAATCGCAGATCCGGCTTCAGCCGGTTTGCCGGCTGAAAGCGCTATT  
TCTTCCAGAATTGCCATGATTTTTTCCCCACGGGAGGCGTCACTGGCTCCCGTGTGTGCGGCAGCTTTGA  
TTCGATAAGCAGCATCGCCTGTTTCAGGCTGTCTATGTGTGACTGTTGAGCTGTAACAAGTTGTCTCAGG  
TGTTCAATTTTCATGTTCTAGTTGCTTTGTTTTACTGGTTTCACCTGTTCTATTAGGTGTTACATGCTGTT  
CATCTGTTACATTGTCGATCTGTTATGGTGAACAGCTTTGAATGCACCAAAAACCTCGTAAAAGCTCTGA  
TGTATCTATCTTTTTTACACCGTTTTTCATCTGTGCATATGGACAGTTTTCCCTTTGATATGTAACGGTGA  
ACAGTTGTTCTACTTTTGTGTTAGTCTTGATGCTTCACTGATAGATACAAGAGCCATAAGAACCTCAG  
ATCCTTCCGTATTTAGCCAGTATGTTCTCTAGTGTGGTTGTTGTTTGTGCGTGAGCCATGAGAACGAAC  
CATTGAGATCATACTTACTTTGCATGTCACTCAAAAATTTTGCCCTCAAAAACCTGGTGAGCTGAATTTTGC  
AGTTAAAGCATCGTGTAGTGTGTTTTCTTAGTCCGTTATGTAGGTAGGAATCTGATGTAATGGTTGTTGGT  
ATTTTGTCAACATTCATTTTTATCTGGTTGTTCTCAAGTTCGGTTACGAGATCCATTTGTCTATCTAGTT  
CAACTTGGAAAATCAACGTATCAGTCGGGCGGCCTCGCTTATCAACCACCAATTTTCATATTGCTGTAAGT  
GTTTAAATCTTTACTTATTGGTTTCAAAACCCATTGGTTAAGCCTTTTAACTCATGGTAGTTATTTTCA  
AGCATTAACATGAACCTAAATTCATCAAGGCTAATCTCTATATTTGCCTTGTTGAGTTTTCTTTTGTGTTA  
GTTCTTTTAAATAACCACTCATAAATCCTCATAGAGTATTTGTTTTCAAAAAGACTTAACATGTTCCAGATT  
ATATTTTATGAATTTTTTAACTGGAAAAGATAAGGCAATATCTCTTCACTAAAACTAATTCTAATTTT  
TCGCTTGAGAACTTGGCATAGTTTGTCCACTGGAAAATCTCAAAGCCTTTAACCAGGATTCTGATTT  
CCAGATTCTCGTCATCAGCTCTCTGGTTGCTTTAGCTAATACACCATAAAGCATTTTCCCTACTGATGTT  
CATCATCTGAGCGTATTGGTTATAAGTGAACGATACCGTCCGTTCTTTCCCTTGTTAGGGTTTTCAATCGGG  
GGGTTGAGTAGTGCCACACAGCATAAAATTAGCTTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATCG  
CTAGTTCAATTTGCTTTGAAAACAATAATTACAGACATACATCTCAATTGGTCTAGGTGATTTTAACTACT  
ATACCAATTGAGATGGGCTAGTCAATGATAAATTACTAGCCTTAATTAAGGTAGTCCCTTTTCCCTTTGAGTT  
GTGGGTATCTGTAAATCTGCTAGACCTTTGCTGGAAAACCTGTAAATCTGCTAGACCTCTGTAAAT  
CCGCTAGACCTTTGTGTGTTTTTTTTTGTGTTATATTCAAGTGGTTATAATTTATAGAATAAAGAAAGAA  
AAAAAAAAGATAAAAAGAATAGATCCAGCCCTGTGTATAACTCACTACTTTAGTCAGTTCCGCAGTATT  
ACAAAAGGATGTCGCAACGCTGTTTGCTCCTCTACAAAACAGACCTTAAACCCCTAAAGGCTTAAGTAG  
CACCTCGCAAGCTCGGGCAAATCGCTGAATATTCCTTTTGTCTCCGACCATCAGGCACCTGAGTCGCTG  
TCTTTTTTCGTGACATTCAGTTTCGCTGCGCTCACGGCTCTGGCAGTGAATGGGGGTAAATGGCACTACAGG  
CGCCTTTTATGGATTTCATGCAAGGAACTACCCATAATACAAGAAAAGCCCGTCACGGGCTTCTCAGGGC  
GTTTTATGGCGGGTCTGCTATGTGGTGCTATCTGACTTTTTGCTGTTTCAAGTTCCTGCCCCTCTGATTT  
TCCAGTCTGACCACTTCGGATTATCCCGTGACAGGTCATTTCAGACTGGCTAATGCACCCAGTAAGGCAGC  
GGTATCATCAACAGGCTTACCCGTCTTACTGTCCCTAGTGCTTGGATTCTCACCAATAAAAACGCCCGG  
CGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCTGAGGTCAATCTGAGTCTATCAACAGGAGTC  
CAAGCGAGCTCTCGAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGCGCATAGAAGGCGATGCGC  
TGCGAATCGGGAGCGGCGATACCGTAAAGCACAGGGAAGCGGTACGCCCATTCGCCGCCAAGCTCTTCAG  
CAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAA  
TCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGCAGGCATCGCCATGAGTCACGACGAGATCC  
TCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCGATGCTCTTCGT  
CCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTG  
GTGGTCAATGGGCAGGTAGCCGGATCAAGCGTATGTCAGCCGCCGCTTGCATCAGCCATGATGGATACT  
TTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCAGCCAGTCCC  
TTCGCCGTTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCG  
CGCTGCCTCGTCTGCACTTCATTCAGGGCACCGGACAGGTGGTCTTGACAAAAAGAACCGGGCGCCCC  
TGCGCTGACAGCCGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCGAATA  
GCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTGTTCATATGCGAAACGATCCTCA  
TCTGTCTCTTGATCAGATCTTGATCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTT  
ACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCGGACGTCTAAGAAACCATTA  
TTATCATGACATTAACCTATAAAAAATAGGCGTATCAG

## 9.1.2. pJl100A Sequence

TTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGCCCGGGGG  
 ATCCGCGGCCGCACTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCCTTTCGTTTTA  
 TCTGTTGTTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCCCTAGACCTAGCCTTAATTAAG  
 GCTAGGGTACGGGTTTTGCTGCCCCGCAACGGGCTGTTCTGGTGTGCTAGTTTTGTTATCAGAATCGCAG  
 ATCCGGCTTCAGCCGGTTTGCCGGCTGAAAGCGCTATTTCTTCCAGAATTGCCATGATTTTTTCCCCACG  
 GGAGGCGTCACTGGCTCCCGTGTGTGCGGAGCTTTGATTGCGATAAGCAGCATCGCCTGTTTCAGGCTGT  
 CTATGTGTGACTGTTGAGCTGTAACAAGTTGTCTCAGGTGTTCAATTTTCATGTTCTAGTTGCTTTGTTTT  
 ACTGGTTTCACCTGTTCTATTAGGTGTTACATGCTGTTTCATCTGTTACATTGTCGATCTGTTTCATGGTGA  
 ACAGCTTTGAATGCACCAAAAACTCGTAAAAGCTCTGATGTATCTATCTTTTTTACACCGTTTTTCATCTG  
 TGCATATGGACAGTTTTCCCTTTTGATATGTAACGGTGAACAGTTGTTCTACTTTTTGTTGTTAGTCTTGA  
 TGCTTCACTGATAGATACAAGAGCCATAAGAACCCTCAGATCCTTCCGTATTTAGCCAGTATGTTCTCTAG  
 TGTGGTTCGTTGTTTTTGGTGAGCCATGAGAACGAACCATTGAGATCATACTTACTTTGCATGTCACTC  
 AAAAATTTTGCCTCAAACTGGTGAGCTGAATTTTTGTCAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTC  
 CGTTATGTAGGTAGGAATCTGATGTAATGGTTGTTGGTATTTTGTCAACATTCAATTTTTATCTGGTTGTT  
 CTCAAGTTCGGTTACGAGATCCATTTGTCTATCTAGTTCAACTTGGAATAACACGTATCAGTCGGGCGG  
 CCTCGCTTATCAACCACCAATTTTCATATTGCTGTTAAAGTGTAAATCTTTACTTATTGGTTTCAAAACCC  
 ATTGGTTAAGCCTTTTAAACTCATGTTAGTTATTTTCAAGCATTAACATGAACCTTAAATTCATCAAGGCT  
 AATCTCTATATTTGCCTTGTGAGTTTTCTTTTGTGTTAGTTCTTTTAATAACCACTCATAAATCCTCATA  
 GAGTATTTGTTTTCAAAAGACTTAACATGTTCCAGATTATATTTTATGAATTTTTTTAACTGGAAAAGAT  
 AAGGCAATATCTCTCACTAAAACTAATTCTAATTTTTCGCTTGAGAACTTGGCATAGTTTGTCCACTG  
 GAAAACTCAAAGCCTTTAACCAAGGATTTCCTGATTTCCACAGTTCTCGTCATCAGCTCTCTGGTTGCT  
 TTAGCTAATACACCATAAGCATTTTCCCTACTGATGTTTCATCATCTGAGCGTATTGGTTATAAGTGAACG  
 ATACCGTCCGTTCTTCCCTTGTAGGGTTTTCAATCGGGGGTTGAGTAGTGCCACACAGATAAAAATTAG  
 CTTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATCGCTAGTTCATTTGCTTTGAAAACAATAATTCA  
 GACATACATCTCAATTGGTCTAGGTGATTTTAATCACTATAACCAATTGAGATGGGCTAGTCAATGATAAT  
 TACTAGCCTTAATTAAGGTAGTCCTTTTCTTTGAGTTGTGGGTATCTGTAAATCTGCTAGACCTTTGCT  
 TGGAAAACCTTGTAATTTCTGCTAGACCTCTGTAAATTCGCTAGACCTTTGTGTGTTTTTTTTTGTGTTA  
 TATTCAAGTGGTTATAATTTATAGAATAAAGAAAGAATAAAAAAAGATAAAAAAGAATAGATCCCAGCCC  
 TGTGTATAACTCACTACTTTAGTCAGTTCCGCAGTATTACAAAAGGATGTCGCAACGCTGTTTGTCTCCT  
 CTACAAAACAGACCTTAAACCCCTAAAGGCTTAAGTAGCACCTCGCAAGCTCGGGCAAAATCGCTGAATA  
 TTCTTTTGTCTCCGACCATCAGGCACCTGAGTCGCTGTCTTTTTTGTGACATTCAAGTTGCTGCGCTCA  
 CGGCTCTGGCAGTGAATGGGGGTAAATGGCACTACAGGCGCCTTTTATGGATTGATGCAAGGAACTACC  
 CATAATACAAGAAAAGCCGTCACGGGCTTCTCAGGGCGTTTTATGGCGGGTCTGCTATGTGGTGTCTATC  
 TGACTTTTTGCTGTTTCAGCAGTTCCCTGCCCTCTGATTTTCCAGTCTGACCACTTCGGATTATCCCGTGAC  
 AGGTCATTTCAGACTGGCTAATGCACCCAGTAAGGCAGCGGTATCATCAACAGGCTTACCCGCTTACTGT  
 CCTAGTGTGCTTGGATTCTCACCAATAAAAAACGCCCGCGGCAACCGAGCGTTCTGAACAAATCCAGATG  
 GAGTCTGAGGTCAATTAAGTATCTATCAACAGGAGTCCAAGCGAGCTCTCGAACCCAGAGTCCCGCTT  
 ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACT  
 CCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA  
 GACCCACGCTCACCAGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTG  
 GTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAGAGTAAGTAGTTTCGCC  
 AGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCTGTTGGTATG  
 GCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG  
 TTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGC  
 AGCACTGCATAATTTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAACC  
 AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCG  
 CGCCACATAGCAGAACTTTAAAAGTGCTCATCATTTGAAAAACGTTCTTCCGGGGCGAAAACCTCTCAAGGAT  
 CTTACCGCTGTTGAGATCCAGTTTCGATGTAACCACTCGTGCACCAACTGATCTTCAGCATCTTTTACT  
 TTCACAGCGTTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACAC  
 GGAAATGTTGAATACTCATACTCTCTCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCAT  
 GAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGGCGGGACTCTGGGGTT  
 CGNGCGAAACGATCCTCATCCGCGAAACGATCCTCATCTGTCTCTTGATCAGATCTTGATCCCCTGCGC  
 CATCAGATCCTTGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCG  
 CCCCAGCTGGCAATTCCGACGTCTAAGAAACCATTA

## 9.1.3. pJl102 Sequence

ATCGTTAATATTGCGAACAATAATTTTCGTGAGTATCACATCTGATAGTATCAGCAGCCAGAGTGGGATTA  
 AGACGGTCGACAAGGCAATTAATAATATGCTCAGGCATTATTTGCCTGACGTTATATTTGTCGAAACATC  
 CGGAGGTAGCCACCCACTGCCATTAATTAAGTATCTGCAGACTCATGAGCGATTACAGCTAAAACACTT  
 TTATCACTTGTGCGATACCGTTATGCTCAATGATGACTATAACGGAGGGAAAGCACTGATACCTGCATTTT  
 AGGAGAATCTTTTCCATGGCAAGCGCAGAATAGAGAATTTACTGGCTGAGCAAATCATGTTTTGTAGCAC  
 TTTGCTGTAAACCAAAAAGGATCGACTGTCATTGACATTGTCACCGATGTGGCGAAAGCAATCCATCCT  
 CTGAATCCTTACGTAAATGTGATAGCTGTGTCATGGGGGAACCTTAACTAGCAGAGTTACTAACCCCTC  
 CGGATTACGATTTTAACCGGGTCGGATTACTTATCAGAGAAGTGAAGACACAATCACAGTAGAAGAAGC  
 AAAAATTTTCAGCCCAAAATGGCGAAATAATTTTCGCGAGTGATTAAAGATGACAGGCCATTCCACCCACA  
 AGATTATGGGAAACCTATCATTATTTATGGGAATGGGTATTTATCGTAGTAAAGGCTTTTTCTGGCTAC  
 CAGGAAGACCTGACATGGCTCTGCTCTGGAACCAGGCCGAGGAAGTATCAATCTGGAGTTTATCAGCTA  
 CTGGAATTCGGGTGTATTGGCAGACCCTGATAATCACCTGACCCACGAAGAGCGCTCTGTTTTGCAGAAA  
 AAAGTCAATAAACTATGGGGCGGTTTGGAGACAGGAGATGCCATTAAACGATTATAGGGCGTTCTGATG  
 AAGTCCATGATTTTACCTCCGCCTTAATTAACCTGTTTTCTTTCTGAAGAAGAAATTGTCTGGTGGCAATC  
 AGGTGGCGTTTTTCGCGGATCCCTGGCCCGTTAATATATCCCGGCTGAAGTATGATTAAACCGCATATTGC  
 AAATATGGCATATGAATCGATCGTAGTTTGGCATTCGAATCGATCATGATTGGGTAAACGTTACGGGA  
 TCCGCGGCCGCAACTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCTTTTCGTTTAT  
 CTGTTGTTTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCCCTAGACCTAGCCTTAATTAAGG  
 CTAGGGTACGGGTTTTGCTGCCCCGAAACGGGCTGTTCTGGTGTGTAGTTTGTATCAGAATCGCAGA  
 TCCGGCTTCAGCCGTTTTGCCGGCTGAAAGCGCTATTTCTCCAGAAATGCCATGATTTTTTCCCCACGG  
 GAGCGCTCACTGGCTCCCGTGTGTGCGGCAGCTTTGATTGCGATAAGCAGCATCGCCTGTTTCAGGCTGTC  
 TATGTGTGACTGTTGAGCTGTAACAAGTTGTCTCAGGTGTTCAATTTTCATGTTCTAGTTGCTTTGTTTTA  
 CTGGTTTTCACTGTCTATTAGGTGTACATGCTGTTTCATCTGTTACATTGTCGATCTGTTATGTTGTTGAA  
 CAGCTTTGAATGCACCAAAAACCTCGTAAAAGCTCTGATGTATCTATCTTTTTTACACCGTTTTTCATCTGT  
 GCATATGGACAGTTTTCCCTTTGATATGTAACGGTGAACAGTTGTTCTACTTTTTGTTTGTAGTCTTGAT  
 GCTTCACTGATAGATACAAGAGCCATAAGAACCCTCAGATCCTTCCGTATTTAGCCAGTATGTTCTCTAGT  
 GTGGTTGCTGTTTTTTCGCTGAGCCATGAGAACGAACCATTGAGATCATACTTACTTTGTCATGTCACTCA  
 AAAATTTTGCTCAAACTGGTGAGCTGAATTTTTCGAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTCC  
 GTTATGTAGGTAGGAATCTGATGTAATGGTTGTTGGTATTTTGTCAACATTCAATTTTATCTGGTTGTTT  
 TCAAGTTTCGGTTACGAGATCCATTTGTCTATCTAGTTTCAACTTGGAAAATCAACGATCAGTCCGGCGGC  
 CTCGCTTATCAACCACCAATTTTCATATTGCTGTAAGTGTTTAAATCTTTACTTATTTGGTTTCAAAACCCA  
 TTGGTTAAGCCTTTTAACTCATGGTAGTTATTTTCAAGCATTAAACATGAACCTAAATTCATCAAGGCTA  
 ATCTCTATATTTGCCTTGTGAGTTTCTTTTTGTGTAGTTCTTTTAAATAACCACTCAAAATCCTCATAG  
 AGTATTTGTTTTCAAAGACTTAACATGTTCCAGATTATATTTTATGAATTTTTTAACTGGAAAAGATA  
 AGGCAATATCTCTTCACTAAAACTAATTTCTAATTTTTCGCTTGAGAACTTGGCAGTATTTGTCCACTGG  
 AAAATCTCAAAGCCTTTAACCAGAGATTCTGTAGTTTCCACAGTTCTCGTCATCAGTCTCGGTTGCTT  
 TAGCTAATACACCATAAGCATTTTCCCTACTGATGTTTCATCATCTGAGCGTATTGGTTATAAGTGAACGA  
 TACCGTCCGTTCTTTCCCTTGTAGGGTTTTCAATCGGGGGTGTAGTAGTGCCACACAGCATAAAATTAGC  
 TTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATCGCTAGTTTCAATTTGCTTTGAAAACAACCTAATTCAG  
 ACATACATCTCAATTTGGTCTAGGTGATTTTAACTACTATACCAATTGAGATGGGCTAGTCAATGATAATT  
 ACTAGCCTTAATTAAGGTAGTCCTTTTCCCTTTGAGTTGTGGGTATCTGTAAATCTGCTAGACCTTTGCT  
 GGAAAACTTGTAATTTCTGCTAGACCCTCTGTAAATTCGCTAGACCTTTGTGTGTTTTTTTTTGTGTTAT  
 ATTCAAAGTGGTTATAATTTTATGAATAAAGAAAGATAAAAAAAGATAAAAAAGATAGATCCGAGCCCT  
 GTGTATAACTCACTACTTTAGTCAGTTCCGCAGTATTACAAAAGGATGTGCGAAACGCTGTTTGCTCCTC  
 TACAAAACAGACCTTAAACCCCTAAAGGCTTAAGTAGCACCTCGCAAGCTCGGGCAAATCGCTGAATAT  
 TCCTTTTGTCTCCGACCATCAGGCACCTGAGTCGTGTCTTTTTTCGTGACATTCAAGTTCGCTGCGCTCAC  
 GGCTCTGGCAGTGAATGGGGGTAATGGCACTACAGGCGCCTTTTATGGATTATGCAAGGAAACTACCC  
 ATAATACAAGAAAGCCCGTCACGGGCTTCTCAGGGCGTTTTATGGCGGGTCTGCTATGTGGTGCTATCT  
 GACTTTTTGCTGTTACAGAGTTTCTGCCCTCTGATTTTCCAGTCTGACCACTTCGGATTATCCCGTGCA  
 GGTCACTTCACTGCTAATGCACCCAGTAAGGCAGCGGTATCATCAACAGGCTTACCCGTCTTACTGTC  
 CTAGTGCTTGGATTCTCACCAATAAAAAACGCCCGCGGCAACCGAGCGTTCTGAACAAATCCAGATGG  
 AGTTCGTAGGTCACTACTGGATCTATCAACAGGAGTCCAAGCGAGCTCTCGAACCCAGAGTCCCGCTTA  
 CCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTC  
 CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG  
 ACCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG  
 TCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAGAGTAAGTACGTTCCGA  
 GTTAATAGTTTTCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGG  
 CTTTCACTCAGCTCCGTTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGT  
 TAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCA  
 GCACTGCATAATTTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA  
 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGC  
 GCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAACGTTCTTCGGGCGAAAACCTCTCAAGGATC  
 TTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT  
 TCACCAGCGTTTTCGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG  
 GAAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATG  
 AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGGCGAAACGATCCTCATC

CTGTCTCTTGATCAGATCTTGATCCCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTAC  
 TTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCCGACGTCTAAGAAACCATTTATT  
 ATCATGACATTAACTATAAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTTCACCTCGAGCCCGGGATT  
 GTGTATAAGAGACAGTACTTATCTGGAGGTAATTGCAATATCTCTGTGAACCTACACAGGGTGGGCTTAC  
 CGCATACACTGACACTTAGCGGATCGACAGAACATTATTAACAGAGCATCACTGAACCGCTACATAATCAG  
 AGTTGCATAAATAAAATGTTATGATATAACATCACAATCACAACATTTGTTTTCGAATGTATATGAATGT  
 AATCAGAAGTGTCAATTTGTACATTAATTATACTTCCGGTGGGATTACAGGCAGCGACCAGTCATTCTTCT  
 ATGGTTAAAGATACAATCACCATTGTCGCGACAGGAAATCAGAACACGGTATTTGAAACGCCGTCGATGG  
 TCAGTGTCTGTCACGAATGACACACCGTGGAGTCAGAATGCGGTTACATCGGCCGGCATGCTGAAAGGTGT  
 TGCCGGTCTCAGCCAGACTGGTGCAGGACGGACCAATGGGCAGACCTTTAATTTACGCGGCTATGACAAA  
 AGCGGGTACTTGTCTTGTGACGGCGTTTCGCCAACTCAGTGACATGGCAAAAAGCAGTGGCACTTTATC  
 TGGATCCGGCACTCGTCAAACGTATCGAAGTTGTCCGCGGGCCAACTCCAGTCTGTACGGCAGTGGCGG  
 GCTGGGAGGTGTAGTGGACTTCAGAAGTGCAGATGCAGCAGATTTTCTTCCCCCGGAGAGACAAACGGT  
 TTAAGTCTGTGGGAAATATCGCCAGTGGTGACCACAGCAGAGGCTCGGGGCTCACCTGGTTTGGTAAAA  
 CTGAAAAACAGATGCGCTCCTTTCTGTCAATATGCGTAAAGAGGTAATATCTATCAAAGTGATGGTGA  
 GCACGCACCTAACAAAGGAAAAACCTGCAGCCCTGTTTTCGAAAGGCTCTGTCTGGTATAACAGACAGTAAC  
 AAAGCAGGTGCCAGCTTGCCTCTTACCGGAATAACACCACTGAACCGGGCAATTCCTACTCAGACACATG  
 GTGACAGCGGCCCTGCGTGACAGAAAAACAGTACAAAATGACGTACAGTTCTGGTACCAGTACGCTCCTGT  
 GGATAACAGCCTCATCAATGTAAAGTCAACGTTATATCTCAGTGATATCACTATCAAGACAAACGGTCAC  
 AAAAAACGGCAGAATGGAGAAACAACAGAACCTCCGGTGTAAATGTTGTCAACAGGAGTCATACTCTGA  
 TTTTTCGGGAGCCATCAGTTAAGTTATGGCGCTGAATATTACCGTCAGCAGCAGAGCCAGAAGGCTC  
 TGCCACACTATATCCGGAAGGAAACATTGACTTTACATCGTTGTATTTCCAGGATGAAATGACAATGAAA  
 AGTACCCGGTTAATATTATCGTCGGTTCCCGCTATGACCGGTACAAGAGCTTCAATCCCCGTGCCGGAG  
 AACTGAAAGCCGAACGCCTGTCCCAAGGGCGGCGATTTTCAGTCTCACCGACAGACTGGCTGATGATGTA  
 CGGCTCCATATCCTCTGCATTCCGAGCGCCACAATGGCAGAAATGTACAGGGATGATGTACATTTTTTAC  
 CGCAAGGGTAACCCAATTACTGGGTTCTTAACCTTAATCTGAAACCAGAAAAATAACATCACCCGTGAGA  
 TTGGCGCAGGTATTCAACTGGATGGCTGCTTACAGACAATGACCGGTGCAGTTAAAGGCGGATATTT  
 CGGAACGGATGCCAGAACTATATTGCCACACGCGTGGATATGAAACGGATGCGTTCTTATTCTTATAAT  
 GTATCCCGGGCCCGTATCTGGGGATGGGATATGCAGGGTAATTACCAGTCTGATTATGTTGACTGGATGC  
 TTTCTTATAACCGGACGGAAAGTATGGATGCCAGCAGCAGGGAATGGCTGGGCTCCGGCAATCCTGACAC  
 ACTATCAGTGACATCAGCATACCTGTTGGTCATAGAGCGTTTATGCCGGATGGCGTGCTGAACTTTCA  
 GCATCAGCCACGCATGTGAAAAAAGCGCATCCCCATCAGGCTGGTTATACCATACATTCTTTTCACTGT  
 CTTATAAGCCTGTAAGTGTTAAAGGCTTTGAGGCGTCAGTAACCTCTGGATAATGCCTTCAACAAGCTTGC  
 CATGAATGGCAAAGGTGTGCCGCTTTCAGGCAGAACTGTGAGTCTTTATACCCGTTATCAGTGGTAACCT  
 CGCAGGAAATAAATCATGATGAAAAATACAGGCTATATCTTAGCTCTTTGTCTGACAGCATCGGGGCATG  
 TCCTAGCCCATGATGTCTGGATTACAGGTAACAGGCAGAGAACACGTTACCGCAGAGATTGGTTATGG  
 TCATAATTTCCCTCAAAGGGGACAATTCTGACAGAAGGGATTTCTTTGAAAATCCCCGGCTTTATAAC  
 GGAAAGAGACAATAAAGCTGAAGCCAGCGTCCACGGATTATGTCTATAAACTGAGTCTGCAAGCAAAG  
 ATAATGGTTACGTTCTGTCAACGTATATGAAACCGGATACTGGTCGAGAACCTCGTCAGGATGGAACCC  
 GGTGACCCGGGAGGGCAGAAATGATGTGGCTTACTGTGAATTTGTCACTAAATATGCAAAATCTTTTATT  
 CCTGGTGAACAGCAGATGCCAGCACAACCTCTATCAGTCTCCAACAGGGCATGAGCTTGAAATCATTCCGT  
 TATCCGATATAAGTCGTTTTCAGTGAAAATGTGAAGCTGAAAGTTCTGTATAAAACGTCCTCCGCTCGCCGG  
 AGCTATCATGGAGCTTGACTCGGTCAGTTATCTGACATCATCCCGTCATACTCATGAGTTGAGCACAAA  
 CATCCTGTTTATAAAGCAGAACTACCTTTGTAACTAATGAGGATGGTATCGTCACAGTACCTTCTCTTC  
 ATATCGGACAGTGGCTGGCGAAAGTCCAAAATAAGAAAAGTTTTCAGGACAAAAGCCTGTGTGATGAAAC  
 TGTCGATGTGGCAACCTTAAGCTTCTCCGAAATTAATGATACTATTCCTGCCCTGTATTTTACGGGCAG  
 GGTATATCAACCGGAAATAGCCTACTGACATCCATGAAAAAACTCATAATTATTTTGTGATGTATCT  
 TACCTAATACATTGATAGCTCCAGTTTATACCAGAGCATTCTCGTTGTCCCTATGACTTGTCTGCAA  
 TATCACTGATATGCAGGAATAATATCTTGGCATGGAGTCAGTTCCCGTCATCTTAGGTATCACTTTAA  
 CTCACAGGTGTCCAGTAAATATGAACAGGATCCAGTAATCATTATCAATGGTTTCTTGGCGCAGGAAA  
 AACCACTCTTATGAAAAACCTGCTCACACAGGCAACAGGAATCACTTAGCTGTATCTGTCAATTGTCAAT  
 GATATGAGCGATCTGGATGTTGATGGTGTACTTATTGCAAATACAGAA



## 9.1.4. pJl300 Sequence

GCGAAACGATCCTCATCTGTCTTTGATCAGATCTTGATCCCCGCGCCATCAGATCCTTGGCGGCAAG  
 AAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCGACG  
 TCTAAGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTCGTCTTCA  
 CCTCGAGCCCGGGGATCCTCTAGTTAGTTAGTAAGGAGTTTAAAAAAACAAATATGGCAAGTAGTGAAG  
 ACGTTATCAAAGAGTTTCATGCGTTTCAAAGTTTCGTATGGAAGGTTCCGTTAACGGTCACGAGTTCGAAAT  
 CGAAGGTGAAGGTGAAGGTTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGTGGT  
 CCGCTGCCGTTTCGCTTGGGACATCCTGTCCCCGCAGTTCAGTACGGTTCCAAAGCTTACGTTAAACACC  
 CGGCTGACATCCCGGACTACCTGAAACTGTCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACCT  
 CGAAGACGGTGGTGTGTTTACCCTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTTCATCTACAAAGTT  
 AAACCTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTGGGAAGCTT  
 CCACCGAACGTATGTACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAATGCGTCTGAAACTGAAAGA  
 CGGTGGTCACTACGACGCTGAAGTTAAACCACCTACATGGCTAAAAAACGGTTCAGCTGCCGGGTGCT  
 TACAAAACCGACATCAAACCTGGACATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAAC  
 GTGCTGAAGGTTCGTCACCTCCACCGGTGCTTAATCTAGTTAGTTAGTAAGGAGTTTACCATGGCAAATATG  
 ACTAAAAAAATTTTCATTTCATTATTAACGGCCAGGTTGAAATCTTTCCCGAAAGTGATGTTTAGTGCAAT  
 CCATTAATTTTGGTGATAATAGTGTTCACCTGCCAATATTGAATGACTCTCATGAAAAACATTATTGA  
 TTGTAATGGAAATAACGAATTACGGTTGCATAACATTGTCAATTTTCTCTATACGGTAGGGCAAAGATTGG  
 AAAAAATGAAGAATACTCAAGACGCAGGACATACATTCTGTGACTTAAAAAAATATATGGGATATTCAGAAG  
 AAATGGCTAAGCTAGAGGCCAATTGGATATCTATGATTTTATGTTCTAAAGGCGGCCCTTTATGATGTTGT  
 AGAAAAATGAACCTTGGTCTCGCCATATCATGGATGAATGGCTACCTCAGGATGAAAGTTATGTTCCGGCT  
 TTTCCGAAAGGTAAATCTGTACATCTGTTGGCAGGTAATGTTCCATTATCTGGGATCATGTCTATATTAC  
 GCGCAATTTTAACTAAGAATCAGTGTATTATAAAAAACATCGTCAACCGATCCTTTTACCCTAATGCATT  
 AGCGTTAAGTTTATTGATGTAGACCCTAATCATCCGATAACGCGCTCTTTATCTTATATATTGGCCC  
 CACCAAGGTGATACATCACTCGCAAAAGAAATTATGCGACATGCGGATGTTATTGTGCGCTTGGGGAGGGC  
 CAGATGCGATTAAATGGGCGGTAGAGCATGCGCCATCTTATGCTGATGTGATTAAATTTGGTTCTAAAAA  
 GAGTCTTTGCATTATCGATAATCCTGTTGATTTGAGCTCCGCAGCGACAGGTGCGGCTCATGATGTTTGT  
 TTTTACGATCAGCGAGCTTGTTTTTCTGCCCAAAACATATATTACATGGGAAATCATTATGAGGAATTTA  
 AGTTAGCGTTGATAGAAAACTTAATCTATATGCGCATATATTACCGAATGCCAAAAAGATTTTGATGA  
 AAAGGCGGCCCTATTCTTTAGTTCAAAAAGAAAGCTTGTTTGCTGGATTAAAGTAGAGGTGGATATTTCAT  
 CAACGTTGGATGATTATTGAGTCAAATGCAGGTGTGCGAATTTAATCAACCATTGGCAGATGTGTGTACC  
 TTCATCACGTCGATAATATTGAGCAAAATATTGCCTTATGTTCAAAAAAATAAGACGCAAAACCATATCTAT  
 TTTTCCTTGGGAGTCATCATTTAAATATCGAGATGCGTTAGCATTAAGGTTGCGGAAAGGATTGTAGAA  
 GCAGGAATGAATAACATATTTTCGAGTTGGTGGATCTCATGACGGAATGAGACCGTTGCAACGATTAGTGA  
 CATATATTTCTCATGAAAGGCCATCTAACTATACGGCTAAGGATGTTGCGGTTGAAATAGAACAGACTCG  
 ATTCCTGGAAGAAGATAAGTTCCCTTGATTTGTCCCATATAGGTAAAGTATGGAAAATGAATCAAAAT  
 ATAAAACATCGACCAGCTTATTTGTGTTGAAGGAATAAAAAAATTCATGTTTGGGAACGTTGCCAGA  
 AGAAAAACAGCCCAAAGAGAAAGAATGCCATTATTATTGCGTCTGGTTTTGCCCCGAGGATGGATCATTTT  
 GCTGGTCTGGCGGAATATTTATCGCGGAATGGATTTTCATGTGATCCGCTATGATTCGCTTACCACGTTG  
 GATTGAGTTTCAGGACAATTGATGAATTTACAATGTCTATAGGAAAGCAGAGCTTGTTAGCAGTGGTTGA  
 TTGGTTAACTACACGAAAAATAAATAAATTCGGTATGTTGGCTTCAAGCTTATCTGCCGGATAGCTTAT  
 GCAAGCCTATCTGAAATCAATGCTTCGTTTTTAATCACCGCAGTCGGTGTTGTTAACTTAAGATATTCCTC  
 TTGAAAGAGCTTTAGGGTTTGATTATCTCAGTCTACCCATTAATGAATTGCCGGATAATCTAGATTTTGA  
 AGCCATGAATTTGGGTGCTGAAGCTTTTGCAGAGATTTGCTTGGATTGTTGGTTGGGAACGTTAGCTTCT  
 ACAATTAATAACATGATGTATCTTGATATACCGTTTATTGCTTTTACTGCAATAACGATAATTGGGTCA  
 AGCAAGATGAAGTTATCACATTGTTATCAAATATTTCGTAGTAATCGATGCAAGATATATTCTTTGTTAGG  
 AAGTTCGCATGACTTGAGTGAAAAATTTAGTGGTCTGCGCAATTTTATCAATCGGTTACGAAAGCCGCT  
 ATCGCGATGGATAATGATCATCTGGATATTGATGTTGATATTACTGAACCGTCATTGGAACATTTAACTA  
 TTGCGACAGTCAATGAACGCCGAATGAGAATTGAGATTGAAAAATCAAGCAATTTCTCTGTCTTAAATCT  
 ATTGAGATATTTCTATCACTCAAATAGCAATATAAGGACTCTCTATGAAATTTGGAAACCTTTTGGCTTACA  
 TACCAACCTCCCCAATTTTCTCAAACAGAGGTAATGAAACGTTTGGTTAAATTAGGTCGCATCTCTGAGG  
 AGTGTGGTTTTGATACCGTATGGTTACTGGAGCATCATTTACGGAGTTTGGTTTGGTTGGTAACCCCTTA  
 TGTCGCTGCTGCATATTTACTTGGCGCGACTAAAAAATTGAATGTAGGAAGTCCGCTATTGTTCTTCCC  
 ACAGCCCATCCAGTACGCCAATTTGAAGATGTGAATTTATTGGATCAAATGTCAAAAGGACGATTTCCGGT  
 TTGGTATTTGCCGAGGGCTTTACAACAAGGACTTTCGCGTATTCGGCACAGATATGAATAACAGTCGCGC  
 CTTAGCGGAATGCTGGTACGGGCTGATAAAGAATGGCATGACAGAGGGATATATGGAAGCTGATAATGAA  
 CATATCAAGTTCCATAAGGTAAAGTAAACCCCGCGGTATAGCAGAGGTGGCGACCCGGTTATGTTGG  
 TGGCTGAATCAGCTTCGACGACTGAGTGGGCTGCTCAATTTGGCCTACCGATGATATTAAGTTGGATTAT  
 AAATACTAACGAAAAAGAAAGCACAACTTGAGCTTTATAATGAAGTGGCTCAAGAATATGGGCACGATATT  
 CATAATATCGACCATTTGCTTATCATATATAACATCTGTAGATCATGACTCAATTAAGCGAAAGAGATTT  
 GCCGGAATTTCTGGGCAATTGGTATGATTCTTATGTGAATGCTACGACTATTTTTGATGATTTCAGACCA  
 AACAAGAGGTTATGATTTCAATAAAGGGCAGTGGCGTGACTTTGTATTAAAGGACATAAAGATACTAAT  
 CGCGTATTGATTACAGTTACGAAATCAATCCCGTGGGAACGCCGAGGAATGTATTGACATAATTCAAA  
 AAGACATTGATGCTACAGGAATATCAAATATTTGTTGTGGATTGAAAGCTAATGGAACAGTAGACGAAAT  
 TATTGCTTCCATGAAGCTCTTCCAGTCTGATGTATGCCATTTCTTAAAGAAAAACAACGTTTCGCTATTA  
 TATTAGCTAAGGAGAAAGAAATGAAATTTGGATTGTTCTTCTTAACTTCATCAATTAACAACGTTTCA  
 AGAACAAAGTATAGTTTCGCATGCAGGAAATAACGGAGTATGTTGATAAGTTGAATTTTGAACAGATTTTA

GTGTATGAAAATCATTTTTTCAGATAATGGTGTGTGTCGGCGCTCCTCTGACTGTTTTCTGGTTTTCTGCTCG  
 GTTTAACAGAGAAAATTTAAATTTGGTTCATTAAATCACATCATTACAACTCATCATCCTGTGCGCCATAGC  
 GGAGGAAGCTTGCTTATTGGATCAGTTAAGTGAAGGGAGATTTATTTTAGGGTTTAGTGATTGCGAAAAA  
 AAGATGAAATGCATTTTTTTAATCGCCCGTTGAATATCAACAGCAACTATTTGAAGAGTGTTATGAAA  
 TCATTAACGATGCTTTAACAACAGGCTATTGTAATCCAGATAACGATTTTTATAGCTTCCCTAAAAATATC  
 TGTAAATCCCCATGCTTATACGCCAGGCGGACCTCGGAAATATGTAACAGCAACCAGTCATCATATTGTT  
 GAGTGGGCGGCCAAAAAGGTATTCTCTCATCTTTAAGTGGGATGATTCTAATGATGTTAGATATGAAT  
 ATGCTGAAAGATATAAAGCCGTTGCGGATAAATATGACGTTGACCTATCAGAGATAGACCATCAGTTAAT  
 GATATTAGTTAACTATAACGAAGATAGTAATAAAGCTAAACAAGAGACGCGTGCATTTATTAGTGATTAT  
 GTTCTTGAAATGCACCCTAATGAAAATTTGAAAAATAACTTGAAGAAATAATTGCAGAAAACGCTGTCG  
 GAAATTATACGGAGTGTATAACTGCGGCTAAGTTGCAATTGAAAAGTGTGGTGCGAAAAGTGATTGCT  
 GTCCTTTGAACCAATGAATGATTTGATGAGCCAAAAAATGTAATCAATATTGTTGATGATAATATTAAG  
 AAGTACCACATGGAATATACCTAATAGATTTTCGAGTTGCAGCGAGGCGGCAAGTGAACGAATCCCCAGGA  
 GCATAGATAACTATGTGACTGGGGTGAGTGAAAGCAGCCAAACAAAGCAGCAGCTTGAAAGATGAAGGGTA  
 TAAAAGAGTATGACAGCAGTGCTGCCATACTTTCTAATATTATCTTGAGGAGTAAACAGGTATGACTTC  
 ATATGTTGATAAACAAGAAATTACAGCAAGCTCAGAAATTGATGATTTGATTTTTTCGAGCGATCCATTA  
 GTGTGGTCTTACGACGAGCAGGAAAAATCAGAAAGAACTTGTGCTTGATGCATTTTCGTAATCATTATA  
 AACATTGTGCGAGAATATCGTCACTGTGTCAGGCACACAAAGTAGATGACAATATTACGGAAATTGATGA  
 CATACCTGTATTCCCAACATCGGTTTTTAAGTTTACTCGCTTATTAACCTCTCAGGAAAACGAGATTGAA  
 AGTTGGTTTTACCAGTAGCGGCACGAATGGTTTTAAAAAGTCAGGTGGCGCGTGACAGATTAAGTATTGAGA  
 GACTCTTAGGCTCTGTGAGTTATGGCATGAAATATGTTGGTAGTTGGTTTGATCATCAAAAGAAATTAGT  
 CAATTTGGGACCAGATAGATTTAATGCTCATAATATTTGGTTTAAATATGTTATGAGTTTGGTGGAATTG  
 TTATATCCTACGACATTTACCGTAACAGAAGAACGAATAGATTTTGTAAACATTTGAATAGTCTTTGAAC  
 GATTAATAAATCAAGGGAAGATCTTTGTCTTATTGGTTTCGCCATACTTTATTTACTCTGCCATTATA  
 TATGAAAGATAAAAAAATCTCATTTTCTGGAGATAAAAGCCTTTATATCATAACCGGAGCGGCTGGAAA  
 AGTTACGAAAAAGAATCTCTGAAACGTGATGATTTCAATCATCTTTTATTTGATACCTTCAATCTCAGTG  
 ATATTAGTCAGATCCGAGATATATTTAATCAAGTTGAACTCAACACTTGTTCCTTTGAGGATGAAATGCA  
 GCGTAAACATGTTCCGCCGTGGGTATATGCGCGAGCGCTTGATCCTGAAACGTTGAAACCTGTACCTGAT  
 GGAACGCCGGGGTTGATGAGTTATATGGATGCGTCAGCAACCAGTTATCCAGCATTTATTTGTTACCGATG  
 ATGTCGGGATAATTAGCAGAGAATATGGTAAGTATCCCGCGCTGCTCGTTGAAATTTTACGTCGCGTCA  
 TACGAGACGCGAGAAAGGGTGTGCTTTAAGCTTAAACCGAAGCGTTTGATAGTTGATATCCTTTGCTTGC  
 GCCGCAACTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCTGTTTTATCTGTTG  
 TTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCCCTAGACCTAGCCTTAATTAAGGCTAGGG  
 TACGGGTTTTGTGCTGCCGCAAACGGGCTGTTCTGGTGTGCTAGTTTGTATCAGAAATCGCAGATCCGGC  
 TTCAGCCGGTTTGCCGGCTGAAAGCGCTATTTCTTCCAGAATTGCCATGATTTTTTCCCCACGGGAGGCG  
 TCACTGGCTCCCGTGTGTGCGGCAGCTTTGATTCGATAAGCAGCATCGCCTGTTTCAAGGCTGTCTATGTG  
 TGACTGTTGAGCTGTAACAAGTTGTCTCAGGTGTTCAATTTTCATGTTCTAGTTGCTTTGTTTTACTGGTT  
 TCACCTGTTTCTATTAGGTGTTACATGCTGTTTCATCTGTTACATTGTGATCTGTTTCAACAGCTT  
 TGAATGCACCAAAACTCGTAAAGCTCTGATGTATCTATCTTTTTTACACCGTTTTTCATCTGTGCATAT  
 GGACAGTTTTTCCCTTTGATATGTAACGGTGAACAGTTGTTCTACTTTTGTGTTGTTAGTCTTGATGCTTCA  
 CTGATAGATACAAGAGCCATAAGAACCCTCAGATCCTTCCGTATTTAGCCAGTATGTTCTCTAGTGTGGTT  
 CGTTGTTTTTGGCTGAGCCATGAGAACGAACCATTGAGATCATACTTACTTTGCATGTCACTCAAAAAT  
 TTGCTCAAAAACCTGGTGAGCTGAATTTTTGTCAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTCCGTTATG  
 TAGGTAGGAATCTGATGTAATGGTTGTTGGTATTTTGTCAACATTCAATTTTTTATGTTGTTTCAAGT  
 TCGGTTACGAGATCCATTTGTCTATCTAGTTCAACTTGGAAAATCAACGTATCAGTCGGGCGGCCTCGCT  
 TATCAACCACCAATTTTCATATTGCTGTAAGTGTTTAAATCTTTACTTATTGGTTTTCAAAACCCATTGGTT  
 AAGCCTTTTTAACTCATGGTAGTTATTTTCAAGCATTAACATGAACCTAAATTCATCAAGGCTAATCTCT  
 ATATTTGCCTTGTGAGTTTTCTTTTGTGTTAGTTCTTTTAAATAACCACTCATAAATCCTCATAGAGTATT  
 TGTTTTCAAAAGACTTAACATGTTCCAGATTATATTTTATGAATTTTTTAACTGGAAAAGATAAGGCAA  
 TATCTCTTCACTAAAACTAATTCTAATTTTTCTGCTGAGAACTTGGCATAGTTTGTCCACTGGAAAAATC  
 TCAAAGCCTTTTAAACCAAGGATTCCTGATTTTCCACAGTTCTCGTCATCAGCTCTCTGTTGCTTTAGCTA  
 ATACACCATAAGCATTTTTCCCTACTGATGTTTCATCATCTGAGCGTATTGGTTATAAGTGAACGATACCGT  
 CCGTCTTTTCCCTTGTAGGGTTTTCAATCGGGGGGTTGAGTAGTGCCACACAGCATAAAATTAGCTTGGTT  
 TCATGCTCCGTTAAGTCATAGCGACTAATCGCTAGTTTCATTTGCTTTGAAAACAACCTAATTCAGACATAC  
 ATCTCAATTGGTCTAGGTGATTTTAACTACTATACCAATTGAGATGGGCTAGTCAATGATAATTACTAGC  
 CTTAATTAAGGTAGTCCTTTTCCCTTTGAGTTGTTGGGTATCTGTAAATCTGCTAGACCTTTGCTGGAAAA  
 CTTGTAAATTTCTGCTAGACCTCTGTAAATTCGCTAGACCTTTGTGTGTTTTTTTTTGTGTTATATTCAA  
 GTGGTTATAATTTATAGAATAAAGAAAAGATAAAAAAAGATAAAAAAGAATAGATCCAGCCCTGTGTAT  
 AACTCACTACTTTAGTCAGTTCCGAGTATTACAAAAGGATGTCGCAACGCTGTTTGCTCCTCTACAAA  
 ACAGACCTTAAACCTTAAAGGCTTAAAGTAGCACCTCGCAAGCTCGGGCAAATCGTGAATATTCCCTTT  
 TGTCTCCGACCATCAGGCACCTGAGTCGCTGTCTTTTTCTGTCGACATTCAGTTTCGCTGCGCTCACGGCTCT  
 GGCAGTGAATGGGGTAAATGGCACTACAGGCGCCTTTTATGGATTTCATGCAAGGAACTACCCATAATA  
 CAAGAAAAGCCGTCACGGGCTTCTCAGGGCGTTTTATGGCGGGTCTGCTATGTGGTGCTATCTGACTTT  
 TTGCTGTTTCAGCAGTTTCTGCTGCCCTGATTTTTCCAGTCTGACCACTTCGGATTATCCCGTGACAGGT  
 TCAGACTGGCTAATGCACCCAGTAAGGCAGCGGTATCATCAACAGGCTTACCCGCTTACTGTCCCTAGT  
 GCTTGGATTCTCACCAATAAAAAACGCCCCGGCGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCT  
 GAGGTCATTACTGGATCTATCAACAGGAGTCCAAGCGAGCTCTCGAACCCAGAGTCCCGCTCAGAAAGAA  
 CTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGCGGATACCGTAAAGCACGAGGAAG

CGGTCAGCCCATTGCGCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGT  
CCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAA  
GCAGGCATCGCCATGAGTCACGACGAGATCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGT  
TCGGCTGGCGCGAGCCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAG  
TACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATGGGCAGGTAGCCGGATCAAGCGTATGCAG  
CCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGC  
CCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAG  
GAACGCCCCGTGCTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTCAGGGCACCGGACAG  
GTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG  
ATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATC  
CATCTTGTTCAATCAT

9.1.5. pDOC-*zntA:rfp* Sequence

AACGCCAGCAACGCGGCCTTTTTACGGTTCTTGCCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTCTG  
CGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCCG  
AACGACCCGAGCGCAGCGAGTCAGTGAGCGAGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCC  
GCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGCAGTGAGCGCA  
ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTAT  
GTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTC  
TAGGGATAACAGGGTAATCGATGGCTGGTGACAGCGAATGCGTTAAGATTGTTGCGCAGGAGAATGGCAA  
GTAGTGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTTCGTATGGAAGGTTCCGTTAACGGTCACGA  
GTTGCAAAATCGAAGGTGAAGGTGAAGGTGCTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACC  
AAAGGTGGTCCGCTGCCGTTTCGCTTGGGACATCCTGTCCCGCAGTTCCAGTACGGTTCCAAAGCTTACG  
TTAAACACCCGGCTGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGT  
TATGAACTTCGAAGACGGTGGTGTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATC  
TACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGTTATGCAGAAAAAACCATGGGTT  
GGGAAGCTTCCACCGAACGTATGTACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAATGCGTCTGAA  
ACTGAAAGACGGTGGTCACTACGACGCTGAAGTTAAACCACCTACATGGCTAAAAAACCGGTTACAGCTG  
CCGGGTGCTTACAAAACCGACATCAAACCTGGACATCACCTCCCAACGAAGACTACACCATCGTTGAAC  
AGTAGCAACGTGCTGAAGGTGCTCACTCCACCGGTCTTAAGAAGTTCCCTATACCTTTCTAGAGAATAGGA  
ACTTCGGAATAGGAACTTCAAGATCCCCACGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGA  
AGCGGAACACGTAGAACTTAAGGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAA  
AGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACA  
GGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGA  
GGCTATTTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTGACG  
GCAGGGGCGCCCGTCTTTTTGTCAAGACCGACTTCCGGTGCCCTGAATGAACGACGAGACGGCA  
GCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGG  
GAAGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTCTCACCTTGCTCCTGCCGA  
GAAAGTATCCATCATGGCTGATGCAATGCGGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGAC  
CACCAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATC  
TGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACCTGTTCCGCCAGGCTCAAGGCGCGCATGCCCGACGG  
CGAGGATCTCGTCGTGACCATGGCGATGCTGCTGTTGCCGAATATCATGGTGGAAAAATGGCCGCTTTTCT  
GGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCCGCTATCAGGACATAGCCTTGCTACCGTGATA  
TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTGCTTTACGGTATCGCCGCTCCCGATT  
GCAGCGCATCGCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTCGAAATGACCG  
ACCAAGCGACGCCCAACCTGCCATCACGAGATTTGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTT  
CGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCC  
CACCACAGCTTCAAAGCGCTCTGAAGTTCTTATACTTTCTAGAGAATAGGAACCTCGGCCAAACCGATCG  
CAACTGATGAGCGCATCGGTCCCTCGTAGGGATAACAGGGTAATGAGCTTGGCACTGGCCGTGCGTTTTTA  
CAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTTCGCCA  
GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATG  
GCGAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAATCAGAACGCAGAAG  
CGGTCTGATAAAACAGAATTTGCCGTGGCGGCAGTAGCGCGGTGGTCCACCTGACCCCATGCCGAACCTCA  
GAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAGTCCAGGCAT  
CAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTGTTTTATCTGTTGTTGTCGGTGAACGCTC  
TCTGATTAGGACAAATCCGCCGGGAGCGGATTGAACGTTGCGAAGCAACGCCGCGGGGTGCGGGC  
AGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCG  
TTTCTACAACTCTTTTTGTTTTATTTTCTAAATACATTCAAATATGCATGCGCCTGATGCGGTATTTTC  
TCCTTACGCATATCGACATCCGCCCTCACCGCCAGAACGCAACCGCAGCCTCATACGCCGGCGCTTCT  
TGGCCGCGCGGGATTCAACCCACTCGGCCAGCTCGTCGGTGTAGCTCTTTGGCATCGTCTCTCGCCTGTC  
CCCTCAGTTTCAGTAATTTCTTGCATTTGCCTGTTTCCAGTCGGTAGATATTCACAAAAACAGCAGGGAAG  
CAGCGCTTTTTCCGCTGCATAACCTGCTTCGGGTCAATTATAGCATTTTTTTCGGTATATCCATCCTTTT  
TCGCACGATATACAGGATTTTGGCCAAAGGTTTCGTAGACTTTCTTGGTGTATCCAAACGGCGTCAGCC  
GGGCAGGATAGGTGAAGTAGGCCACCCGCGAGCGGGTGTTCCTTCTTCACTGTCCCTTATTGCGACCTG  
GCGGTGCTCAACGGGAATCCTGCTCTGCGAGGCTGGCCGGCTACCGCCGGCGTAACAGATGAGGGCAAGC  
GGATGGCTGATGAAACCAAGCCAACAGGAAGGGCAGCCACCTATCAAGGTGTACTGCCTTCCAGACGA  
ACGAAGAGCGATTGAGGAAAAGGCGGCGGCGCGGCATGAGCCTGTCCGCCTACCTGCTGGCCGTCCGC  
CAGGGCTACAAAATCACGGGCGTCTGGACTATGACACAGTCCGCGAGCTGGCCCGCATCAATGGCGACC  
TGGCCGCTGGGCGGCTGCTGAACTCTGGCTCACCGACGACCCGCGCAGCAGCGCGGTTCGGTGATGC  
CACGATCCTCGCCCTGCTGGCGAAGATCGACTCTAGCTAGAGGATCGATCCTTTTTAACCCATCACATAT  
ACCTGCCGTTCACTATTATTTAGTGAAATGAGATATTATGATATTTTCTGAATTGTGATTAAGGCAAG  
CTTTATGCCCATGCAACAGAACTATAAAAAATACAGAGAATGAAAGAAACAGATAGATTTTTTTAGTTTC  
TTTAGGCCCGTAGTCTGCAAATCCTTTTATGATTTTCTATCAAACAAAAGAGGAAAATAGACCAGTTGCA  
ATCCAAACGAGAGTCTAATAGAATGAGGTGCAAAAGTAAATCGCGCGGGTTTGTACTGATAAAGCAGGC  
AAGACCTAAATGTGTAAAGGGCAAGTGATACCTTTGGCGTCACCCCTTACATATTTTAGGTCTTTTTT  
TATTGTGCGTAACCTAAGTGGCATCTTCAAACAGGAGGGCTGGAAGAAGCAGACCGCTAACACAGTACAT  
AAAAAAGGAGACATGAACGATGAACATCAAAAAGTTTGCAAAACAAGCAACAGTATTAACCTTTACTACC  
GCACTGCTGGCAGGAGGCGCAACTCAAGCGTTTGCAGAAAGAACGAACCAAAAGCCATATAAGGAAACAT  
ACGGCATTTCCTCATATTACACGCCATGATATGCTGCAAATCCCTGAACAGCAAAAAATGAAAAATATCA

AGTTCCTGAGTTCGATTCGTCCACAATTAAAAATATCTCTTCTGCAAAAGGCCTGGACGTTTGGGACAGC  
 TGGCCATTACAAAACGCTGACGGCACTGTGCGAACTATCACGGCTACCACATCGTCTTTGCATTAGCCG  
 GAGATCCTAAAAATGCGGATGACACATCGATTTACATGTTCTATCAAAAAGTCGGCGAACTTCTATTGA  
 CAGCTGGAAAAACGCTGGCCGCGTCTTTAAAGACAGCGACAAATTCGATGCAAAATGATTCTATCCTAAAA  
 GACCAAAACACAAGAAATGGTTCAGGTTTCAGCCACATTTACATCTGACGGAAAAATCCGTTTATTCTACACTG  
 ATTTCTCCGGTAAACATTACGGCAAAACAACTGACAACGTCACAAGTTAACGTATCAGCATCAGACAG  
 CTCTTTGAACATCAACGGTGTAGAGGATTATAATCAATCTTTGACGGTGACGGAAAAACGTATCAAAAT  
 GTACAGCAGTTCATCGATGAAGGCAACTACAGCTCAGGCGACAACCATACGCTGAGAGATCCTCACTACG  
 TAGAAGATAAAGGCCACAAATACTTAGTATTTGAAGCAAACTGGAACGAAAGATGGCTACCAAGGCCA  
 AGAATCTTTATTTAAACAAAGCATACTATGGCAAAAGCACATCATTCTCCGTCAAGAAAGTCAAAAACCTT  
 CTGCAAAAGCGATAAAAAACGCACGGCTGAGTTAGCAAAACGGCGCTCTCGGTATGATTGAGCTAAACGATG  
 ATTACACACTGAAAAAAGTGATGAAACCGCTGATTGCATCTAACACAGTAACAGATGAAATTGAACGCGC  
 GAACGTCTTTAAATGAACGGCAATGGTATCTGTTCACTGACTCCCGCGGATCAAAAATGACGATTGAC  
 GGCATTACGTCTAACGATATTTACATGCTTGGTTATGTTTCTAATTCTTTAACTGGCCCATACAAGCCGC  
 TGACAAAACCTGGCCTTGTGTTAAAAATGGATCTTGATCCTAACGATGTAACCTTTACTTACTCACACTT  
 CGCTGTACCTCAAGCGAAAGGAAACAATGTGCTGATTACAAGCTATATGACAAACAGAGGATTCTACGCA  
 GACAAACAATCAACGTTTGCGCCCTAGCTTCTGCTGAACATCAAGGCAAGAAAAACATCTGTTGTCAAAG  
 ACAGCATCCTTGAAACAAGGACAATTAACAGTTAAACAAATAAAAAACGCAAAAGAAAAATGCCGATTATGGTG  
 CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGAC  
 GCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCA  
 TGTGTCAGAGGTTTTCACCGTCATACCCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTT  
 ATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCCGGGAAATGTGCGCGGA  
 ACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAA  
 TGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCCGCTTATTCCCTTTTT  
 TGCGGCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAG  
 TTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCG  
 AAGAACGTTTTTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGC  
 CGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACA  
 GAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACA  
 CTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGG  
 GGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGAC  
 ACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACCTGGCGAACTACTTACTCTAGCTT  
 CCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCC  
 GGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTG  
 GGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAAC  
 GAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTGACACCAAGTTTACTC  
 ATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTTGAT  
 AATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA  
 AAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAAACCACCGCTACC  
 AGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCG  
 CAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAATCTGTAGCACCGC  
 CTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGG  
 GTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAG  
 CCCAGCTTGGAGCGAACGACCTACCCGAACGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGC  
 TTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGA  
 GCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGA  
 TTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAA

## 9.1.6. Sequence of Promoter Inserts for of pJI300 Plasmid Series

*P<sub>zntA</sub>* (pJI301)

TTTGCCGGTCACCTTCCTGATCGTCCGCTCGCTGTATCTCTGATAAACTTGA CTCTGGAGTCGACTCCAG  
AGTGTATCCTTCGGTTAATGAGAAAAA ACTTAACCGGAGGATGCCATGTCTGACTCCTGACAATCACGGCA  
AGAA

*P<sub>znuA</sub>* (pJI303)

GGCGTTGGCCAAAAGAAACCGAGACATTTTCCAGGGAACCGACTTGTCATGTTAATTTTAGTCTTGCA  
GTAGTCATGAAATGTTATAATATCACACTTCTCATATTTCATTACGATTATTGGTCGCATTATGTTACATA  
AAAAAACGCTTCTTTTCGCAGCATTATCCG

*P<sub>znuCB</sub>* (pJI304)

GAGCGGCGGATAATGCTGCGAAAAGAAGCGTTTTTTTTATGTAACATAATGCGACCAATAATCGTAATGAA  
TATGAGAAGTGTGATATTATAACATTTTCATGACTACTGCAAGACTAAAATTAACATGACAAGTCTGGTTT  
CCCTGG

*P<sub>zint</sub>* (pJI305)

ATAAGATAGATAAGTAGAACTGAGAAAGCCATGCTCTCGTTTCCCTAAGAGTTGTTGCATTTTGCTATATG  
TTACAATATAACATTACACATCATATACATTAACCTCGGAGGAACTGTTTTGGCGATTTCGTCTTTACAA  
ACTGGCTGTT

*P<sub>pliG</sub>* (pJI306)

AATGCCGCAACGTGATTTTACCGAACGCAGGAAAGCCACTGGAATTCTCCCGTTAGGGATTGTTATATTA  
TAACAGTTTCATCGTACTCATTCTGAACAGGAGACTACCAATGAAAATCAAGAGCATCAGGAAGGCTGTAT

*P<sub>ykgM</sub>* (pJI307)

GCTAACAATGCCAGAGTTCCCCGTTGCGAAAATTCCACATCCACAAAGAGTCACAGGGATTGAGTGTGGA  
AATGATCCGGATGAGCATGTATCTTTATGGTTATGTTATAACATAACAGGTAAAAATGATGAAGCCCAAT

*P<sub>c1265</sub>* (pJI308)

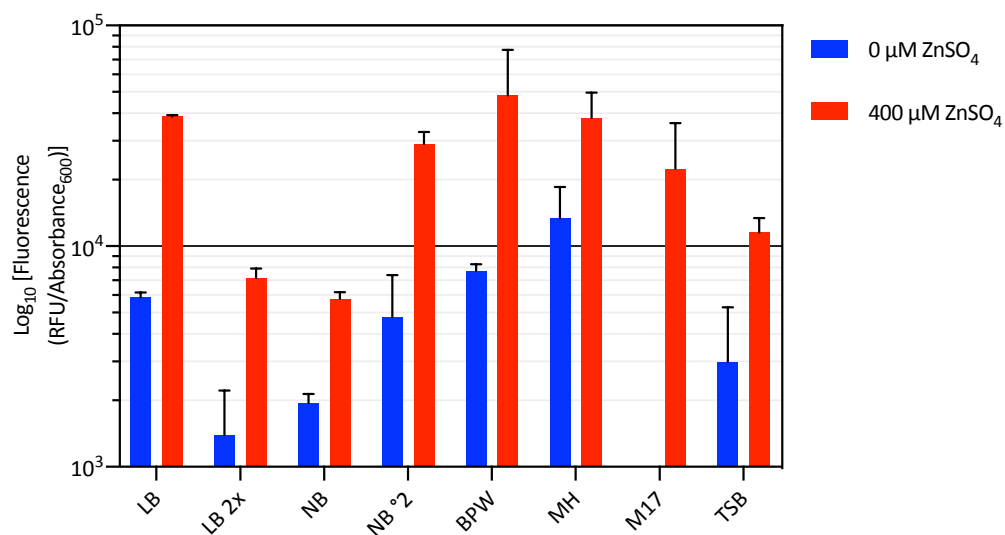
CGACAGAACATTATTAACAGAGCATCACTGAACGCTACATAATCAGAGTTGCATAAAATAAAATGTTATGA  
TATAACATCACAAATCACAACTTTGTTTTCGAATGTATATGAATGTAATCAGAAGTGCATTTGTAC

*P<sub>trc</sub>* (pJI302)

GAGGCCCTTTTCGTCTTCACCGCGCGAAGGCGAAGCGGCATGCATTTACGTTGACACCATCGAATGGTGCA  
AAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCACT  
AACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGGCC  
AGCCACGTTTTCTGCGAAAACGCGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTTCCCAACC  
GCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCA  
CGCGCCGTGCGAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGCG  
ATGGTAGAACGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTG  
GGCTGATCATTAACATATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCC  
GGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACG  
CGACTGGGCGTGGAGCATCTGGTTCGATTGGGTACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTT  
CTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGC  
GGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATC  
GTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCG  
GGCTGCGCGTGGTTCGCGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCC  
GCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACAGCGTGGACCGCTTGTCTGCAACTC  
TCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACACCCCTGG  
CGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTC  
CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCGCGAATTGATCTGGTTTGACAG  
CTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTG  
TGCAGGTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGC  
GCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATCATCCGGCTCGTATA  
ATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAG

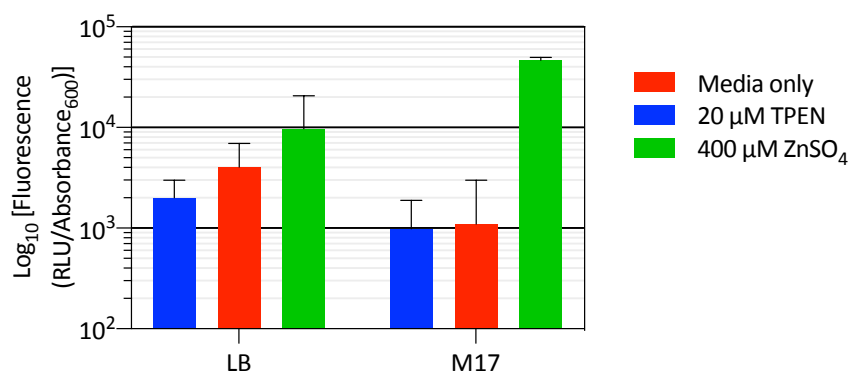
## 9.2. Supplementary Figures

## 9.2.1. Alternative Media



**Supplementary Figure 9.1 End Point Fluorescence Assay of ZntR Regulated  $P_{zntA}$  in Various Standard Growth Mediums**

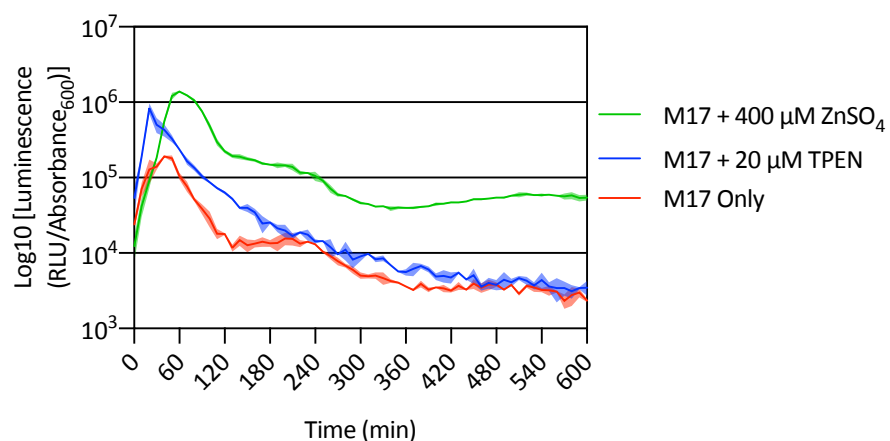
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in various media with 50 μg mL<sup>-1</sup> Kanamycin and with or without 400 μM ZnSO<sub>4</sub>, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D.



**Supplementary Figure 9.2 End Point Fluorescence Assay of ZntR Regulated  $P_{zntA}$  in LB and M17**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in LB or M17 with 50  $\mu\text{g mL}^{-1}$  Kanamycin and with or without 400  $\mu\text{M ZnSO}_4$  and 20  $\mu\text{M TPEN}$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D.

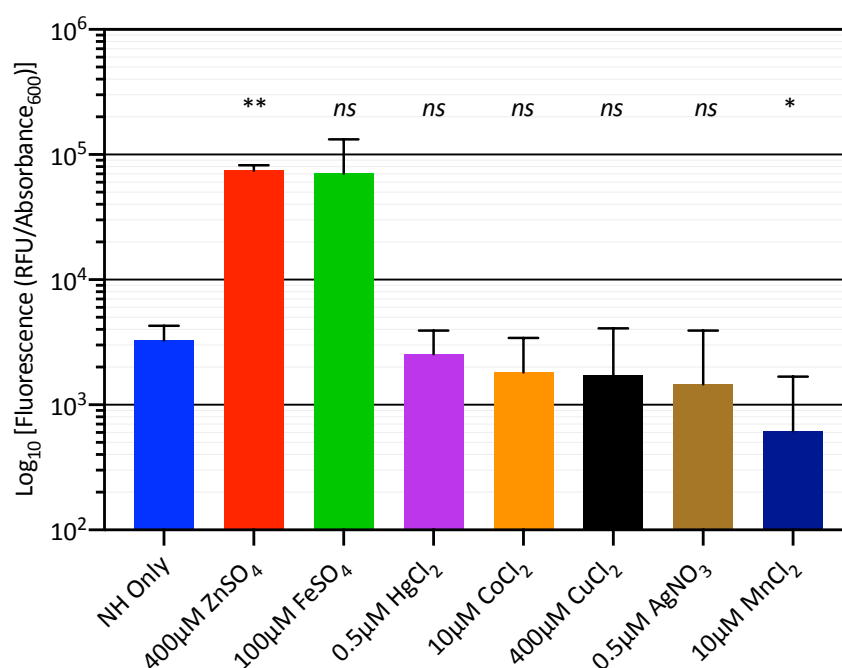




**Supplementary Figure 9.3 Temporal Luminescence of  $P_{zntA}$  in M17**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in M17<sub>kan50</sub>, with either 20  $\mu\text{M}$  TPEN or 400  $\mu\text{M}$  ZnSO<sub>4</sub>, with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 10 hr in a TECAN GENios Pro plate reader.

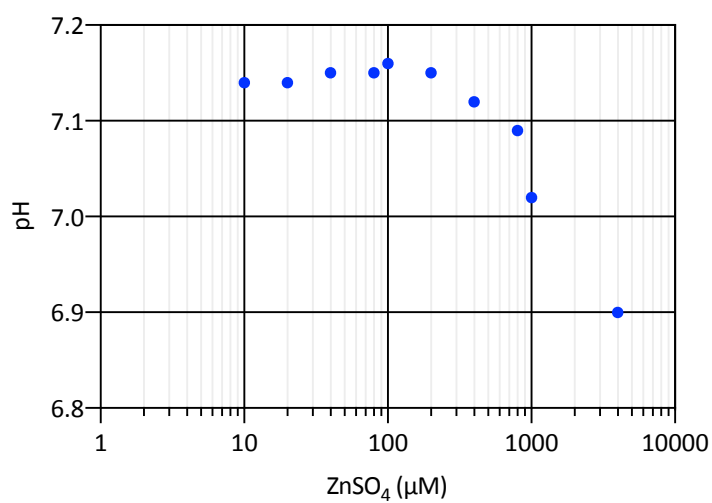
## 9.2.2. Alternative Metal Inducers



**Supplementary Figure 9.4 End Point Fluorescence Assay of ZntR Regulated  $P_{zntA}$  in NH with Various Metal Inducers**

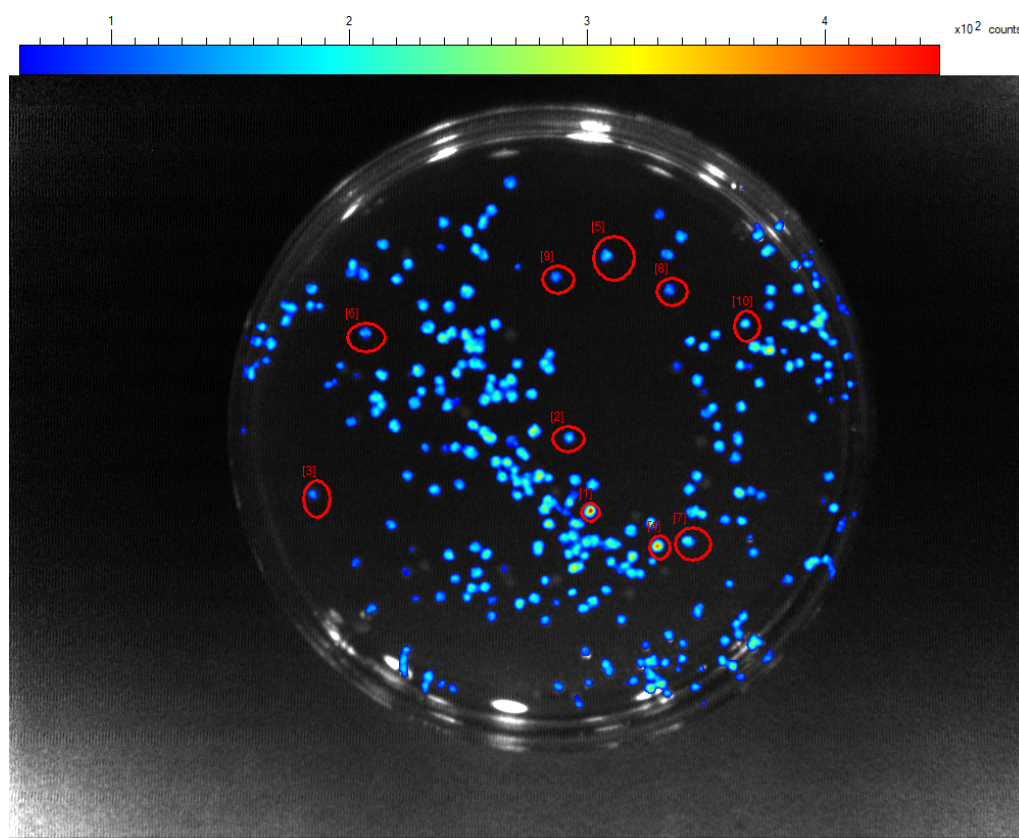
Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $NH_{kan50}$  various metal inducers, with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. \*  $p \leq 0.05$ , (two-tailed, unpaired, t-test) comparison of NH only to alternative metal inducer.

## 9.2.3. pH of NH Media

**Supplementary Figure 9.5 pH of NH with Increasing  $\text{ZnSO}_4$  concentrations**

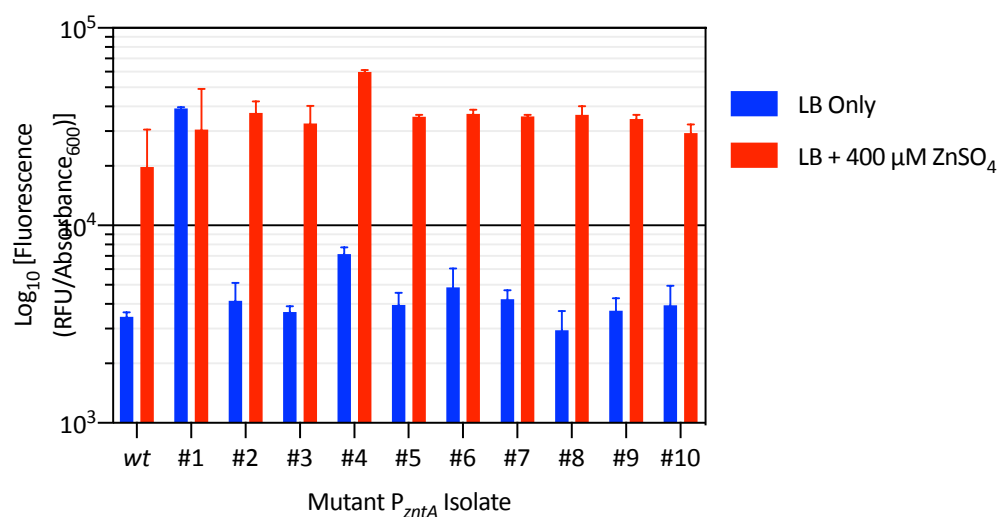
NH Media measure using a pH probe with the increasing concentrations of  $\text{ZnSO}_4$  ( $\mu\text{M}$ )

## 9.2.4. Error Prone PCR



**Supplementary Figure 9.6 Random Mutagenesis of  $P_{zntA}$ , Bioluminescence Image**

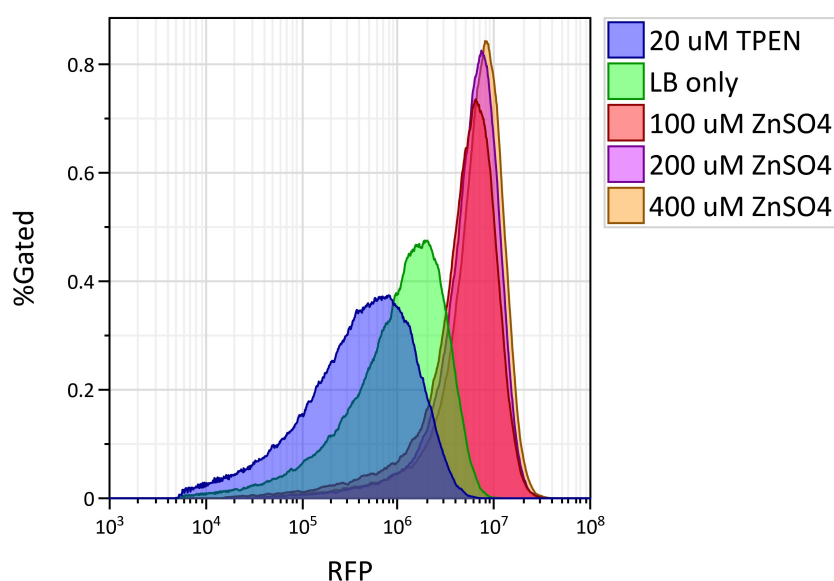
*E. coli* MG1655 transformed with potential randomly mutated  $P_{zntA}$  (pJI301) via error prone PCR and 10  $\mu$ L spread plate onto LB<sub>kan50</sub> + 200  $\mu$ M ZnSO<sub>4</sub> agar. Image and photon count (30-60 s exposure) measured on a Photon IMAGER (Biospace lab). Circled colonies were sub-cultured and end-point fluorescence assay conducted.



**Supplementary Figure 9.7 End Point Fluorescence Assay of Randomly Mutated  $P_{zntA}$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C with shaking (200 RPM) in  $\text{LB}_{\text{kan50}}$  with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. Isolates (#) selected from (Supplementary Figure 9.6) [note, sequence analysis only showed #1 to contain any mutation  $P_{zntA}$  (-22C<T)]

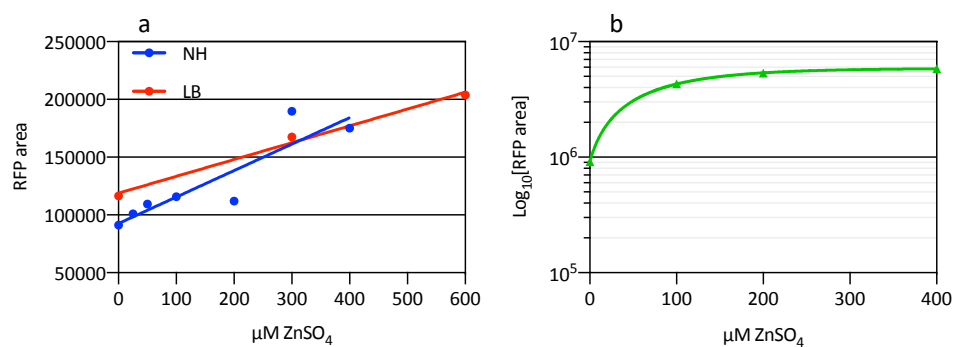
## 9.2.5. Flow Cytometry



	Marker	X-Med	X-Mode	X-GMean
<span style="color: blue;">■</span>	All	446,639	820,748	366,265
<span style="color: green;">■</span>	All	1,184,722	1,890,229	911,997
<span style="color: red;">■</span>	All	5,477,606	6,554,712	4,304,986
<span style="color: purple;">■</span>	All	6,331,291	7,917,442	5,351,654
<span style="color: orange;">■</span>	All	7,101,099	8,170,653	5,827,976

**Supplementary Figure 9.8 Histogram (Flow Cytometry Fluorescence Data) of  $P_{zntA}$  reporter Plasmid (pJI301) Expressed in LB**

$P_{zntA}$  reporter plasmid (pJI301) transformed and expressed in *E. coli* MG1655 with aerobically incubated at 37°C with shaking (200 RPM) in LB<sub>kan50</sub> with various conditions. After 16 hr incubation, 100 µL of bacterial culture was pelleted (13,000 x *g* for 1 min), re-suspended in 1000 µL 4% formaldehyde, and kept on ice. Samples were analysed using a Beckman Coulter Astrios EQ flow cytometer; excitation with a green (561 nm) laser with emission filter (614/620 nm), and an emissions bandwidth (604-624 nm). Cells have been gated for: injection rate; whole cells; single cells; and fluorescence.



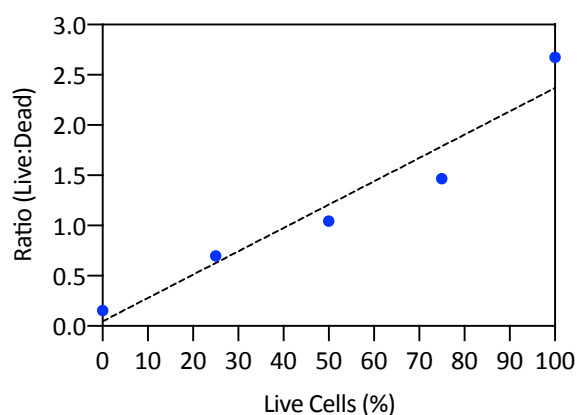
**Supplementary Figure 9.9 Regression Analysis: Flow Cytometry data**

Regression analysis of flow cytometry fluorescence (median) compared to  $\mu\text{M ZnSO}_4$  concentration. (a) *E. coli zntA:rfp* grown in either NH or LB with increasing  $\text{ZnSO}_4$  concentrations. (b) *E. coli wt* transformed with pJI301 ( $P_{zntA}$ ) grown in LB with increasing  $\text{ZnSO}_4$  concentration.

Linear regression of; *E. coli zntA:rfp* in LB  $R^2 = 0.99$ , *E. coli zntA:rfp* in NH  $R^2 = 0.81$ .

Non-linear regression of *E. coli wt* + pJI301 in LB  $R^2 = 1$ .

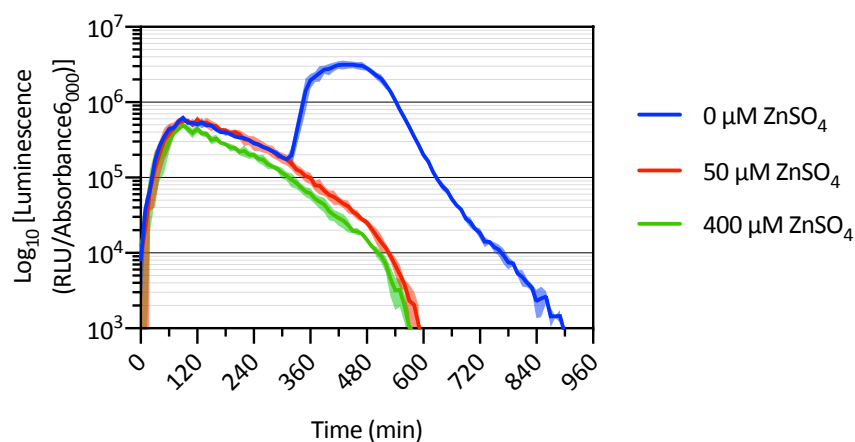
## 9.2.6. BacLight Control

**Supplementary Figure 9.10 BacLight Standard Curve**

*E. coli* MG1655 was inoculated into 20 mL LB broth, with three biological repeats, and incubated at 37°C with shaking (200 RPM) until OD<sub>600</sub> 0.4 was reached and incubated for a further 2 hr. LIVE/DEAD BacLight kit was used to determine live:dead ratio. Standard curve comprised of known ratio of live and dead cells (ethanol treated). Live fluorescence (ex/em 485/535 nm), dead fluorescence (ex/em 485/620 nm), and absorbance (600 nm) was recorded on a TECAN GENious Pro.



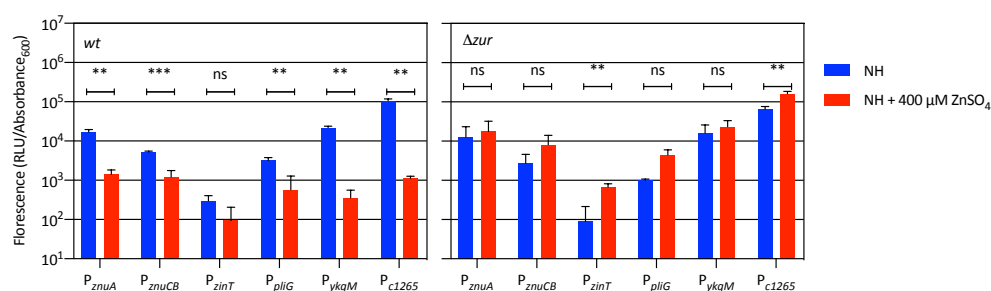
## 9.2.7. Alternative Temporal Luminescence Graphs



**Supplementary Figure 9.11 Temporal Luminescence Assay of  $P_{c1265}$  in  $NH$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 with aerobic incubation at 37°C, with shaking (200 RPM) in  $NH_{kan50}$  with  $ZnSO_4$  (0, 50, 400  $\mu M$ ), with three biological repeats. Absorbance<sub>600</sub> and Luminescence (100 ms integration) was recorded every 10 min for 16 hr in a TECAN GENios Pro plate reader.

## 9.2.8. Additional End Point Fluorescence Assays



**Supplementary Figure 9.12 End Point Fluorescence of Zur Regulated Promoter in NH, Expressed in *E. coli* wt and  $\Delta zur$**

Plasmid constructs transformed and expressed in *E. coli* MG1655 wt and  $\Delta zur$  with aerobic incubation at 37°C with shaking (200 RPM) in NH<sub>kan50</sub> with or without 400  $\mu\text{M}$   $\text{ZnSO}_4$ , with three biological repeats. Absorbance<sub>600</sub> and fluorescence was recorded after 16 hr incubation using a TECAN GENios Pro. Error bars indicated S.D. ns  $p \geq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , (two-tailed, unpaired, t-test).

### 9.3. Manuscript in Production for Submission

Title: The Zur Regulon of Pathogenic and Non-Pathogenic *Escherichia coli*

**Authors:** Selina R. Clayton, Joseph O. Ingram, Chrystala Constantinidou, Mala Patel, Taku Oshima, Tomoyo Watari, Karin Heurlier, Dov J. Stekel, and Jon L. Hobman

#### Abstract

Zinc is an essential micronutrient needed in all forms of life and is a co-factor in all six classes of enzymes. *Escherichia coli* tightly controls its import of zinc through the regulon Zur, an active transcriptional repressor in the presence of zinc, which regulates *znuABC*, *zint*, and *ykgMO*. Genome-wide searches predicted a conserved Zur regulon common between non-pathogenic and pathogenic *E. coli* (K-12, O157:H7 and uropathogenic strain CFT073) which is composed of established Zur regulated genes (*znuABC*, *zint*, and *ykgMO*), and others such as *pliG*, which have been identified in a limited number of previous studies. The Regulatory Sequence Analysis tool (RSAT) Patser predicted undiscovered Zur regulated genes in the urinary pathogenic *E. coli* strain CFT073, *c1265-7*, which are only found in extra-intestinal *E. coli*. Genome-wide transcriptomics data supports Zur regulation of the core Zur regulon and *c1265-7*: the *c1265-7* operon has a significant increase in expression under both zinc-depleted conditions and in  $\Delta$ *zur* strains (up to 61-fold decrease). Zur regulation of *c1265-7* is further confirmed through *c1265* promoter assays, which show significantly increased promoter activity in zinc depleted conditions and  $\Delta$ *zur* strains. *C1265* is predicted to resemble a TonB-dependent outer membrane receptor, whilst *C1267* is a member of the COG0523 family (GTPase).

The Zur regulon of *E. coli* consists of a small conserved group of genes found in both pathogenic and non-pathogenic strains, as well as Zur regulated genes present only in some pathotypes. A novel Zur regulated operon, *c1265-7*, from a uropathogenic *E. coli* has been identified. The *c1265-7* operon may be a specific adaptation that enables ExPEC strains to survive and grow outside of the gut.

#### Introduction

Zinc is an essential micronutrient for all forms of life and is a co-factor for a large number of bacterial enzymes; zinc is present in 4.9% of the bacterial proteome (Andreini *et al.*, 2006). However, at high concentrations it becomes toxic to the cell. In humans an excessive intake of zinc is known to cause copper deficiency; similarly the mechanism of zinc toxicity in bacteria may be due to disruption of the homeostasis of other essential metals (Mcdevitt *et al.*, 2011). In most bacteria the internal zinc concentration is controlled at the level of transcription of zinc export and import systems.

In *Escherichia coli*, there are several known zinc import and export systems. The Zur transcription factor is involved in high affinity uptake of zinc (Patzer and Hantke, 1998) whilst a zinc efflux ATPase, ZntA, exports zinc out of the periplasm (Rensing, Mitra and Rosen, 1997). Subsidiary lower affinity zinc import and export systems are also known; ZupT is a constitutively expressed broad range metal ion importer belonging to the ZIP family of transporters, with a slight preference for  $\text{Zn}^{2+}$  (Grass *et al.*, 2002; Grass, Franke, *et al.*, 2005; Taudte and Grass, 2010). PitA (Beard *et al.*, 2000) and MntH (Makui *et al.*, 2000) are also involved in low affinity zinc uptake. The cation diffusion

facilitator (CDF) family exporter ZitB is a zinc inducible exporter (Grass *et al.*, 2001) whilst another member of the CDF family, YiiP has been implicated in zinc export even though it's primary substrate is  $\text{Fe}^{2+}$  (Grass, Otto, *et al.*, 2005; Lu and Fu, 2007). ZraP is part of the ZraPSR operon, a two-component signal transduction system which responds to zinc overload; it has been shown to bind zinc, but along with ZraSR is not directly involved in zinc resistance, rather it may sense zinc as a marker for envelope stress (Leonhartsberger *et al.*, 2001; Sevcenco *et al.*, 2011; Petit-Härtlein *et al.*, 2015).

The zinc uptake regulator, Zur, is a transcriptional repressor that is active in the presence of zinc and sterically hinders RNA polymerase binding to the promoter region (Outten and O'Halloran, 2001). Zur has been shown to bind to three putative sites in *E. coli* K-12, regulating six genes (*znuABC*, *zinT*, and *ykgMO*) based on a combination of bioinformatics predictions and experimental evidence in *E. coli* and other bacteria (Patzner and Hantke, 1998; Panina, Mironov and Gelfand, 2003; Hemm *et al.*, 2010; Petrarca *et al.*, 2010; Gabbianelli *et al.*, 2011). ZnuABC is a high-affinity ABC-type zinc uptake system, with an auxiliary component, ZinT, which together act as the major source of zinc uptake in the cell (Patzner and Hantke, 1998; Petrarca *et al.*, 2010). YkgMO are alternative zinc-free ribosomal proteins that are thought to replace the zinc-binding ribosomal proteins L31 and L36 under zinc starvation conditions releasing a reservoir of zinc, as is observed in *Bacillus subtilis* (Natori *et al.*, 2006; Gabriel and Helmann, 2009; Hensley, Tierney and Crowder, 2011). When zinc levels in the cell are too high an alternative regulator, ZntR, from the MerR family, activates the expression of a P-type zinc-exporter ZntA (Brocklehurst *et al.*, 1999). An excess of zinc binding sites in the cell is thought to keep the level of free-zinc in the femtomolar range, and therefore the zinc-binding regulators Zur and ZntR must possess an incredibly sensitive zinc-binding ability to maintain the delicate balance between zinc export and import (Outten and O'Halloran, 2001). Recent mathematical modelling of experimental data from several zinc import and export mutants suggests that *E. coli* populations mount a heterogeneous response to zinc stress (Takahashi *et al.*, 2015)

The availability of zinc has been shown to be an important factor in the ability of microorganisms to colonise the host. In many bacteria the loss of ZnuABC reduces pathogenicity and colonisation ability (Campoy *et al.*, 2002; Davis, Kakuda and DiRita, 2009). ZnuABC are also found to be highly expressed in intracellular environments in *Salmonella enterica* serovar Typhimurium, where they contribute to growth in Caco-2 epithelial cells and within phagocytes (Ammendola *et al.*, 2007), and in adhesion to cultured epithelial cells by *E. coli* O157:H7  $\Delta znuABC$  (Gabbianelli *et al.*, 2011). There is also evidence of host restriction of the availability of zinc to invading bacteria, known as nutritional immunity; in response to infection, levels of zinc in the host bloodstream are seen to fall. Additionally, an S100 host protein, calprotectin, which is able to chelate zinc and manganese, and is the most abundant protein found in neutrophils, accumulates in tissue abscesses, and has been shown to restrict the growth of *Staphylococcus aureus* (Corbin *et al.*, 2008). *Salmonella* Typhimurium can overcome the chelating effects of calprotectin by the expression of ZnuABC (Liu *et al.*, 2012). Zur-regulated genes have also been shown to play a role in biofilm formation in *E. coli* O157:H7 and the uropathogenic *E. coli* strain CFT073 (Gunasekera, Herre and Crowder, 2009; Lim *et al.*, 2011). Zur also regulates expression of  $\alpha$ -haemolysin production in a pyelonephritis causing uropathogenic *E. coli* strain J96 (Velasco *et al.*, 2018).

*Escherichia coli* is a diverse species, composed of both commensal/non-pathogenic strains such as the laboratory workhorse K-12 (originally isolated in 1922 and subject to laboratory culture, storage and mutagenesis) and several pathogenic strains that

occupy different niches within the host. There are currently at least seven defined *E. coli* pathotypes that are diarrheagenic in humans: enteropathogenic (EPEC), atypical enteropathogenic (A-EPEC), enteroaggregative (EAEC), enterohaemorrhagic (EHEC), enterotoxigenic (ETEC), enteroinvasive (EIEC), and diffusely adherent (DAEC). Both EPEC and EHEC attach to and efface microvilli within the intestine, with the latter producing a Shiga toxin (Stx). Additionally, there are a number of extraintestinal pathogenic *E. coli* (ExPEC) strains including uropathogenic *E. coli* (UPEC), which is the most common etiological agent of urinary tract infections, and neonatal meningitis *E. coli* (NMEC), which can cause systemic infections in new-borns. ExPEC *E. coli* are the most common cause of blood-borne infections in humans (Laupland, 2013). UPEC can colonise the intestinal tract commensally, but when transferred to the urinary tract can cause serious kidney and bladder infections, which may spread to the bloodstream. Unlike most other *E. coli* strains, some UPEC strains are capable of intracellular replication, which occurs within bladder cells. In addition to these diverse environments within the human host, 'wild' *E. coli* strains can survive outside the body as well as within different animal hosts; enterohaemorrhagic *E. coli* (EHEC) O157:H7, a food-borne pathogen that can cause severe diarrhoeal disease in humans resides commensally in the bovine gastrointestinal tract, and some avian pathogenic *E. coli* (APEC) strains share common traits with NMEC and UPEC strains (Ewers *et al.*, 2007; Johnson *et al.*, 2007). In this diverse range of environments, the concentration of available zinc will vary greatly. In contrast, K-12 laboratory strains are cultured under standardised conditions designed to meet their nutritional needs. This difference in lifestyle may be reflected in the difference in genome size between K-12 MG1655 and W3110 (4.6Mb) (Blattner *et al.*, 1997; Hayashi *et al.*, 2006) and the pathotypes EHEC O157:H7 str. Sakai (5.6Mb) (Hayashi *et al.*, 2001) and UPEC CFT073 (5.6Mb) (Welch *et al.*, 2002) even taking into account the presence of pathogenicity and toxin encoding genes in pathotypes. The repertoire of genes in *E. coli* that contribute to zinc homeostasis may therefore be greater in pathogenic *E. coli* strains than laboratory strains. To test this hypothesis, we defined the Zur regulon of the non-pathogenic laboratory *E. coli* strain K-12, as well as the pathogenic strains EHEC O157:H7 str. Sakai, and UPEC CFT073.

Previous transcriptomics studies on *E. coli* have investigated the effect of zinc depletion in *E. coli* K-12 MG1655 by the use of metal chelators of limited-specificity (Sigdel, Easton and Crowder, 2006) or the use of a zinc-depleted medium (Graham *et al.*, 2009). However, we are aware of no transcriptomics studies that investigate the effect of *zur* deletion in *E. coli*, or to compare the zinc depletion response in both pathogenic and laboratory *E. coli* strains. In this study we investigate the role of Zur in both pathogenic and non-pathogenic *E. coli* strains using a combination of experimental and *in silico* methods to define a conserved Zur regulon as well as pathogen-specific members.

## Materials and Methods

### Bacterial strains and growth conditions

The strains used in this study were; *E. coli* K-12 MG1655 (Blattner *et al.*, 1997), *E. coli* K-12 W3110 (Hayashi *et al.*, 2006), *E. coli* CFT073 (Welch *et al.*, 2002), *E. coli* O157:H7 str. Sakai (Hayashi *et al.*, 2001) (Supplementary Fig. 1).

Strains were routinely grown in Lysogeny Broth (LB) (Bertani, 1951), and Neidhardt's MOPS minimal media (NH) (Neidhardt, Bloch and Smith, 1974) at 37°C with shaking

(200 RPM). NH was prepared using ultrapure water (Millipore Milli-Q) to dissolve analytical grade reagents. To achieve a zinc concentration of <60nM, components used to make NH were treated with Chelex 100 resin (Bio-Rad Laboratories, UK) in polymethyl-pentene (PMP) Erlenmeyer flasks (Thermo Scientific Nalgene) that had been soaked in nitric acid to prevent the leaching of metal ions. A 10x MOPS stock was treated with Chelex 100 resin and stored at -20°C. Glucose and K<sub>2</sub>HPO<sub>4</sub> were treated separately with Chelex 100 resin and added to the stock solution prior to use along with the micronutrient's solution, MgCl<sub>2</sub> and CaCl<sub>2</sub>, which were not subjected to Chelex 100 treatment. All media solutions were filter-sterilised (0.2 µm pore size) prior to use.

#### Inductively Coupled Plasma Mass Spectrometry Analysis of the Zinc Concentration of Growth Medias

Inductively coupled mass spectrometry (ICP-MS) was used to confirm the zinc concentration of growth medias. Samples were applied to the spectrometer (Thermo-Fisher Scientific X-Series II) by an autosampler (Cetac ASX-520) through a concentric glass nebuliser (Thermo-Fisher Scientific; 1 mL min<sup>-1</sup>). Interfering ions were removed by the hexapole collision cell (7% hydrogen in helium). The machine was equilibrated with 20 ppb, 40 ppb and 100 ppb standard solutions of each metal via a T-piece (Claritas-PPT grade CLMS-2 from Certiprep/Fisher). Samples were processed by Plasmalab software (version 2.5.4; Thermo-Fisher Scientific). For zinc-replete NH, ZnSO<sub>4</sub> (Acros Organics, analytical grade) was added to a final concentration of 1 µM.

#### Construction of mutant strains

Deletion of *zur* in *E. coli* MG1655, CFT073 and O157:H7 was achieved by the gene doctoring technique (Lee *et al.*, 2009). *zur* deletion in *E. coli* MG1655 and O157:H7 was achieved using the primers ZurFw and ZurRev (Supplementary Fig. 2). The Zur deletion in CFT073 used ZurFwCFT, in combination with ZurRev. Primers were synthesised by Integrated DNA Technologies (Leuven, Belgium). The kanamycin resistance cassette was retained for selection purposes.

#### Confirmation of *zur* deletions

Deletion of *zur* was initially confirmed by PCR product size using the primers Scan\_ZurFw and Scan\_ZurRev for *E. coli* MG1655 and O157:H7 and Scan\_ZurFwCFT in combination with Scan\_ZurRev for *E. coli* CFT073. PCR products were purified using a Qiagen QIAquick PCR purification kit (Qiagen, Manchester, UK). DNA sequencing of the PCR products was carried out by Eurofins Genomics (UK). Comparative genome hybridisation (CGH) was performed on each strain to confirm deletion of *zur* and the absence of unintended major genome rearrangements (Hobman *et al.*, 2007). *wt* and  $\Delta$ *zur* strains were grown in 50 mL of LB to an OD<sub>600</sub> of 1.7-1.8 and 10 mL of each culture was pelleted by centrifugation at 5,000 x *g*. Genomic DNA was extracted from the *wt* and  $\Delta$ *zur* strains with a Qiagen genomic DNA extraction kit (500/G) (Qiagen Ltd, Manchester UK) according to the manufacturer's instructions, with four additional 70% ethanol washes as described by (Hobman *et al.*, 2007). DNA was sheared by repeated passage (10x) through a sterile 23G hypodermic needle and purity was assessed by HindIII digestion, which is sensitive to NaCl concentration. DNA was quantified using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Loughborough UK). Comparative genome hybridisation (CGH) between *E. coli* MG1655, O157:H7 and CFT073 *E. coli* *wt* strains and their  $\Delta$ *zur* mutants showed a greatly reduced signal at the *zur* locus in the  $\Delta$ *zur* strains compared to the *wt* strains (data not shown). This indicated

successful deletion of *zur*. In contrast, no differences in signal of greater than two-fold were observed for any other genes in each strain, demonstrating that no unintended major chromosomal deletions or duplications had occurred in the  $\Delta zur$  strains.

#### *In silico* prediction of Zur binding sites

Genome-wide searches of the *E. coli* MG1655, CFT073 and O157:H7 genomes were performed using the Regulatory Sequence Analysis Tools (RSAT) (van Helden, 2003). A position-specific scoring matrix (PSSM) was built using known and putative Zur binding sites from  $\gamma$ -proteobacteria using *consensus* (Hertz and Stormo, 1999; Panina, Mironov and Gelfand, 2003; Huang *et al.*, 2006; Novichkov *et al.*, 2009). The retrieve sequence program (RSAT) was used to retrieve sequences up to 200 bp upstream from the start codon for each annotated open reading frame (ORF) in each strain, and novel Zur binding sites were predicted using the *Patser* program in conjunction with the PSSM (Hertz and Stormo, 1999).

#### Microarray design and production

High-density (8 x 15K) DNA microarrays were designed and manufactured for each strain (including plasmids) by Oxford Gene Technology (Oxford, UK) and validated by the University of Birmingham *E. coli* Centre (UBEC) as described previously (Bingle *et al.*, 2014). The 60 bp probes were synthesised *in situ* by inkjet printing technology (Agilent Technologies, UK).

#### RNA extraction and DNA microarray analysis

RNA was harvested for microarray analysis from triplicate independent cultures of *E. coli* MG1655, O157:H7 and CFT073 *wt* and  $\Delta zur$  strains grown at 37°C in 50 mL of NH broth in 250 mL PMP Erlenmeyer flasks. Cultures were harvested during mid-logarithmic growth phase as determined for each strain by growth curve analysis using the equation  $dOD/T = kOD$ , where  $OD = OD_{600}$ ,  $T$  = time and  $k$  = first order rate constant. RNA was stabilised in RNA protect Bacteria Reagent and extracted using a Qiagen RNeasy Mini Kit (Qiagen Ltd, Manchester, U.K). Contaminating DNA was removed with the Qiagen RNase-free DNase set and the purified RNA was stored in RNase-free water at -70°C. The amplification of RNA was performed with a MessageAmp II-Bacteria RNA amplification kit (Ambion, Huntingdon, UK). Amplified RNA was labelled with Cy3 or Cy5 dyes, purified using a Qiagen RNeasy MinElute clean-up kit and quantified using a Nanodrop ND100 spectrophotometer (Thermo Scientific, Loughborough, UK). Hybridisation of RNA to the microarray and washing was performed using Agilent reagents according to the manufacturer's instructions. The microarray slides were scanned using a GenePix 4000 scanner (Molecular Devices, Wokingham, UK) and Genepix Pro v5.1 software. Resulting images were analysed with GenePix software v6 and the data imported into GeneSpring v7 (Agilent Technologies, Stockport, UK).

#### Constructing of Fluorescence reporter plasmids

Zur regulated promoters ( $P_{znuA}$ ,  $P_{znuCB}$ ,  $P_{zinT}$ ,  $P_{pliG}$ ,  $P_{ykgM}$ , and  $P_{c1265-7}$ ) were amplified using Q5 Polymerase using primers in (Supplementary Fig. 2) and cloned into the fluorescent reporter plasmid pJI300 using NEBuilder HiFi assembly according to manufactures instructions. Supplementary Fig. 3 shows list of plasmids.

#### End Point Fluorescent Assay

*E. coli* MG1655 *wt* and  $\Delta$ *zur* were transformed with the fluorescent reporter plasmid pJI300 containing *Zur* regulated promoters, using standard electroporation. Transformants were inoculated into 5 mL of NH media with or without 400  $\mu$ M ZnSO<sub>4</sub>, with 3 biological repeats, and incubated at 37 °C with shaking (200 RPM) for 16 hr. 200  $\mu$ L of culture was transferred to a 96 well plate (porvair, 215006), fluorescence (ex 580, em 620 nm) and absorbance<sub>600</sub> was recorded using a Tecan GENios pro.

#### Phenotypic characterization of mutants

A single colony of freshly grown *wt* and  $\Delta$ *zur* strains of *E. coli* MG1655, O157:H7 and CFT073 were inoculated into 10 mL of LB broth and incubated at 37°C for 12-16 hr, with shaking (200 RPM). The bacterial culture was diluted to an OD<sub>600</sub> of 0.02 in LB with added ZnSO<sub>4</sub>. 300  $\mu$ L of each culture, with three repeats, was transferred into a microtitre plate, and incubated at 37°C for 20 hr. The Absorbance<sub>600</sub> of each culture was measured in a Tecan GenIOS Pro microtitre plate reader (Tecan Group Ltd, Reading, UK).

#### Chromatin affinity purification (ChAP) PCR

*E. coli* W3110 derivatives expressing a C-terminal 12 His-tagged *Zur* were generated as previously described (Kasho *et al.*, 2014). Briefly, to generate a *E. coli* W3110 derivative expressing *Zur* C-terminally tagged with 12 histidines (12His), a modified one-step gene inactivation method (Datsenko and Wanner, 2000) was used. The pSTV28-C-12His plasmid was constructed by inserting a synthetic oligonucleotide coding for 12His and a kanamycin resistant gene derived from pKD4 (Datsenko, Wanner and Beckwith, 2002) into the multiple cloning site of pSTV28 (Takara Bio, Japan). A DNA fragment containing the 12His coding sequence was amplified by PCR using the pSTV28-C-12 His plasmid as template and the *Zur*His\_Fw and *Zur*His\_Rv primer set (Supplementary Fig. 2). To facilitate  $\lambda$ -red mediated insertion of the PCR product into the host strain chromosome, a ~70 bp sequence of the *zur* coding region and its downstream region was added to both primer sequences. The amplified DNA fragment was transformed into *E. coli* BW25113 cells harbouring the pKD46 plasmid encoding  $\lambda$  Red recombinase (Datsenko and Wanner, 2000). Transformants in which the linker and 12His sequences were inserted at the 3' end of the chromosomal *zur* gene, through a double-crossover at the coding and downstream regions of *zur*, were selected using kanamycin resistance to obtain the *E. coli* BW25113 *Zur*-12His strain. The *zur* gene fused with a 12His coding sequence was transferred into the *E. coli* W3110 chromosome, together with the kanamycin resistant gene, via P1 transduction.

Chromatin affinity purification (ChAP) was performed as described previously (Ishikawa *et al.*, 2007). *E. coli* W3110 and its derivative, a *Zur*-His12 strain of W3110, were grown in 200 mL of LB at 37°C with shaking (200 RPM), until logarithmic growth. Formaldehyde was added to the culture at a final concentration of 1% concentration and incubated for 30 min at 37°C. The cross-linked cells were collected by centrifugation (4,200 x *g* for 5 min) and the supernatant was discarded. The cells were washed with 10 mL of TBS (Tris buffered saline), resuspended in UT-buffer (0.1 M HEPES, 0.5 M NaCl, 30 mM Imidazole, 8 M Urea, 1% Triton X-100, 10 mM  $\beta$ -mercaptoethanol, 1 mM Phenyl Methyl Sulphonyl Fluoride), and disrupted by sonication (4 s. of sonication with 10 s. interval, using a Soniprep 150 sonicator (MSE): total sonication time was 10 min.). The sonicated samples were centrifuged (4,200 x *g* for 10 min at 4°C) and supernatant was recovered. 100  $\mu$ L of TALON beads (Invitrogen, Tokyo Japan) was added to the 2 mL of supernatant and mixed on a rotary mixer for



12-16 hr to capture the crosslinked Zur-DNA complexes on the beads. The beads were washed 5 times with 1.5 mL of UT-buffer and the immobilized Zur-DNA complexes were eluted using elution buffer (0.1 M Tris-Cl [pH7.5], 0.5 M imidazole [pH7.5], 1% SDS, 10 mM DTT). The eluate was passed through a Microcon-100 filter (Millipore) to remove non-specifically bound un-crosslinked proteins with a molecular mass lower than 100 kDa. Protein complexes retained on the membrane were washed three times with wash buffer (100 mM Tris-HCl, pH 7.5, 1% SDS, 10 mM DTT), and recovered by the addition of 50  $\mu$ L of the elution buffer. The crosslinking between Zur and DNA fragments was reversed by heating at 65°C for 12-16 hr, and DNA fragments released from crosslinking were purified using the QIAquick PCR purification kit (Qiagen, Tokyo, Japan). The immobilized DNA fragments were eluted with 50  $\mu$ L of nuclease free water. Formation of the Zur-DNA complex was evaluated by PCR using the specific primers to amplify upstream of *gadB* (negative control), *znuC*, *zinT*, *ykgM* and the *pliG* promoters (Supplementary Fig.2) and the purified DNA diluted ( $2^3$  to  $2^6$ ). The PCR reaction was performed using iProof high fidelity DNA polymerase (BioRad Laboratories, Tokyo, Japan) using the following conditions; initial denaturation (98°C for 1 min), cycling reactions (98°C for 10 s, 55°C for 30 s, 72°C for 30 s) x 30 cycles, and final extension (72°C for 5 min). The amplified DNA fragments were detected by electrophoresis using a 2 % agarose TBE gel stained by Ethidium bromide.

## Results

### ICP-MS analysis of zinc concentration in media

The concentration of zinc in NH and LB was determined using inductively coupled plasma mass spectrometry (ICP-MS). The mean concentration of  $Zn^{2+}$  in LB was found to be 12.3  $\mu$ M, whilst the concentration in NH without Chelex 100 treatment was 14.2  $\mu$ M. However, Chelex treatment of NH, which included the glucose and  $K_2HPO_4$  solutions, reduced the concentration of zinc to 54 nM, which is 227x fold lower than LB.

Phenotype of  $\Delta zur$  mutants: *zur* deletion increases the zinc sensitivity of UPEC, but not for K-12 and EHEC)

Growth of the *wt* and  $\Delta zur$  mutants of *E. coli* MG1655, O157:H7 and CFT073 were observed for differences in growth in LB in microtiter plates (Fig. 1). *E. coli* CFT073  $\Delta zur$  strain showed greater sensitivity to zinc than *E. coli* CFT073 *wt* when the media was supplemented with zinc up to 750  $\mu$ M of added  $ZnSO_4$ ; at 1 mM of added  $ZnSO_4$  no growth was observed in any of the three *E. coli* strains, *wt* or  $\Delta zur$ . There was no significant growth difference between *wt* and  $\Delta zur$  strains for *E. coli* MG1655 or O157:H7. Growth of *E. coli* CFT073  $\Delta zur$  was also inhibited compared to *E. coli* CFT073 *wt* when grown in LB in Erlenmeyer flasks supplemented with 1 mM  $ZnSO_4$  (DATA NOT SHOWN).

### Prediction of Zur regulated genes using Regulatory Sequence Analysis Tools (RSAT)

To evaluate the differences in the *zur* regulons in *E. coli* MG1655, O157:H7 and CFT073, predictions of Zur-binding sites were initially made *in silico*, which were then validated by microarray analysis, ChAP-PCR, and end point fluorescent assays. The Zur binding site in *E. coli* MG1655 consists of a 23bp repeat (Patzner and Hantke, 2000), which is highly conserved in other gamma-proteobacteria (Panina *et al.*, 2003). Known and putative Zur binding sites were collected from a variety of sources including a

computational comparative analysis of the Zur regulon of gamma-proteobacteria (Panina, Mironov and Gelfand, 2003), the RegPrecise database (Novichkov *et al.*, 2009) and sites determined in *Xanthomonas campestris* (Huang *et al.*, 2008), as well as the equivalent sites in a number of *E. coli* pathotypes, including O157:H7 and CFT073 (Fig. 2). These sites were aligned and converted to a position-specific scoring matrix (PSSM) by the program *consensus*, and used to search the genomes of the three *E. coli* strains for Zur binding sites using Patser (Hertz and Stormo, 1999). The Patser searches using the Zur binding PSSM accurately predicted Zur binding sites in the three *E. coli* strains upstream of all genes known to be Zur-regulated; *znuA*, *znuCB* and *ykgMO* (Fig. 3). In CFT073 the top-scoring predicted site was upstream of *c1265*, which is annotated as a "heme/hemoglobin receptor" based on sequence homology. This gene is not present in either *E. coli* K-12 strains or O157:H7. A novel Zur binding site was predicted in both *E. coli* K-12 and CFT073 strains situated upstream of *pliG*, encoding a gene annotated as an inhibitor of goose (g-) type lysozyme. Although not detected by Patser, this site was also found to be present in the *E. coli* O157:H7 genome, but because it overlaps with an ORF for a small hypothetical protein (ECs1673) this region was not searched by the program.

#### Transcriptomic response to zinc depletion and *zur* deletion using strain-specific Microarray

Transcriptomics experiments were performed for each strain to investigate the effect of Zur deletion and zinc depletion. A larger number of genes showed greater than two-fold changes in expression in *E. coli* O157:H7 and CFT073 compared to *E. coli* MG1655 under these conditions (Table. 1). In all strains; zinc depletion compared to zinc repletion, and  $\Delta zur$  compared to a *wt* strain, the greatest increases in gene expression within the established Zur regulon were seen in *zinT* and *ykgM*, ranging from a 195-fold to a 2291-fold increase in expression (Fig. 4). Increases in expression of the *znuABC* system were more modest and not all increases were greater than two-fold or statistically significant, although at least one gene from *znuABC* met these criteria for each strain under both test conditions. In *E. coli* CFT073 no transcriptomics data is available on *znuA* expression, as the microarray probe was located outside of the *znuA* coding region due to a mis-annotation in the GenBank entry for the genome sequence of CFT073 (NC\_004431).

*PliG* expression is induced by zinc depletion and *zur* deletion in all strains, and is a conserved new *E. coli* *zur* regulon member.

As predicted by bioinformatics analysis, *pliG* is a target of Zur repression since its expression was significantly up-regulated under both zinc depletion and  $\Delta zur$  deletion conditions in all strains (from 6.63-fold to 39.69-fold) (Fig. 4). Additionally, the binding of Zur to the predicted *pliG* promoter was demonstrated by ChAP-PCR in *E. coli* MG1655, alongside Zur binding to the *znuC*, *zinT* and *ykgM* promoters (Fig. 5). This is in agreement with Reverse Transcriptase (RT)-PCR and *in vitro* Zur  $P_{pliG}$  binding studies (Gilston *et al.*, 2014). Further, end point fluorescent assay shows the promoter ( $P_{pliG}$ ) is significantly repressed when grown NH supplement with 400  $\mu$ M ZnSO<sub>4</sub> compared to NH media only, showing zinc repression of the *pliG* (Fig. 7). *PliG* is a periplasmic protein that has been shown to have inhibitory effects against goose (g-) type lysozyme (Vanderkelen, Van Herreweghe, Vanoirbeek, *et al.*, 2011), but its role in zinc homeostasis is not clear

Changes in expression of low affinity zinc transporters induced by zinc depletion and *zur* deletion.

Transcriptomics data was interrogated for evidence of changes in expression in response to zinc depletion or  $\Delta zur$  for a number of genes: *zupT*, *pitA*, *zitB*, *mntH*, *yjiP* and *zntB*, which are thought to be involved in, or have been shown to play a role, in low affinity zinc transport (Grass *et al.*, 2002). No significant changes in expression of these low affinity transporter genes was seen in either zinc depletion conditions or due to  $\Delta zur$  in any of the strains, suggesting that the genes are constitutively expressed, or may be regulated by an alternative metal ion, which is in agreement with previously published work (data not shown)

Increased expression of *c1265-7* in UPEC CFT073 is induced by zinc depletion and *zur* deletion: *c1265-7* genes are UPEC (and group B2 *E. coli* strains) specific *zur* regulon members.

Three adjacent genes on the *E. coli* CFT073 genome, *c1265-7*, showed significantly increased expression under both zinc depletion conditions and in the  $\Delta zur$  strain (30.98-fold to 60.92-fold) (Fig. 4). Patser genome-wide binding site searches had strongly predicted a Zur binding site upstream of *c1265*. Genome sequence analysis predicts *c1265-7* to form an operon (Taboada, Verde and Merino, 2010). Homologues of these genes could not be identified in *E. coli* K-12 strains or O157:H7; but BLAST searches identified close homologues of *c1265-7* in the genomes of a number of extra-intestinal *E. coli* (ExPEC) strains. The *c1265-7* genes are located within the SerX associated pathogenicity island (PAI), suggestive of acquisition by horizontal gene transfer (HGT), the likelihood of which is corroborated by a difference in GC content (42.4% to 48.4%) from the average of the CFT073 genome (50.5%) (Lloyd, Rasko and Mobley, 2007). The SerX PAI is found in *E. coli* group B2 strains and encodes a cluster of genes, *iroBCDEN*, which encode the salmochelin siderophore system for iron uptake (Fig. 6). Analysis of homologues of these three genes suggests a role in zinc transport and homeostasis. C1265 shows homology to TonB-dependent outer membrane receptors, sharing 41% sequence identity with ShuA from *Shigella dysenteriae*, which transports heme across the outer membrane. C2166 shares 20.48% homology to a cell wall surface anchor family protein, RrgC of *Streptococcus pneumoniae*. The homology is specifically found at the IgG-like domain of RrgC, which is involved in bacterial adhesion (Shaik *et al.*, 2014). Two overlapping domains in the C1266 hypothetical protein were predicted, one of unknown function (DUF418) and one related to CbiK (COG5266), annotated as the periplasmic component of a  $\text{Co}^{2+}$  transport system. Two CbiK homologues are found in *Delsulfovibrio vulgaris Hildenborough*, one of which is believed to act as a periplasmic component of an iron transport system (Lobo *et al.*, 2008), although this shares very little sequence homology with c1266 (3%). c1266 contains a His-rich region between residues 98-109, which might be involved in zinc-binding.

Blast searches using the predicted c1267 protein sequence detected significant homology to cobalamin synthesis proteins; molybdopterin guanine dinucleotide synthesis proteins. C1267 is a member of COG0523, a subset of which are known to be involved in metallocentre synthesis and are linked to zinc homeostasis (Haas *et al.*, 2009). Whilst C1267 shares the common P-loop NTPase domain found in COG0523 members, required for NTPase activity, it is missing the CXCC domain, important in metal-binding, and the C-terminal His-rich region. Analysis by pSORTb v3.0 (Yu *et al.*, 2011), predicts localisation of C1266 within the inner membrane.

To confirm Zur regulation of *c1265-7*, the predicted promoter ( $P_{c1265}$ ) was cloned into a *rfp* fluorescence reporter plasmid (pJI300), transformed into *E. coli* MG1655 *wt* and  $\Delta$ *zur*, and an end point fluorescence assay was conducted (Fig. 7). End point fluorescent assay shows a significant increase in promoter activity of  $P_{c1265}$  when grown in *E. coli* MG1655 *wt* in NH only compared to grown in NH + 400  $\mu$ M ZnSO<sub>4</sub>. This shows  $P_{c1265}$  is clearly repressed with the addition of zinc, as would be expected from a Zur regulated gene. Further, promoter activity of  $P_{c1265}$  in *E. coli* MG1655 *wt* and  $\Delta$ *zur* shows little difference in activity when grown in NH or NH + 400  $\mu$ M ZnSO<sub>4</sub>, which shows similar levels of promoter activity to that of *E. coli* MG1655 *wt* in NH only.

Changes in expression of genes relating to gene mobility in the pathogenic *E. coli* strains

In *E. coli* O157:H7 eight orthologues of two adjacently located genes from COG2963 and COG3436 are up-regulated between 2.02 and 3.98-fold in response to zinc depletion (Supplementary Fig. 10). Both COG2963 and COG3436 are classified as transposases and inactivated derivatives. In *E. coli* O157:H7  $\Delta$ *zur*, all eight of the COG3436 orthologues were also up-regulated under zinc-replete conditions along with two additional members of this group, whilst two of the eight COG2963 members showed up-regulation (Supplementary Fig. 12). Of these sets of orthologous genes two are associated with known pathogenicity islands (PAIs): *Ecs3493/3494* and *Ecs4545/4546*, the latter being located on the locus of enterocyte effacement (LEE). The induction of these genes under both zinc depletion and  $\Delta$ *zur* suggests that Zur may repress their expression in the presence of zinc. In the *E. coli* CFT073  $\Delta$ *zur* nine putative phage genes were down-regulated in the presence of zinc and forty-nine up-regulated (Supplementary Fig. 15-16). The majority of the up-regulated prophage genes are located in the *icdA* PAI, which contains the *sitABCD* iron uptake system. However, only *sitA* expression was significantly up-regulated (2.03-fold). Cryptic prophages have been shown to contribute to host fitness under a range of stress conditions (Wang *et al.*, 2010).

Induction of Fur regulated genes occurs during zinc depletion and *zur* deletion in UPEC CFT073, but not in O157:H7 and K-12

Twelve Fur-regulated genes were up-regulated in the *E. coli* CFT073  $\Delta$ *zur* in zinc-replete conditions (Supplementary Fig. 16). Eight of these (*entF*, *entC*, *fepA*, *ybdZ*, *ybdD*, *fes*, *entE* and *entH*) are involved in enterobactin synthesis, whilst *sufA*, *sufB* and *sufC* are involved in iron-cluster synthesis and *fhuF* is thought to be involved in the use of ferrioxamine B as an iron source (Müller *et al.*, 1998). Two other iron-related genes were up-regulated; *iroB*, which is involved in the glycosylation of enterobactin, and *chuT*, which encodes a hemin-binding protein. Seven Fur-regulated genes were up-regulated in the *E. coli* CFT073 *wt* strain in zinc-deplete conditions, only two of which overlap with the genes up-regulated in the  $\Delta$ *zur* strain: *entC* and *entF*. The five other up-regulated Fur genes were *sodA*, *ycdO/efeO*, *exbD*, *tonB* and *ycdB/efeB*. EfeO and EfeB form part of an acid-induced, high-affinity Fe<sup>2+</sup> uptake system (Große *et al.*, 2006; Cao *et al.*, 2007), which is cryptic in *E. coli* MG1655 but functional in *E. coli* O157:H7 and CFT073. SodA provides protection against superoxide stress. TonB and ExbD form part of a system that energises the transport of exogenous substances across the outer membrane into the periplasmic space. This system usually facilitates iron uptake, via specific receptors that bind substrates such as hemin and various siderophores, but TonB-dependent uptake systems for vitamin B-12 and nickel complexes also exist (Noinaj *et al.*, 2010), and Zur-regulated TonB-dependent zinc uptake receptors have

been identified in *Neisseria meningitidis* and *Caulobacter crescentus* (Stork *et al.*, 2010; Mazzon *et al.*, 2014). De-repression of members of the Fur regulon in the  $\Delta zur$  strain could be due to adventitious zinc binding under zinc replete conditions, particularly as de-repression of members of the Zur regulon will have increased the internal zinc concentration further. However, as some members of the Fur regulon were also derepressed in the absence of zinc there is some suggestion of cross talk between Zur and Fur, evidence of which has also been seen in *Caulobacter crescentus* (Mazzon *et al.*, 2014).

#### Integration of bioinformatics predictions with transcriptomics data

Plotting of patser scores of predicted Zur binding sites against  $\log_2$  changes in expression generated by transcriptomics experiments in  $\Delta zur$  strains shows a clustering of known Zur regulated genes (Fig. 8). In *E. coli* MG1655, the predicted Zur binding site upstream of *pliG* is located within this cluster, as is the case with predicted Zur binding sites upstream of *pliG* and *c1265* in the *E. coli* CFT073 genome. Predicted Zur binding sites with lower patser scores remain separate. This data illustrates the correlation between *in silico* predictions using patser and experimental microarray data, which adds to the overall confidence in each approach.

#### End Point Fluorescence Assay of Zur Regulated Promoters in *wt* and $\Delta zur$

Figure 7 show the promoter activity of known and predicted Zur regulated promoters cloned into the mRFP1 expression plasmid (pJI300) and expressed in *E. coli* MG1655 *wt* and  $\Delta zur$ . There is a significant decrease in expression of all Zur regulated promoters, beside  $P_{zinT}$ , when expressed in *wt* and grown in NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$  compared to NH only. However, when expressed in *E. coli*  $\Delta zur$ , most of the genes show no significant difference when grown in NH only or NH + 400  $\mu\text{M}$   $\text{ZnSO}_4$ .

#### Discussion

Through genome-wide searches we predicted a conserved Zur regulon shared between non-pathogenic and pathogenic *E. coli* composed of the established Zur regulated genes: *znuABC*, *zinT* and *ykgMO*. For *zinT* and *ykgMO* this prediction correlated with transcriptomics data, which saw strong induction of transcription of these genes in response to zinc depletion and *zur* deletion in all strains. The fold-change increase in the expression of the high-affinity zinc uptake system *znuABC* was not large, which may suggest the involvement of other regulators, or that *znuABC* is already being expressed in some or all cells. Previous transcriptomics studies in *E. coli* MG1655 with the addition of the metal-chelator TPEN (N,N,N',N'-tetrakis(2-pyridylmethyl)ethane-1,2-diamine) have also demonstrated greater increases in the expression of Zur regulon members *zinT* and *ykgM* than *znuABC* (Sigdel, Easton and Crowder, 2006; Hensley, Tierney and Crowder, 2011). A chemostat-based zinc starvation study on *E. coli* MG1655, found an approximately 3-fold increase of *zinT* expression relative to *znuA* and *ykgM*, the two other Zur-regulated genes found to be up-regulated (Graham *et al.*, 2009).

A novel member of this conserved *E. coli* Zur regulon, *pliG*, was predicted by Patser in this study, and expression of this gene was seen to be induced in our transcriptomics experiments in all strains both by zinc depletion in the growth media, and in strains where *zur* had been deleted. Expression of *pliG* in *E. coli* MG1655 has been shown to increase 32-fold in response to TPEN treatment in one previous study (Hensley, Tierney

and Crowder, 2011), but not in two other transcriptomics studies looking at either TPEN treatment (Sigdel, Easton and Crowder, 2006) or zinc depletion (Graham *et al.*, 2009). However, a study using purified Zur has shown *in vitro* binding to the *pliG* promoter (Gilston *et al.*, 2014). The role of PliG, which has reported inhibitory effects against g-type lysozyme (Vanderkelen, Van Herreweghe, Callewaert, *et al.*, 2011; Vanderkelen, Van Herreweghe, Vanoirbeek, *et al.*, 2011), in cellular response to zinc depletion is unclear. Lysozymes play an important role in the innate immune system of animals through the lysis of bacterial cells. Most vertebrates, including humans, possess two g-type lysozymes (LYGA1 and LYGA2) (Irwin and Gong, 2003). LYGA1 is expressed in a variety of tissues, whilst LYGA2 is limited to the eye, skin and testis (Irwin, 2014). It is known that metal chelators such as lactoferrin and EDTA can increase the sensitivity of Gram- such chelators may trigger Zur de-repression through a decrease in zinc concentration, resulting in the expression of PliG and its protective effects against g-type lysozyme activity. Alternatively, the g-type lysozymes display a diverse range of functions other than direct lysis of the bacterial cell wall (Irwin, Biegel and Stewart, 2011); lactalbumin for example, catalyses the synthesis of lactose in lactating mammary glands. PliG therefore, might be inhibitory towards a yet uncharacterised member of the g-type lysozyme family that reduces the availability of zinc. Conversely, high levels of added zinc induce expression of proteins that protect against zinc overload, oxidative stress and lysozyme in *Enterococcus faecalis* (Abrantes, de Fátima Lopes and Kok, 2011; Abrantes, Kok and Lopes, 2014) negative bacteria to lysozyme activity due to increased membrane permeability (Vaara, 1992; Branen and Davidson, 2004). In the uropathogenic *E. coli* CFT073 a further three novel members of the Zur regulon have been predicted, *c1265-7*, and this prediction is supported by microarray data for response to zinc depletion and *zur* deletion, as well as data generated by end point fluorescence assay. One of these genes, *c1267*, is part of COG0523, which has Zur regulated members in a number of other bacterial species. *E. coli* K COG0523 members YjiA (Sydor *et al.*, 2013) and YeiR (Blaby-Haas *et al.*, 2012) have been shown to bind zinc and other metals. COG0523 members have also been found downstream of Zur binding sites in a number of organisms, implicating them in some involvement in the zinc starvation response (Gaballa and Helmann, 1998; Chaoprasid *et al.*, 2016; Butof *et al.*, 2017; Jordan *et al.*, 2019; Khemthong *et al.*, 2019). Additionally, the first member of this putative operon, C1265, is predicted to resemble a TonB-dependent outer membrane receptor. Whilst transcriptomics analysis of the CFT073 *zur* mutant did not suggest regulation of the TonB system by Zur, under zinc depletion modest up-regulation of the TonB (2.17-fold), ExbB (1.79-fold) and ExbD (2.48-fold) was observed. Typically, most TonB systems characterised to date have been involved in iron uptake, but more recently a nickel-dependent TonB system has been identified in *Helicobacter pylori* (Schauer *et al.*, 2007) and TonB dependent membrane receptors involved in zinc uptake have been identified in *Neisseria meningitidis* and *Caulobacter crescentus* (Stork *et al.*, 2010; Mazzon *et al.*, 2014). *In silico* analysis of available *E. coli* genomes suggests that *c1265-7* are found primarily in extra-intestinal strains. It is possible that these genes form the whole or part of a novel zinc uptake system that supports growth in the extra-intestinal environment, which includes environments where zinc may be scarce or withheld by calprotectin (Hood *et al.*, 2012). There is increasing evidence of the importance of host-restriction of metal ion availability to microorganisms within the urinary tract, particularly in respect to the availability of zinc ions within this niche (Sabri, Houle and Dozois, 2009; Nielubowicz, Smith and Mobley, 2010). A previous study demonstrated that delegating either *znuABC* or *zupT* from *E. coli* CFT073 it significantly reduced its bacterial count in a kidney infection model, and thus reduced its pathogenicity (Sabri, Houle and Dozois, 2009). *E. coli* CFT073 was isolated from a patient with acute pyelonephritis, so this

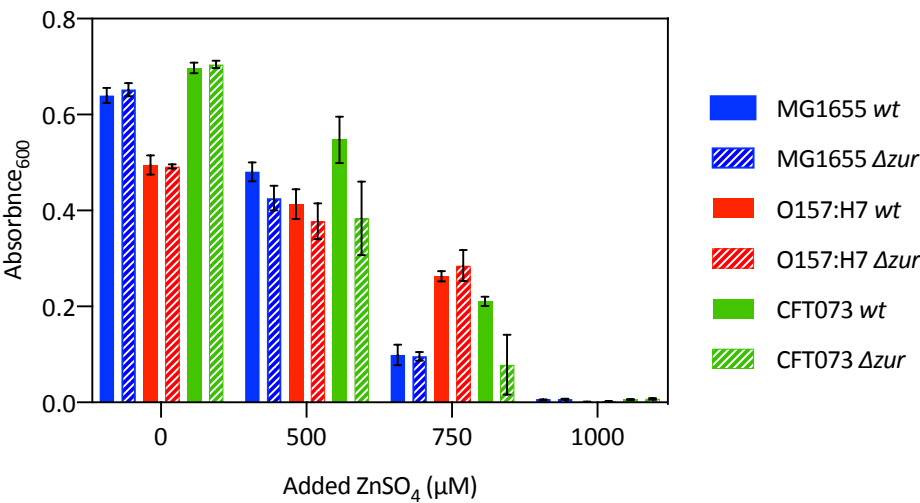
pathogen is capable of causing more serious infections beyond uncomplicated urinary tract infection, by crossing kidney epithelial cells (Mobley *et al.*, 1990). Further evidence that the *E. coli* CFT073 Zur regulon may contain additional zinc uptake systems not present in the other strains included in this study is the increased sensitivity of the *E. coli* CFT073  $\Delta zur$  to zinc supplementation of growth media, which was not observed in *E. coli* O157:H7 or *E. coli* MG1655.

### Conclusions

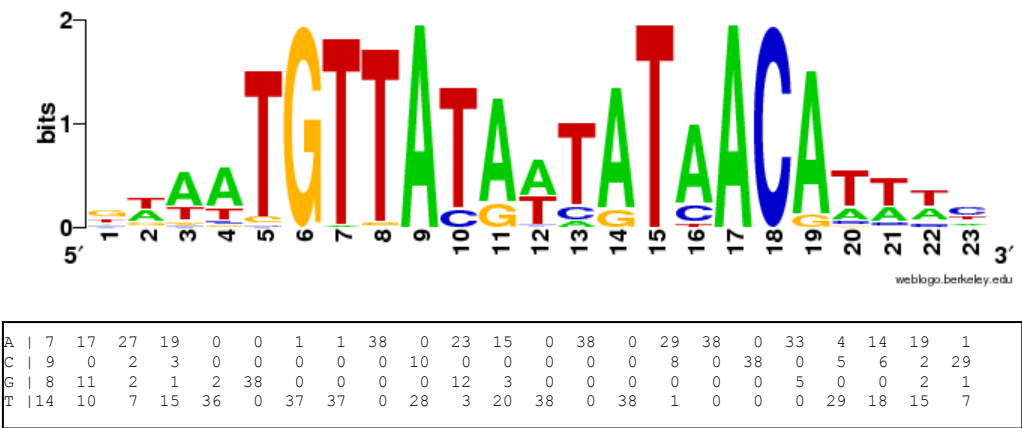
The overall picture of the Zur regulon of *E. coli* from this study is that it consists of a small conserved group of genes shared between the laboratory and pathogenic strains, as well as additional members in the extraintestinal strain CFT073 (with homologues in other ExPEC strains), which may be reflective of the greater diversity of the UPEC lifestyle, or specific adaptations that enable ExPEC strains to survive and grow outside the gut.

It is possible the true phenotype of C1265-7 in CFT073, cannot be observed when other zinc import systems are expressed, such as; ZnuACB, ZupT, ZinT, PliG. Knowing that hots cells use nutritional immunity,  $\Delta znuABC$  show reduced pathogenicity in CFT073 kidney infections, and that C1266 shows minor homology to the bacterial adhesion protein IgG-like domain of RrgC; this could potential suggest that the true phenotype of C1265-7 can only be observed during infection.

Figures



**Fig. 1**  
Growth of *E. coli* strains wt and  $\Delta$ zur mutants in LB media with added ZnSO<sub>4</sub>, with four biological repeats. Bacterial cultures were grown at 37°C. with shaking (200 RPM) for 20 hr in a microtitre plate before reading Absorbance<sub>600</sub>. Error bars indicate 1 S.D.



**Fig. 2**  
A SeqLogo and position-specific scoring matrix (PSSM) was compiled from known and putative Zur-binding sites from within the gammaproteobacteria, including the *E. coli* strains used in this study (Panina, Mironov and Gelfand, 2003; Huang *et al.*, 2008; Novichkov *et al.*, 2009)), using Weblogo version 2.8.2 (Crooks *et al.*, 2004) and *consensus* (Zhao *et al.*, 2012) (van Helden, 2003).



Gene	Start	End	Sequence	Patser Score		
				<i>E. coli</i> MG1655	<i>E. coli</i> O157:H7 str. Sakai	<i>E. coli</i> CFT073
c1265	-44	-22	aaatAAAATGTTATGATATAACATCACAatc	N/A	N/A	19.90
zinT	-56	-34	ttgcTATATGTTACAATATAACATTACacat	19.55	19.52	19.52
znuA	-52	-30	tcatGAAATGTTATAATATCACACTTCcat	18.91	18.91	18.91
znuC	-49	-27	tcatGAAATGTTATAATATCACACTTCcat	18.91	18.91	18.91
ykgM	-34	-12	atctTTATGGTTATGTTATAACATAACaggt	18.13	16.02	16.97
ykgM	-29	-7	ttttTACCTGTTATGTTATAACATAACcata	17.78	17.78	17.78
pliG (ycgK)	-53	-31	ttagGGATTGTTATATTATAACAGTTCatcg	16.54	<10	16.54
ycgL	-178	-156	ttagGGATTGTTATATTATAACAGTTCatcg	<10	<10	16.54
pliG (ycgK)	-53	-31	cgatGAACTGTTATAATATAACAATCCctaa	16.26	<10	16.26
ycgL	-178	-156	cgatGAACTGTTATAATATAACAATCCctaa	<10	<10	16.26
znuA	-52	-30	atgaGAAGTGTGATATTATAACATTTCatga	14.19	14.19	14.19
znuC	-49	-27	atgaGAAGTGTGATATTATAACATTTCatga	14.19	14.19	14.19
zinT	-56	-34	atgtGTAATGTTATATTGTAACATATAgcaa	12.26	12.58	12.58
c1265	-44	-22	gattGTGATGTTATATCATAACATTTTattt	N/A	N/A	12.11
ycgF	-192	-170	aaacAATGTGTTATGATAAAACATATTgcat	9.76	N/A	<10
ycgZ	-144	-122	aaacAATGTGTTATGATAAAACATATTgcat	9.76	N/A	<10
ppiB	-79	-57	gatcGTGCGGTTATGCTATAACACCACccta	8.74	8.74	8.74
cysS	-117	-95	gatcGTGCGGTTATGCTATAACACCACccta	8.74	8.74	<10
ECs1299	-131	-109	gtttATAATGAAATTATATAACACCCCTgga	N/A	7.86	<10
ECs1821	-110	-88	cagaGTACTGTTACACTTTAACAAATCatat	N/A	7.63	N/A
bdm	-136	-114	ttgaTGATTTTATATTTTAACACCATgata	7.61	<10	<10
yeiQ	-54	-32	atttCAATTGTTGTCATATAACTTTACactg	<10	<10	7.31
ybjD	-85	-63	tgctTAAATGTGATGTCATCACGTATTtagca	7.27	7.27	<10
argD	-49	-27	tattTTGTGGTTATAATTCACATTTGttta	7.08	<10	<10
<b>Fig 3</b> Zur-binding sites predicted by Patser in <i>E. coli</i> MG1655, O157:H7 str. Sakai and CFT073 genomes (cut-off score >10). Start and end indicate the position of the site relative to the transcription start site (TSS). N/A Indicates gene not present in the strain						

Genes	Fold Change					
	MG1655		O157:H7		CFT073	
	Zinc - depleted	$\Delta zur$	Zinc - depleted	$\Delta zur$	Zinc - depleted	$\Delta zur$
zinT *	+213.18	+394.57	+401.47	+301.50	+1030.99	+521.30
ykgM*	+195.11	+459.19	N/A	N/A	+2290.97	+887.31
pliG*	+39.69	+23.26	+6.63	+11.26	+24.58	+11.80
znuC*	+3.86	+4.11	+7.30	N/A	+7.48	+7.80
znuA*	N/A	+50.84	+8.73	N/A	N/A	N/A
ECs1820	N/A	N/A	+3.49	N/A	N/A	N/A
c1266	N/A	N/A	N/A	N/A	+60.92	+46.34
c1265	N/A	N/A	N/A	N/A	+41.83	+30.98
c1267	N/A	N/A	N/A	N/A	+31.56	+34.69

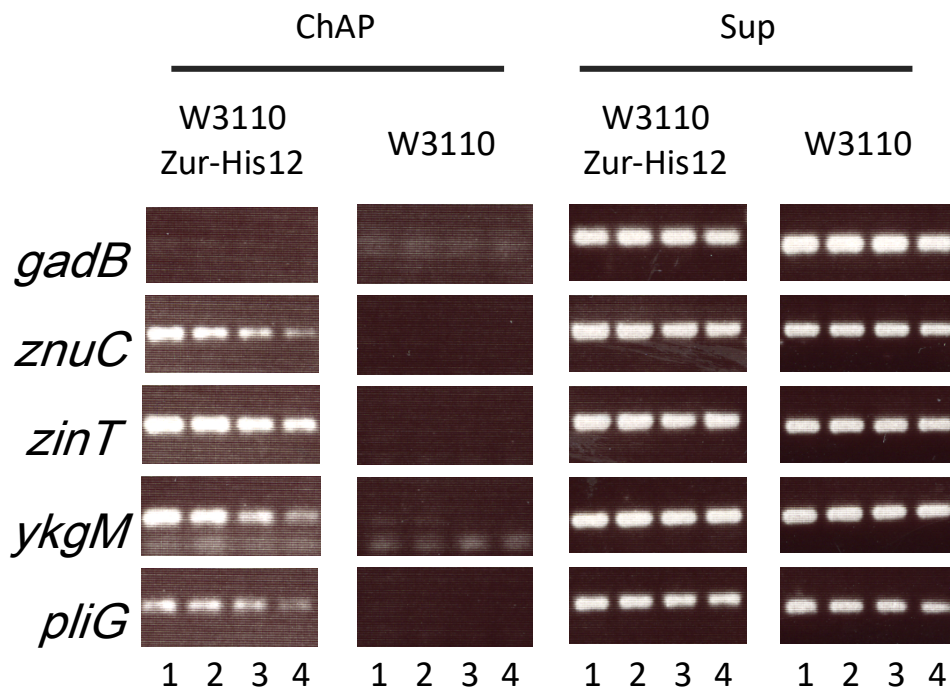
**Fig. 4**

Summary of transcriptomic data showing relevant. Fold change in expression. N/A gene not present in strain. \* Known to be regulated by Zur. All raw data can be seen in Supplementary Fig. 4-16.

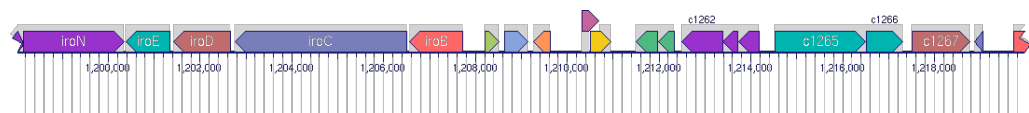
<i>E. coli</i> strain	Effect of zinc depletion ( <i>wt</i> + Zn $\rightarrow$ <i>wt</i> - Zn)		Effect of <i>zur</i> deletion ( <i>wt</i> + Zn $\rightarrow$ $\Delta zur$ + Zn)	
	No. of down-regulated genes	No. of up-regulated genes	No. of down-regulated genes	No. of up-regulated genes
<b>K-12 MG1655</b>	1	7	1	9
<b>O157:H7 str. Sakai</b>	1	31	47	37
<b>CFT073</b>	1	32	34	116

**Table 1.**

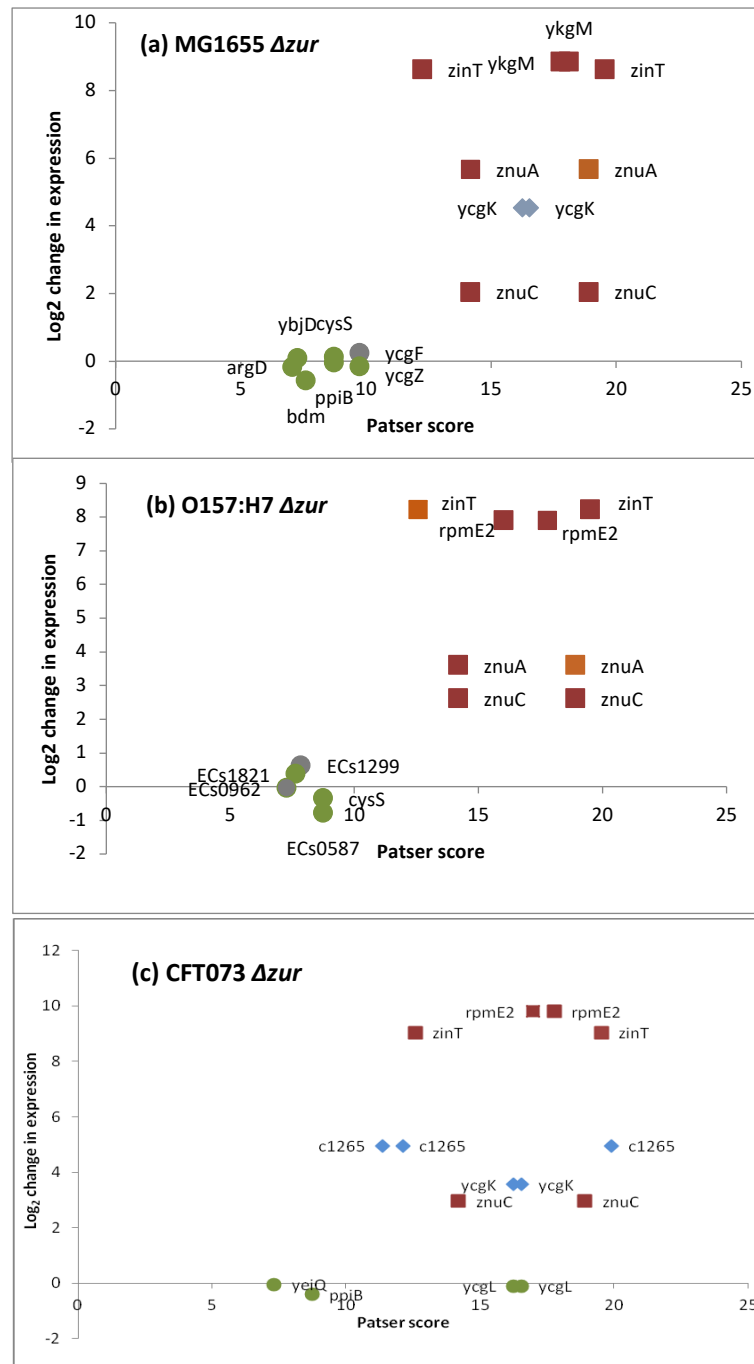
Number of *E. coli* genes either up or down regulated from transcriptional analysis. *E. coli wt* grown in NH media depleted with zinc (*wt* + Zn) was compared to; *E. coli wt* grown in NH media (*wt* - Zn), or *E. coli*  $\Delta zur$  grown in NH media zinc depleted ( $\Delta zur$  + Zn)

**Fig. 5**

Binding of Zur to known and candidate Zur promoters in *E. coli* W3110. W3110 and W3110 Zur-His12 strains were cultivated in LB medium until log growth phase. Formaldehyde was added for protein-DNA cross-linking. Zur-His12-DNA complexes were precipitated with Nickel-agarose resin under denaturing condition. DNA segments corresponding to *gadB* (negative control), *znuC*, *zinT*, *ykgM* and *pliG* promoters were detected by PCR. W3110 was used for the negative control to show the Zur-His12 independent precipitation. The samples were serially diluted to perform PCR (lane 1 =  $2^{-3}$ , lane 2 =  $2^{-4}$ , lane 3 =  $2^{-5}$ , lane 4 =  $2^{-6}$ ).

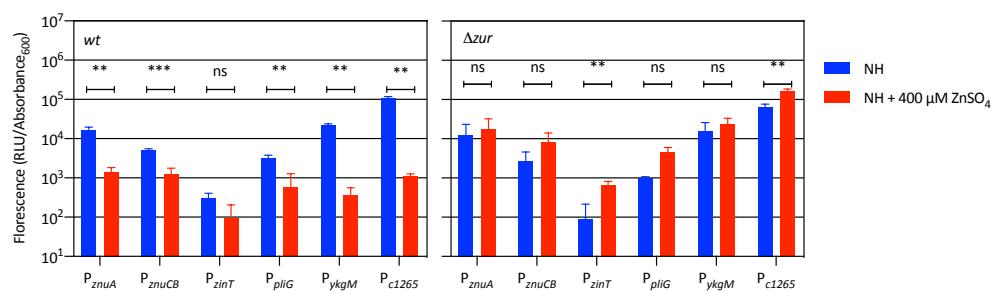
**Fig. 6**

Proximity of *c1265-7* to *iroBCDEN* encoding the samochelin siderophore system in the SerX pathogenicity island of *E. coli* CFT073.



**Fig. 8**

Scatter graph to correlate *patser* score of predicted Zur binding sites against change in expression generated from microarray data in *E. coli* MG1655  $\Delta zur$ , O157:H7 str. Sakai  $\Delta zur$  and CFT073  $\Delta zur$ . Red squares represent known Zur regulated gene. Blue diamond's represent predicted sites with microarray data that shows a >2-fold change in expression and a p-value <0.05. Green circles represent predicted sites with microarray data that shows <2-fold change in expression. **(a)** *E. coli* MG1655  $\Delta zur$ . **(b)** *E. coli* O157:H7 str. Sakai  $\Delta zur$ . **(c)** *E. coli* CFT073  $\Delta zur$ .

**Figure 7**

End point fluorescence assay of Zur regulated promoters in *E. coli* MG1655 *wt* and  $\Delta zur$ . Fluorescent reporter plasmids (pJI300 cloned with *zur* reregulated promoters) were transformed and expressed in *E. coli* MG1655 *wt* and  $\Delta zur$ . Fluorescence/Absorbance<sub>600</sub> recorded after 16 hr of aerobic incubation at 37°C with shaking (200 RPM), with 3 biological repeats, cultured in zinc depleted Neidhardt's MOPs Minimal Media (NH) with or without 400  $\mu$ M ZnSO<sub>4</sub>.

Strain		Source
<i>E. coli</i> MG1655		(Blattner <i>et al.</i> , 1997)
<i>E. coli</i> W3110		(Hayashi <i>et al.</i> , 2006)
<i>E. coli</i> CFT073		(Welch <i>et al.</i> , 2002)
<i>E. coli</i> O157:H7 str. Sakai		(Hayashi <i>et al.</i> , 2001)
<i>E. coli</i> MG1655 $\Delta$ zur		This Study
<i>E. coli</i> O157:H7 str. Sakai $\Delta$ zur		This Study
<i>E. coli</i> CFT073 $\Delta$ zur		This Study

S.1 Table of Strains

Primer Name	Sequence 5'-3'
ZurFw	CTTAACCCCCACTTTGAGGTGCCCGATGATTGGCTGGAGCTGCTTC
ZurRev	GCCCGACGTGTACAAGGATGTACGCCCTTACTCCTTAGTTCCTATTCC
ZurFwCFT	CTTAACCCCCACTGTCGAGGTGCCCAAT GATTGGCTGGAGCTGCTTC
Scan_ZurFw	CGGCAACAATAAGGGTTCTC
Scan_ZurRev	GACCGATGATGATATGACG
Scan_ZurFwCFT	CCATGACACCCTGTTCTC
ZurHis_Fw	GAACAGTGCCAGCATGATCACTCTGTGCAGGTGAAAAAGAAACCGCG TAGAGGATCGCATCACCATCACCATC
ZurHis-Rv	CCCGCAATGAATATCGCTGGTAATTAATCCCTCCTGCCCGACGTGTACA CATATGAATATCCTCCTTAG
zntA_Lux_F	CGGGATCCACATCAGAGAGGACGCGG
zntA_Lux_R	CCGCTCGAGATGCTGCGAAAAAGAAGCG
C1265_lux_F	TATAGGATCCTACCGATGGAAAGAGTCC
C1265_lux_R	CCGTCCATGGGAAAACAAATGTTGTG
pJI300-PzunCB_R	TCCTTACTAACTAACTAGAGCCAGGAAACAGACTTG
pJI300-PzinT_F	GAGGCCCTTTTCGTCTTCACCATAAGATAGATAAGTAGAACTGAG
pJI300-PzinT_R	TCCTTACTAACTAACTAGAGAACAGCCAGTTTGTAAG
pJI300-PpliG_F	GAGGCCCTTTTCGTCTTCACCAATGCCGCAACGTGATTTTAC
pJI300-PpliG_R	TCCTTACTAACTAACTAGAGATACAGCCTTCCTGATGC
pJI300-PykgM_F	GAGGCCCTTTTCGTCTTCACCGCTAACAATGCCAGAGTTC
pJI300-PykgM_R	TCCTTACTAACTAACTAGAGATTGGGCTTCATCATTTTTAC
pJI300-Pc1265_F	GAGGCCCTTTTCGTCTTCACCGACAGAACATTATTAACAGAG
pJI300-Pc1265_R	TCCTTACTAACTAACTAGAGGTACAAATGACAGTTCTGATTAC

S.2 Table of Primers

Name	Description
pJI300	<i>pSC101, mRFP1, luxCDABE, kan<sup>R</sup></i>
pJI300-PznuA	pJI300 with promoter <i>P<sub>znuA</sub></i>
pJI300-PznuCB	pJI300 with promoter <i>P<sub>znuCB</sub></i>
pJI300-PzinT	pJI300 with promoter <i>P<sub>zinT</sub></i>
pJI300-PpliG	pJI300 with promoter <i>P<sub>pliG</sub></i>
pJI300-PykgM	pJI300 with promoter <i>P<sub>ykgM</sub></i>
pJI300-Pc1265	pJI300 with promoter <i>P<sub>c1265</sub></i>
pKD4	
pSTV28	
PSTV28-C-12His	

S.3 Table of Plasmids

Gene	Annotation	Fold change	P-value
<i>fliA</i>	RNA polymerase, sigma 28 (sigma F) factor (flagella biosynthesis)	-2.21	1.35E-04

S.4 Down-regulated genes in *E. coli* MG1655 wt in zinc-deplete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>zinT</i>	conserved metal-binding protein	213.18	6.70E-03
<i>ykgO</i>	rpmJ (L36) paralogue	205.59	4.32E-03
<i>ykgM</i>	rpmJ (L36) paralogue	195.11	6.89E-03
<i>pliG(ycgK)</i>	predicted protein	39.69	1.87E-02
<i>yebA</i>	predicted peptidase	4.92	2.47E-03
<i>rcnA</i>	membrane protein conferring nickel and cobalt resistance	4.23	1.36E-02
<i>znuC</i>	zinc transporter subunit: ATP-binding component of ABC superfamily	3.86	2.24E-02

S.6. Up-regulated genes in *E. coli* MG1655 wt in zinc-deplete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>rpmE</i>	50S ribosomal subunit protein L31	-2.68	2.66E-07

S.7. Down-regulated genes in *E. coli* MG1655  $\Delta$ zur in zinc-replete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>ykgO</i>	rpmJ (L36) paralogue	519.82	4.88E-04
<i>ykgM</i>	rpmJ (L36) paralogue	459.19	8.20E-04
<i>yodA(zinT)</i>	conserved metal-binding protein	394.57	1.28E-03
<i>znuA</i>	zinc transporter subunit: periplasmic-binding component of ABC superfamily	50.84	7.58E-07
<i>pliG(ycgK)</i>	predicted protein	23.26	1.11E-13
<i>yebA</i>	predicted peptidase	5.88	3.35E-10
<i>znuC</i>	zinc transporter subunit: ATP-binding component of ABC superfamily	4.11	9.73E-06
<i>yodB</i>	predicted cytochrome	3.80	3.64E-03
<i>znuB</i>	zinc transporter subunit: membrane component of ABC superfamily	3.35	6.68E-08

**S.8.** Up-regulated genes in *E. coli* MG1655  $\Delta$ *zur* in zinc-replete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>zntA</i>	zinc/cadmium/mercury/lead-transporting ATPase	-7.83935	9.38E-04

**S.9.** Down-regulated genes in *E. coli* O157:H7 str. Sakai *wt* in zinc-deplete Neidhardt's



Gene	Annotation	Fold change	P-value
<i>rpmE2</i>	50S ribosomal protein L31 type B	405.75	4.46E-04
<i>zinT</i>	hypothetical protein	401.47	9.18E-12
<i>znuA</i>	high-affinity zinc transporter periplasmic component	8.73	2.14E-02
<i>znuC</i>	high-affinity zinc transporter ATPase	7.30	8.66E-10
<i>pliG(ycgK)</i>	hypothetical protein	6.63	3.90E-03
<i>znuB</i>	high-affinity zinc transporter membrane component	6.12	7.30E-03
<i>cstA</i>	putative carbon starvation protein	4.27	3.19E-07
<i>ECs2222</i>	hypothetical protein	3.98	1.23E-03
<i>ECs0328</i>	hypothetical protein	3.95	8.96E-08
<i>ECs4545</i>	hypothetical protein	3.63	5.15E-03
<i>ECs1820</i>	hypothetical protein	3.49	6.30E-06
<i>ECs2789</i>	hypothetical protein	3.24	3.54E-03
<i>yebA</i>	hypothetical protein	3.09	6.00E-04
<i>yjiX</i>	hypothetical protein	3.09	2.00E-03
<i>ECs3494</i>	hypothetical protein	3.09	1.11E-04
<i>ECs1102</i>	hypothetical protein	2.87	2.70E-04
<i>ECs1309</i>	hypothetical protein	2.70	1.97E-08
<i>lbpB</i>	heat shock chaperone lbpB	2.48	9.80E-05
<i>ECs4546</i>	hypothetical protein	2.46	1.08E-06
<i>ECs0329</i>	hypothetical protein	2.37	7.90E-05
<i>ECs1103</i>	hypothetical protein	2.30	1.05E-05
<i>ECs3493</i>	hypothetical protein	2.26	1.45E-03
<i>ECs2223</i>	hypothetical protein	2.22	8.00E-06
<i>ECs1659</i>	hypothetical protein	2.14	4.97E-05
<i>ECs1819</i>	hypothetical protein	2.11	1.79E-04
<i>yodB</i>	putative cytochrome	2.08	3.39E-05
<i>ECs1658</i>	hypothetical protein	2.05	4.99E-04
<i>ECs4994</i>	hypothetical protein	2.03	1.69E-03
<i>ECs1308</i>	hypothetical protein	2.02	3.20E-05
<i>ECs2790</i>	hypothetical protein	2.02	1.38E-05
<i>lbpA</i>	heat shock protein lbpA	2.01	1.80E-02

**S.10.** Up-regulated genes in *E. coli* O157:H7 str. Sakai *wt* in zinc-deplete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>IsoA</i>	hypothetical protein	-52.92	1.59E-03
<i>ECs2957</i>	hypothetical protein	-6.97	1.08E-03
<i>cheZ</i>	chemotaxis regulator CheZ	-4.68	1.91E-04
<i>degP</i>	serine endoprotease	-3.68	5.23E-03
<i>gadC</i>	acid sensitivity protein	-3.58	3.24E-05
<i>katE</i>	hydroperoxidase II	-3.56	2.06E-03
<i>yedE</i>	putative inner membrane protein	-3.45	5.96E-04
<i>pepD</i>	hypothetical protein	-3.33	1.26E-02
<i>tktB</i>	transketolase	-3.09	2.96E-04
<i>pfkB</i>	6-phosphofructokinase 2	-3.08	5.64E-03
<i>menB</i>	naphthoate synthase	-2.87	5.20E-03
<i>yhiM</i>	hypothetical protein	-2.82	2.31E-03
<i>hdhA</i>	7-alpha-hydroxysteroid dehydrogenase	-2.69	2.45E-05
<i>folD</i>	bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/ 5,10-methylene-tetrahydrofolate cyclohydrolase	-2.66	1.42E-04
<i>puuA</i>	putative glutamine synthetase	-2.66	5.08E-03
<i>galM</i>	aldose 1-epimerase	-2.59	4.93E-03
<i>yjjK</i>	putative ABC transporter ATP-binding protein	-2.55	2.45E-04
<i>hdeB</i>	acid-resistance protein	-2.53	1.19E-02
<i>sstT</i>	serine/threonine transporter SstT	-2.49	1.19E-03
<i>yhdH</i>	putative dehydrogenase	-2.47	6.57E-04
<i>bet</i>	Bet	-2.47	9.93E-03
<i>nemA</i>	N-ethylmaleimide reductase	-2.46	1.82E-03
<i>dps</i>	DNA starvation/stationary phase protection protein Dps	-2.44	3.92E-03
<i>mscS</i>	mechanosensitive channel MscS	-2.41	4.72E-03
<i>mrp</i>	putative ATPase	-2.36	3.47E-03
<i>nemR</i>	predicted DNA-binding transcriptional regulator	-2.33	2.15E-03
<i>amyA</i>	cytoplasmic alpha-amylase	-2.32	1.97E-05
<i>ECs1075</i>	hypothetical protein	-2.30	7.23E-03
<i>cbpA</i>	curved DNA-binding protein CbpA	-2.30	8.57E-04
<i>poxB</i>	pyruvate dehydrogenase	-2.27	2.60E-04
<i>menC</i>	O-succinylbenzoate synthase	-2.27	3.04E-05
<i>yqjD</i>	hypothetical protein	-2.26	6.40E-06
<i>icd</i>	isocitrate dehydrogenase	-2.25	1.23E-02
<i>ada</i>	O6-methylguanine-DNA methyltransferase	-2.20	5.95E-03
<i>deoD</i>	purine nucleoside phosphorylase	-2.16	1.26E-03
<i>yeiG</i>	putative esterase	-2.16	4.17E-03
<i>yahK</i>	putative oxidoreductase	-2.14	1.55E-05
<i>manA</i>	mannose-6-phosphate isomerase	-2.12	2.33E-03
<i>ycaO</i>	hypothetical protein	-2.09	1.02E-02
<i>aroD</i>	3-dehydroquinate dehydratase	-2.08	2.60E-06
<i>yqhD</i>	putative oxidoreductase	-2.08	1.41E-02
<i>aqpZ</i>	aquaporin Z	-2.07	4.69E-03
<i>mipA</i>	putative scaffolding protein in the formation of a murein-synthesizing holoenzyme	-2.06	2.69E-03
<i>aroG</i>	phospho-2-dehydro-3-deoxyheptonate aldolase	-2.05	1.44E-02
<i>manY</i>	mannose-specific PTS enzyme IIC	-2.04	1.20E-02
<i>purE</i>	phosphoribosylaminoimidazole carboxylase catalytic subunit	-2.04	5.47E-03
<i>ycaR</i>	hypothetical protein	-2.02	1.19E-02

S.11. Down-regulated genes in *E. coli* O157:H7 str. Sakai  $\Delta$ zur in zinc-replete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>zinT</i>	hypothetical protein	301.50	4.26E-10
<i>rpmE2</i>	50S ribosomal protein L31 type B	239.31	3.26E-04
<i>pliG(ycgK)</i>	hypothetical protein	11.26	6.27E-07
<i>znuB</i>	high-affinity zinc transporter membrane component	5.31	8.01E-03
<i>ECs2789</i>	hypothetical protein	3.85	9.13E-03
<i>zntA</i>	zinc/cadmium/mercury/lead-transporting ATPase	2.95	2.88E-08
<i>ECs1319</i>	hypothetical protein	2.88	1.69E-03
<i>ECs1375</i>	hypothetical protein	2.73	6.05E-04
<i>ECs3236</i>	putative replication protein	2.70	9.35E-03
<i>yodB</i>	putative cytochrome	2.68	1.21E-03
<i>racR</i>	putative phage repressor	2.67	8.17E-03
<i>ECs3488</i>	hypothetical protein	2.65	3.95E-03
<i>ECs1762</i>	hypothetical protein	2.58	1.05E-05
<i>ydbL</i>	hypothetical protein	2.54	9.07E-04
<i>yfgG</i>	hypothetical protein	2.49	8.60E-03
<i>ECs1320</i>	hypothetical protein	2.38	9.09E-07
<i>matB</i>	hypothetical protein	2.35	4.55E-03
<i>ECs4546</i>	hypothetical protein	2.34	8.70E-05
<i>ECs1332</i>	putative colicin immunity protein	2.29	3.85E-03
<i>ECs1659</i>	hypothetical protein	2.28	4.03E-07
<i>ECs2790</i>	hypothetical protein	2.27	2.62E-04
<i>ECs3493</i>	hypothetical protein	2.24	7.46E-06
<i>mgtA</i>	magnesium-transporting ATPase MgtA	2.22	8.82E-06
<i>ECs1819</i>	hypothetical protein	2.21	3.04E-05
<i>ECs2223</i>	hypothetical protein	2.21	7.28E-05
<i>ECs1658</i>	hypothetical protein	2.19	1.17E-05
<i>ECs1941</i>	putative transcriptional regulator	2.16	1.63E-05
<i>ECs1103</i>	hypothetical protein	2.16	1.62E-05
<i>ECs1308</i>	hypothetical protein	2.16	7.11E-05
<i>ECs0329</i>	hypothetical protein	2.12	1.50E-04
<i>yhaV</i>	hypothetical protein	2.11	2.61E-03
<i>yajG</i>	hypothetical protein	2.11	9.17E-03
<i>yadD</i>	hypothetical protein	2.08	2.24E-04
<i>ECs1566</i>	hypothetical protein	2.05	4.15E-03
<i>yagW</i>	putative receptor	2.02	5.25E-04
<i>ECs3485</i>	chaperone-like protein	2.00	6.76E-06
<i>insA</i>	hypothetical protein	2.00	2.68E-03

S.12. Up-regulated genes in *E. coli* O157:H7 str. Sakai  $\Delta$ zur in zinc-replete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>zntA</i>	Lead, cadmium, zinc and mercury transporting	-9.74	3.39E-05

S.13. Down-regulated genes in *E. coli* CFT073 wt in zinc-deplete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>ykgM</i>	50S ribosomal protein L31 type B-1	2290.97	2.82E-04
<i>c0406</i>	Putative 50S ribosomal protein L36	1217.93	1.16E-03
<i>yodA(zinT)</i>	Hypothetical protein yodA	1030.99	4.76E-10
<i>c1266</i>	Hypothetical protein	60.92	2.43E-04
<i>c1265</i>	Outer membrane heme/hemoglobin receptor	41.83	1.25E-06
<i>c1267</i>	Hypothetical protein	31.56	3.01E-12
<i>pliG(ycgK)</i>	Protein pliG(ycgK) precursor	24.58	3.50E-10
<i>yhjX</i>	Hypothetical protein yhjX	11.96	4.85E-07
<i>znuC</i>	High-affinity zinc uptake system ATP-binding	7.48	8.98E-06
<i>znuB</i>	High-affinity zinc uptake system membrane	6.60	5.80E-08
<i>ribB</i>	3,4-dihydroxy-2-butanone 4-phosphate synthase	4.72	1.12E-06
<i>c1111</i>	Hypothetical protein	3.68	1.08E-02
<i>c1989</i>	Putative acid shock protein	2.85	5.85E-04
<i>yebA</i>	Hypothetical protein yebA precursor	2.74	4.46E-07
<i>asr</i>	Acid shock protein precursor	2.72	3.76E-04
<i>sodA</i>	Superoxide dismutase (Mn)	2.72	1.09E-05
<i>ycdO</i>	Protein ycdO	2.59	1.14E-05
<i>thrC</i>	Threonine synthase	2.57	1.23E-05
<i>exbD</i>	Biopolymer transport exbD protein	2.48	5.99E-03
<i>ibpB</i>	16 kDa heat shock protein B	2.36	5.29E-04
<i>yjiY</i>	Hypothetical protein yjiY	2.25	1.06E-04
<i>rpsV</i>	30S ribosomal protein S22	2.22	9.04E-05
<i>c3196</i>	Hypothetical protein	2.20	1.33E-02
<i>entF</i>	Enterobactin synthetase component F	2.19	6.30E-03
<i>tonB</i>	TonB protein	2.17	1.43E-04
<i>yfdQ</i>	Hypothetical protein yfdQ	2.15	1.80E-04
<i>thrA</i>	Aspartokinase I	2.14	8.85E-06
<i>ycdB</i>	Hypothetical protein ycdB precursor	2.10	5.64E-04
<i>c3200</i>	Hypothetical protein	2.10	4.15E-05
<i>hdeB</i>	Protein hdeB precursor	2.03	1.15E-03
<i>yecF</i>	Hypothetical protein yecF	2.01	4.60E-03

**S.14.** Up-regulated genes in *E. coli* CFT073 wt in zinc-deplete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>yfiD</i>	Protein yfiD	-23.92	1.37E-04
<i>ndh</i>	NADH dehydrogenase	-9.65	3.33E-06
<i>c2963</i>	Hypothetical protein	-6.17	1.85E-02
<i>cusX</i>	Hypothetical protein cusX precursor	-5.07	8.69E-03
<i>eutH</i>	Ethanolamine utilization protein eutH	-4.37	1.21E-02
<i>cusC</i>	Probable outer membrane lipoprotein cusC	-4.31	1.74E-02
<i>c0550</i>	Hypothetical protein	-4.08	2.41E-02
<i>narI</i>	Respiratory nitrate reductase 1 gamma chain	-3.53	3.56E-02
<i>cusB</i>	Putative copper efflux system protein cusB	-3.02	6.93E-03
<i>fbp</i>	Fructose-1,6-bisphosphatase	-2.91	6.42E-03
<i>c2248</i>	Hypothetical protein	-2.85	3.70E-02
<i>yqhA</i>	Hypothetical protein yqhA	-2.67	3.65E-03
<i>c0965</i>	Phage baseplate assembly protein	-2.47	3.38E-02
<i>c0939</i>	Hypothetical protein	-2.40	4.24E-03
<i>napG</i>	Ferredoxin-type protein napG	-2.40	6.56E-03
<i>yihG</i>	Hypothetical protein yihG	-2.39	1.32E-02
<i>c0956</i>	Possible secretory protein	-2.37	1.81E-02
<i>yciI</i>	Protein yciI	-2.29	1.80E-03
<i>c3178</i>	Hypothetical protein	-2.29	2.13E-02
<i>rpsL</i>	30S ribosomal protein S12	-2.28	2.73E-05
<i>c3153</i>	Putative outer membrane protein of prophage	-2.24	6.48E-03
<i>c0974</i>	Putative phage tail protein	-2.21	6.38E-04
<i>nirB</i>	Nitrite reductase [NAD(P)H] large subunit	-2.21	2.31E-02
<i>c0950</i>	Putative capsid scaffolding protein	-2.17	6.07E-05
<i>inaA</i>	Protein inaA	-2.16	2.10E-06
<i>c0958</i>	Putative membrane protein	-2.14	9.46E-04
<i>c0957</i>	Fels-2 prophage: probable prophage lysozyme	-2.13	2.16E-03
<i>c0954</i>	Putative capsid completion protein	-2.10	1.26E-03
<i>c0962</i>	Putative phage tail protein	-2.07	6.88E-03
<i>c0955</i>	Probable phage tail protein	-2.07	9.37E-03
<i>c2348</i>	Outer membrane porin protein nmpC precursor	-2.04	2.76E-06
<i>artP</i>	Arginine transport ATP-binding protein artP	-2.04	1.25E-04
<i>yiaH</i>	Hypothetical protein yiaH	-2.01	2.23E-02
<i>napF</i>	Ferredoxin-type protein napF	-2.01	3.83E-03

**S.15.** Down-regulated genes in *E. coli* CFT073  $\Delta$ zur in zinc-replete Neidhardt's

Gene	Annotation	Fold change	P-value
<i>ykgM</i>	50S ribosomal protein L31 type B-1	887.31	2.49E-04
<i>c0406</i>	Putative 50S ribosomal protein L36	558.33	4.60E-04
<i>yodA</i>	Hypothetical protein yodA	521.30	6.33E-06
<i>c1544</i>	Hypothetical protein	443.10	8.75E-11
<i>c1545</i>	Hypothetical protein	406.48	3.93E-11
<i>c1540</i>	Lambda Regulatory protein CIII	80.25	9.49E-13
<i>c1523</i>	Unknown protein encoded by bacteriophage	77.83	1.81E-03
<i>c1543</i>	Putative superinfection exclusion protein B of	73.07	1.95E-03
<i>gamW</i>	Putative host-nuclease inhibitor protein Gam of	67.59	3.57E-03
<i>c1548</i>	Putative Regulatory protein CII of bacteriophage	65.72	4.29E-05
<i>c1524</i>	Hypothetical protein	62.98	2.03E-09
<i>c1528</i>	Hypothetical protein	60.30	5.91E-04
<i>c1539</i>	Hypothetical protein	59.65	5.77E-03
<i>c1536</i>	Putative recombination protein Bet of prophage	57.36	2.76E-03
<i>c1531</i>	Hypothetical protein	53.11	2.21E-03
<i>c1546</i>	Repressor protein	52.44	1.05E-05
<i>c1551</i>	Putative exclusion protein ren of prophage	49.57	1.99E-03
<i>c1549</i>	Putative replication protein O of bacteriophage	46.41	4.11E-04
<i>c1266</i>	Hypothetical protein	46.34	7.33E-06
<i>c1527</i>	Hypothetical protein	42.50	2.26E-03
<i>c1541</i>	Putative single-stranded DNA binding protein of	41.91	3.12E-04
<i>c1525</i>	Hypothetical protein	38.07	3.81E-06
<i>c1267</i>	Hypothetical protein	34.69	4.99E-04
<i>c1542</i>	Lambda ant-restriction protein	33.22	2.45E-04
<i>c1533</i>	Hypothetical protein	32.35	2.10E-09
<i>c1265</i>	Outer membrane heme/hemoglobin receptor	30.98	2.52E-05
<i>c1526</i>	Hypothetical protein	29.86	1.94E-05
<i>c1532</i>	Unknown protein encoded within prophage	29.63	4.65E-06
<i>c1550</i>	Putative replication protein P of bacteriophage	22.72	1.35E-08
<i>c1534</i>	Putative exonuclease encoded by prophage	21.10	2.55E-02
<i>c1538</i>	Hypothetical protein	16.99	1.12E-07
<i>c1522</i>	Hypothetical protein	13.94	5.72E-07
<i>pliG(ycgK)</i>	Protein pliG(ycgK) precursor	11.80	4.17E-07
<i>ybcS</i>	Probable lysozyme from lambdoid prophage	11.56	5.58E-11
<i>c1561</i>	Lysis protein S homologue from lambdoid	9.30	4.33E-05
<i>c1572</i>	Putative capsid assembly protein of prophage	9.16	2.80E-04
<i>c1563</i>	Putative Rz endopeptidase from lambdoid	9.13	1.51E-07
<i>c1575</i>	Putative capsid protein of prophage	8.65	2.78E-06
<i>znuC</i>	High-affinity zinc uptake system ATP-binding	7.80	2.53E-06
<i>c1450</i>	Putative capsid protein of prophage	7.19	1.53E-06
<i>znuB</i>	High-affinity zinc uptake system membrane	6.99	3.05E-06
<i>c1574</i>	Putative capsid protein of prophage	6.90	1.70E-08
<i>c1576</i>	Hypothetical protein	6.45	9.93E-08
<i>c1580</i>	Tail protein	6.45	1.06E-07

<i>c1448</i>	Putative capsid assembly protein of prophage	6.25	1.00E-05
<i>c1583</i>	Putative tail component of prophage	5.71	2.31E-07
<i>rus</i>	Crossover junction endodeoxyribonuclease <i>rusA</i>	5.70	9.18E-06
<i>c1591</i>	Hypothetical protein	4.50	1.15E-05
<i>c1585</i>	Putative tail component of prophage	4.39	6.22E-08
<i>ydfD</i>	Hypothetical protein <i>ydfD</i>	4.34	1.90E-06
<i>c1569</i>	Putative DNA packaging protein of prophage;	4.27	1.35E-06

**S.16.** Top fifty up-regulated genes in *E. coli* CFT073  $\Delta$ *zur* in zinc-replete conditions

