



Studying the transcriptional regulator SlyA and the  
global transcriptomic effects of growth rate in  
*Escherichia coli* K-12

**A thesis submitted in part fulfilment for the degree of**

**Doctor of Philosophy**

**Department of Molecular Biology and Biotechnology,**

**The University of Sheffield**

**Thomas David Curran**

**BSc (Hons) University of Sheffield**

**September 2014**

## **Abstract**

The MarR-family regulator SlyA has been shown to directly regulate the expression of two genes in *Escherichia coli* K-12; *hlyE* and *fimB*. In both cases SlyA regulates expression by antagonising the repressive effects of H-NS. This thesis documents the research carried out in order to determine the breadth of the SlyA regulon and whether the corresponding SlyA-regulated promoters were also repressed by H-NS.

In this work it has been demonstrated that the overexpression of *slyA* caused the transcription of 44 genes in *E. coli* to change significantly when analysed via a microarray transcriptomic method. Of these genes, 25% are already known to be repressed by H-NS and, of the 39 SlyA up-regulated genes, 64% have previously been determined to be regulated by another antagonist of H-NS repression, LeuO. Through gel-shift assays, direct binding of SlyA to the promoter regions of nine additional operons has been demonstrated. More in-depth analysis of SlyA binding within the *PmdtM* promoter also suggested that the presence of at least one half-site of the TTA<6nt>TAA consensus sequence was required for strong SlyA binding to target DNA.

The activity of SlyA was not regulated by growth rate, contrary to previous suggestions. However, microarray analyses of glucose limited steady-state chemostat cultures at dilution rates of 0.05, 0.1, 0.2 and 0.5 h<sup>-1</sup> did show that the transcript abundance of 253 genes changed significantly as *E. coli* K-12 doubling time was reduced (86 were up-regulated, 167 were down-regulated). A high proportion of those genes that were down-regulated were associated with secondary metabolism and were regulated by cAMP-CRP. The activities of 167 transcriptional regulators were inferred across the range of growth rates studied, with 38 exhibiting altered activity as growth rate increased.

Transmission electron microscopy and Western blotting were applied to confirm and further characterise the surprising up-regulation of transcripts associated with flagellar biosynthesis at higher growth rates. The decreased expression of the *lsr* operon at higher growth rates was associated with increased rate of autoinducer-2 secretion.

Thus this work has provided new insight into the extent of the *E. coli* K-12 SlyA regulon and the transcriptional reprogramming associated with changes in growth rate.

## **Acknowledgements**

I would like to start by thanking Professor Jeff Green for all his support, advice and patience throughout the last four years. I also express thanks to the late Professor Pete Artymiuk, whose advice and perspective was always greatly valued and appreciated.

Thank you to Chris Hill for all the assistance with the electron microscopy work, Nikki Whiting for providing FlhDC antibodies and Iain Keane for providing the *E. coli* mutant library.

Thanks to everyone (past and present) in F10, you all made the workplace a lovely place to be. Thanks especially to Mel, Laura and Matt. Matt, a lot of the work I have done would not have been possible without your help, thank you so much.

Thanks to Arthur, Ash and Nikki, three brilliant housemates and friends. I was more myself around you than a lot of people, thanks for putting up with that!

Huge thanks go to all my family. Mum and Dad, none of this would have been possible without you, thank you so much for all the support you have given me over the years. Mary and Paul (alphabetically listed to avoid any perceived favouritism), you are quite simply awesome, always there to chat to (usually about rubbish) and laugh with, EPIC thanks. Thanks also to Jen and Isabelle, a superb sister-in-law and an extremely excellent niece.

Thanks also to Helen and Gordon Middleton, my acquired “science-family”. You have always showed an interest in how my work is going and have always made me feel very at home, for which I couldn’t be more grateful.

Catrin and James, thanks for our meet-ups that have never been anything less than brilliant, and the perfect antidote to work-induced lows.

Last, but certainly not least, is Eleri. You’ve seen more behind the scenes than anyone else, thanks for your unending support and level-headedness. You’ve calmed me down and picked me up more times than I can count. I’m lucky to have you, and not a single word of this would have been written without you!

## **Contents**

Abstract .....	I
Acknowledgements .....	II
List of figures .....	VII
List of tables .....	X
Presentations .....	XII
Abbreviations .....	XIII
1. Introduction .....	1
1.1. The RNA polymerase holoenzyme .....	1
1.2. Transcription activators .....	5
1.2.1. Class I mechanisms of transcription activation .....	5
1.2.2. Class II mechanisms of transcription activation .....	7
1.2.3. Class III mechanisms of transcription activation .....	8
1.2.4. Altering promoter DNA conformation .....	9
1.2.5. RNA polymerase appropriation .....	9
1.3. Transcription repressors .....	12
1.3.1. Steric hindrance of RNA polymerase .....	12
1.3.2. DNA looping .....	14
1.3.3. Anti-activation .....	15
1.4. Nucleoid associated proteins and DNA supercoiling .....	15
1.5. Complex activator/repressor systems and integration of regulatory signals .....	16
1.6. The MarR family of transcription regulators .....	19
1.6.1. MarR family proteins as regulators of virulence .....	23
1.7. SlyA of <i>Salmonella enterica</i> Serovar Typhimurium .....	24
1.8. SlyA of <i>Escherichia coli</i> .....	27
1.9. Summary and aims .....	30
2. Materials and methods .....	31
2.1. Strains and plasmids .....	31
2.2. Media .....	32
2.2.1. Rich media .....	32
2.2.2. Minimal media .....	35
2.2.3. Media supplements .....	35
2.3. Growth of bacterial strains .....	36
2.3.1. Chemostat culture of <i>Escherichia coli</i> MG1655 .....	36
2.3.2. Chemostat culture of <i>Escherichia coli</i> mutant library .....	36

2.3.3.	Storage of strains.....	36
2.4.	Bacterial transformation.....	36
2.4.1.	Production of electrically competent cells .....	36
2.4.2.	Transformation of electrically competent cells.....	37
2.5.	Sampling of chemostat cultures .....	37
2.5.1.	Genomic DNA .....	37
2.5.2.	RNA .....	37
2.5.3.	Protein pellets.....	37
2.5.4.	Supernatant .....	38
2.5.5.	Cell dry weight.....	38
2.6.	Nucleic acid methods.....	38
2.6.1.	Primers .....	38
2.6.2.	PCR amplification.....	41
2.6.3.	Colony PCR .....	42
2.6.4.	PCR purification .....	42
2.6.5.	Isolation of RNA from stored pellets .....	43
2.6.6.	Genomic DNA preparation .....	43
2.6.7.	Plasmid purification .....	44
2.6.8.	Quantification of nucleic acid concentration .....	44
2.6.9.	Agarose gel electrophoresis .....	44
2.6.10.	Gel extraction.....	45
2.6.11.	Digestion of DNA with restriction endonucleases.....	45
2.6.12.	Gene deletion using the Lambda-Red system.....	45
2.6.13.	Creation of pET28a: <i>slyA</i> plasmid using the In-Fusion HD cloning system ..	47
2.6.14.	Creation of the pGS2469 plasmid using the aLICator ligation independent cloning and expression system.....	47
2.6.15.	Transcriptomics: Microarray analyses .....	48
2.6.16.	Transposon mediated differential hybridisation (TMDH) arrays .....	51
2.7.	Protein Methods .....	53
2.7.1.	Measurement of protein concentration.....	53
2.7.2.	Denaturing gel electrophoresis (SDS-PAGE).....	53
2.7.3.	Overproduction of SlyA.....	54
2.7.4.	Production of cell-free extracts .....	55
2.7.5.	Purification of recombinant SlyA by affinity chromatography .....	55
2.7.6.	Transfer of proteins onto Hybond-C Extra nitrocellulose membrane.....	56
2.7.7.	Western blotting.....	56

2.7.8.	Electrophoretic Mobility Shift Assays (EMSAs).....	58
2.8.	P1 phage transduction.....	59
2.8.1.	Preparation of transducing lysate.....	59
2.8.2.	Transduction.....	60
2.9.	Transmission electron microscopy (TEM) of <i>Escherichia coli</i> .....	60
2.10.	<i>Vibrio harveyi</i> BB170 autoinducer-2 Bioassay.....	61
2.11.	<sup>1</sup> H-NMR of steady-state culture supernatant samples.....	61
3.	Studying the gene regulator SlyA.....	62
3.1.	Introduction.....	62
3.2.	Gene expression of <i>E. coli</i> MG1655Δ <i>slyA</i> at different growth rates.....	62
3.3.	Gene expression of <i>E. coli</i> MG1655 overproducing SlyA.....	66
3.4.	Selecting targets of interest.....	71
3.5.	Synthesising promoter regions to test for possible SlyA binding.....	77
3.6.	Overexpression and purification of SlyA.....	77
3.7.	Electrophoretic mobility shift assays of SlyA with suspected target promoter regions.....	80
3.8.	Re-analysis of <i>PmdtM</i> , searching for the ATT<6nt>TAA motif and subsequent inference of likely SlyA binding sites in other promoter regions.....	89
3.9.	Discussion.....	91
4.	Studying changes in <i>Escherichia coli</i> gene expression in response to growth rate.....	95
4.1.	Introduction.....	95
4.2.	Chemostat growth parameters and associated data.....	96
4.3.	<i>Escherichia coli</i> strain MG1655 transcript profiles at fixed growth rates.....	99
4.4.	Genes up-regulated at increased growth rate.....	101
4.5.	Genes down-regulated at increased growth rate.....	105
4.6.	Identifying transcription factors playing a role in growth rate adaptation.....	109
4.7.	Quantifying activity of alternative sigma factors FliA and RpoS.....	113
4.8.	Development of a Transposon-Mediated Differential Hybridisation (TMDH) approach in order to test gene fitness over increasing growth rate in <i>E. coli</i> MG1655 ...	114
4.9.	Output from a preliminary TMDH study of <i>E. coli</i> strain MG1655 comparing lower and higher growth rates.....	120
4.10.	Discussion.....	123
5.	Studying <i>Escherichia coli</i> flagellar gene expression and autoinducer-2 production in response to growth rate.....	131
5.1	Introduction.....	131
5.2	Analysis of transcriptomic data attributed to genes associated with motility and flagellar assembly at increased growth rates.....	134

5.3	Quantifying cellular levels of FlhDC, FliA, FlgM and FliC protein over increasing growth rate .....	137
5.4	Transmission electron microscopy of <i>E. coli</i> MG1655 cells cultured at different growth rates.....	142
5.5	Analysis of transcriptomic data in relation to AI-2 synthesis, export and import as growth rate increases.....	143
5.6	Quantifying the extracellular build-up of AI-2 with respect to growth rate .....	147
5.7	Discussion .....	149
6.	General discussion .....	156
7.	Appendices.....	161
8.	References.....	162

## **List of figures**

Figure 1.1 Schematic diagram of the RNA polymerase holoenzyme bound at an ideal promoter.....	3
Figure 1.2 Schematic diagrams of Class I, II and III types of interaction between transcription activators and RNAP.....	6
Figure 1.3 Simple activation by DNA conformational change.....	10
Figure 1.4 Representations of activation by RNA polymerase appropriation.....	11
Figure 1.5 Mechanisms of transcription repression.....	13
Figure 1.6 Examples of complex activator/repressor systems of transcription regulation....	18
Figure 1.7 Structure of an <i>E. coli</i> MarR dimer.....	22
Figure 1.8 Structure of an <i>E. coli</i> SlyA dimer bound to DNA.....	22
Figure 3.1 Schematic diagram of work carried out in Chapter 3 pertaining to transcriptomics undertaken to elucidate the SlyA regulon.....	63
Figure 3.2 Typical result of colony PCR amplification of the <i>slyA</i> region of <i>E. coli</i> MG1655 and <i>E. coli</i> MG1655 $\Delta$ <i>slyA</i> .....	65
Figure 3.3 Western blot showing amount of detectable SlyA in chemostat culture samples of <i>E. coli</i> MG1655(pET28a) and <i>E. coli</i> MG1655(pET28a: <i>slyA</i> ).....	67
Figure 3.4 Fold changes in gene expression levels of <i>E. coli</i> MG1655(pET28a: <i>slyA</i> ) compared to <i>E. coli</i> MG1655(pET28a).....	69
Figure 3.5 Gene expression profiles of operons and individual genes of interest in <i>E. coli</i> MG1655(pET28a) and <i>E. coli</i> MG1655(pET28a: <i>slyA</i> ).....	73
Figure 3.6 Products of PCR amplification of promoter regions possibly bound by SlyA.....	79
Figure 3.7 Typical AKTA trace recording the elution of SlyA from a Hi-Trap chelating affinity column.....	81
Figure 3.8 Coomassie Blue stained SDS-PAGE gel of fractions from SlyA purification....	82
Figure 3.9 EMSA gels showing binding of SlyA to various gene or operon promoter regions.....	83



Figure 3.10 EMSA illustrating lack of SlyA binding to <i>sgcQ</i> - <i>sgcA</i> intergenic region.....	85
Figure 3.11 The presence of the alarmone ppGpp does not alter SlyA binding at the <i>mdtM</i> promoter.....	85
Figure 3.12 Schematic for production of <i>mdtM</i> promoter DNA, searching for possible SlyA binding sites and subsequent production of truncated promoter regions lacking these sites.....	86
Figure 3.13 EMSA of <i>PmdtM</i> and its truncated variants with successive removal of four suspected SlyA binding sites.....	88
Figure 3.14 Electrophoretic Mobility Shift Assay (EMSA) of <i>PmdtM</i> and its truncated derivatives, with and without the presence of 0.1 mg/ml Heparin.....	90
Figure 3.15 Searching for SlyA binding site motif TTA<6nt>TAA within <i>PmdtM</i> .....	90
Figure 4.1 Typical proton-NMR trace of culture supernatants for detection of over-metabolites.....	97
Figure 4.2 Schematic of key reactions related to carbon uptake and metabolism that have been estimated in this work.....	98
Figure 4.3 Graphical representation of all significant gene expression changes plotted against increasing dilution rate ( $h^{-1}$ ) after statistical filtering.....	100
Figure 4.4 Schematic of the core arginine biosynthetic pathway with overlaid data for transcript levels of the corresponding genes at $d= 0.5 h^{-1}$ as a fold change relative to $d= 0.05 h^{-1}$ .....	103
Figure 4.5 Figure 4.5 Graphs illustrating activity trends of transcription factors CdaR, LsrR and CynR which are representative of their respective clusters.....	111
Figure 4.6 Schematic work-flow of the stages leading to eventual data output from TMDH study of <i>E. coli</i> strain MG1655 gene fitness over increased growth rate.....	116
Figure 4.7 Rationale behind crucial criterion for classing probes as Informative or Un-informative.....	119
Figure 5.1 Schematic diagram of hierarchy within flagellar gene expression.....	132
Figure 5.2 Western blot showing amounts of FlhDC detected in samples taken from steady-state cultures grown at different dilution rates.....	139

Figure 5.3 Western blots showing amounts of FliA, FlgM and FliC detected in samples taken from steady-state cultures grown at different dilution rates.....	140
Figure 5.4 Graphs displaying (A) relative amounts of protein detected in Western blots and (B) fold change in associated transcript abundance at different dilution rates for the proteins FlhDC, FliA, FlgM and FliC.....	141
Figure 5.5 Typical transmission electron micrographs of <i>E. coli</i> cells grown at dilution rates of 0.05, 0.1, 0.2 and 0.5 h <sup>-1</sup> .....	144
Figure 5.6 Plot of average <i>E. coli</i> cell length (μm) against dilution rate (h <sup>-1</sup> ).....	145
Figure 5.7 Schematic of the stages in AI-2 synthesis, export and import in <i>E. coli</i> .....	146
Figure 5.8 Relative rates of extracellular AI-2 accumulation plotted against dilution rate..	150
Figure 5.9 Schematic diagram of flagellar gene regulation with overlaid transcriptomic and Western blot data.....	151

## **List of tables**

Table 2.1 Strains of <i>Escherichia coli</i> used in this study.....	31
Table 2.2 Strain of <i>Vibrio harveyi</i> used in this study.....	31
Table 2.3 Plasmids used in this study.....	31
Table 2.4 Primers used in this study.....	38
Table 3.1 Fold change in transcript abundance of <i>E. coli</i> strain MG1655(pET28a: <i>slyA</i> ) compared to expression levels in MG1655(pET28a) .....	70
Table 3.2 Expression data for genes in operons and single genes of interest, when comparing <i>E. coli</i> MG1655(pET28a) and <i>E. coli</i> MG1655(pET28a: <i>slyA</i> ).....	72
Table 3.3 Rationale for production of promoter regions of interest and primers used.....	78
Table 3.4 Possible SlyA binding sites within promoter regions shown to be directly bound by SlyA, based on the binding site motif: TTA<6nt>TAA.....	92
Table 4.1 Calculated rates of glucose consumption, biomass production, acetate production and predicted rate of carbon flux through the Tricarboxylic Acid (TCA) cycle.....	98
Table 4.2 Numbers of genes differentially expressed at different growth rates.....	100
Table 4.3 Transcripts exhibiting increased abundance as growth rate increases, at either $d=0.2$ or $0.5\text{ h}^{-1}$ .....	102
Table 4.4 Transcripts exhibiting decreased abundance as growth rate increases, at either $d=0.2$ or $0.5\text{ h}^{-1}$ .....	106
Table 4.5 Gene regulators exhibiting significant differential activity in response to growth rate.....	110
Table 4.6 Stages in filtering out un-informative probes from subsequent analysis.....	118
Table 4.7 Compilation of genes that were observed to be essential in Biological Replicate #3 of the TMDH study, and were up-regulated as growth rate increased in gene-expression microarrays.....	122
Table 5.1 Fold change in transcript abundance of genes associated with the flagellar regulon, organised by promoter class.....	135

Table 5.2 Observed average cell length and width in TEM micrographs of <i>E. coli</i> cells grown at various dilution rates .....	145
Table 5.3 Relative levels of AI-2 detected in steady-state <i>E. coli</i> MG1655 culture supernatant samples using the <i>V. harveyi</i> strain BB170.....	148

## **Presentations**

### **SGM Autumn Conference - University of Sussex, 2-4 September 2013**

Poster presented: Transcriptomic changes in *Escherichia coli* associated with growth rate, and studying the gene regulator SlyA.

## **Abbreviations**

AB medium	-	Autoinducer bioassay medium
AI-2	-	Autoinducer-2
Amp	-	Ampicillin
APS	-	Ammonium persulfate
ATP	-	Adenosine triphosphate
bp	-	Base pairs
Btn	-	Biotin
cAMP	-	Cyclic-AMP
cDNA	-	Complementary DNA
CDW	-	Cell dry weight
CRISPR	-	Clusters of regularly interspersed short palindromic repeats
CRP	-	cAMP receptor protein
CTP	-	Cytidine triphosphate
dATP	-	Deoxyadenosine triphosphate
dCTP	-	Deoxycytidine triphosphate
dGTP	-	Deoxyguanosine triphosphate
DNA	-	Deoxyribonucleic acid
dNTP	-	Deoxyribonucleotide triphosphate
DPD	-	4,5-dihydroxy-2,3-pentanedione
DTT	-	Dithiothreitol
dTTP	-	Deoxythymidine triphosphate
dUTP	-	Deoxyuridine triphosphate
<i>E. coli</i>	-	<i>Escherichia coli</i>
EDTA	-	Ethylenediaminetetraacetic acid
EHEC	-	Enterohemorrhagic <i>Escherichia coli</i>
EIEC	-	Enteroinvasive <i>Escherichia coli</i>
EMSA	-	Electrophoretic mobility shift assay
FAD <sup>+</sup> /FADH <sub>2</sub>	-	Oxidised and reduced forms of flavin adenine dinucleotide
FMN	-	Flavin mononucleotide
GTP	-	Guanosine triphosphate
IPTG	-	Isopropyl $\beta$ -D-1-thiogalactopyranoside
Kn	-	Kanamycin
LB medium	-	Luria Bertani medium
LD50	-	Median lethal dose
LPS	-	Lipopolysaccharide

mRNA	-	Messenger RNA
NAD <sup>+</sup> /NADH	-	Oxidised and reduced forms of nicotinamide adenine dinucleotide
NAP	-	Nucleoid associated protein
NMR	-	Nuclear magnetic resonance
nt	-	Nucleotides
OD	-	Optical density
PAGE	-	Polyacrylamide gel electrophoresis
PBS	-	Phosphate buffered saline
PCR	-	Polymerase chain reaction
PTS	-	Phosphotransferase system
RNA	-	Ribonucleic acid
RNAP	-	RNA Polymerase
RT	-	Reverse transcription
<i>S. enterica</i>	-	<i>Salmonella enterica</i>
SAH	-	S-adenosyl-L-homocysteine
SAM	-	S-adenosyl-L-methionine
SDS	-	Sodium dodecyl sulphate
SELEX	-	Systematic evolution of ligands by exponential enrichment
SPI-2	-	<i>Salmonella</i> pathogenicity island 2
SRH	-	S-ribosyl-L-homocysteine
sRNA	-	Small RNA
STD	-	Standard deviation
T2SS	-	Type-II secretion system
TCA	-	Tricarboxylic acid
TEM	-	Transmission electron microscopy
TEMED	-	Tetramethylethylenediamine
TMDH	-	Transposon mediated differential hybridisation
TSP	-	Trimethylsilyl propionate
TTP	-	Thymidine triphosphate
UTP	-	Uridine triphosphate
wHTH	-	Winged-helix-turn-helix
$\alpha$ CTD	-	$\alpha$ -N-terminal domain
$\alpha$ NTD	-	$\alpha$ -C-terminal domain

## **1. Introduction**

In the Gram-negative bacterial species *Escherichia coli*, gene regulation is vital for survival and adaptation. An *E. coli* cell is continually receiving and quantifying many environmental and metabolic signals and then activating and/or repressing genes in order to adapt to the changing conditions. Not only is it vital for basic bacterial survival, but is also a pivotal stage in activation of virulence genes in pathogenic strains of *E. coli* wherein the cell can detect when it enters a host environment such as the human intestinal tract (Beltrametti *et al.*, 1999; Abe *et al.*, 2002; Nakanishi *et al.*, 2006; Chekabab *et al.*, 2014). The regulation of protein production can be at the stage of translation, though it is generally considered to be more efficient to control protein levels in the bacterial cell at the level of gene transcription. This review of the current literature covers the essential stages in regulation of gene transcription, the differing methods by which a gene may be regulated and the current knowledge of the *E. coli* gene regulator; SlyA.

### **1.1. The RNA polymerase holoenzyme**

One of the key factors in gene regulation is the interaction of the bacterial enzyme RNA Polymerase (RNAP) with the promoter DNA located upstream of a gene. RNAP is responsible for transcription in all bacteria and was originally characterised in the 1960s (Burgess, 1969). The core RNAP enzyme that is capable of transcription initiation but not specific promoter-directed transcription consists of four subunits,  $\alpha_2\beta\beta'$ , which are encoded by the genes *rpoA*, *rpoB* and *rpoC* respectively. The gene *rpoZ* encodes the additional  $\omega$  subunit that associates with the  $\beta'$  subunit of the core RNAP complex and has been shown to aid in the assembly of functional RNAP (Mukherjee and Chatterji, 1997) and more recently has been linked to roles in the *E. coli* starvation response mediated by ppGpp levels (Vrentas *et al.*, 2005) and altering the  $\sigma$  factor preference of RNAP (Geertz *et al.*, 2011).

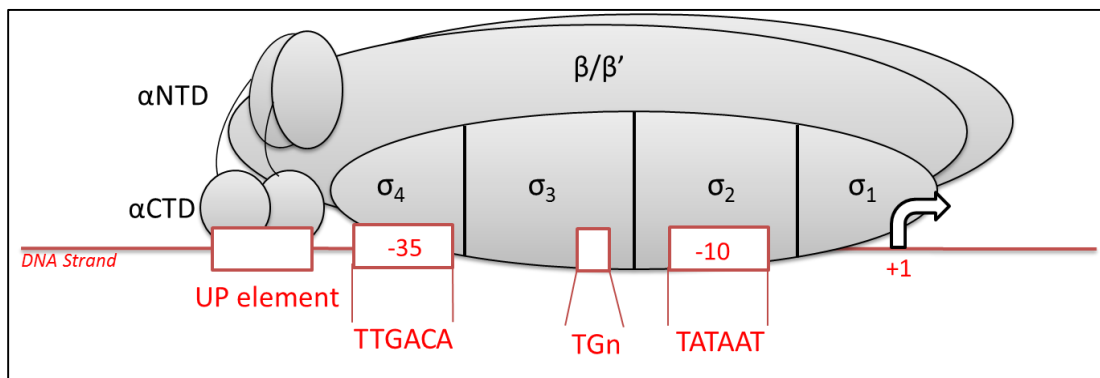
The core RNAP complex is approximately 400 kDa in size and is highly conserved across all bacteria. High resolution structural studies show that it adopts a Crab-Claw structure (Zhang *et al.*, 1999) similar to that of the Yeast RNAP complex (Fu *et al.*, 1999). The  $\beta$  and  $\beta'$  subunits have regions that allow binding of DNA and channelling of the mRNA product (Korzheva *et al.*, 2000), and the two identical  $\alpha$  subunits are each made up of an N-terminal domain ( $\alpha$ NTD) linked to a C-terminal



domain ( $\alpha$ CTD) by a flexible linker region (Blatter *et al.*, 1994). The larger  $\alpha$ NTD allows for the dimerization and also assists in the assembly of  $\beta$  and  $\beta'$  (Zhang and Darst, 1998), whereas the smaller  $\alpha$ CTD has DNA-binding properties and can play an important role in specific promoter binding and activity (Gourse *et al.*, 2000). Detailed structural evidence for the functions of these major RNAP subunits and how they interact with promoter DNA is now available (Murakami *et al.*, 2002a; Murakami *et al.*, 2002c).

For RNAP to begin transcription at any promoter it must first bind a sigma factor ( $\sigma$ ) to form the RNAP holoenzyme ( $\alpha_2\beta\beta'\omega\sigma$ ) (Burgess *et al.*, 1969). The sigma factor has three primary functions: (1) to provide the RNAP enzyme complex with sequence specificity for a promoter/transcription start site; (2) correctly position the RNAP enzyme at a promoter; (3) facilitate the unwinding of the DNA duplex (Wosten, 1998). Most bacteria encode multiple  $\sigma$ -factors which can each bind to the RNAP enzyme complex and provide it with different specificity for different sets of promoters (Gross *et al.*, 1998). All these  $\sigma$  factors, with the exception of  $\sigma^{54}$  (Merrick, 1993), are multi-domain proteins with up to four domains joined by linker regions, these domains are named  $\sigma 1$ ,  $\sigma 2$ ,  $\sigma 3$  and  $\sigma 4$ .

It is the  $\sigma 4$  and  $\sigma 2$  sub-domains that recognise and bind the -35 and -10 hexamer regions that are hallmarks of promoters recognised by the house-keeping sigma factor,  $\sigma^{70}$  in *E. coli*. Consensus sequences have been established for the -35 and -10 regions as TTGACA and TATAAT respectively (Harley and Reynolds, 1987; Murakami *et al.*, 2002b). Additional promoter DNA elements include the extended -10 element (3-4 bp) recognised by  $\sigma 3$  (Sanderson *et al.*, 2003) and the UP element (-20 bp) recognised by  $\alpha$ CTDs (Ross *et al.*, 2001). A schematic diagram of the RNAP holoenzyme bound at an ideal promoter region of DNA can be seen in Figure 1.1. An important feature of the  $\alpha$ CTD domain of RNAP is that it is connected to its respective  $\alpha$ NTD, and hence the core RNAP enzyme, by a flexible linker. This allows for promoter region recognition by  $\alpha$ CTD to occur at a range of distances upstream of the primary promoter elements, though in many cases it binds to the aforementioned UP element and interacts directly with domain  $\sigma 4$  of  $\sigma^{70}$  (Ross *et al.*, 2003).



**Figure 1.1 Schematic diagram of the RNA polymerase holoenzyme bound at an ideal promoter.** RNA polymerase (grey) is split into its key domains. Red line represents a DNA strand, with red boxes highlighting the important recognition sequences detailed in Section 1.1. Large bent black arrow at position +1 is representative of the transcription start site.

The degree of similarity each section of a promoter sequence has to the consensus dictates how strong initial binding of RNAP to the promoter will be. The less a sequence diverges from the established consensus, the stronger the binding. However, if all of the promoter sequence areas were conserved perfectly, binding would be too strong and exit of the RNAP from the promoter region would be inhibited. In fact, it can be said that some promoters do not have elements such as the UP and extended -10 elements for this reason (Hook-Barnard and Hinton, 2007). This variation in the modular assembly of promoters also allows a wide range in the frequency of transcript initiation.

For *E. coli* the primary  $\sigma$  factor is  $\sigma^{70}$ , often referred to as the housekeeping  $\sigma$ -factor as it is responsible for the majority of transcription and is involved in the transcription of all the essential genes of *E. coli* (Ishihama, 2000). A study into the relative amounts of RNAP enzyme and  $\sigma^{70}$  per cell, revealed that several thousand RNAP molecules are present per cell, with an excess of  $\sigma^{70}$  (Grigorova *et al.*, 2006). This suggests that activities of alternative sigma factors, with different promoter region preferences, are limited by competition with  $\sigma^{70}$  which has a higher affinity for core RNAP than  $\sigma^{54}$ ,  $\sigma^{38}$ ,  $\sigma^{32}$ ,  $\sigma^{28}$ ,  $\sigma^{24}$  and  $\sigma^{\text{FecI}}$  (Maeda *et al.*, 2000). The same study showed that the total number of RNA polymerase molecules in a cell exceeds the number that are actively engaged in transcript elongation, with the unemployed RNAP is likely to be sequestered at random chromosomal targets or by RNA, thus acting as a reserve supplying RNAP for targeted transcription activation when necessary (Grigorova *et al.*, 2006).

After the initial binding of the RNAP holoenzyme to the promoter region, the DNA between bases -10 and +2 (in relation to the first base of the transcript) unwinds to form a bubble, wherein the template and non-template strand are separated and the enzymatic activity of RNAP can commence into the elongation phase of transcription (Tsujikawa *et al.*, 2002).

The transcription of many bacterial genes however is under more complex control than simply promoter sequence variation and presence of promoter elements. This extra level of regulation can come in the form of transcription regulator proteins which fall into two broad categories; activators and repressors. Some of these regulators (a.k.a. global regulators) control the expression of a large number of genes

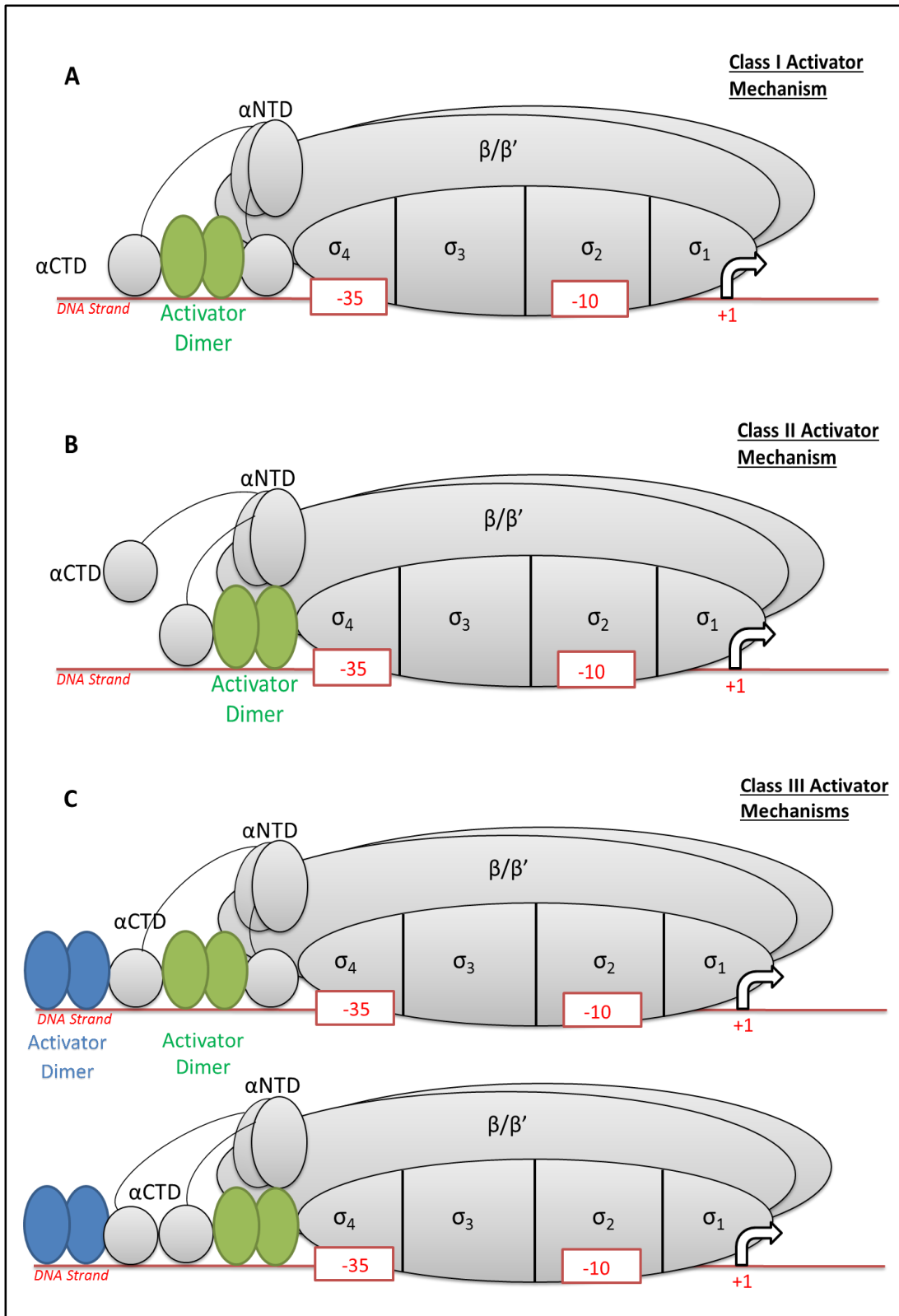
(e.g. CRP) (Kolb *et al.*, 1993), where others may specifically regulate only one or two (e.g. the *lac* repressor) (Muller-Hill, 1996). In fact it has been estimated that just seven transcription regulators control the expression of 50% of *E. coli*: CRP, FNR, IHF, Fis, ArcA, NarL and Lrp (Martinez-Antonio and Collado-Vides, 2003). Further detail on the mechanisms by which these regulators may activate or repress gene transcription will be discussed herein.

## **1.2. Transcription activators**

The more simple methods by which many transcription regulators aid in the recruitment of RNAP to promoter regions and activate transcription can be grouped into three classes; Class I, Class II and Class III (Figure 1.2). The terms Class I and II are more widely used and were introduced based on the differing mechanisms by which the cAMP receptor protein (CRP) has been observed to activate transcription; Class I refers to cases where an interaction between upstream bound CRP and the  $\alpha$ CTD of RNAP are sufficient (Ebright, 1993) and Class II refers to CRP bound to the promoter such that it overlaps the -35 region and has a direct interaction with RNAP independent of  $\alpha$ CTD (Igarashi *et al.*, 1991). Class III refers to those promoters wherein two transcription factors make independent contacts with RNAP (Scott *et al.*, 1995). CRP was originally discovered to be the activator of the *E. coli lac* operon and was subsequently found to be a global regulator affecting hundreds of genes (Zheng *et al.*, 2004). CRP, along with the homologous regulator FNR (Shaw *et al.*, 1983), have since become useful paradigms for understanding transcription activation and will be referred to often in the following overviews.

### **1.2.1. Class I mechanisms of transcription activation**

CRP activates transcription from the Class I *lac* promoter (Figure 1.2A). CRP binds 61.5 base pairs (bp) upstream of the *lac* transcriptional start site (DNA target is centered between bases 61 and 62) and it was observed that transcription of *lac* was abolished on removal of the  $\alpha$ CTD of RNAP (Igarashi *et al.*, 1991), suggesting that transcription activation at this specific promoter was dependent on interaction between CRP and  $\alpha$ CTD. This was rationalised by the subsequent finding that the  $\alpha$ CTD was connected to  $\alpha$ NTD via a flexible linker (Blatter *et al.*, 1994), which would allow for  $\alpha$ CTD to 'reach out' and interact with CRP bound upstream of the -35 element (at -61.5). This arrangement has been confirmed by a combination of



**Figure 1.2 Schematic diagrams of Class I, II and III types of interaction between transcription activators and RNAP.** Activator dimers are presented as green or blue circles. For Class III interactions (C), organisations wherein two Class I RNAP contacts or a combination of one Class I contact and one Class II contact are presented (top and bottom respectively).

three dimensional EM and high resolution X-ray structures (Hudson *et al.*, 2009), wherein one of the  $\alpha$ CTD subunits is observed to be sandwiched between the bound CRP dimer and the  $\sigma 4$  region of the RNAP holoenzyme. The position of the second  $\alpha$ CTD subunit is less well defined in this model system, though it is likely to bind further upstream and have no functional requirement within this Class I system (Lloyd *et al.*, 2002).

Although the recruitment of RNAP holoenzyme by CRP binding within the *lac* promoter is likely compensating for the absence of desired promoter elements such as the UP element, the position of the activator is flexible thanks to the aforementioned unstructured linker in the  $\alpha$  subunit of RNAP. This has been demonstrated by altering the position of the CRP DNA-binding site, such that it is 11 bp further upstream of the *lac* promoter (-72.5), and still observing an activatory interaction between CRP and RNAP. However, if only half a helical turn is inserted (i.e. 5 bp) this interaction is abolished (Straney *et al.*, 1989). Further to this, studies utilising a synthetic promoter that requires CRP binding for activation of a *lacZ* structural gene in  $\beta$ -galactosidase assays, have shown that CRP can induce transcription when positioned as far as 92.5 bp from the transcription start site, though only at 10-11 bp intervals wherein CRP is bound to the same helical face as the RNAP (Ushida and Aiba, 1990).

### **1.2.2. Class II mechanisms of transcription activation**

Upon removal of  $\alpha$ CTD from RNAP activation of some promoters by CRP was still observed (Igarashi *et al.*, 1991). These instances were examples of Class II activation of transcription (Figure 1.2B) wherein the CRP binding site overlaps the -35 promoter region and CRP interacts directly with the  $\alpha$ NTD subunit in the RNAP holoenzyme (Busby and Ebright, 1997). For Class II promoters the  $\alpha$ CTD domains bind DNA upstream of the CRP and form additional interactions, and while these interactions further aid transcription initiation, they are not required. It has also been observed that on substitution of the specific CRP residue, K52, a region of CRP is unmasked that allows for alternative direct binding to RNAP via the  $\sigma 4$  domain (Rhodius and Busby, 2000).

This interaction between an activatory protein overlapping the -35 region and the  $\sigma 4$  domain of RNAP is attributed to a multitude of activators. One of the most

heavily studied examples of this mechanism of activation is the activation of the  $P_{RM}$  promoter by the bacteriophage  $\lambda$  cI protein (Nickels *et al.*, 2002), with structural studies confirming an activator-RNAP interaction, mediated by clusters of acidic side chains on both cI and  $\sigma_4$ , takes place when both factors are DNA bound (Jain *et al.*, 2004). Also of note, these studies show little to no conformational changes take place in either constituent, suggesting that the cI protein is acting only as an aid in promoter binding which is sufficient to subsequently activate transcription.

Whilst binding of Class I activators can occur at a range of distances from the transcription start site, due to the nature of the direct binding between activator and the  $\sigma_4$  region of RNAP (or  $\alpha$ NTD in the case of CRP), Class II regulators must bind in a region that overlaps the -35 promoter site (Dove *et al.*, 2003). Examples of those regulators that exhibit similar activity include multiple members of the AraC family of transcriptional regulators. Specifically, MelR and RhaS have both been shown via genetic analyses to interact directly with domain 4 of  $\sigma^{70}$  (Grainger *et al.*, 2004; Bhende and Egan, 2000).

### **1.2.3. Class III mechanisms of transcription activation**

The study of CRP mechanisms of transcription activation has highlighted examples wherein transcription is dependent or enhanced on CRP binding at multiple sites within the same promoter, with each bound CRP making independent contacts with the RNAP holoenzyme (Figure 1.2C). Within this class there are two sub-classes. Firstly, a system wherein one CRP dimer overlaps the -35 region and acts via a Class II mechanism operate in tandem with a CRP dimer bound further upstream that contacts the spare  $\alpha$ CTD via a Class I mechanism (Belyaeva *et al.*, 1998). Alternatively, both CRP dimers can bind upstream of the -35 element and contact individual  $\alpha$ CTD subunits of RNAP, hence both dimers operate by a Class I mechanism (Beatty *et al.*, 2003). Class III mechanisms of transcriptional activation are widespread (Keseler *et al.*, 2013) and can involve two different regulatory proteins acting synergistically, exemplified by the activation of *ansB* transcription when CRP and FNR are bound in tandem (Scott *et al.*, 1995).

#### **1.2.4. Altering promoter DNA conformation**

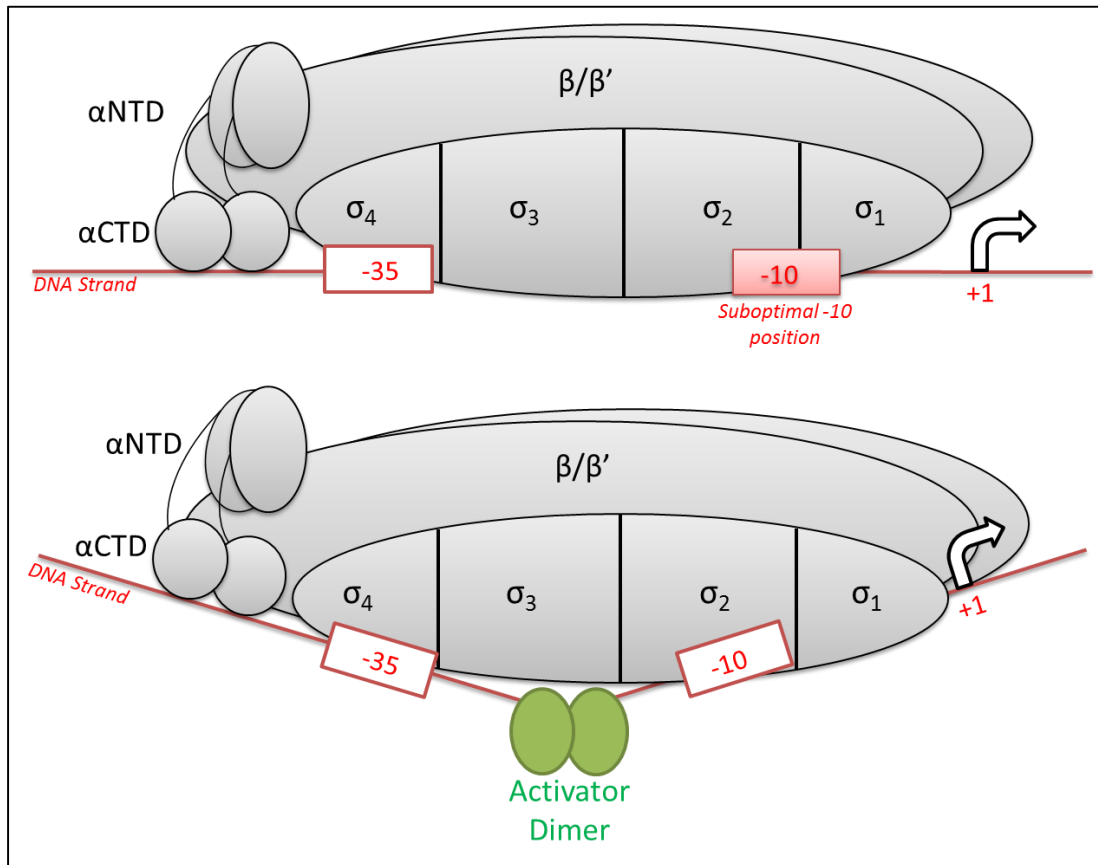
A method by which some bacterial gene regulators activate transcription is to improve the spacing between the -35 and -10 promoter elements by inducing a conformational change in the DNA within those promoters where the spacing is sub-optimal (Figure 1.3). The best understood examples of this regulatory mechanism belong to the MerR family of transcription factors which mostly bind between the -35 and -10 elements (Brown *et al.*, 2003). This localised distortion in the DNA has been shown through high-resolution structural studies of the *Bacillus subtilis* MerR family protein MtaN (Newberry and Brennan, 2004). Examples of MerR family proteins that exist in *E. coli* include SoxR and the two metal-dependent regulators ZntR and CueR (Brown *et al.*, 2003). It was originally thought that this mechanism of promoter activation was restricted to the MerR family of proteins, although a similar mechanism is seemingly used by the unrelated GrlA protein in activation of genes in the locus of erythrocyte effacement of enterohaemorrhagic *E. coli* (Islam *et al.*, 2011).

#### **1.2.5. RNA polymerase appropriation**

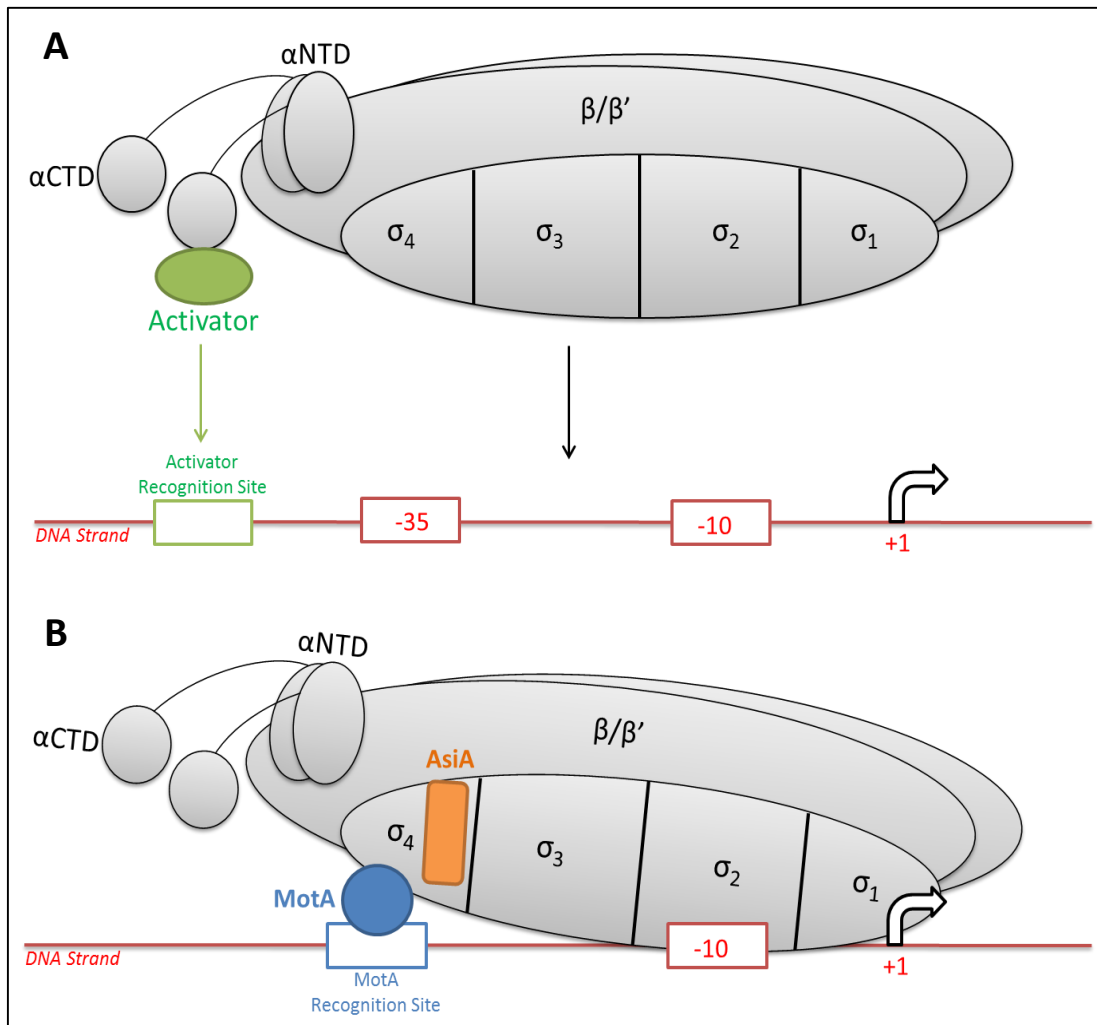
Alternative to the mechanisms of activation discussed thus far, where transcription factors bind to DNA and improve the attractiveness of promoters to RNAP; there are examples that involve altering the promoter preference by direct binding to RNAP *away* from the promoter region. An example of this, as previously mentioned, is the binding of alternative  $\sigma$ -factors to the RNAP core enzyme (Gruber and Gross, 2003). Hence, the gene *rpoS* encodes the alternative sigma factor  $\sigma^S$  and is upregulated in response to multiple stresses including nutrient starvation, thus larger proportions of RNAP incorporate this sigma factor and transcription is focused to promoters of genes vital for stress resistance (Battesti *et al.*, 2011). However, it must be stated that many genes can be activated by RNAPs carrying various sigma factors, as demonstrated in ChIP-chip studies wherein a significant overlap between the  $\sigma^{34}$  and  $\sigma^{70}$  regulons was recorded (Wade *et al.*, 2006).

Other examples of this RNAP appropriation are provided by SoxS and MarA of *E. coli* (Figure 1.4A). These regulators are both of the AraC family of transcription regulators, though they contain only the DNA-binding domain of this protein family and are hence controlled predominantly by their relative concentration





**Figure 1.3 Simple activation by DNA conformational change.** Activator (green circles) is shown as a dimer binding between the -35 and -10 promoter elements, causing a conformational change and bringing the -10 element into register with the  $\sigma_2$  sub-domain.



**Figure 1.4 Representations of activation by RNA polymerase appropriation.** (A) An activator protein (green) can bind to an  $\alpha$ CTD domain of RNAP away from DNA and direct holoenzyme binding to promoters that contain DNA sequences homologous to the activator's DNA-binding recognition site. (B) In the specific case of the T4 phage protein AsiA binding to RNAP, the  $\sigma_4$  subdomain is remodelled such that it no longer recognises -35 promoter elements, and is instead directed towards DNA bound MotA protein.

within the cell rather than activation by binding of a specific signal. MarA and SoxS induce expression of large stress regulons associated with antibiotic resistance and superoxide stress respectively, and bind to a common DNA target known as the Mar box (Martin *et al.*, 2008). They were originally thought to activate transcription via Class I and Class II mechanisms, though it has now been suggested that they bind to RNAP holoenzyme whilst it is not associated with a promoter. Specifically, it has been shown that these regulators bind directly to the DNA-binding  $\alpha$ CTD domain of RNAP, and redirect the binding preference away from potential UP-elements towards their desired Mar box targets (Dangi *et al.*, 2004; Shah and Wolf, 2004).

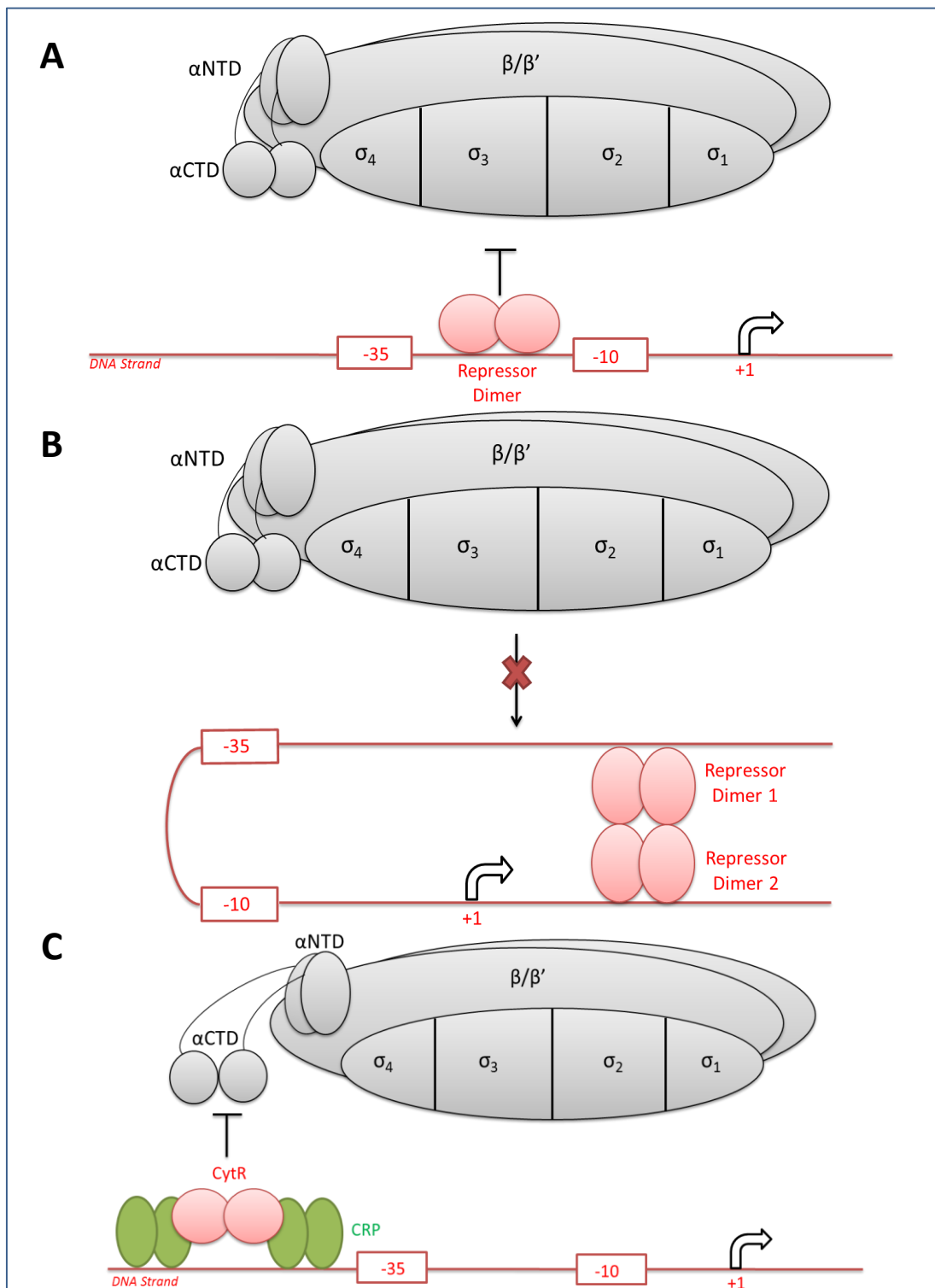
Another model of how RNAP appropriation may occur is given by the T4 bacteriophage protein AsiA and its co-activator protein MotA (Figure 1.4B). It has been shown that AsiA has the ability to bind directly to  $\sigma^{70}$  of the RNAP holoenzyme and remodel the  $\sigma 4$  domain. This results in RNAP no longer recognising the -35 promoter element (recognition of the -10 element is maintained) and instead being targeted to DNA regions bound by the MotA activator protein leading to expression of the T4 genes associated with the middle-phase of infection (Hinton *et al.*, 2005).

### **1.3. Transcription repressors**

Whilst there are many transcription factors that aid in the recruitment of RNAP to gene promoters, and subsequently promote the transcription of particular genes, there are also many examples of proteins that have a negative impact on a subset of gene targets when expressed. Much like simple activation of promoters, simple repression often works by the interaction of a single repressor molecule with the promoter region thereby preventing transcription of the gene in question. Three general mechanisms by which this may occur will be discussed here (Figure 1.5).

#### **1.3.1. Steric hindrance of RNA polymerase**

An example of steric hindrance (Figure 1.5A) is given by one of the most thoroughly studied bacterial repressors, the Lac-repressor LacI, which represses the genes of the *lac* operon associated with the transport and catabolism of lactose. When grown in the absence of rapidly metabolisable glucose, *E. coli* CRP can activate transcription of the *lacZYA* operon but only in the presence of the alternative



**Figure 1.5 Mechanisms of transcription repression.** Displayed are schematic representations of transcription repression by (A) steric hindrance, (B) DNA looping and (C) the CytR mediated paradigm of anti-activation. Repressors are shown as red circles in all cases, with activators shown in green.

carbon source; lactose (Lewis, 2005). In the absence of lactose, and its isomeric form allolactose, the *lacZYA* operon is repressed by LacI. This repression is imposed by LacI binding primarily to a 21 bp operator sequence, starting at what would be the first transcribed nucleotide of the *lac* operon. This DNA binding then inhibits the action of RNAP by physically blocking the formation of a closed complex (i.e. inhibiting the binding of RNAP to the promoter region) or inhibiting the isomerisation from a closed complex to a transcription-capable open complex (Sanchez *et al.*, 2011). When allolactose is bound by LacI, a conformational change occurs such that LacI can no longer bind DNA, and expression of the operon is permitted (Lewis, 2005).

### **1.3.2. DNA looping**

There is a family of transcriptional regulators with a high degree of sequence homology to the LacI repressor, and these are collectively referred to as the LacI/GalR family of transcriptional regulators (Weickert and Adhya, 1992). GalR is particularly similar to LacI in not only its structure but also its function, in that it regulates the *E. coli* genes required for utilisation of the carbon source D-galactose in the absence of glucose (Semsey *et al.*, 2007). GalR specifically represses the *galETKM* operon when it is not bound to D-galactose. It does this by binding two 16-bp operator sequences;  $O_E$  centred at -61.5 upstream of the *P\_I* promoter of *galETKM* and  $O_I$  113 bp downstream of  $O_E$ . On binding of one GalR dimer to each of these sequences, a DNA-loop encompassing 113 bp can form due to interaction between the two separated dimers (Figure 1.5B). This loop formation is dependent on binding at the apex by the nucleoid associated protein HU (Semsey *et al.*, 2002; Semsey *et al.*, 2004). Hence, the formation of this DNA loop denies RNAP access to the promoter region. Much like LacI, on reception of the inducer D-galactose GalR undergoes a conformational change which reduces its DNA-binding affinity, followed by de-repression of the genes necessary for the utilisation of this alternative carbon-source. It should be noted that the *lac* operon, repressed by LacI, has been shown to have multiple operator sequences specific to LacI binding, and DNA-looping has been suggested to enhance repression when these sites are occupied (Lewis, 2005).

### **1.3.3. Anti-activation**

The third mechanism by which a protein may act as a repressor is through anti-activation, the best example of which is CytR repression at promoters dependent on CRP activation (Figure 1.5C). As has been previously described, many promoters are activated by binding of one or more CRP dimers. In some cases these activatory complexes of CRP:DNA are specifically recognised by the regulatory protein CytR and activatory contacts between CRP and RNAP are inhibited (Meibom *et al.*, 2000). Many promoters that experience anti-activation by CytR have two binding sites for CRP centred at positions 93.5 and 41.5 upstream of the transcription start site, with the CytR binding region being sandwiched between the two, as shown in the case of the P2 promoter for the *deoCADB* operon (Shin *et al.*, 2001). On binding of cytidine by CytR, the CytR/CRP cooperativity is perturbed and anti-activation is inhibited, hence activation of the target operon by CRP is permitted (Pedersen *et al.*, 1991).

### **1.4. Nucleoid associated proteins and DNA supercoiling**

Bacterial chromosomes also form interactions with RNA and proteins, and undergo supercoiling. The family of nucleoid associated proteins (NAPs) are known to create tightly packed DNA secondary structures and coat up to half of the chromosome (Ishihama, 2000); hence this provides another opportunity for regulation of gene transcription. *Escherichia coli* has a dozen proteins which aid chromosomal compaction, including: Fis, IHF, H-NS and HU, StpA and Dps. Although the activity of such proteins is not regulated by one specific event, they are recruited to several promoter loci and can affect both transcription activation and repression on a global scale (Dillon and Dorman, 2010).

One of the best understood members of the NAP family is H-NS, mostly associated with forming extended nucleoprotein complexes which can completely silence gene expression, as is the case for the cryptic adenine deaminase gene of *E. coli*, *adeD* (Petersen *et al.*, 2002). In fact, it has been suggested that H-NS alone may play a role in regulation of up to 5% of the *E. coli* genome (Hommais *et al.*, 2001). There are also multiple examples of different NAPs repressing gene transcription through a coordinated sequestration of target DNA; as is the case for the repression of FNR-mediated activation of *nir*, wherein transcription is silenced by the binding of Fis, IHF and H-NS to the DNA to form a highly ordered nucleoprotein complex

(Browning *et al.*, 2000). Examples of NAPs having a positive effect on gene expression are also numerous (McLeod and Johnson, 2001), for example the positive influence Fis has on the expression of *proP* by direct interaction with RNAP holoenzyme when bound to DNA overlapping the -35 promoter element (Xu and Johnson, 1995).

### **1.5. Complex activator/repressor systems and integration of regulatory signals**

Many bacterial promoters are controlled by more than one specific activator and/or repressor, and therefore can respond to multiple environmental cues. Some promoters, however, ‘integrate’ multiple signals at the same time in order to achieve transcription initiation. These complex regulatory systems often involve a global regulator activating a subset of genes, and then more specific regulators controlling the individual genes within that group. A good illustration of this has already been given in the *lac* operon; which is firstly activated by CRP (a global regulator responding to glucose starvation) and then specifically regulated by the *lac* repressor (responding to the specific presence of allolactose).

Some complex regulatory systems exist that incorporate just repressors, though most involve the concerted activity of both repressors and activators. In the majority of such systems this involves the repressor and activator proteins operating independently, though systems such as CytR repression of *deoP2* involve direct interaction between the CRP transcriptional activator and CytR, causing active repression of the gene.

Four general mechanisms for the concerted activity of two or more activators can occur. These involve independent activator-RNAP interactions, activator repositioning, cooperative activator interactions and anti-repression by an activator.

Independent activator-RNAP interactions refers to the previously described Class III mechanism of activation, specifically the case in which CRP and FNR both act independently in activating transcription of *ansB* (Scott *et al.*, 1995). A similar situation has been observed in the regulation of the plasmid encoded enteroaggregative *E. coli* toxin gene; *pet*. It was shown that sub-optimal positioning of a CRP site overlapping the -35 region of the *pet* promoter was compensated for by

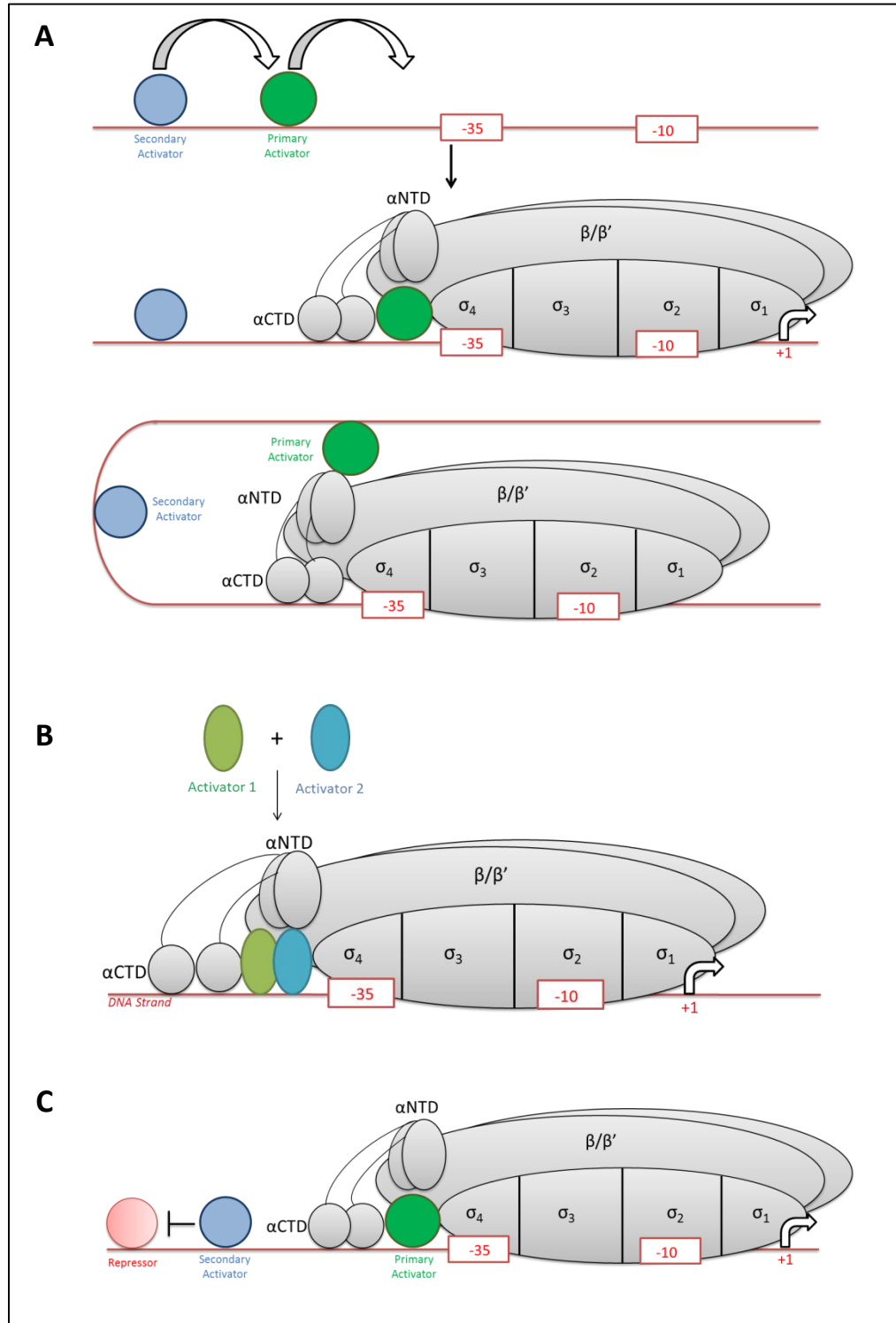
additional direct interaction between Fis and RNAP, likely via the  $\alpha$ CTD (Rossiter *et al.*, 2011).

An example of activator repositioning (Figure 1.6A) is the action of MalT at the *malk* promoter. Within this promoter region there are multiple overlapping MalT-binding sites, and in the absence of CRP, MalT preferentially binds the higher affinity sites which are positioned too far from the -35 and -10 promoter elements to activate transcription. However, in the presence of CRP binding to its own sites within the promoter, MalT is repositioned to the lower affinity sites closer to the RNAP binding region and transcription can then proceed (Richet *et al.*, 1991). Alternatively, at the *narG* promoter, which is activated by the presence of FNR overlapping the -35 promoter element, enhanced activation by the NarL transcription activator binding 125 bp upstream of the transcription start site requires the presence of IHF binding at an intermediate site. This causes a bend in the DNA such that NarL can also directly interact with the RNAP holoenzyme.

Co-operative activator binding (Figure 1.6B) refers to those cases wherein one or both of a pair of activators cannot bind to the promoter DNA without the presence of the other. For example the activation of the *melAB* operon by the binding of MelR and CRP (Wade *et al.*, 2001). Here, binding of MelR to distal sites upstream of the promoter (-120.5 and -100.5) allows CRP to then bind the promoter (-81.5), which subsequently leads to increased occupation of additional downstream MelR sites (-62.5 and -42.5) required for transcriptional activation. More recently activation of the *gadA* promoter and expression of associated acid stress resistance genes has been shown to require cooperativity between GadE and RcsB (Castanie-Cornet *et al.*, 2010). Basal activity of RcsB on its own actually leads to repression of the *gadA* operon, but in the presence of increased GadE in response to multiple environmental signals a GadE/RcsB heterodimer is formed which binds to the upstream GAD box (-62.5) and activates transcription.

Finally, anti-repression (Figure 1.6C) is when a single activator is sufficient for transcription initiation but its activity is inhibited due to a repressor, hence a secondary activator is required in order to relieve this repression. An example of this is the regulation of the *nir* promoter. In this case activation of the *nir* operon by FNR is silenced by a highly ordered nucleoprotein complex (Browning *et al.*, 2000).





**Figure 1.6 Examples of complex activator/repressor systems of transcription regulation.** (A) Two different mechanisms of activator repositioning. (B) Cooperative activator binding. (C) Anti-repression. Repressors are shown in red. In the case of cooperative binding, binding of either Activator 1 (green) or Activator 2 (blue) requires binding of the other. For activator repositioning and anti-repression, transcription activation by direct contact between the primary activator (green) and RNAP requires the additional action of a secondary activator (blue).

However, the binding of either the nitrate-activated transcription factor NarL within the promoter region can interfere with IHF binding and therefore disrupt the repressive effect of the nucleoprotein complex on FNR activation (Browning *et al.*, 2004). Interestingly, it has been shown that activation of gene transcription by anti-repression could be mediated by the most heavily-studied repressor of all, LacI, if its binding site is present in an appropriate context. This was demonstrated by introduction of LacI binding sites in the *bgl* promoter, such that H-NS silencing was disrupted (Caramel and Schnetz, 1998).

As has already been stated, in addition to the variation in how different regulators activate or repress transcription via their interaction with their relevant promoter regions, the methods by which their own activity can be controlled must also be considered. Several mechanisms exist that regulate the activity of these transcription regulators. Some are regulated in response to binding of a specific ligand wherein this may act to promote binding to the promoter DNA element or cause dissociation, e.g. CRP binding of cAMP (Kolb *et al.*, 1993). Others may be regulated by covalent modification in the form of phosphorylation by a membrane-bound kinase, which links gene regulation directly to signals being received outside the cell (e.g. NarL, being controlled by NarX and NarQ) (Stock *et al.*, 2000). Alternatively, transcription regulator activity can be determined by their intracellular concentration which is controlled by protein synthesis or turnover (e.g. MarA and SoxS) (Martin *et al.*, 2008).

#### **1.6. The MarR family of transcription regulators**

The gene transcription regulator MarR (**M**ultiple **a**ntibiotic **r**esistance **R**egulator) was first characterised in *E. coli* wherein it negatively regulates the *marRAB* gene locus responsible for resistance to not only antibiotics, but also organic solvents, oxidative stress agents and even household disinfectants (Martin and Rosner, 1995; Alekshun and Levy, 1999b). Since then many structurally similar transcription regulators have been characterised, and the MarR family of transcription regulators is considered to be widespread throughout the Bacterial and Archaeal kingdoms. In fact, a genome database search highlighting any proteins that are likely to contain structural similarity to known MarR-family proteins (Pfam: PF01047) produces a list of over 19,000 possible MarR-like proteins spread over

3,773 species; 3677 bacterial species and 96 archaeal species (Finn *et al.*, 2014). The MarR family is defined by the presence of a characteristic DNA-binding domain known as the **winged-helix-turn-helix** (wHTH), which averages approximately 135 amino acids in length. These MarR family gene regulators exist as homodimers, which allows them to bind to sequences that are largely palindromic, resulting in either repression or activation of their target genes (Perera and Grove, 2010). A trend observed for many members of this family of transcription regulators is that they respond and bind to specific ligands. In the absence of ligand they are able to bind to their DNA targets, often resulting in repression, however when ligand is bound the ability to bind DNA is lost and so too is the repressive effect on the target gene (Wilkinson and Grove, 2006). While this ligand binding characteristic may be true for most MarR family regulators, in many cases the natural ligand is unknown.

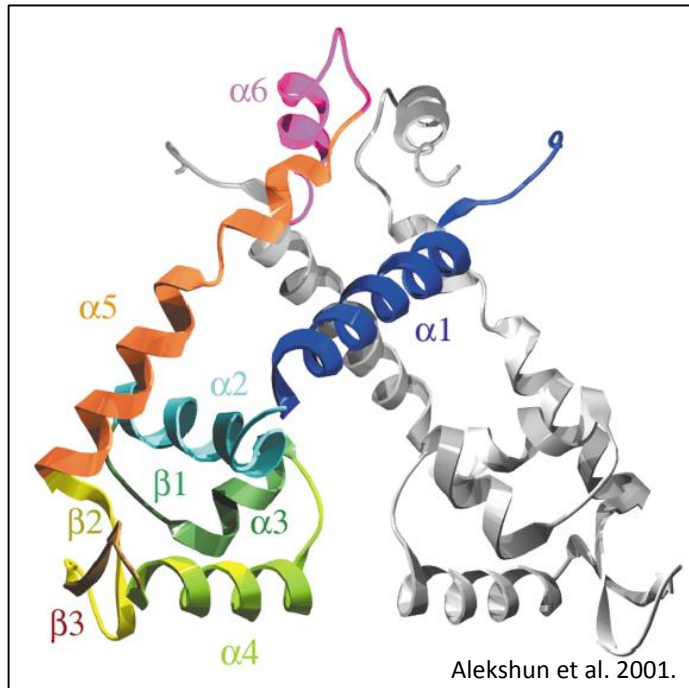
The method by which a MarR family regulator may repress gene expression is illustrated by the mechanism of *E. coli* MarR itself acting at the *marRAB* locus. By binding to two palindromic sequences present between *marRAB* and the divergently transcribed *marC*, MarR obstructs the -35 and -10 promoter elements of the *marRAB* operon, hence denying access to RNAP (Martin and Rosner, 1995). On the other hand it has been shown that a MarR family transcription regulator can act as an activator by stabilising the RNAP-promoter DNA interaction. In the case of *Streptomyces coelicolor* OhrR, which controls the transcription of genes necessary for organic hydroperoxide resistance, in its reduced form it binds cooperatively to multiple sites within the *ohrA* promoter region and abuts the -35 and -10 promoter elements, much like MarR in *E. coli*. However, the oxidised form of OhrR has reduced DNA-binding affinity, but can still bind a lone site within the target promoter. Binding at this singular site actually enhances transcription initiation, likely through stabilising the RNAP-DNA interaction (Oh *et al.*, 2007). Alternatively, MarR family proteins have been shown to act as activators by competing with repressor proteins. For example, it has been demonstrated that RovA in *Yersinia enterocolitica* has multiple target gene promoters, a high proportion of which are subject to repression by H-NS at sites that overlapped the sequence that RovA also recognised. These data therefore suggest that the principle function of RovA in *Y. enterocolitica* is to act as an antagonist of H-NS-mediated repression (Cathelyn *et al.*, 2007). As has already been discussed in the general overview of bacterial gene

transcription regulation, the ability of a regulator to aid in either activation or repression of transcription initiation is almost entirely dependent on the position of regulator binding relative to the target gene promoter elements.

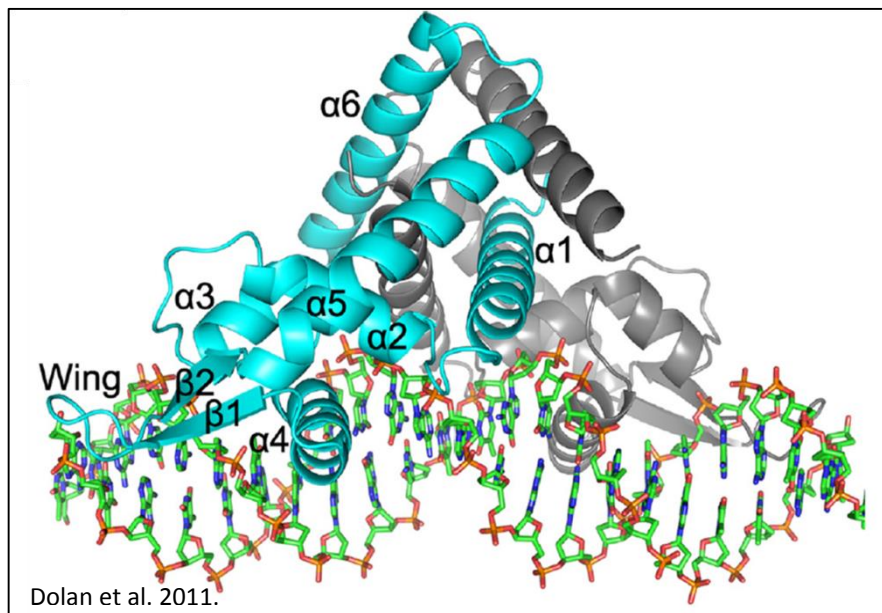
The common structural features of members within the MarR family of transcription regulators have been summarised (Perera and Grove, 2010). Most MarR homologues form homodimers *in vivo* which assume a roughly triangular shape with a pseudo-2-fold symmetry. Alignment of multiple MarR homologue sequences indicates a common secondary structure organisation of six  $\alpha$ -helices and three  $\beta$ -strands in the following order:  $\alpha 1$ - $\alpha 2$ - $\beta 1$ - $\alpha 3$ - $\alpha 4$ - $\beta 2$ - $\beta 3$ - $\alpha 5$ - $\alpha 6$ , with a homodimer forming the tertiary structure shown in Figure 1.7 (Alekhshun *et al.*, 2001). The  $\alpha$ -helices situated at both the N- and C-terminal regions of each monomer form the dimerisation domain, with dimerisation occurring by interdigitation of the  $\alpha$ -helices between the two monomers consisting of hydrophobic interactions and intermolecular hydrogen bonds (Alekhshun *et al.*, 2001). Crystal structures also reveal the presence of the characteristic MarR family wHTH motif, a variant of the more common helix-turn-helix DNA-binding domain (Gajiwala *et al.*, 2000).

In the case of the MarR family regulator SlyA of *Salmonella enterica* Serovar Typhimurium the structure has been determined when bound to DNA, and is typical of MarR family regulators (Dolan *et al.*, 2011). This structure is shown in Figure 1.8. The  $\alpha$ -helices  $\alpha 3$  and  $\alpha 4$  make up the helix-turn-helix portion of the DNA binding domain, with  $\alpha 4$  comprising the primary DNA sequence recognition helix. In the case of this specific structure the wing portion consists of two anti-parallel  $\beta$ -strands ( $\beta 1$  and  $\beta 2$ , equivalent to  $\beta 2$  and  $\beta 3$  in the common secondary structure order detailed above) and a connecting loop; in some MarR homologues the presence of a third  $\beta$ -strand can provide additional stabilisation (Perera and Grove, 2010). In terms of DNA-binding, the recognition helix of the wHTH domain binds the major groove of DNA, whilst the wing portion makes contact with the adjacent minor groove.

Several mutational studies have shown how the wing portion of the wHTH domain is vital for DNA-binding. For example, in MarR of *E. coli* mutation of the Arg94 residue situated at the tip of the wing to Cys actually abolishes the repressor



**Figure 1.7 Structure of an *E. coli* MarR dimer.** Schematic structure shown is as determined and presented in Alekshun *et al.* 2001. One monomer is shown in colour, with different colours used to differentiate the secondary structure elements.



**Figure 1.8 Structure of a *Salmonella enterica* Serovar Typhimurium SlyA dimer bound to DNA.** Schematic SlyA structure bound to DNA is as determined and presented in Dolan *et al.* 2011. One subunit is shown in cyan, and the other in grey. Secondary structure subunits are labelled. Bound DNA is shown in stick form.

function of the regulator (Alekhshun *et al.*, 2000), whereas mutation of Gly95 to Ser increases DNA affinity of MarR up to 30-fold (Alekhshun and Levy, 1999a).

### **1.6.1. MarR family proteins as regulators of virulence**

The MarR family members RovA, PecS and SlyA are all examples of regulators that are involved in the regulation of virulence genes, all having been shown to affect expression in both positive and negative ways.

It has been shown that expression of the gene *inv* is positively regulated by RovA in *Y. enterocolitica* and *Yersinia pseudotuberculosis*, with *inv* encoding an outer membrane protein required for aiding the translocation of the bacterium across intestinal M cells, and thus aiding in the invasion of the intestinal epithelium (Revell and Miller, 2000; Nagel *et al.*, 2001). Evidence has shown that RovA positively impacts the expression of *inv* by displacing the repressive nucleoprotein complex formed by H-NS and YmoA (Ellison and Miller, 2006). More recently it has been shown that RovA also positively regulates the operon *yaxAB*, which encodes two proteins necessary for cytotoxic attack and lysis of mammalian cells, via a similar anti-H-NS mechanism (Wagner *et al.*, 2013).

PecS is a gene regulator of the plant pathogen *Dikeya dedantii* (formerly known as *Erwinia chrysanthemi*) that is essential for expression of virulence genes in this enterobacterium. It was found to negatively regulate the pectin lyase genes *pelD* and *pelE* by competing with the positive regulator CRP for DNA-binding at overlapping sites (Rouanet *et al.*, 1999). PecS has also been shown to repress the expression of the genes necessary for the biosynthesis of indigoidine, a blue pigment important in resistance to products of the oxidative burst including hydrogen peroxide (Reverchon *et al.*, 2002) and genes necessary for flagellar biosynthesis (Rouanet *et al.*, 2004). Conversely, PecS can also act as an activator of virulence genes as demonstrated by its ability to act as a derepressor of the *peh* genes encoding polygalacturonases which, like the pectin lyases, aid in the degradation of pectin in plant cell walls (Nasser *et al.*, 1999).

One of the most studied examples of a MarR family regulator playing a role in regulation of virulence genes is SlyA of *S. enterica* Serovar Typhimurium.

### 1.7. SlyA of *Salmonella enterica* Serovar Typhimurium

SlyA in *S. enterica* Serovar Typhimurium is one of the most extensively studied examples of a MarR family virulence regulator because it was found that, upon deletion, virulence was significantly attenuated within the mouse model of infection. This was characterised by a >1000-fold increase in LD<sub>50</sub> for all routes of infection (Libby *et al.*, 1994). It was found that without *slyA*, *S. enterica* Serovar Typhimurium could no longer survive within macrophages and were in fact hypersusceptible to the reactive oxygen species such as H<sub>2</sub>O<sub>2</sub>, which would be found in such cells (Buchmeier *et al.*, 1997).

A consensus binding site has been suggested for the 17 kDa SlyA protein in *S. enterica* Serovar Typhimurium. It is a 12 bp sequence with a near perfect inverted repeat: TTAGCAAGCTAA. SlyA recognises this site upstream of the *slyA* promoter, along with four other repeats of related sequences. Occupation of these sites covered the -10 to -35 region of the *slyA* promoter, preventing open complex formation, thereby creating a feedback inhibition system (Stapleton *et al.*, 2002). In this same study supernatant fractions and outer membrane proteins were compared between wild type and a *slyA* mutant of *S. enterica* Serovar Typhimurium. It was found that FliC, Iron, PagC and other outer membrane proteins (OmpC, OmpF and OmpA) were all differentially expressed in the absence of SlyA, and the proteins FliC, Iron, PagC and OmpC had all previously been shown to be important in virulence and survival within macrophages. Amounts of FliC were observed to be significantly reduced in the absence of SlyA, whereas PagC, for example, was increased in its absence. This is evidence that SlyA plays a role as both a repressor and activator of transcription (Stapleton *et al.*, 2002).

Various proteomic and transcriptomic studies have been carried out on wild type and *slyA* strains of *S. enterica* Serovar Typhimurium in order to determine which genes may be regulated by the SlyA protein. Proteomic studies have shown differences in protein expression between the strains, with evidence suggesting that SlyA can both positively and negatively regulate protein expression. One such study showed 12 proteins to be up-regulated while 11 were down-regulated with many of them not appearing to be virulence factors (Spory *et al.*, 2002). In related transcriptomic studies similar results were found wherein SlyA was found to up-

regulate 23 target genes, while down-regulating 8 others (Navarre *et al.*, 2005). The differences in number of genes differentially expressed between the transcriptomic study and proteomic study are likely due to varying experimental conditions, but it is clear that SlyA has the ability to both activate and repress genes and affect the subsequent protein expression profile. The transcriptomic study was specifically concerned with the investigation of the cross over between the PhoP/Q two-component regulatory system and the SlyA regulon. It was found that there was quite a significant overlap between the two regulons, including *pagC*, *pagD*, *ugtL*, *mig-14*, *virK*, *phoN*, *pgtE*, *pipB2*, *sopD2*, *pagJ* and *pagK*. Many of these genes are associated with the bacterial envelope and some are even directly involved with virulence and resistance to anti-microbial peptides. Of these genes *pagC*, which encodes an outer membrane protein, was one example that was also found in the proteomic studies, and has been confirmed as being directly regulated by SlyA (Navarre *et al.*, 2005). Interestingly, it was determined that whilst SlyA directly binds to the promoter region of *pagC*, no such PhoP site exists. It was therefore logical to assume that this co-dependence on both regulators is likely caused by PhoP up-regulation of SlyA, an assumption that is supported by the work of Norte *et al.* wherein it is suggested that PhoP up-regulates an unknown “Factor X” that subsequently activates *slyA* transcription (Norte *et al.*, 2003). Other DNA footprinting analyses seem to show direct binding of PhoP upstream of the SlyA promoter (Shi *et al.*, 2004). However, this is contradicted by the apparent observation that levels of SlyA protein and mRNA are not noticeably affected by PhoP inducing conditions of Mg(II) starvation (Navarre *et al.*, 2005), though this may be a product of different experimental conditions. Possible mechanisms of how the PhoP and SlyA systems may interact indirectly have been suggested, one of which puts forward the idea that PhoP may instead activate the production of an unknown soluble ligand which activates SlyA protein that is already present (Navarre *et al.*, 2005). Though it has been demonstrated that SlyA can bind DNA *in vitro*, presumably without any ligand (Stapleton *et al.*, 2002), which may suggest it could regulate alternate regulons depending on whether a ligand is bound or not and therefore have a regulatory shift depending on extracellular signals.

In the Navarre *et al.* (2005) study it is stated that PhoP binding sites existed in the promoter for the antimicrobial resistance gene *ugtL* in addition to SlyA



binding sites, a trait not present in the *pagC* promoter. The PhoP binding site was found to be situated upstream of the *ugtL* promoter, whereas the SlyA binding site was found to be downstream of the +1 transcription start site. Binding of a transcription activator downstream of a promoter is unusual, and it was suggested that SlyA may aid in transcriptional activation by antagonising the binding of H-NS, modifying the local nucleoprotein structure and thus aiding classical gene activation by PhoP (Shi *et al.*, 2004). More recently, binding of both SlyA and PhoP to the promoter region of *pagC* has been demonstrated, with SlyA being required for anti-silencing from H-NS, and PhoP being required for subsequent transcriptional activation (Perez *et al.*, 2008). This same study confirmed that a similar method of transcriptional control was in place for expression of *ugtL*.

A mutational analysis of SlyA determined amino acid residues that were important for DNA-binding (Okada *et al.*, 2007). It was found that Leu-63, Val-64, Arg-65, Leu-67, Leu-70, Arg-86 and Lys-88 in the winged-helix region were vital for DNA binding, and residues Leu-12 and Leu-126 within the  $\alpha$ -helices of the N-terminal and C-terminal regions were required for efficient dimer formation. At this point, the residues were mapped to structural features based on modelling to the elucidated structure of the *Bacillus subtilis* homologue, YusO. In this same study they confirmed that SlyA regulates transcription of the sensor kinase SsrA by binding the *ssrA* promoter. This sensor kinase activates expression of the *Salmonella* pathogenicity island 2 (SPI-2) which is vital for intracellular survival and replication (Okada *et al.*, 2007). In 2011 the structure of *S. enterica* Serovar Typhimurium SlyA was published, in its apo-form, and when bound to DNA. Structural details have already been discussed (Section 1.6, Figure 1.8), and it confirmed the importance of residues Arg-65 and Arg-86 in DNA-binding to specific sequences, which mapped to the DNA-recognition helix and the wing portion of the wHTH respectively (Dolan *et al.*, 2011).

A link between the stringent response molecule ppGpp and the transcription activation activity of *S. enterica* Serovar Typhimurium SlyA has been suggested. This alarmone signal has been observed to regulate a multitude of processes in bacteria, namely directing modification of RNAP promoter preference or the alteration of  $\sigma$ -factor preference away from  $\sigma^{70}$  (Dalebroux and Swanson, 2012). It has been demonstrated that ppGpp may bind to the SlyA protein and enhances its

affinity for DNA (Zhao *et al.*, 2008); an example of ppGpp imposing an effect through direct interaction with a protein other than RNAP which is not without precedence (Dalebroux and Swanson, 2012).

It is clear that SlyA is important in the regulation of genes in *S. enterica* Serovar Typhimurium and contributes to its virulence, whether it be through regulation of genes of SPI-2 via activation of *ssrA* expression or activation of horizontally acquired genes not within SPI-2 in the case of *ugtL* and *pagC* (Perez *et al.*, 2008). There is also a multitude of evidence for the SlyA regulon being linked to that of the PhoP/Q system, with coordinated binding having been demonstrated at the promoters of both *pagC* and *ugtL*. How the PhoP/Q system directly influences SlyA activity is less clear, with contradictory evidence suggesting that PhoP may only regulate DNA-binding activity of SlyA indirectly (Navarre *et al.*, 2005) or that PhoP possibly regulates the expression of *slyA* directly (Shi *et al.*, 2004).

### **1.8. SlyA of *Escherichia coli***

When *S. enterica* Serovar Typhimurium SlyA was first being studied it was overexpressed in *E. coli*, which resulted in a haemolytic phenotype. At first this led researchers to believe that SlyA was itself a haemolysin (Libby *et al.*, 1994). As discussed above it was eventually found to be a gene regulator, and the haemolytic phenotype was due to the activation of a gene called *hlyE*. This gene encoded a novel haemolysin, and it was subsequently found that *E. coli* encodes its own SlyA homologue with 89% sequence homology to the *S. enterica* Serovar Typhimurium SlyA (Oscarsson *et al.*, 1996). It was eventually demonstrated through  $\beta$ -galactosidase assays, gel-shifts and DNase I footprinting that regulation of *hlyE* by *E. coli* SlyA occurred by the antagonism of H-NS silencing of the *hlyE* promoter wherein H-NS was preventing transcription initiation by denying access to CRP or FNR, both of which can enhance *hlyE* expression (Wyborn *et al.*, 2004).

A more recent study attributed a novel interaction between *E. coli* SlyA and H-NS in the activation of transcription of genes in the K5 capsule gene operon from the PR1 promoter at 37°C (Corbett *et al.*, 2007). Electrophoretic mobility shift assays and DNase I footprinting demonstrated that at 20°C H-NS repressed open complex formation at PR1, this was alleviated at 37°C in a process involving SlyA. Maximal gene transcription, even at 37°C, would appear to require a complex

interplay between H-NS and SlyA, suggesting that the presence of both creates an ideal environment for open complex formation or binding of a further Class I or II activator. This study also showed that expression of *slyA* in *E. coli* is positively auto-regulated, with purified SlyA binding at three sites within *PslyA*; at positions between -404 and -361 and between -141 and -108 upstream of a proposed transcription start point (57 bp upstream of the initiating TTG codon), with an additional site situated between +20 and +54 downstream (Corbett *et al.*, 2007). It was also observed that *slyA* was expressed maximally earlier in the growth cycle (Corbett *et al.*, 2007) than was observed for *S. enterica* Serovar Typhimurium, which was negatively autoregulated and expressed maximally in stationary phase growth (Buchmeier *et al.*, 1997; Stapleton *et al.*, 2002). A similar relationship between SlyA and H-NS has been found for the PR3 promoter that exists at the other end of the K5 capsule gene operon (Xue *et al.*, 2009). Interestingly, in the case of both PR1 and PR3, SlyA can activate transcription to a limited extent in the absence of H-NS suggesting its role is not limited to H-NS anti-repression.

Most recently, SlyA of *E. coli* has been directly linked to the activation of *fimB* expression, one of the genes responsible for catalysing the inversion of a short segment of DNA (*fimS*) that leads to expression of the fimbrial structural operon (McVicker *et al.*, 2011). Though present in both non-pathogenic and pathogenic strains of *E. coli*, expression of Type 1 fimbriae is an important virulence factor for uropathogenic *E. coli* infection of the urinary tract (Wright *et al.*, 2007). It was determined that, once again, SlyA enhances expression of *fimB* by antagonising the repressive effects of H-NS binding within the promoter region. Two H-NS binding sites within the promoter region (H-NS2 and H-NS3) overlap with two SlyA operator sites ( $O_{SA1}$  and  $O_{SA2}$ ), and SlyA was shown to displace H-NS from its binding sites *in vitro*. In the case of binding to the promoter region of *fimB*, no effect of ppGpp on the binding activity of SlyA was observed, contrary to observations with *S. enterica* Serovar Typhimurium SlyA binding to the *pagD-pagC* intergenic region (McVicker *et al.*, 2011; Zhao *et al.*, 2008). Of note, in contrast with the observations in the K5 capsular gene promoters in an *hns* mutant background, SlyA binding to the promoter region of *fimB* appeared to have an inhibitory effect *in vitro* when H-NS was not present. However, it is rightly stated that “SlyA has a net

activating effect on *fimB* expression in the wild type background” by antagonizing H-NS (McVicker *et al.*, 2011).

Further studies into the role of SlyA in the complex regulatory network attributed to *fimB* expression have suggested an additional role for the MarR family regulator. It is hypothesised that SlyA may activate the expression of an unknown protein factor that regulates an additional unknown stress response pathway, ensuring that on reception of a stress signal (represented experimentally by procaine) *fimB* expression is inhibited. In the absence of SlyA, this stress response pathway will constitutively repress *fimB* expression, regardless of stress signal (Moores *et al.*, 2014).

In a proteomic study comparing *S. enterica* Serovar Typhimurium and enteroinvasive *E. coli* (EIEC) bacterial species, the effect of a *slyA* deletion was investigated. It was found that, as in *S. enterica* Serovar Typhimurium, SlyA of *E. coli* seems to have both positive and negative regulatory effects on multiple genes. Genes that appeared to be regulated by *E. coli* SlyA included several molecular chaperones (including GroEL, GroES, DnaK, GrpE and CbpA), a few proteins involved in acid resistance (HdeA, HdeB and GadA) and five of the eight histidine biosynthesis enzymes (HisA,B,D,F and G) (Spory *et al.*, 2002). Significantly, with the exception of GroEL, the regulons of *E. coli* SlyA and *S. enterica* Serovar Typhimurium SlyA differed completely, and though they specifically studied EIEC it appeared that SlyA played no role in regulation of *E. coli* virulence genes as all affected proteins are present in *E. coli* K-12.

*Escherichia coli slyA* expression is maximal during phases of growth prior to stationary phase (Corbett *et al.*, 2007) and it was hypothesised that SlyA may be preferentially expressed at slower rates of growth due to the *slyA* open reading frame starting with the unusual UUG codon. It has been shown that modification of the first codon to the more common AUG increased SlyA expression by over 5-fold (McVicker *et al.*, 2011). Given that it had been previously shown that expression of poorly-translated proteins generally rises as growth rate is reduced (Liang *et al.*, 2000), it was suggested that SlyA expression would be increased with slower growth in minimal medium compared with relatively fast growth in rich medium. This was

supported by Western blot analysis wherein 21% more SlyA was observed in cells grown in minimal medium (McVicker *et al.*, 2011).

### **1.9. Summary and aims**

Regulation of bacterial gene expression is complex, involving the input of multiple factors, including promoter consensus sequence similarity, sigma factor preference of RNAP, dependence on transcription activator and/or repressor proteins, the nucleoprotein architecture imposed by NAPs and specific environmental signal response mediated by any of the above. The MarR family of transcription regulators is involved in a wide range of regulatory roles. Within this family, the mechanism by which gene transcription is activated by derepression of the H-NS nucleoprotein complex is a recurring theme, often leading to expression of genes important for virulence in pathogenic species.

Despite the apparent importance of the MarR family regulator SlyA in *S. enterica* Serovar Typhimurium, direct interaction of regulator with gene promoter regions has been demonstrated in very few cases. Even less is known about the breadth of the regulatory role of SlyA in *E. coli* where, to date, *hlyE* and *fimB* are the only genes confirmed to be directly regulated by SlyA in *E. coli* K-12 (in addition to positive auto-regulation at *PslyA*). The initial aims of the work described here were to: (1) establish whether the regulatory role of SlyA is significantly affected by growth rate in *E. coli* when other growth conditions are kept constant; (2) further determine whether ppGpp plays a role in modulating the ability of SlyA to bind promoter DNA; and, perhaps most importantly, (3) identify the SlyA regulon through microarray analysis of *E. coli* K-12 lacking or overexpressing *slyA* and infer whether the mechanism of H-NS antagonism may be a common theme.

## 2. Materials and methods

### 2.1. Strains and plasmids

The bacterial strains used in this study are listed in Table 2.1 and Table 2.2, with the plasmids listed in Table 2.3.

**Table 2.1. Strains of *Escherichia coli* used in this study**

Strain	Relevant Characteristics	Source or Reference
DH5 $\alpha$	<i>supE44, lacU169</i> ( $\phi$ 80 <i>lacZ</i> M15), <i>hsdR17, recA1, endA1, gyrA96, lthE-1, relA1</i>	Lab collection
MG1655	F <sup>-</sup> $\lambda$ <i>ilvG- rfb-50 rph-1</i>	Lab collection
MG1655 $\Delta$ <i>slyA</i>	MG1655 <i>slyA</i>	This work
BL21 $\lambda$ (DE3)	<i>Escherichia coli</i> BL21 $\lambda$ (DE3) lysogen carrying a copy of the T7 RNA polymerase under the control of the IPTG-inducible <i>lacUV5</i> promoter	Lab collection
Stellar Competent Cells	F <sup>-</sup> , <i>endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, <math>\Phi</math>80d lacZ<math>\Delta</math> M15, <math>\Delta</math>(<i>lacZYA - argF</i>) U169, <math>\Delta</math>(<i>mrr - hsdRMS - mcrBC</i>), <math>\Delta</math><i>mcrA, <math>\lambda</math></i></i>	Clontech
MG1655(pKD46)	<i>Escherichia coli</i> MG1655 carrying the lambda-red recombinase encoding plasmid; pKD46.	Lab collection

**Table 2.2. Strain of *Vibrio harveyi* used in this study**

Strain	Relevant Characteristics	Source or Reference
BB170	Km <sup>R</sup> ; <i>luxN::Tn5</i>	Klaus Winzer

**Table 2.3. Plasmids used in this study**

Plasmid	Relevant characteristics	Source or Reference
pET28a	Multi-copy vector carrying kanamycin resistance	Novagen

pET28a: <i>slyA</i>	Kn <sup>R</sup> , pET-28a derivative containing <i>slyA</i> gene under the control of its own promoter.	This work
pLATE-51	Amp <sup>R</sup> , Low copy number, linearised expression vector with 14 nt 3'-5' sticky end (3'-CCACTACTACTACT-5') and 13 nt 5'-3' sticky end (5'-ACTTCCCATCTCC-3'). Places gene under the control of the bacteriophage T7 promoter. Produces protein with an N-terminal 6x His tag.	Thermo Scientific
pGS2469	Amp <sup>R</sup> , derivative of pLATE-51 for production of SlyA with an N-terminal 6x His tag.	This work
pKD4	Amp <sup>R</sup> , plasmid encoding kanamycin resistance cassette for use in the lambda-red recombination method of gene deletion.	Lab collection
pKD46	Amp <sup>R</sup> , plasmid encoding lambda-red recombinase under the control of an L-arabinose inducible promoter. Has a temperature sensitive origin such that it replicates at 30°C but not at 37°C	Lab collection

## 2.2. Media

### 2.2.1. Rich media

*Escherichia coli* was cultured in Luria Bertani (LB) medium.

<u>LB medium</u>	<u>LB</u> (g l <sup>-1</sup> )	<u>LB Agar</u> (g l <sup>-1</sup> )
Tryptone	10	10
NaCl	10	10
Yeast extract	5	5
Agar bacteriological	-	13

Auto Induction Medium was used for growth of *E. coli* strain BL21λ(DE3)(pGS2469) to be used for SlyA protein expression as described in (Studier, 2005), and was made as follows:

<u>ZYP-20052S Medium (500 ml)</u>	
Casamino acids	5 g
Yeast extract	2.5 g

After addition of 435 ml dH<sub>2</sub>O and autoclaving, the following sterile solutions were added:

Buffer P (20X stock)	25 ml
Solution 5052 (50X stock)	10 ml
Trace elements (1000X stock)	100 µl
1 M MgSO <sub>4</sub>	1 ml
1 M Sodium succinate	12.5 ml
50% Glycerol	15 ml
20 mM Na <sub>2</sub> SeO <sub>3</sub>	10 µl

The various stock solutions involved were prepared as follows and autoclaved:

<u>20X Buffer P</u>	<u>g in 500 ml</u>
1 M NaHPO <sub>4</sub>	70.98
1 M KH <sub>2</sub> PO <sub>4</sub>	68.05
0.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	33.04

<u>50X Solution 5052</u>	
Glycerol	25% (v/v)
Glucose	2.5% (w/v)
Alpha-lactose monohydrate	10% (w/v)



<u>1000X Trace elements</u>	<u>g in 100 ml</u>
50 mM FeCl <sub>3</sub> .6H <sub>2</sub> O	1.350
20 mM CaCl <sub>2</sub> .2H <sub>2</sub> O	0.290
10 mM MnCl <sub>2</sub>	0.200
10 mM ZnSO <sub>4</sub>	0.290
2 mM CoCl <sub>2</sub>	0.048
2 mM CuCl <sub>2</sub>	0.034
2 mM NiCl <sub>2</sub>	0.048
2 mM Na <sub>2</sub> MoO <sub>4</sub>	0.048
2 mM H <sub>3</sub> BO <sub>3</sub>	0.012
Made up to 100 ml with 60 mM HCl	

Autoinducer Bioassay (AB) medium was used in the autoinducer-2 bioassay with *Vibrio harveyi*, and contained the following:

<u>AB medium</u>	<u>g l<sup>-1</sup></u>
NaCl	17.5
MgSO <sub>4</sub> .7H <sub>2</sub> O	12.3
Vitamin-free casamino acids	2

The pH of the solution was adjusted to 7.5 before autoclaving. Immediately prior to use, medium was supplemented with the following:

Glycerol	1% (v/v)
L-Arginine	1 mM
KH <sub>2</sub> PO <sub>4</sub>	10 mM

### 2.2.2. Minimal media

<u>Evans' minimal medium (20 litres)</u>	
2 M NaH <sub>2</sub> PO <sub>4</sub> .2H <sub>2</sub> O	100 ml
2 M KCl	100 ml
0.25 M MgCl <sub>2</sub>	100 ml
4M NH <sub>4</sub> Cl	500 ml
0.4M Na <sub>2</sub> SO <sub>4</sub>	100 ml
0.004M CaCl <sub>2</sub> .2H <sub>2</sub> O	100 ml
Trace elements*	100 ml
C <sub>6</sub> H <sub>6</sub> NNa <sub>3</sub> O <sub>6</sub>	7.6 g

<u>*Trace elements (2.5 litres)</u>	
ZnO	1.03 g
FeCl <sub>3</sub> .6H <sub>2</sub> O	13.5 g
MnCl <sub>2</sub> .4H <sub>2</sub> O	5 g
CuCl <sub>2</sub> .2H <sub>2</sub> O	0.43 g
CoCl <sub>2</sub> .6H <sub>2</sub> O	1.19 g
H <sub>3</sub> BO <sub>3</sub>	0.16 g
Na <sub>2</sub> MoO <sub>4</sub> .H <sub>2</sub> O	0.01 g
HCl (37%)	20 ml

Evans' minimal medium was adjusted to pH 6.95 prior to autoclaving, after which the medium was supplemented with 20 mM sterile glucose and 30 µg ml<sup>-1</sup> of Na<sub>2</sub>SeO<sub>3</sub>.5H<sub>2</sub>O before use.

### 2.2.3. Media supplements

Antibiotics were added to the autoclaved media when required, at the following concentrations:

Ampicillin	100 µg ml <sup>-1</sup>
Kanamycin	20-30 µg ml <sup>-1</sup>

### **2.3. Growth of bacterial strains**

#### **2.3.1. Chemostat culture of *Escherichia coli* MG1655**

Unless otherwise stated, 10 ml of the appropriate overnight culture grown in Evans' minimal medium was injected into a 2 l Labfors 3 chemostat vessel (Infors-HT, Switzerland) with a 1 l working volume of glucose-limited Evans' minimal medium. The steady state culture was then maintained aerobically at dilution rates of 0.05 h<sup>-1</sup>, 0.1 h<sup>-1</sup>, 0.2 h<sup>-1</sup> or 0.5 h<sup>-1</sup>. A temperature of 37°C and pH of 6.9 were maintained throughout growth, with constant agitation by a 400 rpm stirrer and aeration with filtered air at a rate of 1 l min<sup>-1</sup>.

#### **2.3.2. Chemostat culture of *Escherichia coli* mutant library.**

Growth conditions were the same as above, the only difference being that the starter culture was grown in L-broth, and dilution rate was raised incrementally through d= 0.05, 0.1, 0.2 and 0.5 h<sup>-1</sup> with a minimum of four vessel volumes being passed through between increments.

#### **2.3.3. Storage of strains**

Strains were stored on solid media at 4 °C for up to one month. For long term storage strains were kept in the form of glycerol stocks. These stocks were made as follows; 5 ml overnight cultures were centrifuged and cell plates were resuspended in 1.25 ml LB, 1 ml sterile 80% (v/v) glycerol and the appropriate antibiotics. These were stored at -20 °C.

### **2.4. Bacterial transformation**

#### **2.4.1. Production of electrically competent cells**

LB (50 ml) was inoculated with an overnight culture (1%) and grown at 37°C with shaking (250 rpm) until the OD<sub>600</sub> reached 0.6. Cells were pelleted at 4°C (4,020 xg for 10 min) and washed three times in sterile 10% (v/v) glycerol. The cell pellet was then resuspended in 300 µl of 10% glycerol, and stored as 50 µl aliquots at -70°C.

#### **2.4.2. Transformation of electrically competent cells**

Unless otherwise stated, plasmid DNA (up to 1  $\mu\text{g}$ ) was added directly to 50  $\mu\text{l}$  aliquots of electrically competent cells. The mixtures were kept on ice for 10 min and then transformed by electroporation using a Hybaid Cell Shock unit (1800 V, 1 mm path length) and a chilled electroporation cuvette (Cell Projects, EP-101). LB (1 ml) was added and the cells were incubated at 37°C for 1 h. Cells were then plated onto LB agar containing the correct antibiotic.

#### **2.5. Sampling of chemostat cultures**

##### **2.5.1. Genomic DNA**

Cell culture samples (10 ml) were taken and divided into 750  $\mu\text{l}$  aliquots. Genomic DNA was then purified from these samples using the DNeasy Blood and Tissue Kit (QIAgen, 69504) (Section 2.6.6).

##### **2.5.2. RNA**

Samples for RNA isolation were gathered by addition of 1 volume of bacterial culture to 2 volumes of RNAprotect Bacteria Reagent (QIAgen). This was immediately vortexed for 5 sec and incubated at room temperature for 5 min. The sample was then centrifuged in a cooled centrifuge (4°C) for 10 min at 3,380  $\times\text{g}$ . The supernatant was subsequently poured off and the pellet resuspended in any residual solution. This mixture was transferred to a 1.5 ml eppendorf tube and centrifuged for 10 min at 20,000  $\times\text{g}$  before removing the liquid and storing the pellet at -80°C.

##### **2.5.3. Protein pellets**

New protein synthesis was prevented by addition of 20  $\mu\text{l}$  of a 25  $\text{mg ml}^{-1}$  solution of chloramphenicol to 6 ml of culture. The mixture was divided into 500  $\mu\text{l}$  aliquots in 1.5 ml eppendorf tubes and centrifuged for 3 min at 10,000  $\times\text{g}$ . Pellets were then stored at -80°C.

#### 2.5.4. Supernatant

Samples were prepared by centrifugation of 6 ml of culture at 3,380 xg for 10 mins. The supernatant was then filtered using a single use 0.2 µm Minisart syringe filter (Sartorius Stedim biotech). Samples were stored in 1.5 ml eppendorf tubes at -80°C.

#### 2.5.5. Cell dry weight

Determination of the total cell dry weight of chemostat cultures required the initial weighing of eight empty 15 ml Falcon tubes to a milligram degree of accuracy. To four of these tubes, 10 ml of culture sample was added followed by centrifugation at 3,380 xg for 10 min. The supernatant was then poured off, and the pellet resuspended in 10 ml of MilliQ water followed by further centrifugation at 3,380 xg for 10 min. After removal of the supernatant, all four tubes and the four tubes containing no sample were dried in an oven overnight. The following day the tubes were all re-weighed; weight change of tubes containing sample was equivalent to cell dry weight after addition of observed weight loss due to heating of the four tubes that lacked sample.

### 2.6. Nucleic acid methods

#### 2.6.1. Primers

Primers to be used in PCR amplification were synthesised by and purchased from Eurofins, and are listed in Table 2.4, with their associated sequences and melting temperature values ( $T_m$ ). Primers were dissolved in molecular biology grade water (Sigma, W4502) to a stock concentration of 100 pmoles µl<sup>-1</sup> and stored at -20°C.

**Table 2.4 Primers used in this study**

Primer	Sequence	$T_m$ (°C)	Function
TC7	TAAAAGCCGCATAATA TCTTAGCAAGCTAATT ATAAGGAGATTACACG TCTTGAGCGATT	53	Forward Primer for knockout of <i>slyA</i> in <i>E. coli</i> MG1655. Region in black has homology to pre- <i>slyA</i> sequence, red region is homologous to Kn-resistance cassette start (pKD4).

TC8	TTGCGTGTGGTCAGGT TACTGACCACACGCCC CCTTCATT <b>CATATGAA</b> <b>TATCCTCCTTAG</b>	51	Reverse Primer for knockout of <i>slyA</i> in <i>E. coli</i> MG1655. Region in black has homology to post- <i>slyA</i> sequence, red region is homologous to Kn-resistance cassette end (pKD4).
TC9	CTGACGGTAACCAAAT GCAG	57	Forward checking primer for <i>slyA</i> knockout.
TC10	TTTGCCTGTGGTCAGG TTAC	57	Reverse checking primer for <i>slyA</i> knockout.
TC43	GACGGAGCTCGAATTA <b>TCCAAACGCGAATGCT</b> <b>TTG</b>	55	Forward primer for InFusion insertion of <i>slyA</i> (with additional 300bp upstream and 185bp downstream) into pET28a. Black section is homologous to pET8a insertion site, red section is homologous to <i>slyA</i> coding region.
TC44	TCGCGGATCCGAATTA <b>GGGTGTCGAGCTGGAA</b> <b>ATT</b>	57	Reverse primer for InFusion insertion of <i>slyA</i> (with additional 300bp upstream and 185bp downstream) into pET28a. Black section is homologous to pET8a insertion site, red section is homologous to <i>slyA</i> coding region.
TC45	CTTTCGGGCTTTGTTAG CAG	57	Forward checking primer of <i>slyA</i> insertion into pET28a.
TC46	CAGCAGCCATCATCAT CATC	57	Reverse checking primer of <i>slyA</i> insertion into pET28a.
TC47	GGTGATGATGATGACA <b>AGGAATCGCCACTAGG</b> <b>TTCTGATC</b>	60	Forward primer for producing <i>slyA</i> with overhangs necessary for incorporation into pLATE-51. First part is specific to pLATE51, red part is homologous to start of <i>SlyA</i> coding region.
TC48	GGAGATGGGAAGTCAT <b>TACCCTTTGGCCTGTA</b> <b>ACTCAATG</b>	60	Reverse primer for producing <i>slyA</i> with overhangs necessary for incorporation into pLATE-51. First part is specific to pLATE51, red part is homologous to end of <i>SlyA</i> coding region.
TC49	[Btn]ACTCTCTCCTTAT AACCAATTG	54	Forward primer for production of the 5' Biotin-labelled 355 bp intergenic region between <i>ssuE</i> and <i>elfA</i> .
TC50	CGTTATCATCCTGATCT CTT	53	Reverse primer for production of the 5' Biotin-labelled 355 bp intergenic region between <i>ssuE</i> and <i>elfA</i> .
TC51	[Btn]TGGTGAATATTAT TGATCAATTAAT	51	Forward primer for production of the 5' Biotin-labelled 344 bp intergenic region between <i>leuO</i> and <i>leuLABCD</i> .
TC52	ACTTAACTCCACTGTC ACACTTAA	57	Reverse primer for production of the 5' Biotin-labelled 344 bp intergenic region between <i>leuO</i> and <i>leuLABCD</i> .
TC53	[Btn]TTGTTCTCCTTCAT ATGCTC	53	Forward primer for production of the 5' Biotin-labelled 414 bp intergenic region between <i>casA</i> and <i>cas3</i> .
TC54	CTTCGGGAATGATTGT TATC	53	Reverse primer for production of the 5' Biotin-labelled 414 bp intergenic region between <i>casA</i> and <i>cas3</i> .

TC55	[Btn]TGTTGCTAATAGT TAAATCGC	52	Forward primer for production of the 5' Biotin-labelled 257 bp intergenic region between <i>paaA</i> and <i>paaZ</i> .
TC56	GTCATCACCTTTACGA TTCC	55	Reverse primer for production of the 5' Biotin-labelled 257 bp intergenic region between <i>paaA</i> and <i>paaZ</i> .
TC57	[Btn]AACAAACAACCTCC TTGTCCG	55	Forward primer for production of the 5' Biotin-labelled 400 bp promoter region of <i>mdtM</i> .
TC58	CCCCGAGGCGCTTTCC AGGC	67	Reverse primer for production of the 5' Biotin-labelled 400 bp promoter region of <i>mdtM</i> .
TC59	[Btn]AGAACTTCTGTT TTAATTATTG	51	Forward primer for production of the 5' Biotin-labelled 179 bp intergenic region between <i>gspA</i> and <i>gspC</i> .
TC60	GATGTATGTTCTAATA AAATAGATTG	53	Reverse primer for production of the 5' Biotin-labelled 179 bp intergenic region between <i>gspA</i> and <i>gspC</i> .
TC61	[Btn]CCGTCGTTGACTC CATGC	58	Forward primer for production of the 5' Biotin-labelled 130 bp intergenic region between <i>sgcA</i> and <i>sgcQ</i> .
TC62	GATGGGGATAAGCAG AGC	56	Reverse primer for production of the 5' Biotin-labelled 130 bp intergenic region between <i>sgcA</i> and <i>sgcQ</i> .
TC63	[Btn]GCGGAGTGCATCA AAAGT	53	Forward primer for production of the 5' Biotin labelled 291 bp intergenic region between <i>fecI</i> and <i>insA-7</i> .
TC64	GCAAGCACCTTAAAAT CAC	52	Reverse primer for production of the 5' Biotin labelled 291 bp intergenic region between <i>fecI</i> and <i>insA-7</i> .
TC65	[Btn]TTTCATCTCCTTAT AATTAGCTT	51	Forward primer for production of the 5' Biotin labelled 200 bp intergenic region between <i>slyA</i> and <i>ydhI</i> .
TC66	AAAGTAGATTCCTTTA CGACC	54	Reverse primer for production of the 5' Biotin labelled 200 bp intergenic region between <i>slyA</i> and <i>ydhI</i> .
TC67	TTTTATTCTTCTGCAAA CGAT	50	Reverse primer for production of the 5' Biotin labelled 377 bp truncated <i>PmdtM</i> region, <i>PmdtM(A)</i> .
TC68	ACATTTTTCCGGAAC AAGC	53	Reverse primer for production of the 5' Biotin labelled 356 bp promoter region of <i>mdtM</i> , truncated <i>PmdtM</i> region, <i>PmdtM(B)</i> .
TC69	CAATTCAAGAGGTGTA ATGT	51	Reverse primer for production of the 5' Biotin labelled 226 bp promoter region of <i>mdtM</i> , truncated <i>PmdtM</i> region, <i>PmdtM(C)</i> .
TC70	[Btn]AGCTATCTCCGTA GACCGT	56	Forward primer for production of the 5' Biotin labelled 400 bp promoter region of <i>sgcX</i> .
TC71	GATTATCTATACTCCCT CTGAATC	57	Reverse primer for production of the 5' Biotin labelled 400 bp promoter region of <i>sgcX</i> .

TC72	TGACTATTTGTAATCG TTATACATTC	55	Reverse primer for production of the 5' Biotin labelled 201 bp promoter region of <i>mdtM</i> , truncated <i>PmdtM</i> region, <i>PmdtM(D)</i> .
------	--------------------------------	----	--------------------------------------------------------------------------------------------------------------------------------------------------

### 2.6.2. PCR amplification


The polymerase chain reaction (PCR) enables *in vitro* amplification of specific target DNA sequences. When producing products that required high fidelity; PCR reactions were carried out using Extensor Hi-Fidelity PCR Master Mix (Thermo Scientific) that contains the Extensor Long Range PCR Enzyme Blend, dNTPs, Extensor Long Range PCR Buffer and MgCl<sub>2</sub> in a 2x concentrated mix.

DNA templates used for PCR amplification were either *E. coli* MG1655 genomic DNA or the appropriate plasmid.

The standard 20 µl reaction mixture is shown here:

Extensor Hi-Fidelity PCR Master Mix (2x)	10 µl
Forward Primer (100 pmol µl <sup>-1</sup> )	1 µl
Reverse Primer (100 pmol µl <sup>-1</sup> )	1 µl
Template (gDNA or plasmid)	2 µl
Nuclease-free water	7 µl

Reactions were amplified using the conditions outlined below on a Techne TC-3000 Thermal Cycler.

Initial Denaturation	95°C	7 min	 X30 Cycles
Denaturation	94°C	1 min	
Annealing	Variable	45 s	
Elongation	68°C	Variable	
Final Elongation	68°C	5 min	

Elongation times were selected allowing 1 min per kilobase-pair of the expected PCR product. Annealing temperature was dependent on the lower T<sub>m</sub> of the primer pair used (Table 2.4).



### 2.6.3. Colony PCR

Colony PCR can be used for fast amplification of a specific DNA fragment straight from a bacterial colony and subsequent analysis by agarose gel electrophoresis. These PCR reactions were carried out using ReddyMix PCR Master Mix (Thermo Scientific) that contains the Thermoprime Plus DNA Polymerase, dNTPs, reaction buffer and MgCl<sub>2</sub> in a 2x concentrated mix. ReddyMix Master Mix also contains a dye and precipitant to facilitate gel loading.


The DNA template for colony PCR is a crude single-colony scraping from an LB agar plate.

The standard 20 µl reaction mixture is shown here:

ReddyMix PCR Master Mix (2x)	18 µl
Forward Primer (100 pmol µl <sup>-1</sup> )	1 µl
Reverse Primer (100 pmol µl <sup>-1</sup> )	1 µl
Template (colony scraping)	N/A

Reactions were amplified using the conditions outlined below on a Techne TC-3000 Thermal Cycler.

Initial Denaturation	95°C	7 min
Denaturation	94°C	30 s
Annealing	Variable	1 min
Elongation	68°C	Variable
Final Elongation	68°C	5 min



X30 Cycles

Elongation times were selected allowing 1 min per kilobase-pair of the expected PCR product. Annealing temperature was dependent on the lower T<sub>m</sub> of the primer pair used (Table 2.4).

### 2.6.4. PCR purification

PCR products were purified using QIAquick PCR Purification kit (QIAGEN, 28104) following the manufacturer's instructions. Washed DNA was eluted into 50

$\mu\text{l}$  of molecular biology grade water (Sigma W4502). Purified PCR products were routinely stored at  $-20^{\circ}\text{C}$ .

#### **2.6.5. Isolation of RNA from stored pellets**

RNA sample pellets were resuspended in 400  $\mu\text{l}$  of TE buffer (10 mM Tris-Cl, 1 mM EDTA pH 8.0) containing lysozyme (15 mg/ml) and 200  $\mu\text{l}$  of this sample was used. Enzymatic lysis of bacteria was carried out according to the manufacturer's instructions of RNAprotect Bacteria Reagent Handbook (QIAGEN, 76506), utilising the RNeasy Mini Kit (QIAGEN, 74106). Samples were eluted in 50  $\mu\text{l}$  of molecular biology grade water (Sigma, W4502).

#### **2.6.6. Genomic DNA preparation**

Genomic DNA preparation followed a protocol based on the DNeasy Blood and Tissue Handbook, using solutions from the DNeasy Blood and Tissue Kit (QIAGEN, 69504). This was carried out using 750  $\mu\text{l}$  of overnight cell culture or steady-state chemostat culture.

- Cells were harvested by centrifugation in a microcentrifuge for 10 min at 5,000  $\times g$ . Supernatant was discarded.
- Pellet was resuspended in 180  $\mu\text{l}$  Buffer ATL.
- 20  $\mu\text{l}$  of supplied proteinase K (15 mg  $\text{ml}^{-1}$ ) was added and mixed thoroughly by vortexing.
- Sample was then incubated at  $56^{\circ}\text{C}$  for a minimum of 1 h with vortex mixing at least once every 30 min) until the tissue was completely lysed. This was followed by vortexing for 15 sec.
- In a separate tube, 200  $\mu\text{l}$  Buffer AL was mixed with 200  $\mu\text{l}$  ethanol (99%) by vortexing. This 400  $\mu\text{l}$  mix was then added to the sample, followed by vortexing.
- Mixture was pipetted into a DNeasy Mini spin column placed in a 2 ml collection tube, and centrifuged at 6,000  $\times g$  for 1 min. Flow through and collection tube was discarded.
- DNeasy Mini spin column was placed in a new 2 ml collection tube, 500  $\mu\text{l}$  AW1 was applied to the membrane, followed by centrifugation at 20,000  $\times g$  for 1 min. Flow through and collection tube was discarded.

- DNeasy Mini spin column was placed in a new 2 ml collection tube, 500  $\mu$ l AW2 was applied to the membrane, followed by centrifugation at 20,000  $\times g$  for 3 min. Flow through and collection tube was discarded.
- DNeasy Mini spin column was placed into a standard clean 1.5 ml microcentrifuge tube and 200  $\mu$ l of molecular biology grade water (Sigma, W4502) was applied to the membrane. This was incubated at room temperature for 1 minute, followed by centrifugation for 1 min at 6,000  $\times g$  to elute the desired sample.
- For maximum DNA yield, a fresh 200  $\mu$ l of molecular biology grade water (Sigma, W4502) was applied to the membrane, elution was repeated and samples were combined.

#### **2.6.7. Plasmid purification**

Plasmid DNA was typically purified from 5 ml overnight cultures of *E. coli* using the Qiagen QIAprep Spin Miniprep kit, in accordance with manufacturer's instructions.

#### **2.6.8. Quantification of nucleic acid concentration**

For measurement of nucleic acid concentrations a NanoDrop ND-1000 Spectrophotometer was used. This measured the concentration in  $\text{ng } \mu\text{l}^{-1}$  from 2  $\mu$ l of undiluted sample.

#### **2.6.9. Agarose gel electrophoresis**

Agarose gel electrophoresis enables the separation and visualisation of nucleic acid fragments based on their size. Agarose (1% or 1.5% (w/v)) was added to 1X TAE buffer and dissolved by heating in a microwave oven. Once sufficiently cooled, GelRed solution (Biotium) was added to the gel in a 1:10000 dilution before casting. Before loading, DNA samples were mixed with 6X loading dye (Fermentas). A 1 kb ladder 'HyperLadder I' (Bioline) or a 100 b ladder 'HyperLadder IV' (Bioline) was used for size calibration. Gels were electrophoresed in 1X TAE buffer at 100 V for 1 h. DNA fragments were subsequently visualised using an UVitech photodocumentation system.

50X TAE Buffer	
Tris	242 g
Glacial acetic acid	57.1 ml
EDTA (0.5 M, pH 8)	100 ml
dH <sub>2</sub> O to 1000 ml	

#### **2.6.10. Gel extraction**

Specific DNA fragments were removed from agarose gels and purified using the QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions.

#### **2.6.11. Digestion of DNA with restriction endonucleases**

Restriction enzymes were used in conjunction with the appropriate 10X buffer in accordance with the manufacturer's instructions. Digests were usually carried out in 20 µl reaction volumes containing the DNA to be digested and the enzyme at 1:10 of the reaction volume. Reactions were incubated at 37°C, usually for 2 h. The desired DNA fragments were then purified by PCR purification (Section 2.6.4) and analysed by agarose gel electrophoresis (Section 2.6.9).

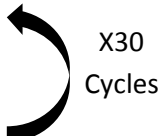
In the case of digestion of genomic DNA samples in the transposon mediated differential hybridisation (TMDH) study, 5 µg of purified genomic DNA was incubated with AluI (1:20 of the reaction volume) and 10X Tango buffer (1:10 of the reaction volume) (Thermo Scientific, ER0011). Reactions were carried out in a 20 µl volume and left to digest at 37°C for 16 h.

#### **2.6.12. Gene deletion using the Lambda-Red system**

Gene knockouts were carried out as described by Datsenko and Wanner (2000). Firstly, an antibiotic resistance cassette was produced by PCR, utilising the primers TC7 and TC8 with the plasmid pKD4 as the template. This produced a kanamycin resistance cassette flanked by 40 bp homologous to the *slmA* gene region targeted for deletion. PCR reactions were prepared as detailed below in a 300 µl volume, which was then split into 3 x 100 µl volumes for amplification in a Techne TC-3000 Thermal Cycler. HotStarTaq Mastermix (Qiagen, 203443) was used as

source of DNA polymerase. The 100 pmoles  $\mu\text{l}^{-1}$  primer stocks underwent a further 1:10 dilution (10 pmoles  $\mu\text{l}^{-1}$ ), and the pKD4 template was added as a 1:1000 dilution of a plasmid prep.

HotStarTaq Mastermix (2x)	150 $\mu\text{l}$
TC7 Forward Primer	24 $\mu\text{l}$
TC8 Reverse Primer	24 $\mu\text{l}$
Template (pKD4)	24 $\mu\text{l}$
Nuclease free water	78 $\mu\text{l}$

HotStart step	95°C	15 min	
Denaturation	94°C	45 s	
Annealing	56°C	1 min	
Elongation	72°C	2 min	
Final Elongation	72°C	10 min	

The PCR reactions (300  $\mu\text{l}$ ) were then cleaned up and eluted into 50  $\mu\text{l}$  nuclease free water as described in Section 2.6.4. In order to remove any remaining plasmid template from the PCR product, DNA was digested with DpnI (Promega, R6231) in a 100  $\mu\text{l}$  reaction volume and incubated at 37°C for 4 h. After incubation, the reactions were purified using a QiaQuick PCR cleanup column (Section 2.6.4).

The purified PCR product (5  $\mu\text{g}$ ) was introduced into *E. coli* MG1655(pKD46) cells expressing the lambda-red recombinase by electroporation (Section 2.4). The pKD46 plasmid contains the lambda-red recombinase under the control of an L-arabinose inducible promoter, an ampicillin resistance gene to aid in selection for the plasmid and a temperature sensitive origin of replication that permitted replication at 30°C but not at 37°C. Hence, electrically competent cells were prepared by growth in LB in the presence of 1 mM L-arabinose and 100  $\mu\text{g ml}^{-1}$  ampicillin at 30°C prior to harvest and preparation as described (Section 2.4.1). After electroporation, cells were recovered in 1 ml LB and incubated at 37°C for 1 h followed by plating of 100  $\mu\text{l}$  of these cells onto LB agar containing kanamycin (30  $\mu\text{g/ml}$ ). Transformants containing the kanamycin resistance cassette in place of the *slyA* gene region were confirmed by colony PCR (Section 2.6.3) using the primers

TC9 and TC10. In order to reduce the risk of secondary mutations or chromosomal deletions that may develop in this procedure, the desired *slyA* deletion mutation was transduced with P1 into wild-type *E. coli* MG1655 (Section 2.8).

#### **2.6.13. Creation of pET28a:*slyA* plasmid using the In-Fusion HD cloning system**

A purified PCR fragment encoding *slyA* and its own promoter region with appropriate 15 bp overhangs (synthesised with primers TC43 and TC44) was incubated with a sample of linearised pET28a vector (digested with EcoRI as described in Section 2.6.11) according to the In-Fusion HD cloning system manufacturer's instructions (Clontech, 639642). The reaction was carried out in a 10 µl volume:

5X Infusion HD Enzyme Premix	2 µl
Linearised pET28a vector	50 ng
<i>slyA</i> PCR product with 15 bp overhangs	80 ng
Nuclease-free water	To 10 µl final volume

A portion of the total reaction mix (2.5 µl) was then transformed into Stellar Competent Cells according to the manufacturer's instructions (Clontech, PT5055-2), which was subsequently plated on a kanamycin containing LB agar plate. Kanamycin resistant colonies were selected, restreaked onto fresh plates, and insertion of *slyA* fragment was confirmed by colony PCR with primers TC45 and TC46 (Section 2.6.3). The DNA sequence of the *slyA* insertion was verified (University of Sheffield Core Genomic Facility), and *E. coli* MG1655 was subsequently transformed with pET28a:*slyA* (Section 2.4.).

#### **2.6.14. Creation of the pGS2469 plasmid using the aLICator ligation independent cloning and expression system**

A purified PCR fragment coding for SlyA with appropriate 18 bp overhangs (synthesised with primers TC47 and TC48) was incubated with linearised pLATE51 vector according to the manufacturer's instructions (Thermo Scientific, K1251). Reactions were carried out in a 10 µl volume containing:

5X LIC Buffer	2 $\mu$ l
Purified <i>slyA</i> PCR product	32.5 ng
Nuclease-free water	To 10 $\mu$ l final volume
T4 DNA Polymerase (1 u $\mu$ l <sup>-1</sup> )	1 $\mu$ l

Reaction mix was incubated at room temperature for 5 min, followed by addition of 0.6  $\mu$ l EDTA (0.5 M) to stop the reaction. To this, 1  $\mu$ l of linearised pLATE51 vector (60 ng) was added. After mixing, this mixture was incubated at room temperature for a further 5 min by which point the annealing of the synthesised SlyA encoding fragment into the pLATE51 vector was complete. A portion of this reaction mixture (1  $\mu$ l) was then used to transform *E. coli* DH5 $\alpha$  by electroporation (Section 2.4), with subsequent plating onto LB agar plates containing ampicillin.

Presence of *slyA* insertion in pLATE51 for those colonies that grew on LB agar plates was confirmed by colony PCR using primers provided in the aLICator ligation kit (Section 2.6.3). The DNA sequence of the *slyA* insertion was verified (University of Sheffield Core Genomic Facility), and *E. coli* BL21(DE3) was subsequently transformed with pLATE51:*slyA*, now designated pGS2469 (Section 2.4).

### **2.6.15. Transcriptomics: Microarray analyses**

An indirect method of microarray analysis was used in all cases, with RNA samples always being labelled with fluorophore Cy5 and *E. coli* MG1655 genomic DNA labelled with the fluorophore Cy3 as a reference. This indirect approach has the advantage that dye-swap experiments are not required, the comparison between many RNA samples requires fewer microarrays and that multiple datasets for which genomic DNA has been used as a reference can be compared (Yang and Speed, 2002).

#### **2.6.15.1. Direct labelling of RNA**

RNA labelling was carried out using SuperScript III reverse transcriptase (Invitrogen, 18080-044) with the Cy5-dCTP included in the dNTP mixture. Random priming reverse transcription reactions were set up in 0.5 ml microcentrifuge tubes, as follows:

- 8 µg of total purified RNA was adjusted to 6.15 µl with molecular biology grade water (Sigma, W4502).
- This was followed by addition of 2.5 µg random primers (0.85 µl of 3 mg/ml stock (Invitrogen, 48190-011)).
- The mixture was incubated at 72°C for 10 min, placed on ice for a further 10 min and then centrifuged briefly to collect liquid together.
- To this, 6.25 µl of ‘Reverse Transcription (RT) mix’ was added.

<u>RT mix (6.25 µl)</u>	
5X First Strand Buffer	3 µl
0.1 M DTT	1.5 µl
50X dNTP mix	0.3 µl
Molecular biology grade water	1.45 µl

50X dNTP mix contained 25 mM each of dATP, dTTP, dGTP and 10 mM dCTP.

- This was followed by addition of 1 µl of 1 mM Cy5-dCTP (GE Healthcare, PA55021) and 0.75 µl reverse transcriptase (200 U µl<sup>-1</sup>, SuperScript III).
- This was incubated at 25°C for 5 min followed by overnight at 50°C.
- 7.5 µl of fresh 0.1 M NaOH was added and left at 72°C for 10 min in order to hydrolyse the RNA. Following this, 7.5 µl of fresh 0.1 M HCl was added to neutralise the alkali.
- Labelling reactions were then cleaned up by PCR purification (Section 2.7.4.).
- Concentration of cDNA and labelling efficiency was measured by NanoDrop (Section 2.7.8.). Before testing concentration, 2 µl of sample to be tested was denatured at 100°C for 2 min.

#### **2.6.15.2. Direct labelling of MG1655 genomic DNA**

The fluorophore Cy3 was incorporated into an *E. coli* MG1655 genomic profile, the BioPrime DNA Labelling System (Invitrogen, 18094-011) was used as a source of random primers (octamers), Klenow polymerase and reaction buffer. Reactions were set up as follows:

- 2 µg genomic DNA was made up to 21 µl with molecular biology grade water (Sigma, W4502) in a 1.5 ml microcentrifuge tube.



- This was followed by addition of 20  $\mu$ l 2.5X random primers/reaction buffer mix from the BioPrime kit, boiling for 5 min and placing on ice for 5 min.
- While on ice, 5  $\mu$ l of 10X dNTP mix (1.2 mM each dATP, dGTP, dTTP; 0.6 mM dCTP; 10 mM Tris pH 8.0; 1 mM EDTA) was added to the mixture. The dNTP mix from the BioPrime kit was not used as it contained biotinylated dCTP.
- To this 3  $\mu$ l of 1 mM Cy3-dCTP (GE Healthcare, PA53021) and 1  $\mu$ l of Klenow from the BioPrime kit was added.
- This was centrifuged briefly and incubated at 37°C overnight in the dark.
- Labelling reactions were cleaned up by PCR purification (Section 2.6.4), washing twice with Buffer PE.
- Concentration of cDNA and labelling efficiency was measured by NanoDrop (Section 2.7.8.). Before testing concentration, 2  $\mu$ l of sample to be checked was denatured at 100°C for 2 min.

#### **2.6.15.3. Microarray hybridisation**

Microarray hybridisation and subsequent washing of array slides was carried out as per manufacturer's instructions in the Agilent "Two-Color Microarray-Based Prokaryotic Analysis (FairPlay III Labeling) Protocol"; 300 ng of Cy5-labelled cDNA representing sample to be investigated was hybridised in tandem with 300 ng of Cy3-labelled cDNA representing a whole genomic profile of *E. coli* MG1655. Arrays used were of a custom design based on the Agilent *E. coli* Gene Expression Microarrays (Agilent, G413A) with additional probes added representing sRNAs (Agilent Design ID: 029412). Hybridisation was carried out at 65°C for 17 h in an Agilent Hybridisation Oven (Agilent, G2545A).

#### **2.6.15.4. Microarray scanning and data analysis**

Arrays were scanned in an Agilent Microarray Scanner using associated software (Agilent, G2565CA). Data were extracted from .tif files using Agilent Feature Extraction 11.5 software and analysed using Agilent GeneSpring 7.3.1. Signal intensity data for each array was divided by the signal obtained from the control channel (i.e. Cy3-labelled MG1655 genomic cDNA) with a median shift being applied across all samples included in each comparison.

In the case of the comparison between *E. coli* MG1655 $\Delta$ *slyA* and wild type MG1655, data obtained for the mutant strain was normalised to the amount of transcript detected in the wild type sample. This comparison was carried out separately for samples associated with growth at each dilution rate tested.

A similar comparison was carried out between *E. coli* MG1655(pET28a:*slyA*) and wild type MG1655, with all data normalised to the amount of transcript detected in the wild type sample. This comparison was only carried out for samples grown at  $d = 0.2 \text{ h}^{-1}$

For the growth rate study described in Chapter 4, data associated with wild type *E. coli* MG1655 grown at each dilution rate was normalised to transcript abundances detected in samples grown at the lowest dilution rate ( $d = 0.05 \text{ h}^{-1}$ ).

#### **2.6.16. Transposon mediated differential hybridisation (TMDH) arrays**

For the TMDH study, genomic DNA of the transposon mutant library was purified from an LB overnight inoculum, a chemostat grown batch culture and the subsequent steady-state culture samples grown at increasing dilution rates (Section 2.6.6). In the case of the *E. coli* MG1655 control sample, genomic DNA was purified from a 750  $\mu\text{l}$  culture sample grown overnight in LB at 37 °C. All genomic DNA samples were digested with AluI restriction enzyme (Section 2.6.11), the product of which was run on a 1.5X agarose gel, visualised on a UV lightbox, and any digested DNA product in the range of 100 to 200 bp was purified by Gel Extraction (Section 2.6.10) and eluted in 200  $\mu\text{l}$  of nuclease-free water.

The digested genomic DNA was then phenol:chloroform extracted and ethanol precipitated as follows. The sample was mixed with phenol:chloroform (1:1, v/v), followed by vortexing and then centrifuged at 20,000  $\times g$  at 4°C for 30 min, thus separating the aqueous and organic phases. The top aqueous phase was then removed. To the aqueous phase, 10% (v/v) of 3 M sodium acetate (pH 5.2) was added, followed by 1  $\mu\text{l}$  of 20 mg  $\text{ml}^{-1}$  glycogen. Cold ethanol (-20°C) was then added (2:1, v/v), mixed by vortexing and left at -20°C overnight. This sample was then centrifuged at 20,000  $\times g$ , at 4°C, for 30 min and the resulting supernatant was removed. Any remaining ethanol was removed by centrifugation in a heated vacuum

centrifuge for 30 min. Sample was then resuspended in 5  $\mu$ l molecular biology grade water (Sigma, W4502) and quantified by NanoDrop (Section 2.6.8).

#### **2.6.16.1. Synthesis of labelled T7-promoter derived cRNA products**

The desired labelled cRNA products were synthesised from T7-promoters present in digested genomic DNA using the MEGAscript High Yield Transcription Kit (Ambion, AM1354), and Cy5-UTP (GEHealthcare, PA55026). Reactions were set up as follows:

Nuclease-free Water	To 10 $\mu$ l final volume
T7 10X Reaction Buffer	1 $\mu$ l
15 mM ATP	1 $\mu$ l
15 mM CTP	1 $\mu$ l
15 mM GTP	1 $\mu$ l
15 mM UTP	0.75 $\mu$ l
2 mM Cy5-UTP	1.875 $\mu$ l
Digested genomic DNA (Template)	500 ng
T7 Enzyme Mix	1 $\mu$ l

Reactions were incubated at 37°C overnight. Samples were then treated with 0.5  $\mu$ l TURBO DNase and incubated at 37°C for a further 15 min. The cRNA products were then purified using the RNeasy Mini Kit (QIAGEN, 74106), according to the manufacturer's instructions. Concentration of cRNA and labelling efficiency was measured by NanoDrop (Section 2.6.8).

#### **2.6.16.2. Microarray hybridisation**

Hybridisation of samples to arrays and subsequent washing of slides was based on the protocols described in the Agilent "One-Color Microarray-Based Prokaryote Analysis (FairPlay III Labeling) Protocol", though in this case 200 ng of Cy5-labelled sample was hybridised to each array. Arrays used represented the entire 4.6 Mb genome of *E. coli* MG1655 (NCBI RefSeq: NC\_000913.2) and were designed by Genotypic Technology Ltd. Each probe was 60 bases in length, spaced approximately 150 bp apart on the genome, on both the sense and anti-sense strands (Agilent Design ID: 059665).

### **2.6.16.3. Array scanning and data analysis**

Arrays were scanned in an Agilent Microarray Scanner using associated software (Agilent, G2565CA). Data were extracted from .tif files using Agilent Feature Extraction 11.5 software. Spot intensity data was then analysed as described in Sections 4.8 and 4.9.

## **2.7. Protein Methods**

### **2.7.1. Measurement of protein concentration**

Protein concentration was typically estimated using the Bio-Rad reagent (Bradford, 1976). A protein assay standard curve for a 5 ml reaction volume was used with absorbance measured at 595 nm using a Unicam HELIOS spectrophotometer.

### **2.7.2. Denaturing gel electrophoresis (SDS-PAGE)**

SDS-PAGE was performed as described by (Laemmli, 1970). Analysis was completed using 12%, 17.5% or 20% (w/v) resolving gel, and 6% (w/v) stacking gel as shown below:

	<u>Resolving gel</u>			<u>Stacking gel</u>
	<u>12% (ml):</u>	<u>17.5% (ml):</u>	<u>20% (ml):</u>	<u>5% (ml):</u>
30% (w/v) Acrylamide	3.2	4.65	5.33	0.75
1.5 M Tris-HCl (pH 8.8)	2	2	2	-
1 M Tris-HCl (pH 6.8)	-	-	-	0.47
dH <sub>2</sub> O	2.6	1.17	0.5	2.46
10% (w/v) SDS	0.08	0.08	0.08	0.037
10% (w/v) APS	0.08	0.08	0.08	0.037
TEMED	0.003	0.003	0.003	0.0037

Glass plates with spacers were set up and clamped together, ready for pouring the gel. The resolving gel was poured, leaving space for the stacking gel to be poured on top later. Isopropanol was applied to the top of the gel to prevent drying out. Once the gel had set, the Isopropanol was rinsed off with dH<sub>2</sub>O and the

stacking gel was applied and a comb inserted to form the wells. This comb was removed once the entire gel had set.

The gel, sandwiched between the glass casting plates, was placed within a tank with 1X SDS running buffer. Samples were mixed in a 1:1 ratio with loading buffer and boiled for 10 min at 95°C before loading. Protein size markers were also loaded in order to determine molecular weight of visualised proteins (Precision Plus Protein Standards from Bio-Rad, 161-0363). Electrophoresis was carried out at 200 V for approximately 50 min, or until the dye front reached the base of the gel.

<u>2X SDS Loading buffer:</u>		<u>1X SDS Running buffer:</u>	
Glycerol	20% (v/v)	Tris	3 g
Tris-HCl	100 mM	SDS	1 g
SDS	4% (v/v)	Glycine	14.4 g
Bromophenol blue	0.02% (v/v)	dH <sub>2</sub> O to 1000 ml	
2-Mercaptoethanol	200 mM		

After electrophoresis the gel was removed from its glass plates and immersed in Coomassie Blue stain overnight, followed by destain for 1 to 2 h.

	<u>Coomassie Blue stain</u> <u>(1 litre):</u>	<u>Destain</u> <u>(1 litre):</u>
Coomassie Brilliant Blue (R250)	1.15 g	-
Methanol	400 ml	400 ml
Acetic acid	100 ml	100 ml

### **2.7.3. Overproduction of SlyA**

SlyA was typically overproduced by culturing the host strain in a 2 l flask containing 500 ml of autoinduction medium and ampicillin, inoculated at 1:100 from an overnight culture. This culture was grown at 37°C with shaking (250 rpm) for 24 h. Cells were then pelleted by centrifugation at 13,433 xg for 30 min at 4°C. Pellets were either frozen overnight at -20°C or used immediately to produce cell-free extracts (Section 2.7.4).

#### **2.7.4. Production of cell-free extracts**

Cell pellets were typically resuspended in 15 ml of breakage buffer. Resuspended cells were lysed by two passages through a French pressure cell at 16,000 psi. The soluble and insoluble fractions were separated by centrifugation at 27,216 xg for 15 min at 4°C. The soluble fraction was used immediately for purification of proteins.

<u>Breakage buffer (pH 7.5):</u>	
Tris	20 mM
NaCl	500 mM
Glycerol	5% (v/v)
dH <sub>2</sub> O to 100 ml	

#### **2.7.5. Purification of recombinant SlyA by affinity chromatography**

SlyA was purified using the engineered histidine tag. Purification used the soluble fractions after the French press step. Purification was carried out using the His-tag purification programme on the AKTA prime machine (GE Healthcare), and a 1 ml HiTrap chelating column (GE Healthcare). The appropriate reagents were prepared and the machine was configured according to the manufacturer's instructions. The 10 ml cell-free extract was injected and the 1 ml fractions eluting from the column were collected. The fractions which corresponded to the second absorption peak (280 nm) after addition of the elution buffer were collected and samples were separated by SDS-PAGE to locate and confirm presence of the target protein. Fractions were stored at 4°C until required.

<u>Binding buffer (pH 7.4):</u>	
Tris	20 mM
NaCl	500 mM
Imidazole	20 mM
dH <sub>2</sub> O to 1 litre	

<u>Elution buffer (pH 7.4):</u>	
Tris	20 mM
NaCl	500 mM
Imidazole	500 mM
dH <sub>2</sub> O to 1000 ml	

<u>Ni-loading eluent:</u>	
NiSO <sub>4</sub>	100 mM

Before use, purified protein was dialysed using a Vivaspin 6 column (Sartorius Stedim biotech, VS0601), with a 10 kDa Molecular Weight Cut-off, into fresh Dialysis buffer:

<u>Dialysis buffer (pH 7.4):</u>	
Tris-HCl	20 mM
NaCl	200 mM

#### **2.7.6. Transfer of proteins onto Hybond-C Extra nitrocellulose membrane**

Proteins were first separated using SDS-PAGE (Section 2.8.2.), before blotting onto the Hybond-C Extra nitrocellulose membrane (GE Healthcare, RPN203E). The gel was soaked for 20 min in transfer buffer, along with blotting paper and sponge layers required for assembly of the transblotting sandwich. The membrane was soaked in dH<sub>2</sub>O for 10 min followed by soaking in Transfer buffer for 10 min. The transblotting sandwich was assembled, the tank filled with Transfer buffer and an ice block inserted before electroblotting was carried out at 100 V for 1 h.

<u>Transfer buffer</u>	
Tris	5.8 g
Glycine	2.9 g
Methanol	200 ml
10% SDS (w/v)	3.7 ml
dH <sub>2</sub> O up to 1000 ml	

#### **2.7.7. Western blotting**

Once polypeptides had been transferred onto the Hybond-C Extra membrane, the membrane was placed in a square petri dish, such that the side which was in contact with the SDS-PAGE gel during transfer was now face up. This was then blocked with 50 ml Blocking solution (5% dried skimmed milk (w/v), 1X PBS and 0.05% Tween 20 (v/v)) overnight at 4°C.

<u>10X PBS (pH 7.4)</u>	
NaCl	80 g
KCl	2 g
Na <sub>2</sub> HPO <sub>4</sub>	14.4 g
KH <sub>2</sub> PO <sub>4</sub>	2.4 g
dH <sub>2</sub> O up to 1000 ml	

The following day, the Blocking solution was removed and the membrane was washed in 1X PBS with 0.05% Tween (v/v), once for 15 min followed by three 5 min washes, each wash in fresh solution. The membrane was then soaked for 1 h in Blocking solution which contained the primary antibody at the dilution listed in Table 2.5. This was followed by repetition of the four wash phases with 1X PBS containing 0.05% Tween (v/v) and soaking for 1 h in Blocking solution containing the secondary antibody (listed in Table 2.5) at a 1:1000 dilution.

**Table 2.5. Antibodies utilised in this work**

<u>Primary Antibody Target</u>	<u>Source of Primary Antibody</u>	<u>Dilution of Primary Antibody</u>	<u>Secondary antibody Required</u>
FlgM	Kelly Hughes	1:10000	Anti-Rabbit
FlhDC	Lab stock	1:5000	Anti-Rabbit
FliA	Abcam	1:1000	Anti-Mouse
FliC	Lab stock	1:3000	Anti-Rabbit
RpoS	Abcam	1:1000	Anti-Mouse
SlyA	Ian Blomfield	1:1000	Anti-Rabbit

The membrane was then soaked in Pierce ECL Western Blotting Substrate (Thermo Scientific, 32106) followed by X-ray film exposure according to the manufacturer's instructions.



### 2.7.8. Electrophoretic Mobility Shift Assays (EMSAs)

Visualisation of protein binding activity to a Biotin-labelled substrate was carried out using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific, 20148), following the manufacturer's instructions.

Binding reactions were set up using the LightShift EMSA Optimization and Control Kit (Thermo Scientific, 20148X). All reactions were carried out in a 20  $\mu$ l volume, incubated at room temperature for 30 min, and contained:

<u>EMSA binding reaction (20 <math>\mu</math>l)</u>	
10X Binding buffer	2 $\mu$ l
1 $\mu$ g $\mu$ l <sup>-1</sup> Poly (dI•dC)	1 $\mu$ l
DTT (50 mM)	2 $\mu$ l

Where specified, reactions also contained 0.1 mg ml<sup>-1</sup> of Heparin. Biotin-labelled promoter DNA being studied and SlyA were added in the amounts described in Section 3.7.

Samples were separated on a 6% Native-PAGE gel (recipe below) for 1 hour followed by transfer onto Hybond-N+ nylon membrane (GE Healthcare, RPN303B) and crosslinking at 120 mJ cm<sup>2</sup> <sup>-1</sup> using a commercial UV-light crosslinking instrument (Amersham, 80-6222-31) for 60 s. The Nucleic Acid Detection Module (Thermo Scientific, 89880) was then used for visualising as described in the manufacturer's instructions, with X-ray film being exposed for 30-60 s.

	<u>6% Resolving gel</u>	<u>5% Stacking gel</u>
	<u>(ml)</u>	<u>(ml)</u>
30% (w/v) Acrylamide	1.6	0.68
1.5 M Tris-HCl (pH 8.8)	2	-
1 M Tris-HCl (pH 6.8)	-	0.5
dH <sub>2</sub> O	4.32	2.76
10% (w/v) APS	0.08	0.04
TEMED	0.006	0.004

## **2.8. P1 phage transduction**

### **2.8.1. Preparation of transducing lysate**

LB agar plates (1.5% agar bacteriological) containing glucose and  $\text{Ca}^{2+}$  were prepared. To 100 ml molten agar 0.56 ml 1 M glucose and 0.5 ml 0.5 M  $\text{CaCl}_2$  (filter sterilised) was added before pouring and these were then pre-warmed at  $37^\circ\text{C}$  before use.

A stock of P1 grown up on wild type *E. coli* MG1655 was provided by Dr Matt Rolfe, from which decimal dilutions (from  $10^{-1}$  down to  $10^{-5}$ ) were prepared in LB containing 2.5 mM  $\text{CaCl}_2$  and 0.1% glucose (P1 LB). The presence of  $\text{CaCl}_2$  provides  $\text{Ca}^{2+}$  which is necessary for P1 phage infectivity.

A 10  $\mu\text{l}$  sample of each dilution of P1 phage stock was mixed with 200  $\mu\text{l}$  of overnight donor culture in a 1.5 ml microcentrifuge tube and incubated for 1 h at room temperature. Overnight donor culture (5 ml) of *E. coli* MG1655 was prepared in P1 LB containing kanamycin (30  $\mu\text{g}/\text{ml}$ ).

LB soft-top agar (0.6% agar bacteriological) containing (per 100 ml) 0.56 ml 1 M glucose and 0.5 ml 0.5 M  $\text{CaCl}_2$  was prepared, with 3 ml aliquots of this mixture added to 6 ml glass tubes in a hot block ( $50^\circ\text{C}$ ). The cell/phage mixtures were then added to these aliquots of molten soft-top agar, mixed briefly by hand-rolling the tube and then poured onto the pre-warmed agar plates ensuring a complete lawn of top agar was produced. Once set, plates were inverted and incubated at  $37^\circ\text{C}$  overnight.

The following day, phage lysate plates were examined, and the one in which plaques had almost reached complete confluency was selected. Phage was harvested from the plate by adding 3 ml of P1 LB to the surface of the plate and incubating at room temperature for 2 to 3 h, with gentle agitation on a rotary shaker. This P1 LB was then poured into a 50 ml Falcon tube in addition to scrapings of the upper layer of soft-top agar from the plates. Addition of 200  $\mu\text{l}$  chloroform ensured killing of any remaining viable cells. The Falcon tube was then spun in a centrifuge at 4,020  $\times g$  for 10 min and supernatant was poured into a sterile glass bijou containing 200  $\mu\text{l}$  chloroform. This mixture was the donor lysate to be used in subsequent transductions.

### **2.8.2. Transduction**

Serial dilutions of donor lysate were prepared in LB over the range of  $10^{-1}$  to  $10^{-4}$ . To 1.5 ml microcentrifuge tubes, 100  $\mu$ l of overnight culture of recipient strain grown in P1 LB was added and incubated at room temperature for 30 min. To each tube 100  $\mu$ l of a donor P1 lysate dilution was added and incubated at 37°C for 20 min. These cultures were then pelleted by centrifugation at 20,000  $\times g$ , with cells being resuspended in 100  $\mu$ l LB and plated onto agar plates (containing no  $\text{CaCl}_2$  or glucose) prepared with kanamycin (30  $\mu\text{g/ml}$ ) and sodium citrate (0.125 mM). Presence of sodium citrate acted to chelate  $\text{Ca}^{2+}$  thus preventing further transduction events on plating of the culture. These plates were incubated overnight at 37°C.

Control samples included cells in the absence of any P1 lysate, and also P1 lysate in the absence of cells.

Following overnight incubation, any transductant colonies that grew in the presence of kanamycin were streaked individually onto fresh LB plates containing kanamycin. Presence of kanamycin cassette in the desired gene region was confirmed by colony PCR (Section 2.6.3) and cross-streaking of culture with P1 ensured that the transduced cells could be re-infected by P1 and thus P1 was no longer present within cells. This involved streaking of a transductant colony across a sample of P1 phage stock on an LB agar plate containing glucose and  $\text{CaCl}_2$  as previously described. This was incubated at 37°C overnight, and observed culture death was confirmation of P1 phage infection.

### **2.9. Transmission electron microscopy (TEM) of *Escherichia coli***

Cell culture samples from chemostats were taken and diluted to an  $\text{OD}_{600}$  of approximately 0.5. These samples were then mixed in a 1:1 ratio with a solution of 3% glutaraldehyde (v/v) in 0.1 M cacodylate buffer and left at 4°C until visualisation.

For TEM visualisation, the fixed samples were vortexed and 10  $\mu$ l was applied to a 400-mesh copper grid (Agar Scientific, AGG2400C) coated with formvar film for 30 s before excess solution was drawn off with blotting paper. Then 10  $\mu$ l of 1% phosphotungstic acid (w/v) was applied to the same grid for a further 30 s in order to stain the sample. Once excess liquid had been removed with blotting paper the grid was washed in a sequential  $\text{dH}_2\text{O}$  series. Samples were then examined

and micrographs taken in a FEI G<sup>2</sup> Bio-twin Tecnai 120 Kv (Hillsboro, or, USA). All TEM work was carried out with assistance from Chris Hill (University of Sheffield).

### **2.10. *Vibrio harveyi* BB170 autoinducer-2 Bioassay**

This assay was used in order to determine relative levels of autoinducer-2 present in chemostat culture supernatant samples. It is based on the work by Surette and Bassler (Surette and Bassler, 1998) and was carried out as follows:

2 ml of LB medium (supplemented with kanamycin) was inoculated with *Vibrio harveyi* BB170 taken from a fresh plate (no older than 4 days). This was grown overnight at 30°C for approximately 16 h. AB medium (Section 2.2.1) was inoculated with this pre-culture in a 1:4000 ratio and vortexed. In a 96 well plate 20 µl of supernatant sample, 36 µl dH<sub>2</sub>O and 144 µl of inoculated AB medium were added per well. Three technical replicates were carried out per sample including a control sample which consisted of sterile Evans' minimal medium containing 20 mM glucose in place of a culture supernatant sample. This was then incubated at 30°C in a plate reader which was used to take both fluorescence readings (420 nm) and OD<sub>600</sub> readings every 30 min for 6 h. Average output at 420 nm of three empty wells was subtracted from all values.

In order to calculate AI-2 activity; light output of a given sample was divided by that obtained for the control sample. Thus, AI-2 activity is expressed as a fold increase of bioluminescence at 420 nm relative to the negative control.

### **2.11. <sup>1</sup>H-NMR of steady-state culture supernatant samples**

A 550 µl sample consisting of 495 µl supernatant sample (Section 2.5.4.) and 55 µl 10 mM trimethylsilyl propionate (TSP) was pipetted into glass NMR tubes. TSP (10 mM) was made up in D<sub>2</sub>O (Sigma, 151882). All spectra were acquired, with help from Dr Matt Rolfe (University of Sheffield), on a Bruker DRX-500 spectrometer operating at 500 MHz. Spectra were processed and peaks quantified by integration using Topspin. Concentrations were established by reference to TSP.

### **3. Studying the gene regulator SlyA**

#### **Main Findings**

- No significant changes in gene expression between wild type *Escherichia coli* and an isogenic *slyA* deletion strain
- Overproduction of SlyA resulted in 39 up-regulated and 5 down-regulated genes
- SlyA bound directly to nine selected promoter regions, one of which was *PslyA*, with the rest being newly discovered members of the SlyA regulon

#### **3.1. Introduction**

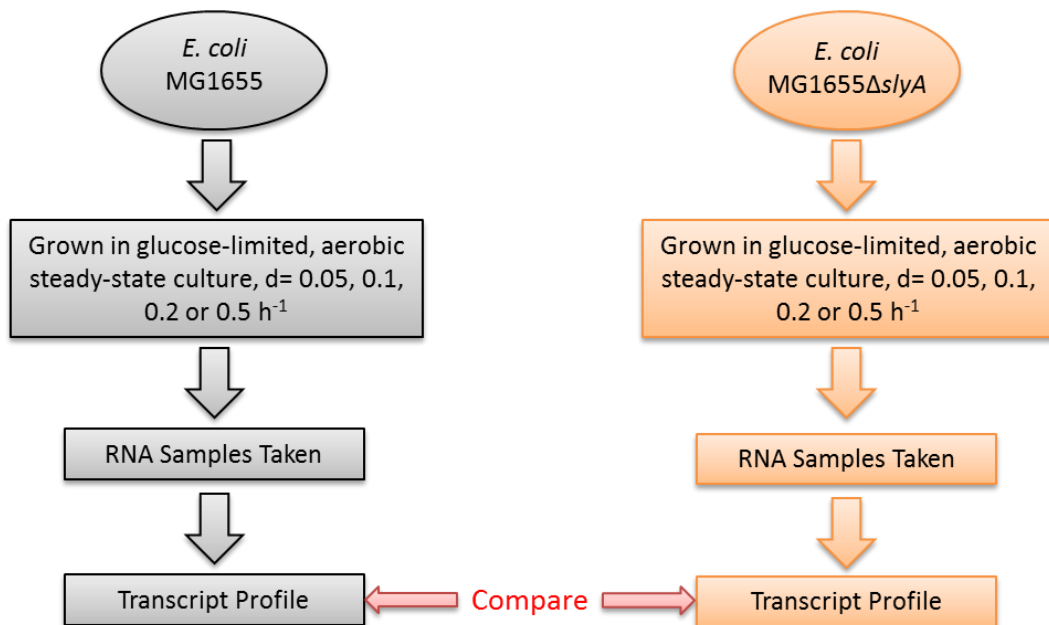
SlyA has been shown to directly regulate the expression of two genes in *E. coli* K-12; *hlyE* and *fimB* (Wyborn *et al.*, 2004; McVicker *et al.*, 2011). In both cases SlyA regulates expression by antagonising the repressive effects of H-NS. It was an aim of this project to elucidate whether there are more genes under the control of SlyA and whether the corresponding SlyA-regulated promoters were also repressed by H-NS. Also, the effect of growth rate on SlyA expression was investigated based on previous observations wherein increased levels of SlyA protein were observed when *E. coli* was grown in minimal medium compared to levels observed when grown in rich medium (McVicker *et al.*, 2011). This was attributed to the presence of the rarely used UUG start codon and the observation that the expression of poorly translated proteins tends to increase as growth rate decreases due to fewer resources being dedicated to the translation of ribosomal proteins (Liang *et al.*, 2000).

The initial work pertaining to transcriptomic comparisons between various strains is summarised schematically in Figure 3.1, and encapsulates the work carried out in Sections 3.2 and 3.3.

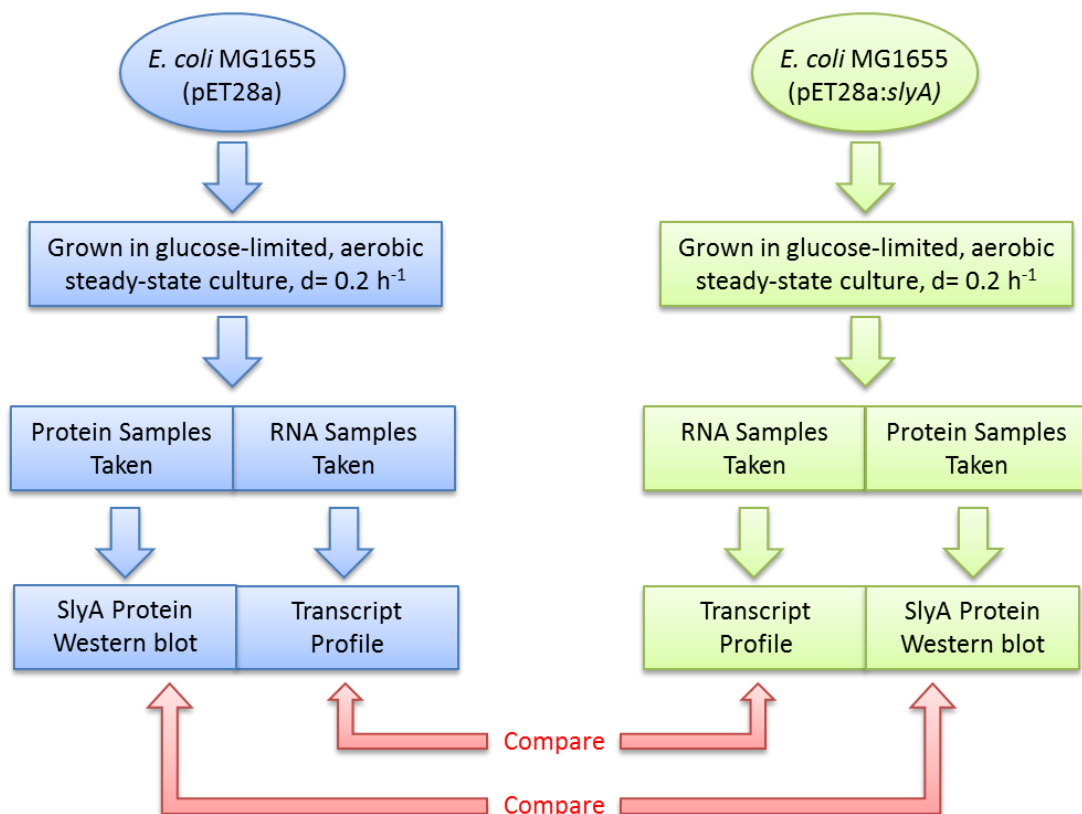
#### **3.2. Gene expression of *E. coli* MG1655 $\Delta$ *slyA* at different growth rates**

The first stage of this investigation was to create a strain of *E. coli* MG1655 that lacked the gene of interest, *slyA*. A *slyA* deletion mutant was produced using the Lambda-Red system (Datsenko and Wanner, 2000), with the entire protein coding

### Work in Section 3.2



### Work in Section 3.3

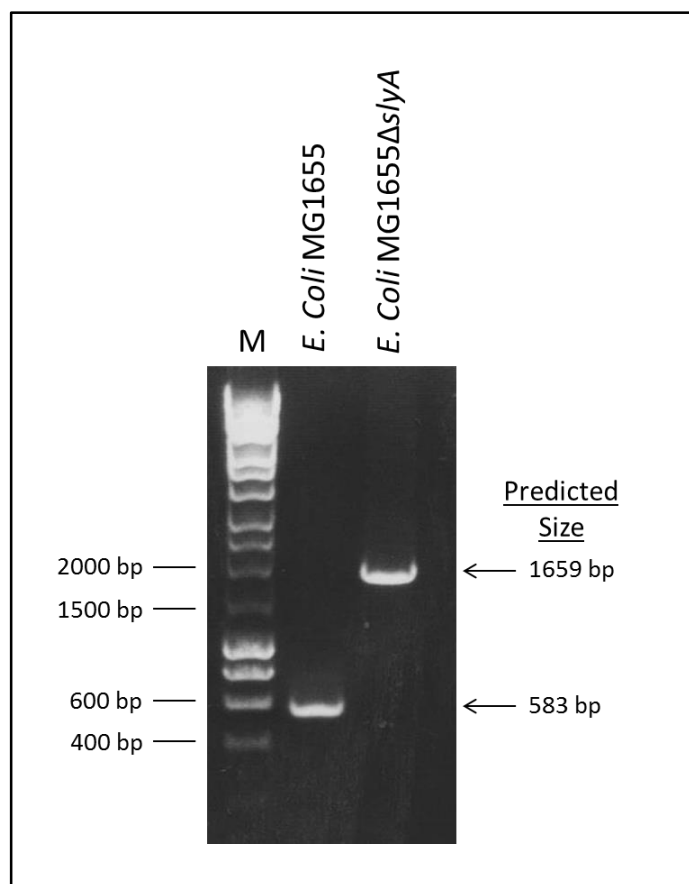


**Figure 3.1 Schematic diagram of work carried out in Chapter 5 pertaining to transcriptomics undertaken to elucidate the SlyA regulon.** Diagram specifically details the work that is relevant to Sections 3.2 and 3.3.

region of the gene replaced by a kanamycin antibiotic resistance cassette. A PCR fragment was generated, using pKD4 as a template, which contained the antibiotic resistance gene flanked by 40 bp that were homologous to the *slyA* gene region targeted for deletion. This PCR product was then purified and electroporated into *E. coli* MG1655 expressing the Lambda Red recombinase from the plasmid pKD46. The subsequent homologous recombination into the genome caused loss of the *slyA* gene and acquisition of kanamycin resistance. This resistance trait was selected for on agar plates. Transfer of the disrupted genomic region into a clean *E. coli* MG1655 background was then carried out by P1 phage transduction. Mutant verification was achieved by PCR amplification and a typical result illustrating the presence of a kanamycin resistance cassette within the *slyA* coding region is shown in Figure 3.2.

The two strains, *E. coli* MG1655 and *E. coli* MG1655 $\Delta$ *slyA* were grown individually in a chemostat vessel supplied with glucose-limited Evans' minimal medium at a dilution rate of 0.05, 0.1, 0.2 or 0.5 l h<sup>-1</sup>. Once steady-state growth had been established, samples were taken for RNA purification as described in Sections 2.5.2 and 2.6.5. Two biological replicates were carried out per *E. coli* strain (Figure 3.1).

RNA labelling and microarray hybridisations were carried out as described in Section 2.6.15, with data filtered using GeneSpring 7.3.1 to identify genes that showed statistically significant changes in expression between the two strains. A separate comparison was carried out for samples grown at each of the four dilution rates. Statistically significant changes were defined as those genes whose expression-level change passed a *t*-test ( $p \leq 0.05$ , Benjamani & Hochberg multiple testing correction) and also changes more than 2-fold. However, no genes passed this statistical filtering and, in fact, none passed the *t*-test prior to filtering for significant fold-changes. Not even the transcript level of *slyA* itself was seen to change, suggesting that its level of expression in wild type *E. coli* is very low under the conditions used here. This was the case for all the growth rates tested. These data suggested that, under the conditions tested, when the *slyA* gene was removed no significant change in gene expression occurred and there was no significant response to a change in growth rate.



**Figure 3.2 Typical result of colony PCR amplification of the *slyA* region of *E. coli* MG1655 and *E. coli* MG1655Δ*slyA*.** Colony-PCR products were mixed with 5X GelPilot Loading Dye (Qiagen) and run on a 1.5X agarose gel containing GelRed solution (Biotium) in 1X TAE buffer (Qiagen) and visualised with a UV lightbox. The same set of PCR primers (TC9 and TC10) were used for both Colony-PCR reactions, and hybridised to the genomic DNA flanking *slyA*. Change in product size is due to the presence of a 1517 bp kanamycin resistance cassette in MG1655Δ*slyA* instead of the *slyA* open reading frame (441 bp).



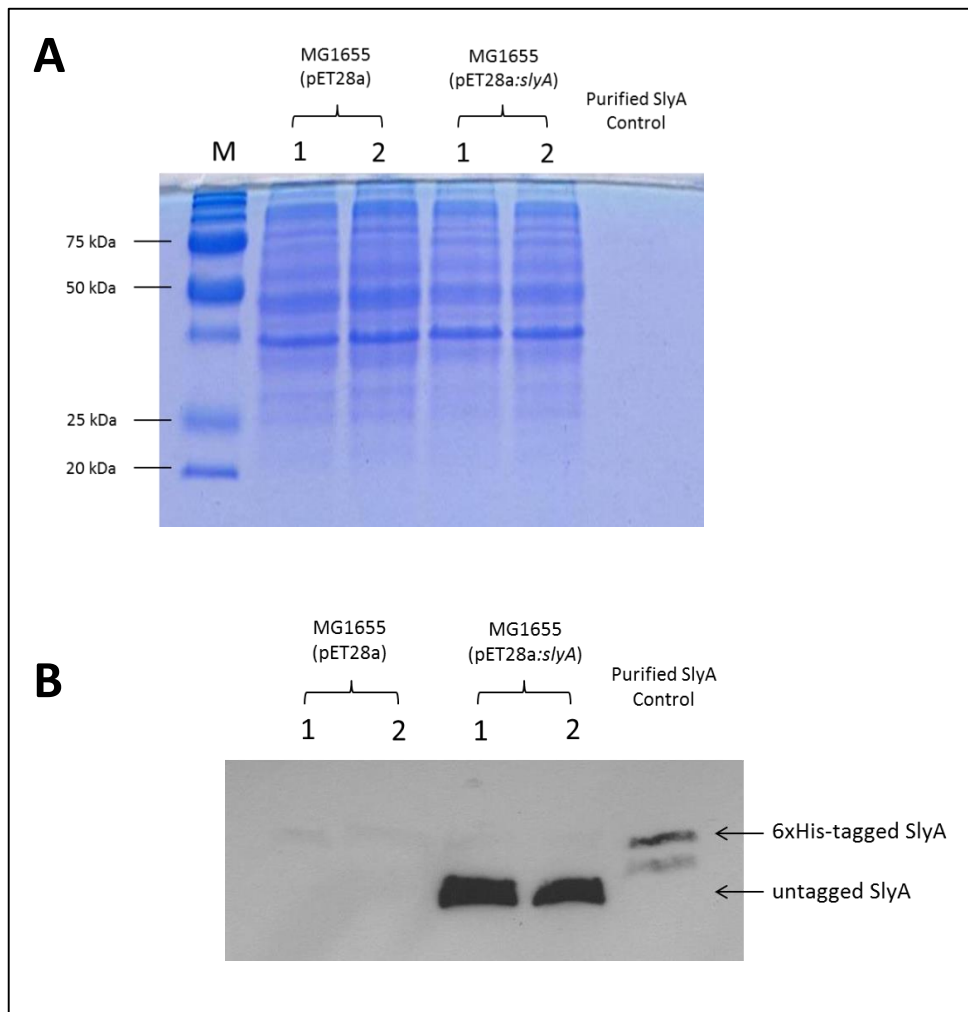
### **3.3. Gene expression of *E. coli* MG1655 overproducing SlyA**

An alternative approach to identifying genes that may be under the control of SlyA was necessary. This was because there was no significant difference in gene expression upon deletion of *slyA*, under the conditions tested, indicating that SlyA was inactive or not expressed. Therefore transcript levels of a strain producing wild type levels of SlyA and a strain overproducing SlyA were compared.

Firstly, a strain of *E. coli* that overproduces SlyA was created. Using the In-Fusion cloning system (Section 2.6.13), the *slyA* gene under the control of its own promoter was introduced into the multi-copy vector; pET28a. This vector was chosen as it harboured kanamycin resistance, vital for the subsequent culturing in a chemostat because the mechanism of kanamycin resistance does not degrade the extracellular levels of antibiotic. Thus, the selective pressure remains throughout the relatively long culturing time required for chemostat work. This engineered pET-28a:*slyA* plasmid was then used to transform *E. coli* MG1655. As a comparator strain, the pET-28a vector was used to transform *E. coli* MG1655.

These two strains were grown individually in a chemostat vessel supplied with glucose-limited Evans' minimal medium at a dilution rate of 0.2 h<sup>-1</sup>. Once steady-state growth had been established, samples were taken for RNA purification as described in Sections 2.5.2 and 2.6.5. Two biological replicates were carried out per *E. coli* strain (Figure 3.1).

A Western blot on samples of the two strains showed that *E. coli* MG1655(pET28a:*slyA*) was producing a relatively high amount of SlyA protein, despite *slyA* being under the control of its native promoter (Figure 3.3). After separation on a 17.5% SDS-PAGE gel, protein was transferred to a Hybond-C Extra membrane as described in Section 2.7.6. After protein transfer, the membrane was probed with a 1:1000 dilution of primary antibody, followed by a 1:1000 dilution of secondary anti-rabbit antibody and then visualised following the method described in Section 2.7.7. The blots showed that MG1655(pET28a:*slyA*) produced far more SlyA than the control *E. coli* strain, MG1655(pET28a) (Figure 3.3B). In fact, the control strain produced no detectable SlyA, in agreement with the determined lack of *slyA* expression in wild type *E. coli* MG1655 observed in the work carried out in Section 3.2. The amounts of SlyA detected were also consistent between



**Figure 3.3 Western blot showing amount of detectable SlyA in chemostat culture samples of *E. coli* MG1655(pET28a) and *E. coli* MG1655(pET28a:slyA).** Gel (A) shows relative amounts of total loaded protein per lane and gel (B) shows the Western blot result. Samples were separated on identical 17.5 % SDS-PAGE gels. The gel shown in (A) was stained with Coomassie Blue. Lane M represents the Precision Plus Protein™ Molecular Weight marker (BioRad) with molecular weights shown. (1) and (2) indicate independent biological replicates. Purified SlyA Control was 8.6 ng of purified SlyA, which was too little to be visualised on SDS-PAGE gel stained with Coomassie Blue. The disparity in size compared to culture samples is due to the engineered 6x His-tag used in recombinant SlyA purification.

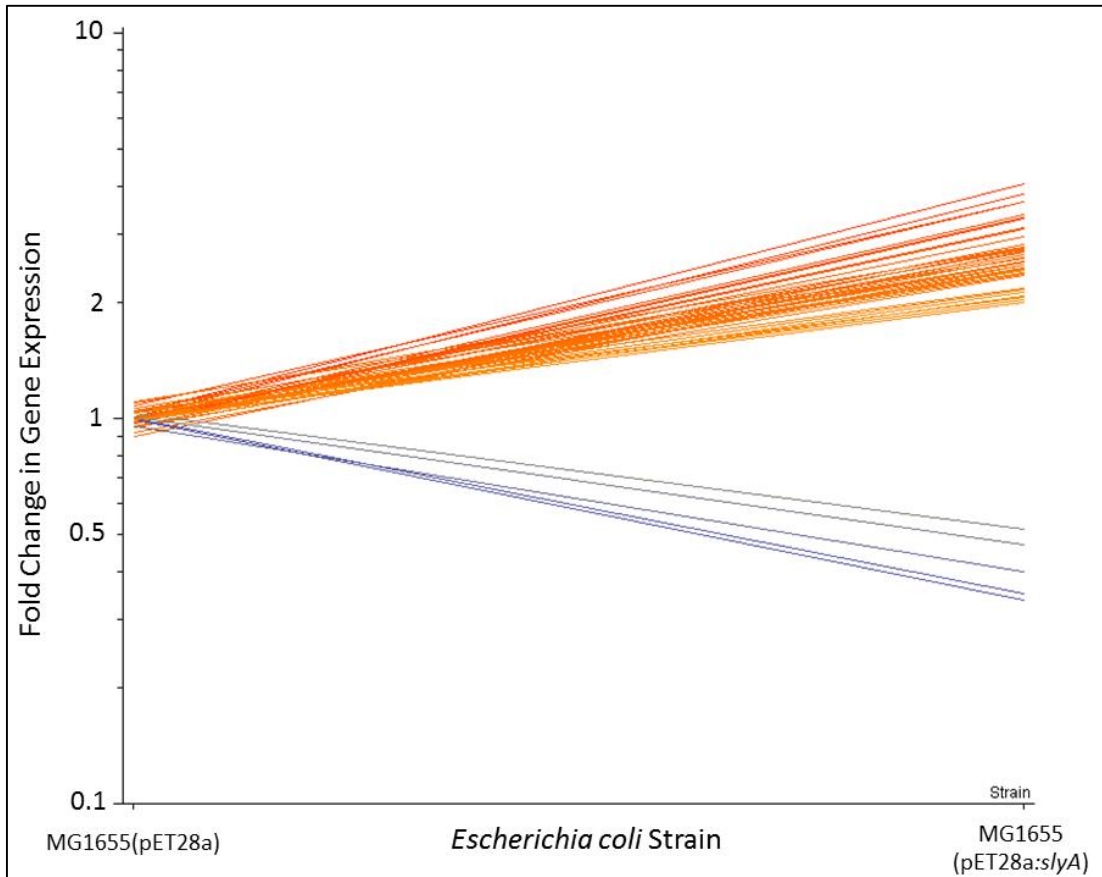
In the case of culture samples, 1 ml of chemostat culture was pelleted and resuspended in 50 µl, and mixed 1:1 with SDS Loading Buffer containing 1.2 M DTT. 8 µl of each sample was then loaded per lane.

independent biological replicates. A known quantity of purified SlyA protein was included in this Western blot (details on SlyA purification in Section 3.6) and enabled, with the use of ImageJ software (Schneider *et al.*, 2012), the estimation of the amount of SlyA protein in 1 ml of culture to be 216 ng (+/- 28 ng), equivalent to ~1.7  $\mu$ M SlyA dimer in the bacterial cytoplasm. This estimation is based on the observed dry cell weight of culture being 1.4 g l<sup>-1</sup> with a standard deviation of 0.2, and assumes the dry weight of one cell is 3x10<sup>-13</sup> g and the aqueous volume is 7x10<sup>-13</sup> ml (Sundararaj *et al.*, 2004).

RNA labelling and microarray hybridisations were carried out as described in Section 2.6.15, with data filtered using GeneSpring 7.3.1 to identify genes that showed statistically significant changes in expression between the two strains (wild type plus vector and wild type plus *slyA* overexpression plasmid). These are defined as those genes whose expression-level change passes a *t*-test ( $p \leq 0.05$ , Benjamani & Hochberg multiple testing correction) and also changed more than 2-fold. Only the genes that passed this rigorous filtering are mentioned in this chapter, unless otherwise stated. A graphical representation of all the genes that passed the statistical filtering can be seen in Figure 3.4.

The most striking observation is that the vast majority of gene expression changes were positive, i.e. the presence of SlyA has a positive effect on transcription. In fact, of the 44 transcripts (genes) changing significantly in abundance (expression); 39 were up-regulated (88.6%) and only 5 were down-regulated (11.4%). If the primary mechanism by which SlyA operates is by antagonising H-NS silencing as has been previously stated, then this widespread positive effect on expression is to be expected.

Analysis of the list of genes deemed to be SlyA-regulated (Table 3.1) revealed that: (1) the known SlyA-activated gene *hlyE* was up-regulated 2-fold; (2) a 3-fold increase in the expression of *slyA* itself was most likely a direct effect of the plasmid-based *slyA* overexpression, but mild (2-fold) positive autoregulation of *slyA* has been previously reported (Corbett *et al.*, 2007). However, taking into account the large amounts of SlyA protein observed in Western blots compared to the relatively small fold change in *slyA* transcript abundance of *E. coli* MG1655(pET28a:*slyA*) compared to *E. coli* MG1655(pET28a), it is likely that any regulation of SlyA is



**Figure 3.4 Fold changes in gene expression levels of *E. coli* MG1655(pET28a:*slyA*) compared to *E. coli* MG1655(pET28a).** Comparing the fold changes in gene expression between the *slyA* overexpression strain and a control strain carrying the vector showed that SlyA directly or indirectly regulated 44 genes. Each line in the graph represents an individual gene's expression profile between the two strains. Results have been filtered for those that show a minimum of a 2-fold change in expression level and p-value  $\leq 0.05$  (Benjamani & Hochberg multiple testing correction,  $n = 4$ ).

**Table 3.1 Fold change in transcript abundance of *E. coli* strain MG1655(pET28a:slyA) compared to expression levels in MG1655(pET28a).** Left and right tables show up-regulated and down-regulated gene lists respectively, in response to increased SlyA. Those that are listed as having an “Unknown Function” are those that are found to have no clear function when searched for on Ecocy.org (Keseler et al., 2013) or return no results when a simple literature search is performed. From information available through Ecocyc.org, evidence of confirmed H-NS regulation was highlighted. “LeuO” refers to whether a gene listed here was also found to be up/downstream of a LeuO binding site in the SELEX-chip study carried out by Shimada *et al.* (Shimada et al., 2011). This includes those genes that are found in the same operon as one identified in their study.

Genes up-regulated on over-production of SlyA					Genes down-regulated on over-production of SlyA				
Fold Change	Gene Name	Unknown Function	Regulation		Fold Change	Gene Name	Unknown Function	Regulation	
			H-NS	LeuO				H-NS	LeuO
4.06	<i>ybeT</i>	X		X	0.51	<i>yecH</i>	X		
3.83	<i>trkG</i>				0.47	<i>sgcC</i>			
3.64	<i>ssuA</i>			X	0.4	<i>fecR</i>			
3.64	<i>yehD</i>		X	X	0.35	<i>sgcQ</i>			
3.36	<i>mngA</i>			X	0.34	<i>sgcB</i>			
3.31	<i>ssuD</i>			X					
3.28	<i>casA</i>		X	X					
3.11	<i>yghS</i>	X		X					
2.95	<i>slyA</i>								
2.82	<i>yfbN</i>	X		X					
2.78	<i>casC</i>		X	X					
2.77	<i>paal</i>								
2.76	<i>ybeU</i>	X		X					
2.74	<i>elfA</i>		X	X					
2.72	<i>paaG</i>			X					
2.69	<i>paaA</i>			X					
2.66	<i>ygeG</i>	X							
2.62	<i>yjdA</i>	X							
2.6	<i>elfD</i>		X	X					
2.55	<i>casB</i>		X	X					
2.55	<i>sfmH</i>		X						
2.54	<i>agaB</i>			X					
2.54	<i>ydhV</i>								
2.51	<i>yiiE</i>	X							
2.5	<i>mdtM</i>								
2.46	<i>leuO</i>		X	X					
2.44	<i>C0299</i>								
2.42	<i>ycjN</i>			X					
2.41	<i>yadN</i>			X					
2.39	<i>gspE</i>		X	X					
2.37	<i>gspC</i>		X	X					
2.35	<i>paaB</i>			X					
2.18	<i>ydhJ</i>	X							
2.15	<i>paaC</i>			X					
2.13	<i>ydhI</i>	X							
2.07	<i>yfdM</i>	X							
2.07	<i>paaD</i>			X					
2.03	<i>ssuC</i>			X					
1.99	<i>hlyE</i>		X						

predominantly post-transcriptional; (3) eleven of the genes are known to be negatively regulated by H-NS accounting for 25% of the entire list of genes, a significant overrepresentation because H-NS regulates only approximately 5% of the entire *E. coli* genome (Hommais *et al.*, 2001). It must also be stated that many of the genes in the list are of unknown function or have not been heavily studied (11 out of 44 genes; 25%); hence there is a strong possibility that they too may be regulated by H-NS.

One of the genes up-regulated upon SlyA overproduction, *leuO*, encodes a transcriptional regulator that has also been found to operate by antagonising H-NS regulation (Stratmann *et al.*, 2008; Shimada *et al.*, 2011). Therefore, the list of genes affected by SlyA overproduction was compared to a list of genes adjacent to LeuO binding sites identified in the SELEX-chip study of Shimada *et al.* (2011). Twenty five of the 39 SlyA up-regulated genes (64%) were also present in the LeuO study. This is a very strong correlation and is likely due to two possibilities; (1) the positive effect SlyA has on the expression of LeuO results in an increase in expression of the entire LeuO regulon; or (2) SlyA and LeuO have overlapping regulons due to the fact they both operate by antagonising H-NS mediated repression.

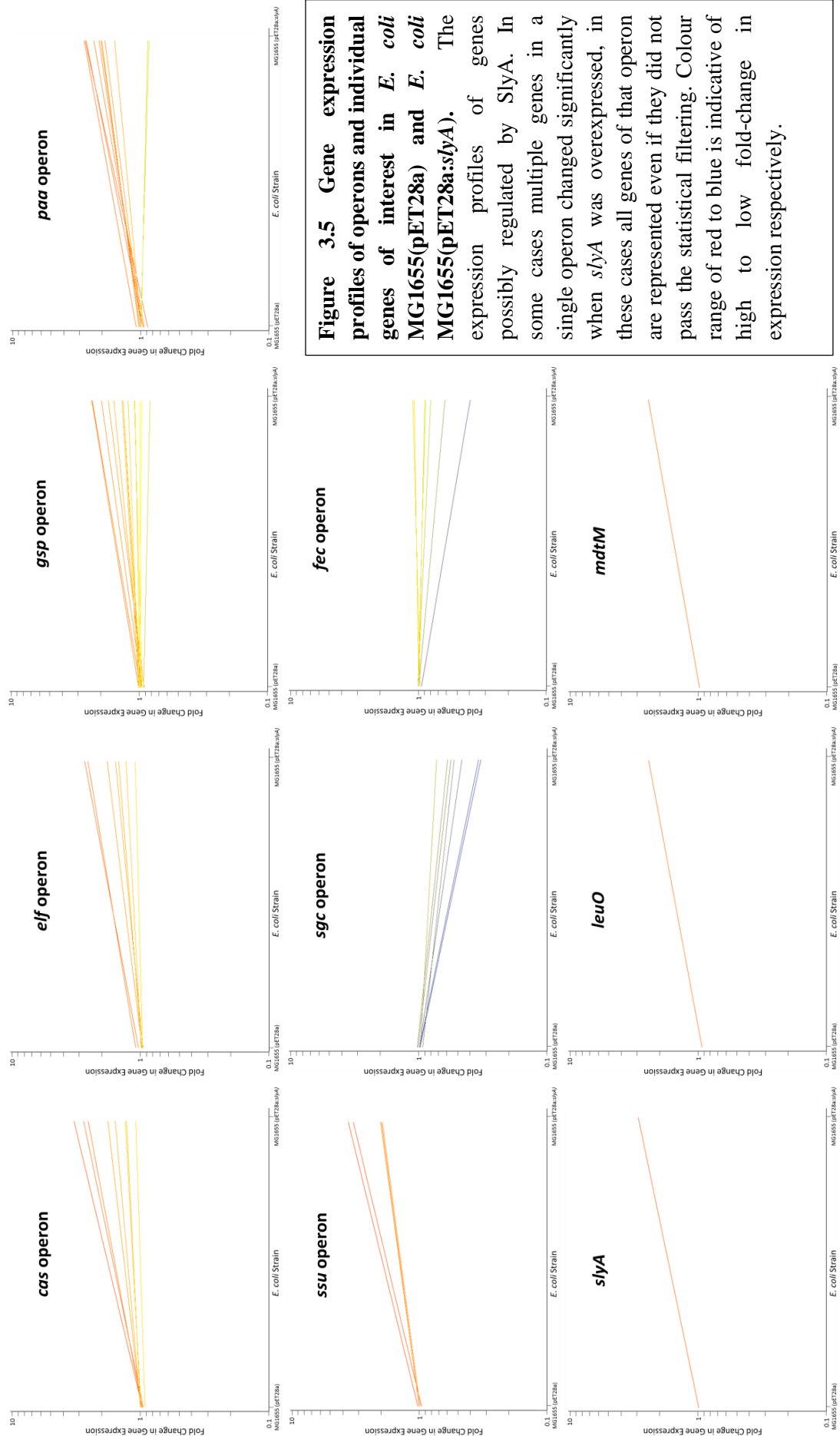
In conclusion, a strain of *E. coli* MG1655 that overexpressed *slyA* mRNA approximately 3-fold, and SlyA protein to  $\mu\text{M}$  concentrations in the cell compared to levels in the wild-type, exhibited significant changes in the expression of 44 genes. Amongst those, H-NS regulated genes were overrepresented suggesting that SlyA broadly acts by antagonising H-NS activity in *E. coli*. A significant overlap with the regulon of another regulator known to antagonise H-NS, LeuO, was also observed.

#### **3.4. Selecting targets of interest**

In several cases, more than one gene in an operon was represented in the list of putative SlyA-regulated genes (Table 3.1). The data for these operons of interest are shown in Table 3.2. The same data for both operons of interest and individual genes can be seen in graphical form in Figure 3.5. The fact that multiple genes in an operon may be affected by the overexpression of SlyA suggests that SlyA may be directly influencing promoter activity; such promoter regions were therefore good targets for subsequent Electrophoretic Mobility Shift Assays (EMSAs) to determine whether the observed SlyA-mediated regulation is direct or indirect. In almost every

**Table 3.2 Expression data for genes in operons and single genes of interest, when comparing *E. coli* MG1655(pET28a) and *E. coli* MG1655(pET28a:slvA).** Gene expression changes for each gene are displayed as a fold change in transcript level for MG1655(pET28a:slvA) compared to MG1655(pET28a). In the case of the operons shown, all genes are represented and not only those that passed the statistical filtering of the initial data analysis. Red signifies that a gene was seen to be significantly up-regulated in the initial analysis with blue signifying that a gene was seen to be significantly down-regulated.

<b>Operons of Interest</b>	<i>casABCDE12</i> and <i>cas3</i> Fold Change on o/e of SlvA	<i>casA</i> 3.28	<i>casB</i> 2.55	<i>casC</i> 2.78	<i>casD</i> 1.57	<i>casE</i> 1.80	<i>cas1</i> 1.28	<i>cas2</i> 1.31	<i>cas3</i> 1.09													
	<i>elfADCG-ycbUVF</i> Fold Change on o/e of SlvA	<i>elfA</i> 2.74	<i>elfD</i> 2.60	<i>elfC</i> 1.84	<i>elfG</i> 1.58	<i>ycbU</i> 1.32	<i>ycbV</i> 1.51	<i>ycbF</i> 1.12														
	<i>gspAB</i> and <i>gspCDEFGHIJKLMO</i> Fold Change on o/e of SlvA	<i>gspA</i> 1.24	<i>gspB</i> 1.00	<i>gspC</i> 2.37	<i>gspD</i> 1.76	<i>gspE</i> 2.39	<i>gspF</i> 2.00	<i>gspG</i> 1.60	<i>gspH</i> 1.11	<i>gspI</i> 1.37	<i>gspJ</i> 1.10	<i>gspK</i> 1.35	<i>gspL</i> 1.35	<i>gspM</i> 0.84	<i>gspO</i> 0.97							
	<i>paaABCDEFGHIJK, paaXY</i> and <i>paaZ</i> Fold Change on o/e of SlvA	<i>paaA</i> 2.69	<i>paaB</i> 2.35	<i>paaC</i> 2.15	<i>paaD</i> 2.07	<i>paaE</i> 2.03	<i>paaF</i> 1.92	<i>paaG</i> 2.73	<i>paaH</i> 1.61	<i>paaI</i> 2.77	<i>paaJ</i> 0.90	<i>paaK</i> 0.87	<i>paaL</i> 0.99	<i>paaM</i> 0.96	<i>paaN</i> 1.03							
	<i>ssuEADCB</i> Fold Change on o/e of SlvA	<i>ssuE</i> 2.00	<i>ssuA</i> 3.64	<i>ssuD</i> 3.31	<i>ssuC</i> 2.03	<i>ssuB</i> 1.97																
	<i>sgcXBCQAER</i> Fold Change on o/e of SlvA	<i>sgcX</i> 0.54	<i>sgcB</i> 0.34	<i>sgcC</i> 0.47	<i>sgcQ</i> 0.35	<i>sgcA</i> 0.58	<i>sgcE</i> 0.61	<i>sgcR</i> 0.74														
	<i>fecABCDE</i> and <i>fecLR</i> Fold Change on o/e of SlvA	<i>fecA</i> 0.90	<i>fecB</i> 1.09	<i>fecC</i> 0.81	<i>fecD</i> 1.12	<i>fecE</i> 0.89	<i>fecI</i> 0.63	<i>fecR</i> 0.40														
	<b>Single genes of Interest</b>	<i>slvA</i> 2.95	<i>leuO</i> 2.46	<i>mdtM</i> 2.50																		



**Figure 3.5 Gene expression profiles of operons and individual genes of interest in *E. coli* MG1655(pET28a) and *E. coli* MG1655(pET28a:slyA). The expression profiles of genes possibly regulated by SlyA. In some cases multiple genes in a single operon changed significantly when *slyA* was overexpressed, in these cases all genes of that operon are represented even if they did not pass the statistical filtering. Colour range of red to blue is indicative of high to low fold-change in expression respectively.**



case, perhaps with the exception of the *fecIR* operon, the genes closest to the known or predicted operon promoter region were the most significantly affected by the presence of increased SlyA with each successive gene showing a diminished response. This is most apparent in the *casABCD12* and *paaA-K* operons. It should also be noted that the operons for *ssuEADCB* and *elfADCG* are divergently transcribed and situated within 400 bp of each other and are, therefore, likely to be influenced by overlapping regulatory elements.

The functions of the SlyA-regulated operons span a range of cellular roles. One of the most intriguing, due to its topical relevance, was the *casABCD12* operon for its role in regulating the CRISPR system employed by *E. coli* (Horvath and Barrangou, 2010). The gene designation of *cas* is derived from “CRISPR-associated genes”, with CRISPR standing for Clusters of Regularly Interspersed Short Palindromic Repeats. This operon encodes a range of proteins involved in maintaining and utilising the library of foreign genetic elements interspersed between CRISPR sequences which act as the immune system memory of Bacteria and Archaea (Horvath and Barrangou, 2010). CRISPR loci, in general, consist of closely spaced direct repeats separated by short spacer regions of variable sequence. Spacer regions mostly correspond to sections of foreign plasmid or viral sequences which have been integrated. The CRISPR loci are found adjacent to the *casABCD12* operon, hence the fact that the *casABCD12* operon was significantly affected by the overexpression of SlyA suggests that this regulator may contribute to viral resistance and immunity in *E. coli*.

The *elfADCG-ycbUVF* operon was also significantly affected by the overexpression of *slyA*. This is a putative, and cryptically expressed, chaperone-usher fimbrial operon that has been shown to promote adhesion to abiotic surfaces via the production of observable surface fimbrial structures (Korea *et al.*, 2010).

Elements of the divergently transcribed *gspAB* and *gspCDEFGHIJKLMO* operons were also affected by the relative abundance of SlyA, especially the *gspC-O* operon. This operon has been shown to encode a Type-II Secretion System (T2SS) for the export of endogenous proteins (Francetic and Pugsley, 1996; Francetic *et al.*, 2000). Though *E. coli* has genes encoding for this machinery, these genes are not usually expressed in laboratory conditions, and only in the absence of H-NS have

they been seen to be transcribed (Francetic *et al.*, 2000). Once the Gsp secretion is in place it is said to aid in the export of the endogenous endochitinase ChiA, whose gene transcription is also silenced by H-NS (Francetic *et al.*, 2000).

The fourth of the operons of interest is *paaABCDEFGHIJK* (flanked by the operons *paaXY* and *paaZ*), which encodes the machinery necessary for the catabolism of phenylacetic acid (Ismail *et al.*, 2003). Aromatic organic compounds are an abundant class of environmental pollutant that are a challenge to utilise due to their very stable aromatic ring structures. The system employed by *E. coli* combines features of both aerobic and anaerobic strategies for the utilisation of aromatic compounds. Like anaerobic metabolism, phenylacetic acid is attached to Coenzyme A by PaaK to form a CoA derivative (Ferrandez *et al.*, 1998). This is subsequently catabolised by the phenylacetyl-CoA monooxygenase complex encoded by *paaA,B,C,D* and *E* (Grishin *et al.*, 2011). Subsequent isomerisation and ring opening is carried out by PaaG and PaaZ respectively, with a combination of PaaF, PaaG, PaaH and PaaJ carrying out the final  $\beta$ -oxidation steps leading to acetyl-CoA and succinyl-CoA production (Teufel *et al.*, 2010). It has been shown that PaaX acts as a transcriptional repressor of the *paa* operon, with phenylacetyl-CoA as a specific inducer that prevents PaaX binding to its DNA target (Ferrandez *et al.*, 2000). Finally PaaY has a suggested, and as yet unconfirmed, regulatory role in inactivating PaaK through acetylation (Teufel *et al.*, 2010). With the significantly positive effect that SlyA appears to have on the expression of this operon, it is highly possible that SlyA has an important role to play in the utilisation of the growth substrate, phenylacetic acid.

The *ssuEADCB* operon, the fifth set of genes that was up-regulated by SlyA, is involved in the acquisition and utilisation of alkanesulfonates as alternative sulfur sources (van der Ploeg *et al.*, 1999; Eichhorn *et al.*, 2000; Eichhorn *et al.*, 1999). The *ssuA*, *B* and *C* are components of an ABC transport system for the uptake of alkanesulfonates, with SsuE and *D* encoding an NAD(P)H-dependent flavin mononucleotide (FMN) reductase and monooxygenase respectively. SsuE provides SsuD with reduced flavin for the desulfonation of alkanesulfonates when in the presence of oxygen.

The sixth set of genes that was significantly affected was those of the *sgc* operon; *sgcXBCQAER*, which contains three out of five of all the negatively affected genes in the transcriptomic data. This is a poorly understood operon wherein *sgcA*, *B* and *C* are thought to encode a sugar transporting phosphotransferase system (PTS) with an unknown sugar specificity. This was established due to sequence similarity between the *sgcABC* encoded Enzyme II PTS subunits and those of other, better established, PTS systems (Reizer *et al.*, 1994). Evidence suggests that the operon plays a role in the transport and phosphorylation of 5-carbon sugars (Reizer *et al.*, 1997).

The final operons of interest that were investigated were the *fecIR* and *fecABCDE* operons attributed with the uptake of ferric citrate from the environment into the cell (Pressler *et al.*, 1988; Harle *et al.*, 1995). In the transcriptomics the *fecIR* operon was negatively affected by the overproduction of SlyA. The gene *fecA* encodes the outer membrane ferric citrate uptake receptor that on binding of ferric citrate can transport it across the outer membrane and also transmits a signal across the periplasm, to *fecR* in the cytoplasmic membrane. This, in turn, transmits the signal to *fecI* ( $\sigma^{19}$ ) which activates transcription of the *fecABCDE* operon wherein *fecBCDE* encode components of a cytoplasmic membrane bound ferric citrate uptake system (Braun *et al.*, 2006). It has also been shown that the TonB energy transducing system found in the cytoplasmic membrane is required for providing energy to *fecA* for the import of ferric citrate across the outer membrane (Braun, 1995).

When selecting promoter regions for further investigation some single genes were also chosen. Firstly, the promoter region for *slyA* itself in *S. enterica* Serovar Typhimurium contains multiple SlyA binding sites (Stapleton *et al.*, 2002), with three SlyA binding sites having been identified by footprinting within the *E. coli* *slyA* promoter (Corbett *et al.*, 2007), and a basic pattern search shows a binding site consensus sequence (TTAGCAAGCTAA) located 28 bp upstream of the start codon. For these reasons, *P<sub>slyA</sub>* presents itself as a good control. The gene *leuO* is a known transcription regulator of the *leuABCD* leucine biosynthesis operon (Chen *et al.*, 2005) and has also been found to operate as a global transcription regulator via antagonism of H-NS-mediated gene silencing like SlyA (Westra *et al.*, 2010; Shimada *et al.*, 2011). Lastly, *mdtM* encodes a multidrug transporter associated with resistance to ethidium bromide and chloramphenicol (Holdsworth and Law, 2012),

and has more recently been attributed to alkaline pH homeostasis (Holdsworth and Law, 2013).

### **3.5. Synthesising promoter regions to test for possible SlyA binding**

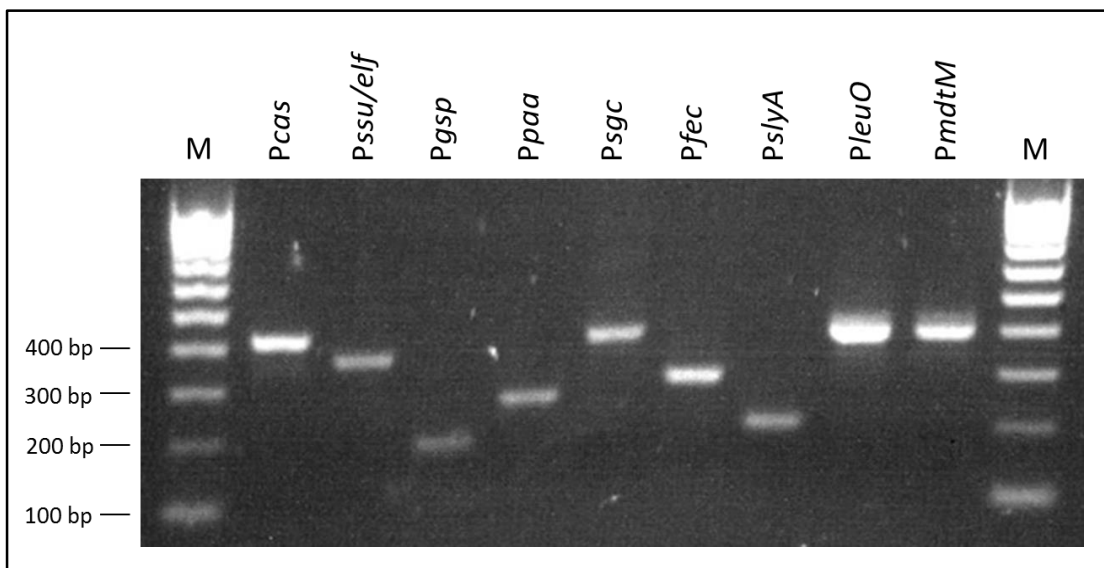
Table 3.3 shows the oligonucleotide primers used for producing the promoter regions of the genes and operons listed above and the rationale behind the choice of region. It should be noted that the *ssu* and *elf* operons share the same promoter region. If there was significantly more than 400 bp between the operon/gene of interest and the next downstream gene, then 400 bp was deemed sufficient. The regions listed were synthesised by PCR and purified by Gel Extraction (following the protocols outlined in Section 2.6.2 and 2.6.10 respectively). In all cases, the forward primer was modified with a 5'-Biotin label to be used for visualisation in subsequent assays. Figure 3.6 shows the successful synthesis of all nine of the selected SlyA target promoter regions.

### **3.6. Overexpression and purification of SlyA**

In order to investigate the affinity of SlyA for the synthesised promoter regions purification of SlyA was necessary. The DNA coding for SlyA was synthesised from a genomic DNA template by PCR using primers TC47 and TC48. These primers amplified a DNA product consisting of *slyA* with overhangs that allowed incorporation into the pLATE51 protein expression vector following the protocol outlined in Section 2.6.14. The resulting expression plasmid (pGS2469) encodes an N-terminally 6x His-tagged SlyA protein under the control of T7 RNA polymerase. The *E. coli* strain BL21(DE3) was transformed with pGS2469 and grown in autoinduction medium. The use of autoinduction medium is based on the work by Studier (2005). Autoinduction medium contains a limited amount of glucose for initial culture growth and then an excess of lactose which is utilised after glucose depletion. Once metabolising lactose, cells produce allolactose which in turn relieves repression of the *lacUV5* promoter associated with T7 RNA polymerase expression in *E. coli* BL21(DE3). T7 RNA polymerase then activates expression of *slyA* in pGS2469. This yielded improved SlyA overproduction compared to the standard method of growth in LB with subsequent induction by the addition of Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), a molecular mimic of allolactose (not shown).

**Table 3.3 Rationale for production of promoter regions of interest and primers used.** Table lists the promoter regions produced for EMSA experiments. The total size of the produced intergenic region is listed centrally; with flanking gene regions listed either side. “Description” provides details on why this region was chosen, and highlights where a total base pair size cut-off of 400 bp was used. In all cases the Forward Primer was 5’-Biotin labelled to allow visualisation in EMSA assays.

<u>Region Name</u>	<u>Forward Primer</u>	<u>Left Gene</u>	<u>Region Size</u>	<u>Right Gene</u>	<u>Reverse Primer</u>	<u>Description</u>
<i>Pcas</i>	TC53	<i>casABCDEI2</i> <	414 bp	< <i>cas3</i>	TC54	Entire intergenic region between both <i>cas</i> operon regions
<i>Pssu/elf</i>	TC49	<i>ssuEADCB</i> <	355 bp	> <i>elfADCG-ycbUVF</i>	TC50	Entire intergenic region between both separate operons of interest
<i>Pgsp</i>	TC59	<i>gspAB</i> <	179 bp	> <i>gspCDEFGHIJKLMNO</i>	TC60	Entire intergenic region between both <i>gsp</i> operon regions
<i>Ppaa</i>	TC55	<i>paaZ</i> <	257 bp	> <i>paaABCDEFGHIJK, paaXY</i>	TC56	Entire intergenic region between <i>paaZ</i> and the <i>paaABCDEFGHIJK</i> operon
<i>Psgc</i>	TC70	<i>sgcXBCQAER</i> <	400 bp	> <i>insA-7</i>	TC71	400 bp region upstream of the <i>sgcXBCQAER</i> operon (768 bp to nearest gene)
<i>Pfec</i>	TC63	<i>fecABCDE, fecIR</i> <	291 bp	> <i>insA-7</i>	TC64	Entire intergenic region between <i>fecIR</i> and <i>insA-7</i>
<i>PslyA</i>	TC65	<i>slyA</i> <	200 bp	> <i>ydhl</i>	TC66	Entire intergenic region between <i>slyA</i> and <i>ydhl</i>
<i>PleuO</i>	TC51	<i>leuLABCD</i> <	344 bp	> <i>leuO</i>	TC52	Entire intergenic region between <i>leuO</i> and <i>leuLABCD</i>
<i>PmdtM</i>	TC57	<i>mdtM</i> <	400 bp		TC58	400 bp region upstream of <i>mdtM</i> (479 bp to nearest gene)



**Figure 3.6 Products of PCR amplification of promoter regions possibly bound by SlyA.** Table 3.3 lists the primers used to amplify each promoter region of interest and the expected size of PCR product (in base pairs (bp)). PCR products were mixed with 5X GelPilot Loading Dye (Qiagen) and run on a 1.5X agarose gel containing GelRed solution (Biotium) in 1X TAE buffer (Qiagen), the band corresponding to expected product size was gel extracted using a QIAquick Gel Extraction kit (Qiagen). The gel shown here is of PCR products after gel extraction. Lane M contains Hyperladder IV (Bioline) with fragment size shown. DNA fragments were visualised with a UV lightbox.

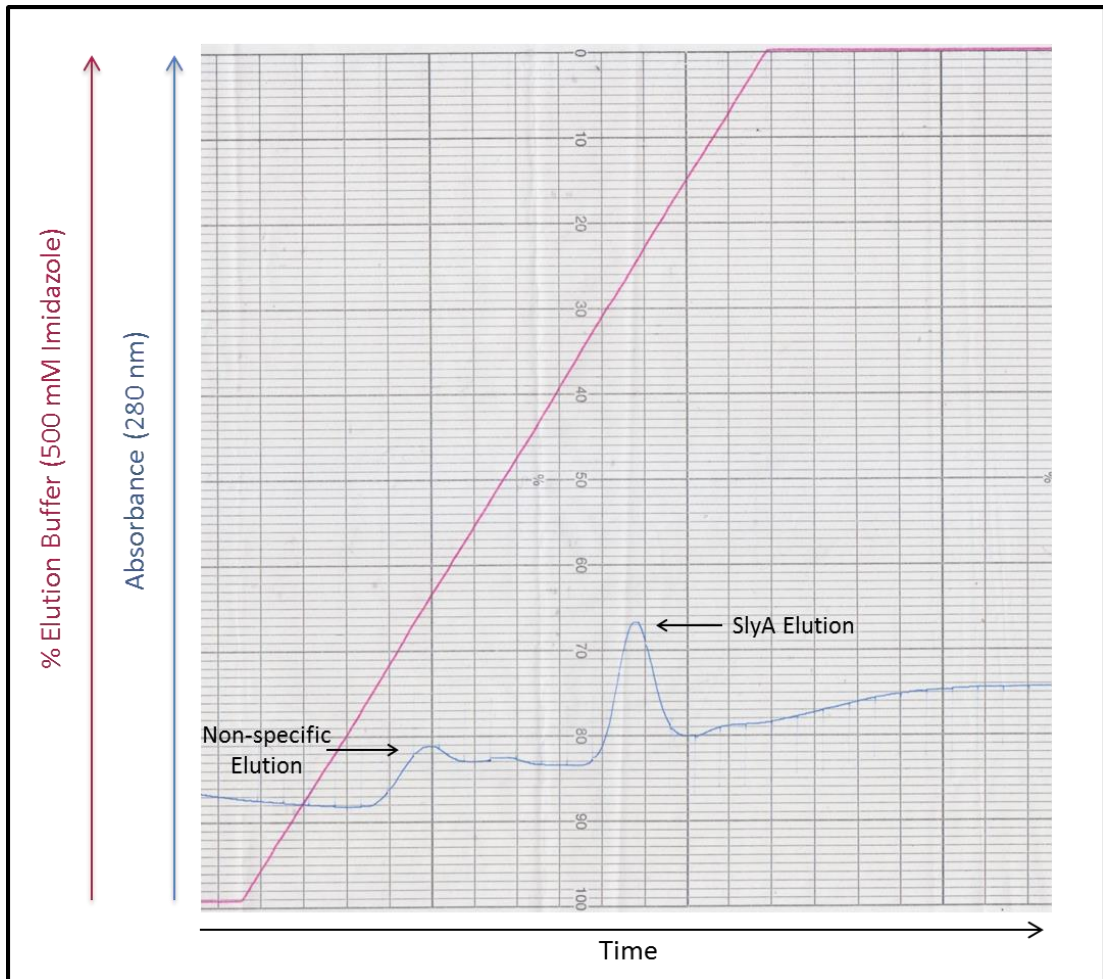
After growth of *E. coli* BL21(DE3)(pGS2469) culture in autoinduction medium for 24 h, cell-free extracts were made and the SlyA protein was purified by affinity chromatography as detailed in Sections 2.7.5. A typical AKTA trace associated with SlyA purification is shown in Figure 3.7, wherein the His-tagged SlyA was eluted at approximately 70% Elution buffer. Multiple fractions associated with the range in which SlyA was eluted were separated on 20% SDS-PAGE gels and stained with Coomassie Blue. A typical gel can be seen in Figure 3.8, where (A) and (B) show separation of fractions without and with addition of DTT respectively. Addition of excess DTT reducing agent inhibits the formation of SlyA dimers, and resolves the visualised proteins into a single species. Once the purified protein had been dialysed into a buffer containing 20 mM Tris-HCl, pH 7.4 and 200 mM NaCl, it was ready for use in subsequent assays.

### **3.7. Electrophoretic mobility shift assays of SlyA with suspected target promoter regions**

To test whether SlyA had any affinity for the promoter regions associated with some of the more interesting transcriptional changes Electrophoretic Mobility Shift Assays (EMSAs) were carried out. Using this approach, if the Biotin labelled double-stranded DNA target was bound by purified SlyA a shift in mobility could be detected upon native gel electrophoresis.

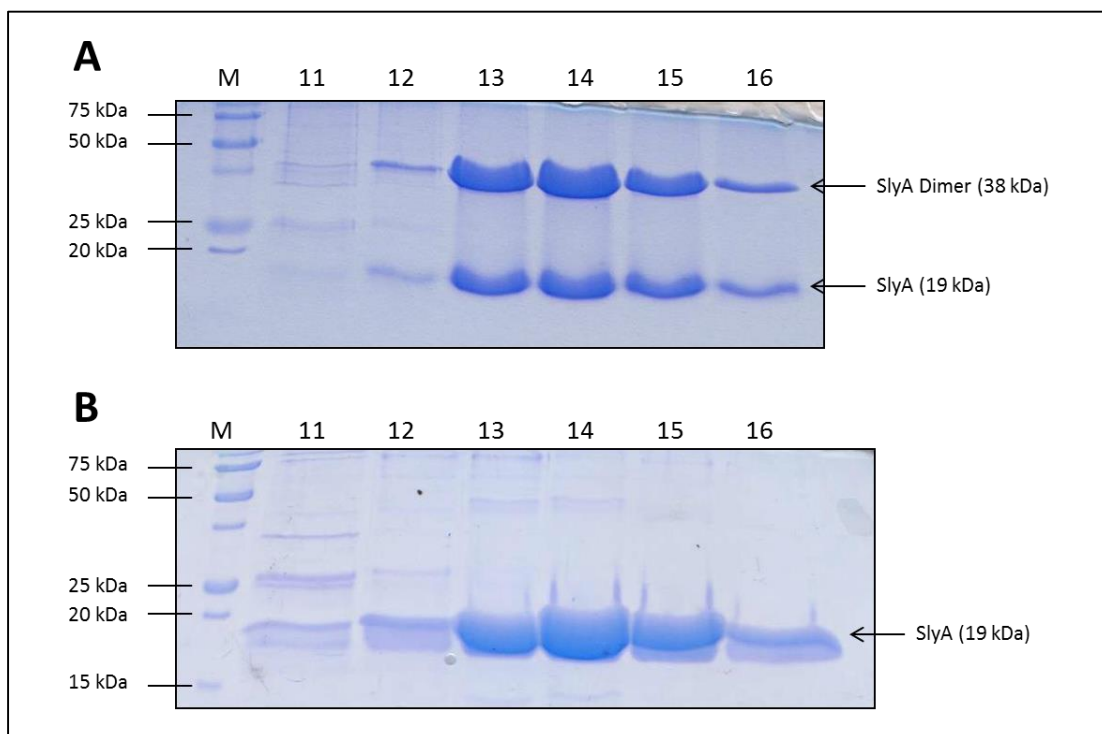
EMSAs were carried out as described in Section 2.7.8 and the results can be seen in Figure 3.9. In all cases, a range of SlyA concentrations comprising 0, 1, 5, 10, 50, 100, 200 and 500 nM was used. This protein concentration range was tested against a femtomolar amount of the promoter region in question, and in all cases SlyA binding was seen. In all cases, other than *PsgcXBCQAER*, a complete shift of the labelled DNA was seen at 100-200 nM SlyA, with *PsgcXBCQAER* exhibiting a partial shift at 500 nM SlyA.

Of note, it has been shown that both *E. coli* and *S. enterica* Serovar Typhimurium *slyA* promoters have multiple binding sites (Stapleton *et al.*, 2002; Corbett *et al.*, 2007); in the gel shifts presented here, at 50 nM SlyA, the *E. coli slyA* promoter region (*PslyA*) showed three retarded species; consistent with the three sites identified by Corbett *et al.* (Corbett *et al.*, 2007). The shifts attributed to the *Pgsp* target showed evidence of two discernible sites. The other targets might also

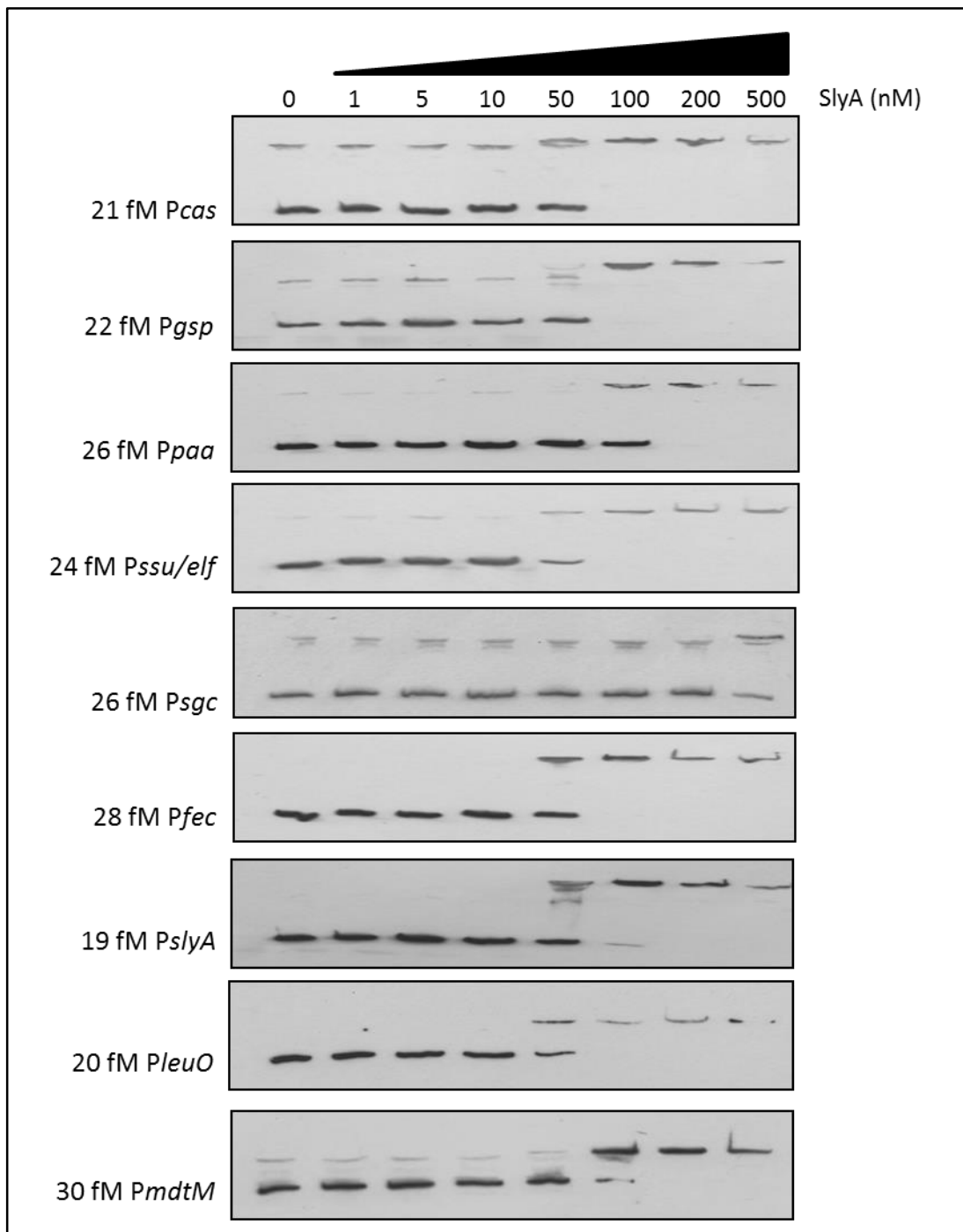


**Figure 3.7 Typical AKTA trace recording the elution of SlyA from a Hi-Trap chelating affinity column.** Blue trace represents absorbance at 280 nm as protein is eluted. Second peak (Fractions 13-15) represents His-tagged SlyA. The percentage of Elution buffer (20 - 500 mM imidazole) was increased as a linear gradient represented by the purple trace. Fractions (1 ml) were collected every minute.





**Figure 3.8 Coomassie Blue stained SDS-PAGE gel of fractions from SlyA purification.** Aliquots (10  $\mu$ l) of the indicated fractions from the elution profile in Figure 5.7 were mixed with 10  $\mu$ l of Loading Buffer and separated on a 20% SDS-PAGE gel and then stained with Coomassie Blue. Lane M represents the Precision Plus Protein<sup>TM</sup> Molecular Weight marker (BioRad) with molecular weights shown. Lanes 11 – 16 signify elution fractions 11 – 16. ‘A’ and ‘B’ show the same samples run without and with 1.2 M DTT present in loading buffer respectively.



**Figure 3.9 EMSA gels showing binding of SlyA to various gene or operon promoter regions.** Binding reactions and subsequent visualisations were carried out as described in Section 2.7.8 with samples separated on a 6% Native-PAGE gel.

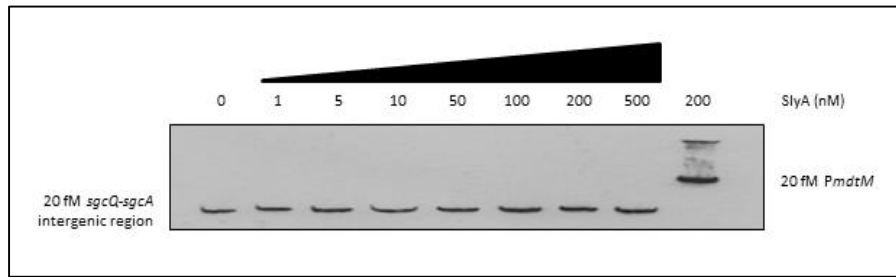
In all cases a range of 0 to 500 nM SlyA was tested against a femtomolar amount of target DNA (precise amounts of each target are shown). On the left are those promoter regions associated with operons, whilst on the right are those that are associated with individual genes of interest.

possess multiple binding sites, but those were not evident in the titration range used here.

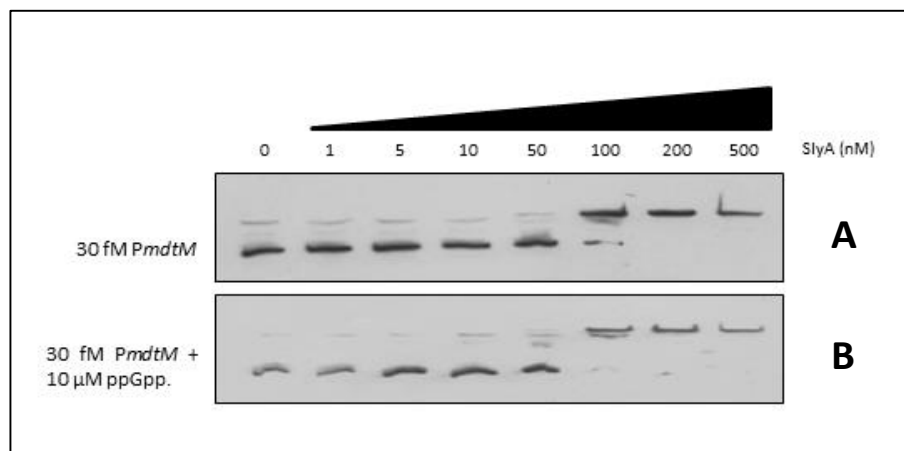
In order to determine whether SlyA may be binding non-specifically a similar SlyA titration carried out against the 130 bp intergenic length of DNA between *sgcQ* and *sgcA*. This 130 bp fragment of DNA was produced by PCR as before, using the primers TC61 and TC62. The results of this control can be seen in Figure 3.10, wherein no binding between SlyA and the DNA was observed over the entire range of protein concentration. A positive control consisting of 200 nM SlyA and 20 fM *PmdtM* was included to show that the protein sample used in this reaction was active. From these data it is suggested, with some confidence, that SlyA does indeed bind to the promoter regions associated with the operons and genes of interest identified by the transcriptomic analyses of a SlyA overproduction strain, and is therefore likely to be a direct regulator of these genes.

It has been suggested that the transcription regulation activity of SlyA in *S. enterica* Serovar Typhimurium is modulated by the binding of ppGpp (Zhao *et al.*, 2008), an alarmone signalling molecule heavily involved in the stringent response that is also found in *E. coli* (Potrykus and Cashel, 2008). Because of this, the presence of ppGpp on the binding affinity of SlyA to one of the promoter DNA targets was tested. A titration EMSA with increasing SlyA concentration against 30 fM *PmdtM*, in the presence of an excess of ppGpp (10  $\mu$ M), can be seen in Figure 3.11. The addition of ppGpp had little to no effect on the binding affinity of SlyA, with a complete shift still occurring at approximately 100 nM SlyA, this is in agreement with previous work (McVicker *et al.*, 2011). This may suggest that either ppGpp has no effect on the activity of *E. coli* SlyA or that ppGpp may modulate SlyA to bind targets in addition to those that it binds without supplementation.

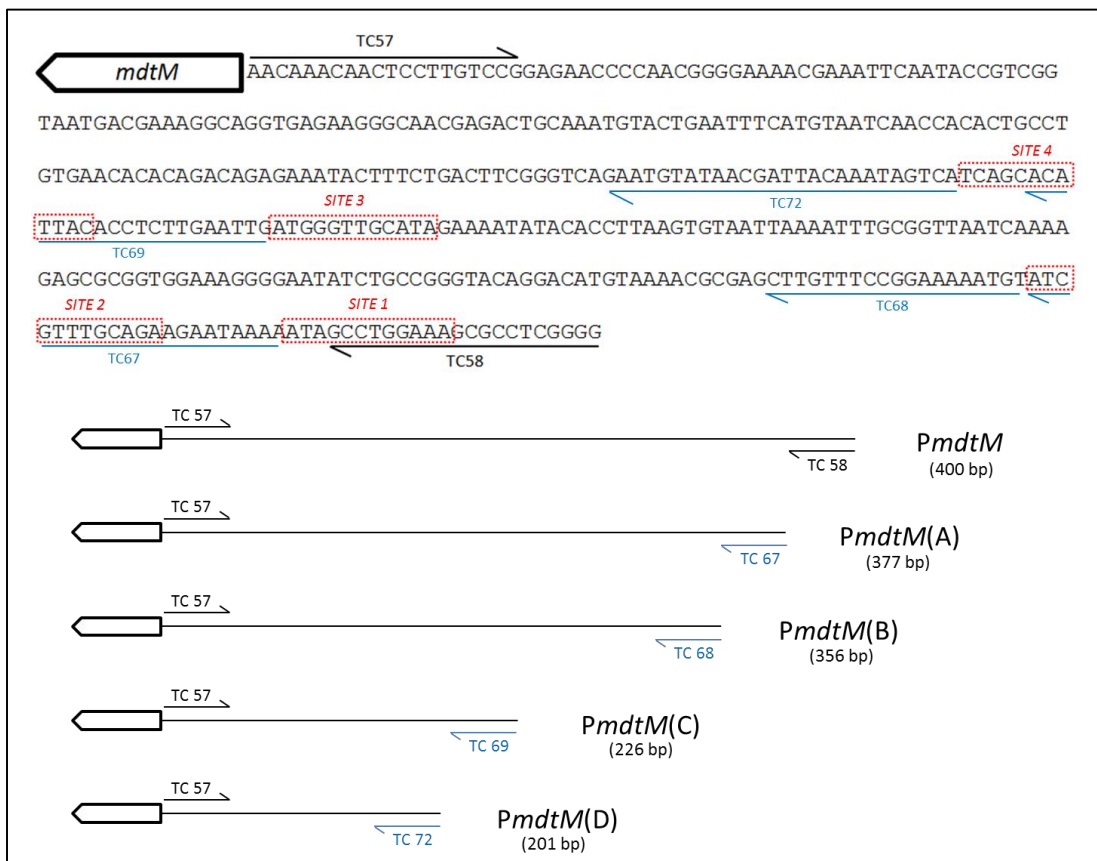
In order to further analyse the specificity of SlyA binding to the newly discovered SlyA regulon members, *PmdtM* was selected for analysis. The SlyA binding consensus sequence of *S. enterica* Serovar Typhimurium was used to search for likely recognition sites. This established 12 bp consensus sequence, TTAGCAAGCTAA (Stapleton *et al.*, 2002), is a near-perfect inverted repeat and four sites within *PmdtM* that show some similarity are summarised in Figure 3.12. The suggested sites 1, 2, 3 and 4 show a base pair match of 7 out of 12, 5 out of 12, 5



**Figure 3.10 EMSA illustrating lack of SlyA binding to *sgcQ-sgcA* intergenic region.** The 130 bp intergenic region between *sgcQ* and *sgcA* (produced using primers TC61 and TC62) showed no shift over the entire range of SlyA concentrations used in previous EMSAs. SlyA (200 nM) was enough to cause a shift of 20 fM of the *PmdtM* region on the same gel.



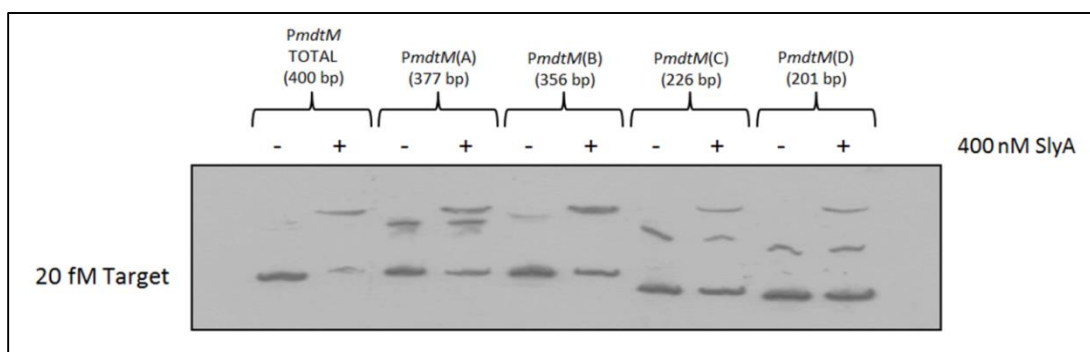
**Figure 3.11 The presence of the alarmone ppGpp does not alter SlyA binding at the *mdtM* promoter.** EMSAs were carried out as described in Section 2.7.8 with samples separated on a 6% Native-PAGE gel. For (B) 10  $\mu$ M ppGpp was included in the binding reaction mixture and incubated with SlyA for 30 min at room temperature, followed by addition of *PmdtM* (30 fM) and a further 20 min incubation at room temperature. Sample separation on Native-PAGE gel and visualisation were carried out as normal (Section 2.7.8).



**Figure 3.12 Schematic for production of *mdtM* promoter DNA, searching for possible SlyA binding sites and subsequent production of truncated promoter regions lacking these sites.** Primers TC57 and TC58 were used to produce the 400 bp fragment that preceded the *mdtM* start codon; this was termed the *mdtM* promoter region (*PmdtM*). Red-dashed boxes represent possible SlyA binding sites based on their similarity to the consensus sequence elucidated in *S. enterica* Serovar Typhimurium (TTAGCAAGCTAA) (Stapleton et al., 2002). Primers TC67, 68, 69 and 72 represent those used to produce increasingly truncated promoter regions lacking these suspected binding sites. These truncated promoter regions were named *PmdtM*(A), *PmdtM*(B), *PmdtM*(C) and *PmdtM*(D) and were 377 bp, 356 bp, 226 bp and 201 bp in length, respectively.

out of 12 and 7 out of 12 respectively when compared with the consensus sequence. This is not too dissimilar to the number and homology of sites found in the *PslyA* promoter of *S. enterica* Serovar Typhimurium via DNase I footprinting, wherein the sequence of weaker *PslyA* II, III, IV and V sites show a sequence similarity of 7 out of 12, 8 out of 12, 9 out of 12 and 6 out of 12 respectively, compared to the *PslyA* I site (TTAGCAAGCTAA). Considering this, there is a strong likelihood that *most* SlyA regulated promoter regions will have multiple binding sites with flexible degrees of consensus similarity. Figure 3.12 illustrates how four increasingly truncated *PmdtM* regions were synthesised, each lacking one more suspected binding site than the last. Each one of these truncated *PmdtM* targets was applied to an EMSA as before, alongside the total *PmdtM* region. As can be seen in Figure 3.13, in the presence of 400 nM SlyA all of the targets were shifted despite removal of putative binding sites. This suggests the presence of additional binding sites, and it has to be taken into account that the consensus sequence being used for locating possible sites is that of a homologue to the protein being investigated, albeit with 89% sequence homology (Oscarsson *et al.*, 1996). The consensus sequence being used was also ultimately born of a combination of results from a SELEX strategy and binding sites found through DNase I footprinting, taken in isolation these strategies showed variations in their elucidated consensus sequence (t/gTg/aGCAAGCTAA and TTAGCAAg/tCa/tAA respectively), this leaves scope for variations on the SlyA binding site consensus sequence.

The truncated promoter regions synthesised in order to search for possible SlyA binding sites were then utilised to test whether SlyA binding affinity was compromised on loss of a particular region. This was carried out by EMSA with the addition of heparin. Heparin, being a negatively charged, sugar-containing molecule with a helical structure is a DNA molecular mimetic and is bound strongly by DNA binding proteins. In this assay 40 fM of DNA target was incubated with 400 nM SlyA with and without the addition of 0.1 mg/ml heparin. Heparin was added to the binding buffer and incubated alongside SlyA and the DNA target at room temperature for 30 min, before separation on a Native-PAGE gel and subsequent visualisation as described in Section 2.7.8. This assay was carried out on the total *PmdtM* region, *PmdtM*(A), *PmdtM*(B) and *PmdtM*(C). The result showed that *PmdtM*,



**Figure 3.13 EMSA of *PmdtM* and its truncated variants with successive removal of four suspected SlyA binding sites.** In all cases 20 fM of DNA target was used and is shown here with and without the presence of 400 nM SlyA. DNA target and SlyA were incubated together at room temperature for 30 min, as detailed in Section 2.7.8. All targets tested showed shifts in the presence of SlyA, even when all four suspected SlyA binding sites were removed. These suspected sites are shown in Figure 3.12.

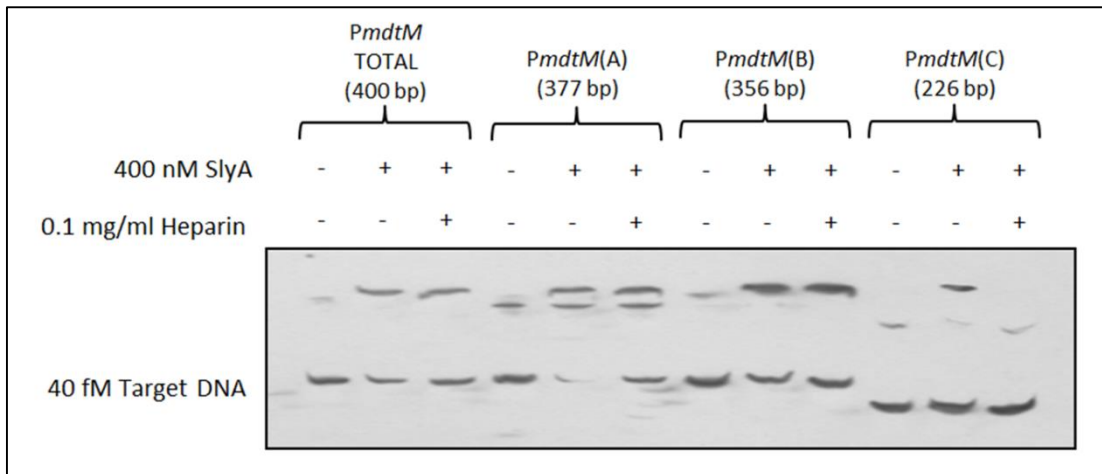
*PmdtM(A)* and *PmdtM(B)* were still bound by SlyA in the presence of heparin (Figure 3.14). However, the affinity of SlyA for *PmdtM(C)* was no longer sufficient to compete with the presence of heparin. These data demonstrate that SlyA likely has low specificity for the target motif to which it binds but it exhibits increased affinity to particular regions, in this case any regions of increased SlyA affinity were lost on truncation of *PmdtM(B)* to *PmdtM(C)*.

### **3.8. Re-analysis of *PmdtM*, searching for the ATT<6nt>TAA motif and subsequent inference of likely SlyA binding sites in other promoter regions**

Taking into account the data from the EMSAs carried out in Section 3.7, little correlation between binding affinity of SlyA to promoter region truncations and the location of regions with the most homology to the TTAGCAAGCTAA SlyA binding motif was observed. Thus, the observation by Haider *et al.* was applied to a new search for possible binding sites, wherein no single base pair in the SlyA palindrome is required for binding but the most important base pairs for DNA recognition by SlyA are situated at the 5' and 3' regions of the motif. It was noted that the motif of TTA<6nt>TAA is highly conserved between already elucidated SlyA binding sites, with little conservation in any of the six central nucleotides (Haider *et al.*, 2008). Applying a TTA<6nt>TAA motif search to the *PmdtM* region, with a minimum match of 4 out of the 6 bases, possible sites were observed to be evenly distributed throughout the DNA fragment. However, upon application of a rule that at least one of the half sites must be represented in its entirety, the distribution of possible SlyA binding sites changes dramatically. This site distribution is illustrated in Figure 3.15 and goes some way towards explaining why, on truncation of *PmdtM* down to *PmdtM(C)*, affinity of SlyA for the remaining promoter DNA is diminished; as all binding sites with at least one complete half-site are lost or disrupted by this truncation. By including additional information from the TTAGCAAGCTAA consensus, the best *PmdtM* SlyA sites according to these criteria are GCAGAAGAATAA and TATACACCTTAA (sites A and D respectively, Figure 3.15).

This new motif search was then applied to all the other promoter regions tested for direct binding by SlyA in Section 3.5, with suggestions for SlyA binding





**Figure 3.14 Electrophoretic Mobility Shift Assay (EMSA) of *PmdtM* and its truncated derivatives, with and without the presence of 0.1 mg/ml Heparin.** EMSA binding reactions were set up as described in Section 2.7.8, in this particular case 0.1 mg/ml Heparin was also present in binding buffer in which SlyA (400 nM) and DNA target was incubated for 30 min at room temperature. Samples were then separated on a 6% Native-PAGE gel and visualised as normal. Target produced by PCR with TC57 and TC69 (*PmdtM*(C)) experienced no shift in the presence of Heparin.



**Figure 3.15 Searching for SlyA binding site motif TTA<6nt>TAA within *PmdtM*.** Following determination of strong SlyA binding being lost on truncation of *PmdtM* down to *PmdtM*(C), as shown in Figure 3.14 (portion that *PmdtM*(C) lacks is highlighted here in yellow), the total *PmdtM* sequence was probed for sites homologous to the TTA<6nt>TAA SlyA binding motif as determined by Haider et al. (Haider et al., 2008). Sites A, B, C and D shown here are those that: (1) had a sequence similarity of at least 4 out of the 6 bases said to be most important for SlyA binding; and (2) had at least one complete half-site, i.e. TTA<9nt>TAA. Individual site sequences are listed at the bottom, with bold bases highlighting those that are homologous to TTA<6nt>TAA and underlined bases highlighting those that have additional homology to TTAGCAAGCTAA motif elucidated by Stapleton et al. (Stapleton et al., 2002).

sites summarised in Table 3.4. All the promoter regions listed had at least one site with a minimum of 5 out of the 6 consensus sequence nucleotides present, *PmdtM* is in fact the only promoter region studied that has a TTA<6nt>TAA homology of only 4 out of 6. Regions *Pcas*, *Pgsp*, *PslyA* and *PleuO* each contained at least one possible binding site with complete consensus homology. Given the data produced in studying the promoter region for *mdtM*, it is suggested that the SlyA binding site consensus motif of TTA<6nt>TAA be prioritised in searching for new SlyA regulated regions in *E. coli* and, if a cut-off as low as 4 out of 6 bases is used, that the presence of at least one complete three-nucleotide half-site is required. This leads to the hypothesis that the presence of at least one half-site is necessary for SlyA binding and is sufficient for binding in the presence of excess heparin.

### **3.9. Discussion**

In this chapter it was demonstrated through microarray analysis of a wild type *E. coli* MG1655 strain compared to an *E. coli* MG1655 $\Delta$ *slyA* mutant strain grown at different growth rates that *slyA* expression and its subsequent transcription regulatory effect was not significantly controlled by growth rate.

It has also been demonstrated that the overexpression of *slyA* caused the transcription of 44 genes in *E. coli* to change significantly when analysed via a microarray transcriptomic method. Of these genes, 89% were positively affected by the presence of SlyA, in agreement with the hypothesis that SlyA, like its *S. enterica* Serovar Typhimurium homologue, is an antagonist of H-NS mediated repression of transcription. Here we have seen that 25% of the total gene list produced is already known to be repressed by H-NS.

In *S. enterica* Serovar Typhimurium the majority of genes whose transcription is affected by SlyA code for proteins that are associated with the bacterial cell envelope and are important for virulence and survival within murine macrophages. Though it has been previously shown that the majority of genes observed to be regulated by SlyA in *S. enterica* Serovar Typhimurium are not present in *E. coli* K-12 (Spory *et al.*, 2002; Navarre *et al.*, 2005), a similar propensity for cell envelope proteins being regulated by the *E. coli* homologue of SlyA has been demonstrated here. In fact, considering only those genes with a known or predicted function, 42% (14 out of 33) encode proteins thought to reside in either the inner

**Table 3.4 Possible SlyA binding sites within promoter regions shown to be directly bound by SlyA, based on the binding site motif: TTA<6nt>TAA.** Displayed are those sites within each promoter that had the greatest degree of homology to the SlyA binding site consensus TTA<6nt>TAA (Haider et al., 2008). In cases where more than one site in a promoter region had the same degree of homology, sites which had greatest overall homology to the partially palindromic sequence TTAGCAAGCTAA (Stapleton et al., 2002) were prioritised. Only those sites which had the greatest overall homology are shown, though the total numbers of sites with the same level of TTA<6n>TAA similarity per promoter region are listed in column B. If a site had all six of the important bases of TTA<6n>TAA present, it is present regardless of the degree of homology in the central nucleotides. Location of a given site is expressed as: number of base pairs from the start codon of the specified gene to the centre of the proposed binding site. All sites are given in a 5' to 3' orientation.

Promoter Region	Possible SlyA binding sites	Location of Site relative to gene start codon	A	B
<i>Pcas</i>	<b>TTATTGAATTAA</b>	100 bp upstream of <i>casA</i>	6	1
<i>Pssu/elf</i>	<b>TCAGGATGATAA</b>	8 bp upstream of <i>elfA</i>	5	12
<i>Pgsp</i>	<b>TTATATTAGTAA</b>	79 bp upstream of <i>gspA</i>	6	1
<i>Ppaa</i>	<b>TTAAATCGCGAA</b>	239 bp upstream of <i>paaA</i>	5	7
	<b>TTATAAAAATAG</b>	136 bp upstream of <i>paaA</i>		
	<b>TTACTTAACTAT</b>	81 bp upstream of <i>paaA</i>		
<i>Psgc</i>	<b>TTATGCTGGGAA</b>	336 bp upstream of <i>sgcX</i>	5	2
	<b>TTTCAACCATAA</b>	188 bp upstream of <i>sgcX</i>		
<i>Pfec</i>	<b>TTAGAAAAACAA</b>	109 bp upstream of <i>fecl</i>	5	7
<i>PslyA</i>	<b>TTAGCAAGCTAA</b>	22 bp upstream of <i>slyA</i>	6	2
	<b>TTAGATTAATAA</b>	161 bp upstream of <i>slyA</i>		
<i>PleuO</i>	<b>TTAATGCATTAA</b>	305 bp upstream of <i>leuO</i>	6	2
	<b>TTAAATATATAA</b>	297 bp upstream of <i>leuO</i>		

**Column A** = Level of ATT<6nt>TAA homology (out of 6)

**Column B** = Total number of sites with same level of ATT<6nt>TAA homology

membrane, outer membrane or the periplasmic space with 4 separate operons represented that encode fimbrial-like adhesins (*yehDCBA*, *elfADCGycbUVF*, *sfmACDHF* and *yadCKLMNhtrEyadVN*).

A strong correlation between the *E. coli* SlyA regulon identified here and that of another antagonist of H-NS repression, LeuO, is further evidence that *E. coli* SlyA acts by relieving repression by H-NS. It is possible that this correlation occurs because SlyA has a positive effect on *leuO* expression leading to an indirect effect on the genes within the LeuO regulon. However, this can be countered, at least partially, by the observed direct binding of SlyA at *Pssu*, *Pcas*, *Ppaa*, *Pelf*, *PleuO* and *Pgsp*, all of which are promoter regions of genes or operons proposed to be part of the LeuO regulon. This suggests that due to the similarity in their mode of action, i.e. H-NS de-repression, SlyA and LeuO exhibit a substantial overlap in the genes they regulate such that on activation by their respective signals a similar effect is elicited.

It was observed under the conditions used in this work that *slyA* expression in wild type *E. coli* MG1655 was at undetectable levels and as yet, the signal that may activate the transcription of *E. coli slyA* is not known, although H-NS is not thought to be involved (Corbett *et al.*, 2007). The suggestion that regulation of the intracellular concentration of SlyA is predominantly post-transcriptional has also been made in this work, due to the disparity between transcript level fold change (2.95-fold) and protein level fold change (>>25-fold) observed between a SlyA overproducing *E. coli* strain and a control strain.

It has previously been shown that in *S. enterica* Serovar Typhimurium the presence of ppGpp stimulates the regulatory activity of SlyA (Zhao *et al.*, 2008). This trait that was not observed here, in that ppGpp was not required for SlyA DNA binding and its presence had no effect on binding affinity when tested against *PmdtM*. This was also found to be the case with *PfimB* (McVicker *et al.*, 2011). Interestingly, the expression of *leuO* in *E. coli* has previously been found to be enhanced by the presence of ppGpp during the transition from exponential phase to stationary phase growth (Fang *et al.*, 2000). It is tempting to speculate that in *E. coli* the presence of ppGpp may activate the transcription of both *leuO* and *slyA* which then go on to affect the transcription of many genes, some of which are unique to each regulator and some of which are regulated by the presence of either.

This work has expanded on the number of confirmed genes and operons that are under the influence of SlyA in *E. coli* K-12. Where previously *hlyE* and *fimB* were the only confirmed genes whose promoters were directly regulated by SlyA, binding of SlyA to *Pcas*, *Pelf*, *Pgsp*, *Ppaa*, *Pssu*, *Psgc*, *Pfec*, *PleuO* and *PmdtM* has now been shown. It must be noted that these do not represent all the genes that were seen to be affected in the transcriptomic experiments, and there is scope for more SlyA targets to be confirmed. Interestingly no significant effect on the expression of *fimB* was seen in the transcriptomics (1.16 fold increase,  $p$ -value=0.008), despite it having been shown that SlyA binds the *fimB* promoter region and, again, antagonises H-NS silencing in *E. coli* (McVicker *et al.*, 2011). However, it is clearly stated in that study that *fimB* is “under complex and independent regulation by multiple factors” other than SlyA; namely NanR and NagC which activate transcription unless in the presence of *N*-acetylneuraminate or *N*-acetylglucosamine 6-phosphate respectively. Thus it is presumed that under the conditions studied here transcription of *fimB* was not favoured perhaps due to lack of NanR and/or NagC activity.

Purified SlyA protein was further used to examine possible binding sites within *PmdtM* and test its specificity to the TTAGCAAGCTAA consensus sequence elucidated for the *S. enterica* Serovar Typhimurium homologue (Stapleton *et al.*, 2002). Four sites were targeted, within the selected 400 bp *PmdtM* region, each bearing some similarity to the aforementioned consensus sequence. On removal of the portion containing these sites, SlyA was seen to still bind the remaining DNA. However, this led to an alternative binding site motif search (TTA<6nt>TAA) based on observations by Haider *et al.* (Haider *et al.*, 2008), and the suggestion that the presence of at least one complete half-site is necessary and sufficient for SlyA binding in the presence of excess heparin. Gene transcription regulation in the presence of only one complete half-site has been observed elsewhere (Zhou *et al.*, 1993; Sawers *et al.*, 1997). As a result of the observed binding motif requirements, a list of the most likely SlyA binding sites within the promoter regions tested for direct SlyA binding in this work has been compiled.

To conclude, these data demonstrate how SlyA may bind DNA with relatively low sequence specificity, but certain sites are bound preferentially. These traits, taken together, are appropriate for a regulator that operates by competing with the widespread, unspecific nucleoid structuring protein: H-NS.

#### **4. Studying changes in *Escherichia coli* gene expression in response to growth rate**

##### **Main findings**

- Transcript abundance of 253 genes changed significantly in response to increasing growth rate (86 genes were up-regulated and 167 were down-regulated)
- Genes associated with flagella assembly and motility were unexpectedly up-regulated as growth rate increased
- A high proportion of down-regulated genes were associated with secondary metabolism and were regulated by cAMP-CRP
- Activities of 167 transcriptional regulators were simultaneously inferred, as well as the activity of the alternative sigma factors FliA and RpoS
- Thirty eight transcriptional regulators exhibited altered activity as growth rate increased
- Preliminary steps were made in developing the TMDH method for analysis for *Escherichia coli* gene fitness in response to growth rate

##### **4.1. Introduction**

Though SlyA expression was not observed to be significantly influenced by growth rate, comparing samples of the same wild type *E. coli* MG1655 strain cultured at different dilution rates did show that growth rate had a significant effect on a number of other genes. To date adaptation of *E. coli*, and the bacterial domain as a whole, to changes in growth rate have focused on general cell parameters such as cell size, molecular composition (Schaechter *et al.*, 1958; Kubitschek and Freedman, 1971; Klumpp *et al.*, 2009) and more recently mRNA half-life (Esquerre *et al.*, 2014). Here, global transcriptional changes taking place in response to growth rate, in a controlled and defined environment, were measured to shed new light on processes required to sustain both relatively rapid and slow growth rates.

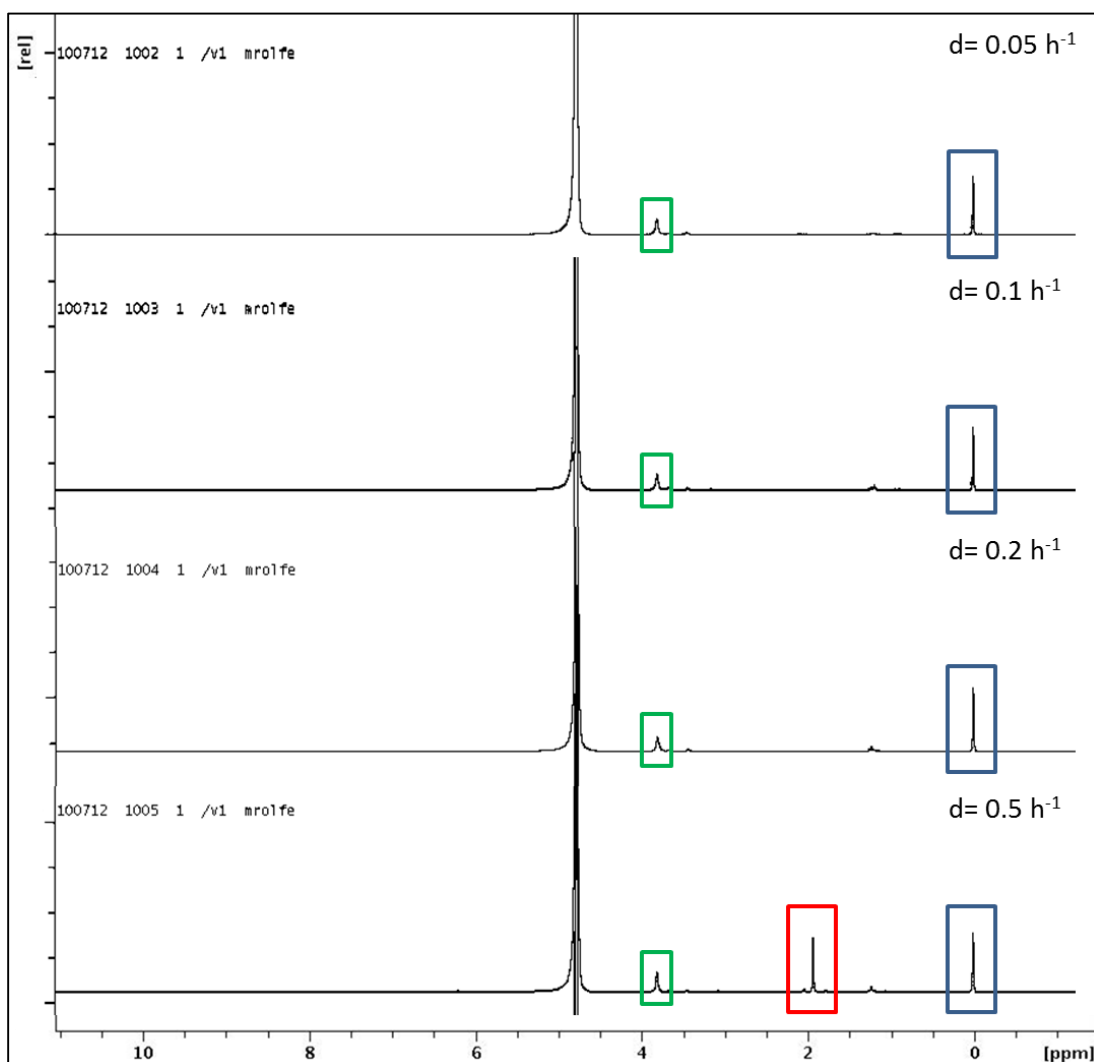
## **4.2. Chemostat growth parameters and associated data**

Cultures of *E. coli* K-12 strain MG1655 were grown aerobically in glucose-limited Evans' minimal medium in a 2 l chemostat vessel at a working volume of 1 l. Growth rate was controlled by varying dilution rate of fresh medium containing 20 mM glucose, rates of 0.05 h<sup>-1</sup>, 0.1 h<sup>-1</sup>, 0.2 h<sup>-1</sup> and 0.5 h<sup>-1</sup> were used which equate to *E. coli* doubling times of 13.9 h, 6.9 h, 3.5 h and 1.4 h respectively. When cultures reached steady state growth, samples were taken and RNA was purified as described in Sections 2.5 and 2.6.5. Independent, duplicate cultures were analysed.

Temperature and pH were maintained at 37°C and 6.9 respectively, the culture was also continuously aerated with filtered air at a rate of 1 l h<sup>-1</sup>. NMR analysis of culture supernatants indicated that all the glucose supplied was utilised. Specific rates of glucose consumption were calculated to be 0.91, 1.81, 5.73 and 8.86 mmol gCDW<sup>-1</sup> h<sup>-1</sup> at dilution rates of 0.05, 0.1, 0.2 and 0.5 h<sup>-1</sup> respectively, with a maximum observed biomass yield of 0.057 gCDW mmol<sup>-1</sup> glucose and a minimum of 0.625 mmol glucose gCDW<sup>-1</sup> h<sup>-1</sup> required for cell growth.

Further analysis of culture supernatants by proton NMR showed that no over-metabolites were detectable at the three lowest dilution rates, indicating that the bacteria were aerobically respiring (Figure 4.1). However, at the highest dilution rate (0.5 h<sup>-1</sup>) acetate (2.4 mM +/- 0.9, n=2) was detected in culture supernatants. This was in agreement with previous findings (Vemuri *et al.*, 2006; Nahku *et al.*, 2010) and is likely due to limitations in the flux through the Tricarboxylic Acid (TCA) cycle at higher growth rates. Taking into account the dilution rate and observed cell dry weight, the specific rate of acetate production at d= 0.5 h<sup>-1</sup> was calculated to be 1.1 mmol gCDW<sup>-1</sup> h<sup>-1</sup>.

These data then allowed calculation of carbon flux through the TCA cycle after glycolysis, accounting for the fact that a portion of carbon is directed to acetate synthesis at d= 0.5 h<sup>-1</sup>. In Table 4.1 rates of glucose consumption, biomass synthesis and acetate synthesis have been calculated in terms of mmoles of carbon used, per litre, per hour. Subsequently, the rate of carbon flux through the TCA cycle was estimated (Table 4.1, "RateC<sup>TCA</sup>"). Rates of glucose consumption and biomass production increased in relation to dilution rate. However, rate of carbon flux through the TCA cycle was seen to reach a maximum of 28.6 mmoles carbon l<sup>-1</sup> h<sup>-1</sup>

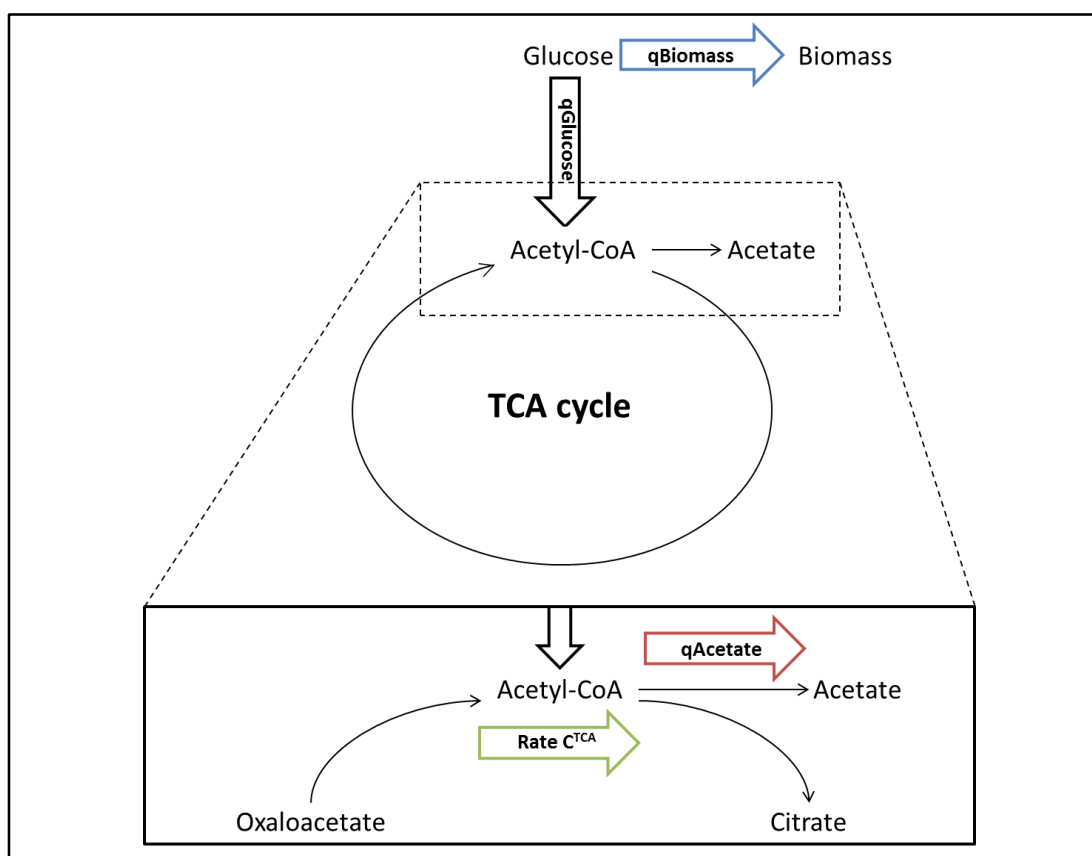


**Figure 4.1 Typical proton-NMR trace of culture supernatants for detection of over-metabolites.** Culture supernatants from steady-state chemostat cultures grown at various dilution rates ( $d = 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  from top to bottom respectively) were analysed by NMR. This was carried out twice, once per biological replicate. In both cases the same trend was observed. Blue boxes highlight 1 mM of trimethylsilyl propionate (TSP) used for calibration; as it is at a defined concentration and it is known to have 9 protons. The red box highlights detection of acetate (1.92 ppm) at  $d = 0.5 \text{ h}^{-1}$ . The area under the acetate peak relative to TSP-standard was measured and allowing for the fact that acetate has 3 protons, acetate concentration across both biological repeats was calculated to be 2.4 mM (St. Dev = 0.9,  $n = 2$ ). Green boxes highlight peaks representing nitrilotriacetic acid, a component of Evans' minimal medium.



**Table 4.1 Calculated rates of glucose consumption, biomass production, acetate production and predicted rate of carbon flux through the Tricarboxylic Acid (TCA) cycle.** All calculated rate values are displayed as mmoles of carbon  $l^{-1} h^{-1}$ . Glucose consumption, biomass production and acetate production are displayed in columns “qGlc”, “qBiomass” and “qAcetate” respectively. Calculation of qBiomass in terms of mmoles of carbon was based on the measurement that 50% of biomass produced is carbon (w/w) (Neidhardt et al., 1990). Rate of carbon flux through the TCA cycle is given by the estimated rate at which acetyl-CoA would have been metabolised to citrate, “rateC<sup>TCA</sup>”, calculated by subtracting amounts of carbon consumed in biomass and acetate production from the total amount of carbon input (qGlucose). Reactions considered are represented in Figure 4.2.

Dilution Rate ( $h^{-1}$ )	qGlucose	qBiomass	qAcetate	rateC <sup>TCA</sup>
	mmol carbon $L^{-1} h^{-1}$			
0.05	5.4	2.3	0.0	3.2
0.1	10.9	4.6	0.0	6.3
0.2	34.4	5.8	0.0	28.6
0.5	53.1	23.5	2.1	27.5



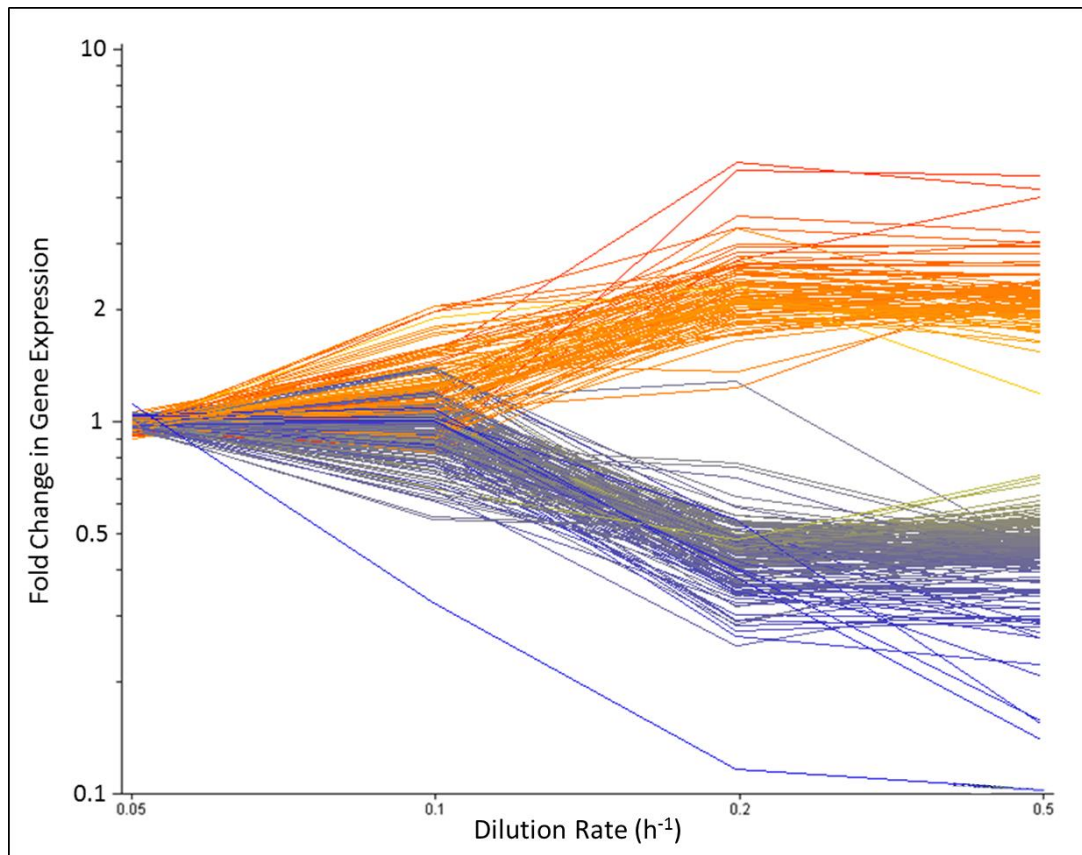
**Figure 4.2 Schematic of key reactions related to carbon uptake and metabolism that have been estimated in this work.** Large arrows highlight four of the key reactions that are considered in Table 4.1. Dashed lines enclose the region of TCA cycle that has been expanded on in the solid-outlined box.

when cultures were grown at a dilution rate of  $0.2 \text{ h}^{-1}$ , with a slight decrease to  $27.5 \text{ mmol carbon l}^{-1} \text{ h}^{-1}$  at  $d= 0.5 \text{ h}^{-1}$  to account for the observed production of acetate. This suggests that flux through the TCA cycle is near maximal at  $d= 0.2 \text{ h}^{-1}$ , with any additional influx of carbon being directed to acetate production which yields only one ATP molecule. The estimated maximal rate of carbon flux through the TCA cycle ( $\sim 28.6 \text{ mmol C l}^{-1} \text{ h}^{-1}$ ) was in agreement with the calculated maximum flux when *E. coli* strain ML308 was grown on acetate as a sole carbon source ( $14.4 \text{ mmol acetyl-CoA converted to citrate l}^{-1} \text{ h}^{-1}$ , equivalent to  $28.8 \text{ mmol C l}^{-1} \text{ h}^{-1}$ ).

#### **4.3. *Escherichia coli* strain MG1655 transcript profiles at fixed growth rates**

RNA labelling, microarray hybridisations and scanning were carried out on chemostat-derived samples as described in Section 2.6.15, with data filtered using GeneSpring 7.3.1 to identify genes that showed statistically significant changes in expression. These were defined as those genes whose expression-level changed more than 2-fold in at least one of the dilution rates studied and passed a *t*-test ( $p \leq 0.05$ , Benjamani & Hochberg multiple testing correction).

The transcript profiles for all 253 genes that passed the statistical filtering are shown in Figure 4.3. The data pertaining to this transcriptomic experiment, and subsequently referred to through the majority of this chapter, is an expression of transcript level as a fold change normalised to the transcript abundance at  $d= 0.05 \text{ h}^{-1}$  (i.e. relative to the slowest growth rate). There was no significant change in the expression profile of any genes at  $d= 0.1 \text{ h}^{-1}$  when compared to transcript levels at  $d= 0.05 \text{ h}^{-1}$ , however a total of 253 genes were seen to have a minimum of a 2-fold change in expression at either  $d= 0.2 \text{ h}^{-1}$  or  $d= 0.5 \text{ h}^{-1}$ . Of these genes, 86 (34 %) were up-regulated and 167 (66 %) were down-regulated. A more detailed breakdown of the distribution of significant changes is provided in Table 4.2. Upon transition between the dilution rates  $0.2 \text{ h}^{-1}$  and  $0.5 \text{ h}^{-1}$  very little further change in transcript abundance occurred. This is perhaps an indication that adaptation of the *E. coli* MG1655 transcriptome, in order to support growth at higher rates of carbon-supply, mostly takes place when approaching a doubling time of 3.5 h and any further decrease in doubling time does not require further responses at the transcriptional level.



**Figure 4.3 Graphical representation of all significant gene expression changes plotted against increasing dilution rate ( $h^{-1}$ ) after statistical filtering.** Each line represents the expression profile of an individual gene as a fold change relative to the level of expression at  $d= 0.05 h^{-1}$ , with dilution rate increasing from left to right. All expression profiles shown are those of genes that passed statistical filtering in at least one condition ( $p \leq 0.05$  and fold change  $\geq 2$ ).

**Table 4.2 Numbers of genes differentially expressed at different growth rates.** Values indicate the number of genes that pass statistical filtering ( $p \leq 0.05$  and fold changes  $\geq 2$ -fold) when total gene expression data for each individual dilution rate was compared to the transcript level at a dilution rate of  $0.05 h^{-1}$ .

Dilution Rate ( $h^{-1}$ )	Comparison to $d=0.05 h^{-1}$	
	Up-regulated	Down-regulated
<b>0.1</b>	0	0
<b>0.2</b>	66	138
<b>0.5</b>	60	119

#### **4.4. Genes up-regulated at increased growth rate**

A total of 86 genes exhibited at least a 2-fold increase in abundance in cultures grown at dilution rates of 0.2 or 0.5 h<sup>-1</sup>, i.e. increased significantly in expression at a doubling time of  $\leq 3.5$  h relative to a doubling time of 13.9 h. Table 4.3 is a list of the 62 genes within that group that have a known function, any genes with an as yet unknown function were removed, though a full list can be seen in Supplementary Data 4.2. Functional groups represented include those associated with adhesion (5%), amino acid biosynthesis (11%), DNA replication and cell division (5%), lipopolysaccharide biosynthesis (8%), metabolism (8%), motility and chemotaxis (18%), nucleotide metabolism (5%), stress response (8%), translation (8%), transport (5%) and those that were not categorised in any of the above (18%).

The top five genes which express the largest fold-change in response to growth rate are, in descending order; *azuC*, *yfdI*, *yecH*, *yrbN* and *yjdM*, all of which are of unknown function. These are therefore interesting targets for further studies into their specific function, and determining why they may be important in growth rate adaptation.

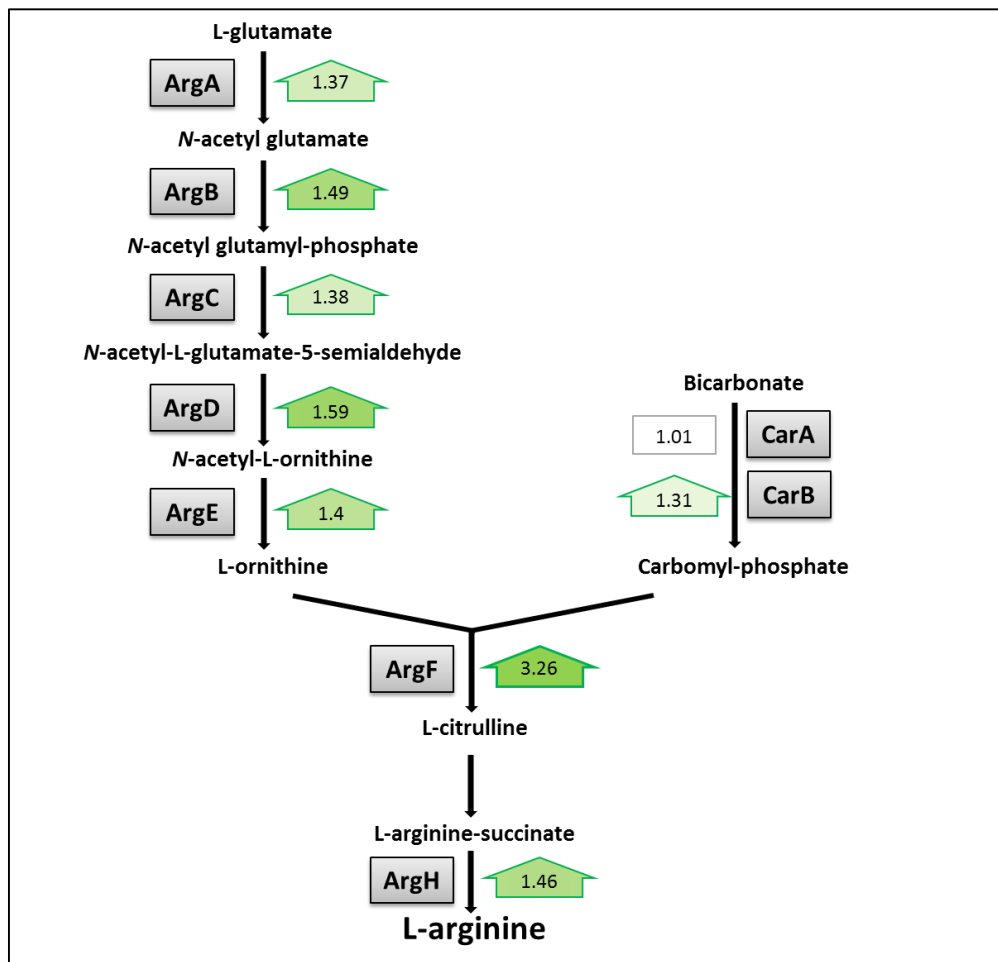
As would be expected due to the decrease in doubling time and hence protein translation and turnover, multiple genes associated with the biosynthesis of amino acids (histidine, methionine, proline, valine, glutamic acid and aspartic acid) are up-regulated as growth rate increased. Though only a small percentage of the total genes in this functional group are represented, when each gene is viewed on an individual basis (Supplementary Data 4.1) they often follow the expected trend as growth rate increased, as illustrated in Figure 4.4. Here, genes associated with the core biosynthetic pathway for arginine are all seen to be up-regulated as growth rate increased even though they did not pass the statistical filtering stated previously. Given that all cultures were grown in a minimal medium, necessitating amino acid biosynthesis regardless of dilution rate, the lack of statistical significance in transcript up-regulation for genes within this functional group is evidence that the rate of transcription is not significantly limiting growth rate.

Perhaps less expected is the observation that transcripts associated with chemotaxis and motility exhibited increasing abundance as the growth rate increased. Two operons in particular, *motAB-cheAW* and *tar-tap-cheRBYZ*, were almost

**Table 4.3** Transcripts exhibiting increased abundance as growth rate increases, at either  $d= 0.2$  or  $0.5 \text{ h}^{-1}$ . All data are expressed as fold changes in transcript abundance relative to  $d= 0.05 \text{ h}^{-1}$ . Only those genes that passed statistical filtering are shown ( $p \leq 0.05$  with  $\geq 2$ -fold change). Gene regulation by RpoS, FliA or CRP was identified using the Ecocyc database (Keseler et al., 2013). In the case of CRP regulation; “0” indicates dual regulation, “+” indicates positive regulation and “-” indicates negative regulation. Table is continued overleaf.

Fold change in gene expression relative to $d= 0.05 \text{ h}^{-1}$			Gene Name	RpoS Regulation	FliA Regulation	CRP Regulation
$d= 0.1 \text{ h}^{-1}$	$d= 0.2 \text{ h}^{-1}$	$d= 0.5 \text{ h}^{-1}$				
<b>Adhesion</b>						
1.15	1.95	2.26	<i>fimA</i>			
1.27	1.92	2.1	<i>fimC</i>			
1.20	1.80	2.09	<i>fimF</i>			
<b>Amino Acid Biosynthesis and Related</b>						
1.55	1.96	2.38	<i>gltP</i>			
1.06	1.74	1.99	<i>hisA</i>			
1.99	2.52	2.51	<i>mmuM</i>			
1.79	2.4	2.15	<i>mmuP</i>			
0.97	2	2.34	<i>proA</i>			
1.6	3.03	2.98	<i>proP</i>	X		0
1.41	2.18	2.39	<i>ygaH</i>			
<b>DNA Replication and Cell Division</b>						
1.22	2.24	2.16	<i>dnaB</i>			
1.59	2.63	2.51	<i>yjdA</i>		X	
1.34	2.10	1.94	<i>recF</i>	X		
<b>LPS Biosynthesis</b>						
0.92	2.27	2.13	<i>kdsD</i>			
1.13	2.01	2.26	<i>rfaZ</i>			
1.17	2.82	2.67	<i>rfbC</i>			
1.03	1.94	2.05	<i>rffG</i>			
1.47	2.66	2.65	<i>yijP</i>			
<b>Metabolism</b>						
1.14	2.05	2.00	<i>dld</i>			
1.24	1.85	1.80	<i>glpR</i>			+
1.02	2.35	2.27	<i>gpmM</i>	X		
1.56	2.32	1.73	<i>puuB</i>	X		-
0.91	1.83	1.99	<i>ptsG</i>			+
<b>Motility / Chemotaxis</b>						
1.38	2.10	1.93	<i>cheA</i>		X	
1.52	2.51	2.27	<i>cheB</i>		X	
1.99	3.35	3.05	<i>cheR</i>		X	
1.17	2.36	2.1	<i>cheW</i>		X	
1.27	2.37	2.25	<i>flgK</i>		X	
1.31	2.15	1.83	<i>fliC</i>		X	
1.10	2.10	1.78	<i>fliJ</i>		X	
1.52	2.03	1.94	<i>motB</i>		X	
1.53	2.7	2.36	<i>tap</i>		X	
1.42	2.13	1.91	<i>tar</i>		X	
1.76	2.13	2.04	<i>tsr</i>		X	
<b>Nucleotide Metabolism</b>						
1.3	2.5	2.52	<i>cdh</i>			
1.07	2.38	2.23	<i>purC</i>			
0.95	2.18	2.24	<i>queA</i>			
<b>Other</b>						
0.93	1.66	2.05	<i>ftnA</i>			
2.05	2.22	2.00	<i>lit</i>			
1.61	2.34	2.2	<i>mcrB</i>			
1.44	2.88	2.85	<i>nei</i>			
1.08	1.24	2.43	<i>ompW</i>			
1.28	1.73	2.17	<i>phoA</i>			
0.93	2.14	2.21	<i>speG</i>			
1.59	2.08	1.85	<i>yeiG</i>			
1.39	1.86	1.97	<i>ykgJ</i>			
1.44	1.37	2.18	<i>hybB</i>			
1.52	1.71	2.19	<i>hybC</i>			

Fold change in gene expression relative to $d= 0.05 \text{ h}^{-1}$			Gene Name	RpoS Regulation	FliA Regulation	CRP Regulation
$d= 0.1 \text{ h}^{-1}$	$d= 0.2 \text{ h}^{-1}$	$d= 0.5 \text{ h}^{-1}$				
<b>Stress Response</b>						
1.1	2.24	2.38	<i>sodB</i>			0
1.14	1.85	2.08	<i>hslU</i>			
1.21	2.15	2.21	<i>htpG</i>			
2.05	2.63	2.7	<i>ybbM</i>			
1.21	1.74	2.26	<i>evgA</i>	X		
<b>Translation</b>						
1.25	2.35	2.12	<i>deaD</i>			
1.26	2.04	1.95	<i>rlmF</i>			
1.03	2.11	2.08	<i>ybaK</i>			
1.10	2.16	1.91	<i>yciH</i>			
0.96	1.88	2.01	<i>yjjK</i>	X		
<b>Transport</b>						
1.19	2.09	2.07	<i>dppF</i>			
1.92	2.35	1.20	<i>kdpF</i>			
0.83	1.93	2.06	<i>mscS</i>			
1	2.57	2.17	<i>ydeA</i>			



**Figure 4.4 Schematic of the core arginine biosynthetic pathway with overlaid data for transcript levels of the corresponding genes at  $d= 0.5 \text{ h}^{-1}$  as a fold change relative to  $d= 0.05 \text{ h}^{-1}$ . All data are from the total data set (Supplementary Data 4.1) before statistical filtering. All genes, but *carA*, show an up regulation in expression at  $d= 0.5 \text{ h}^{-1}$  relative to  $d= 0.05 \text{ h}^{-1}$ .**

represented in their entirety and exhibited increases in their component transcripts of up to 2.36-fold and 3.35-fold respectively. This observation was unexpected as chemotaxis and motility are two traits not obviously necessary to allow more rapid growth in a chemostat vessel wherein the culture volume is undergoing constant agitation and mixing. Whilst the growth medium is carbon-limited, this is true across the whole range of dilution rates. It may therefore be a combination of the sustained carbon limitation and the relatively high rate of cell division at cell doubling times  $\leq$  3.5 h that the bacteria perceive as necessitating increased motility to access additional nutrients.

Also exhibiting significantly increased transcript abundance as doubling time was reduced were those associated with Type 1 fimbriae expression (*fimA*, *fimC* and *fimF*). Although it has previously been found that exponential phase growth in aerobic cultures favours afimbriate cells over ones expressing Type-1 fimbriae (Gally *et al.*, 1993), the increase in expression seen here for an aerobic culture growing at an increased rate could be explained by the observation by Adiciptaningrum *et al.* (2009), who showed that the inversion of *fimS* and the subsequent OFF-to-ON phase variation of *fim* genes occurred preferentially at the beginning of the cell cycle, an event that would be occurring at an increased rate at the higher dilution rates studied. This, coupled with an increase in the occurrence of overlapping rounds of DNA replication per cell (Cooper and Helmstetter, 1968) may all contribute to a relative increase in *fim* gene expression.

At higher rates of growth it was also seen that five separate genes (*kdsD*, *rfaZ*, *rfbC*, *rffG* and *yijP*) associated with the production and maintenance of the lipopolysaccharide (LPS) layer of *E. coli* MG1655 were upregulated. The LPS is an essential outer membrane glycolipid, and is also important for pathogenicity and stress resistance, including withstanding antibiotic activity (Raetz and Whitfield, 2002). Here, a correlation between decreased cell doubling time increased expression of specific LPS biosynthesis genes has been observed, suggesting that this process may be growth rate-limiting.

#### **4.5. Genes down-regulated at increased growth rate**

A total of 167 genes were observed to decrease significantly in expression as growth rate was increased i.e. they exhibited at least a 2-fold decrease in abundance at a doubling time of  $\leq 3.5$  h relative to a doubling time of 13.9 h. Table 4.4 is a list of the 108 genes within that group that have a known function, any genes with an as yet unknown function were removed, though a full list can be seen in Supplementary Data 4.3. Functional groups represented in this gene list include metabolism (30%), small RNAs (6%), stress response (3%), gene regulation (18%), translation (1%), transport (30%) and those that are not categorised in any of the above (12%).

An over-representation of genes that are known to be regulated by the cyclic-AMP-Receptor Protein (CRP) was observed in the down-regulated genes of the metabolism, gene regulation and transport groups (44%, 50% and 40% of the genes within those functional groups respectively). On binding of cyclic-AMP (cAMP), CRP has been shown to influence the expression of over 260 transcriptional units (data from ecocyc.org, (Keseler *et al.*, 2013)) many of which are known to encode the machinery necessary for the catabolism of secondary carbon sources in the absence of a rapidly metabolisable substrate such as glucose (Fic *et al.*, 2009; Gorke and Stulke, 2008). The majority of genes that were observed to decrease in transcript abundance as growth rate increased, and are regulated by cAMP-CRP, fit into this category of secondary metabolism of alternative carbon sources. It must also be noted that of all the genes regulated by CRP-cAMP in this list 90% of them are known to be positively influenced. Taken as a whole, the data agree with the previous finding that the intracellular concentration of cAMP decreased as growth rate increased in a glucose-limited medium (Kuo *et al.*, 2003). As a result of this, positive regulation of a number of genes associated with alternative carbon source utilisation would be diminished as the dilution rate was increased.

The most significantly down-regulated gene in the dataset was that encoding the Ribosome Modulation Factor, *rmf*. Its function involves converting active 70S ribosomes into inactive 100S dimeric forms in order to limit translation in cells experiencing stationary phase conditions (Wada *et al.*, 1995). The lower transcription was thus expected, as it has already been shown that *rmf* expression is inversely proportional to growth rate (Yamagishi *et al.*, 1993) and the *rmf* response,



**Table 4.4 Transcripts exhibiting decreased abundance as growth rate increases, at either  $d= 0.2$  or  $0.5 \text{ h}^{-1}$ .** All data are expressed as a fold change in transcript level relative to  $d= 0.05 \text{ h}^{-1}$ . Only those genes that passed statistical filtering are shown ( $p \leq 0.05$ ,  $\geq 2$ -fold change). Gene regulation by RpoS, FliA or CRP was identified using the Ecocyc database (Keseler et al., 2013). In the case of CRP regulation; “+” indicates positive regulation and “-” indicates negative regulation. Table is continued overleaf.

Fold change in gene expression relative to $d= 0.05 \text{ h}^{-1}$			Gene Name	RpoS Regulation	FliA Regulation	CRP Regulation
$d= 0.1 \text{ h}^{-1}$	$d= 0.2 \text{ h}^{-1}$	$d= 0.5 \text{ h}^{-1}$				
<b>Metabolism</b>						
1.02	0.35	0.26	<i>aldB</i>	X		+
1.34	0.46	0.63	<i>astC</i>	X		
0.94	0.48	0.45	<i>atoD</i>			
0.95	0.35	0.31	<i>ddpX</i>	X		
1.39	0.45	0.43	<i>dgoA</i>			
0.96	0.44	0.53	<i>fadA</i>			
1.06	0.54	0.49	<i>fadH</i>			+
0.89	0.53	0.46	<i>fdrA</i>			
0.66	0.25	0.35	<i>gadA</i>	X		-
0.78	0.45	0.58	<i>gadB</i>	X		-
0.93	0.53	0.48	<i>idnD</i>			+
0.94	0.45	0.42	<i>narY</i>			
1.08	0.44	0.40	<i>paaB</i>			+
1.06	0.41	0.42	<i>paaC</i>			+
1.07	0.44	0.48	<i>paaD</i>			+
1.04	0.44	0.42	<i>paaE</i>			+
0.88	0.48	0.53	<i>paal</i>			+
1.17	0.45	0.46	<i>paaK</i>			+
0.92	0.47	0.54	<i>poxB</i>	X		
1.20	0.47	0.29	<i>prpB</i>			+
1.41	0.53	0.26	<i>prpC</i>			+
1.40	0.59	0.42	<i>prpE</i>			+
0.93	0.36	0.41	<i>rbsD</i>			+
0.76	0.46	0.56	<i>rutA</i>			
0.99	0.48	0.44	<i>scpA</i>			
1.08	0.56	0.51	<i>tynA</i>			
1.23	0.46	0.53	<i>ulaG</i>			
0.88	0.38	0.41	<i>wcaB</i>			
1.06	0.41	0.54	<i>xdhB</i>			
1.04	0.44	0.56	<i>ybhO</i>			
1.04	0.41	0.48	<i>ygeX</i>			
1.07	0.47	0.54	<i>yihT</i>			
<b>Other</b>						
1.20	0.44	0.41	<i>cspD</i>			
0.62	0.42	0.48	<i>cspI</i>			
0.64	0.48	0.55	<i>ecnB</i>	X		
0.95	0.42	0.44	<i>hyaB</i>	X		
0.63	0.42	0.43	<i>pinE</i>			
0.87	0.71	0.40	<i>sufA</i>			
0.95	0.47	0.49	<i>ybdK</i>			
1.15	1.29	0.45	<i>ybdZ</i>			
1.13	0.39	0.45	<i>yeiC</i>			
0.90	0.45	0.40	<i>ygfJ</i>			
1.14	0.54	0.50	<i>hyfD</i>			+
0.69	0.36	0.34	<i>yjfO</i>			
0.93	0.46	0.58	<i>yqjG</i>			
<b>sRNA</b>						
0.90	0.48	0.40	<i>isrA</i>			
0.55	0.52	0.45	<i>omrA</i>			
0.73	0.47	0.43	<i>rprA</i>			
0.77	0.46	0.41	<i>rybB</i>			
0.82	0.76	0.47	<i>ryhB</i>			
0.68	0.33	0.31	<i>ryjA</i>			
<b>Stress</b>						
0.74	0.46	0.56	<i>hdeB</i>	X		
0.89	0.46	0.58	<i>hdeD</i>			
0.71	0.45	0.47	<i>yodD</i>			

Fold change in gene expression relative to d= 0.05 h <sup>-1</sup>			Gene Name	RpoS Regulation	FliA Regulation	CRP Regulation
d= 0.1 h <sup>-1</sup>	d= 0.2 h <sup>-1</sup>	d= 0.5 h <sup>-1</sup>				
<b>Regulation</b>						
1.00	0.50	0.55	<i>araC</i>			+
0.67	0.48	0.72	<i>bssR</i>			
0.91	0.51	0.57	<i>cdaR</i>			
1.06	0.35	0.35	<i>feaR</i>			+
0.74	0.29	0.45	<i>gadE</i>	X		-
1.05	0.37	0.46	<i>galS</i>			+
0.96	0.29	0.42	<i>hcaR</i>			
0.80	0.53	0.48	<i>IsrK</i>			+
0.79	0.34	0.45	<i>IsrR</i>			+
1.07	0.47	0.45	<i>melR</i>			+
1.03	0.46	0.71	<i>mhpR</i>			+
0.95	0.39	0.42	<i>mlrA</i>	X		
1.01	0.36	0.28	<i>rhaS</i>	X		+
1.13	0.60	0.48	<i>sgcR</i>			
1.04	0.48	0.45	<i>tdcA</i>			+
1.05	0.32	0.47	<i>yeaT</i>			
0.99	0.36	0.38	<i>yehU</i>			
0.94	0.39	0.47	<i>ygeV</i>			
0.85	0.37	0.44	<i>ygiV</i>			
1.13	0.47	0.52	<i>yidF</i>			
<b>Translation</b>						
0.33	0.12	0.10	<i>rmf</i>			+
<b>Transport</b>						
1.13	0.53	0.49	<i>ascF</i>			+
1.03	0.48	0.46	<i>atoE</i>			
1.03	0.41	0.47	<i>dcuB</i>			+
0.99	0.47	0.47	<i>ddpA</i>	X		
1.05	0.42	0.44	<i>frlA</i>			
0.98	0.44	0.44	<i>frvB</i>			
1.06	0.44	0.45	<i>fucP</i>			+
0.85	0.35	0.35	<i>gspG</i>			
1.05	0.44	0.61	<i>lldP</i>			
0.81	0.49	0.46	<i>IsrA</i>	X		+
0.91	0.28	0.30	<i>IsrC</i>	X		+
1.06	0.35	0.37	<i>IsrD</i>	X		+
0.98	0.50	0.47	<i>IsrF</i>	X		+
0.85	0.47	0.49	<i>nhaA</i>	X		
0.87	0.39	0.43	<i>rbsA</i>			+
1.09	0.38	0.14	<i>sgcB</i>			
1.01	0.40	0.16	<i>sgcC</i>			
0.98	0.49	0.35	<i>sgcE</i>			
1.03	0.55	0.16	<i>sgcQ</i>			
1.00	0.40	0.21	<i>sgcX</i>			
1.41	0.53	0.49	<i>treB</i>			-
1.19	0.50	0.59	<i>ugpA</i>			+
1.27	0.63	0.49	<i>uhpT</i>			+
1.14	0.43	0.37	<i>xylF</i>			+
1.37	0.41	0.39	<i>xylG</i>			+
0.91	0.49	0.53	<i>ybaE</i>			
0.87	0.26	0.22	<i>ydcS</i>	X		
0.97	0.29	0.35	<i>ydcT</i>	X		
1.06	0.36	0.33	<i>ydcV</i>	X		
1.23	0.48	0.45	<i>yicO</i>			
0.94	0.41	0.50	<i>yidK</i>			
0.97	0.44	0.42	<i>ytfQ</i>			
0.93	0.45	0.47	<i>ytfR</i>			

along with the predictable cAMP-CRP observations, are useful benchmarks for the quality and usefulness of the dataset produced here. What is perhaps more interesting is that *rmf*, like many other genes, seems to experience very little change in gene expression level between doubling times of 3.5 h and 1.4 h, suggesting that once a doubling time of approximately 3.5 h has been reached this particular route to enhanced growth rates is no longer limiting.

In the down-regulated gene list, as well as *rmf*, there were four other small proteins with attributed functions. These were *cspD* (74 amino acids (aa)), *cspI* (70 aa), *ecnB* (47 aa) and *yodD* (75 aa). Notable amongst these is *cspD* which although it has high sequence similarity to a cold shock inducible gene, *cspA*, it does not share the same trait of induction on a shift from 37°C to 15°C (Lee *et al.*, 1994), but is an inhibitor of DNA replication that is induced during stationary phase growth (Yamanaka *et al.*, 2001). The data presented here suggest that *cspD* expression may be less regulated by a distinct growth phase but instead undergoes a reduction in expression at doubling times <6.9 h. Given its function, CspD may play an important role in limiting DNA replication, and therefore the cell division rate of *E. coli*, at the lower dilution rates tested here.

As illustrated by *cspD* expression relative to growth rate, the distinction between the slow growth rate conditions studied here and stationary phase growth must be made. Under the slow growth rate conditions bacteria are still proliferating at a rate that exceeds cell death. This highlights how a gene may be described as being induced in stationary-phase growth, and while this may be when it is maximally expressed, it may also be expressed in conditions where cells are undergoing steady-state growth in an environment with a slow rate of nutrient feed. The total dataset produced here (Supplementary Data 4.1) provides a good source for identifying genes that follow this pattern.

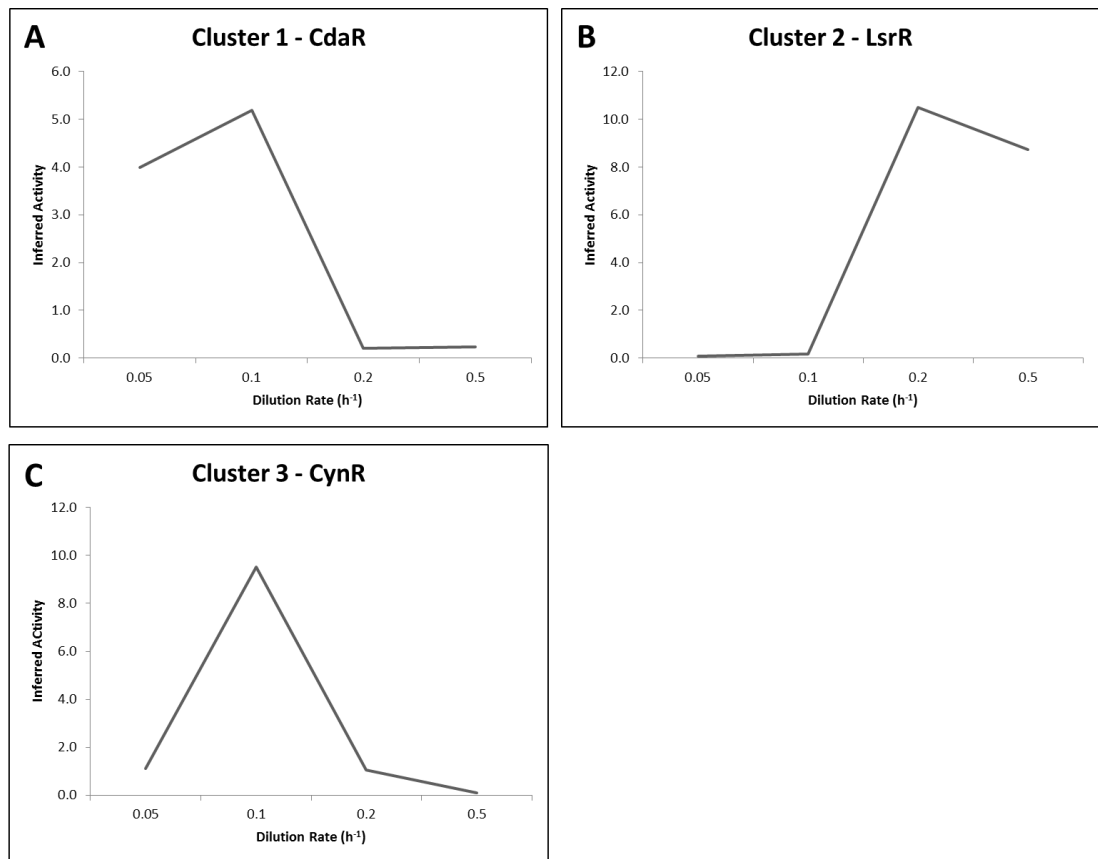
In the down-regulated gene dataset *lsrA*, *lsrC*, *lsrD* and *lsrF*, along with the regulator *lsrR*, are all present although cell density remains very similar at all growth rates studied here. The *lsr* genes are associated with the uptake of the quorum sensing molecule AI-2 into the *E. coli* cell (Xavier and Bassler, 2005) and will be discussed in more detail in Chapter 5.

#### **4.6. Identifying transcription factors playing a role in growth rate adaptation**

Changes in the abundance of transcripts are mostly attributable to changes in the activities of transcription factors. In order to identify the transcription regulators that play roles in the adaptation to changes in growth rate, the complete transcriptional datasets produced from the microarray analyses (Supplementary Data 4.1) were further analysed with TFInfer (Asif *et al.*, 2010). This tool infers the activity of a transcription factor by utilising a connectivity matrix which links each regulator to its known target genes and their determined transcriptional levels. Using this programme, 167 transcriptional regulators were examined simultaneously at each of the growth rates and 38 were predicted to exhibit a significant change in activity (signal to noise ratio >2.5) as growth rate increased; these are listed in Table 4.5. In this analysis the term “activity” refers to the regulator’s ability to bind DNA and affect gene transcription. Three response classes were evident (Figure 4.5). There are those that have increased activity at the lower dilution rates of 0.05 h<sup>-1</sup> and 0.1 h<sup>-1</sup> and then show decreased activity at the higher dilution rates of 0.2 h<sup>-1</sup> and 0.5 h<sup>-1</sup> (Cluster 1; represented by CdaR in Figure 4.5A). Those that show the inverse (Cluster 2; represented by LsrR in Figure 4.5B) and those that show significantly increased activity at only one of the dilution rates tested (Cluster 3; represented by CynR in Figure 4.5C). Of the responsive transcription factors, 18 were assigned to Cluster 1, 14 were assigned to Cluster 2 and 6 were assigned to Cluster 3. This analysis shows that, of the transcription regulators that are responsive, 84% (32 out of 38) showed a clear change in activity on the transition between  $d=0.1\text{ h}^{-1}$  and  $0.2\text{ h}^{-1}$  (doubling times of 6.9 and 3.5 hours respectively). Of these 32 regulators, 19 regulate genes involved in metabolism of alternative carbon or nitrogen sources (AgaR, AllR, AscG, CdaR, CRP, DpiA, FeaR, FhlA, GalS, HcaR, NtrC, PaaX, PspF, RbsR, RhaR, TreR, UlaR, XylR and YiaJ), and in all cases, whether they are repressors or activators of their primary target operons, their action resulted in a reduction of the expression of genes required for the uptake and/or metabolism of their target substrates as growth rate was increased. For example, *rbsR* encodes a transcriptional repressor for the *rbsDACBKR* operon associated with transport and utilisation of ribose as a carbon-source (Lopilato *et al.*, 1984; Mauzy and Hermodson, 1992). Comparing  $d=0.1\text{ h}^{-1}$  and  $d=0.2\text{ h}^{-1}$  cultures, the regulatory activity of RbsR was increased (an inferred activity of 0.2 compared to 7.9 respectively) and remained

**Table 4.5 Gene regulators exhibiting significant differential activity in response to growth rate.** Data presented as a list of gene regulators alongside their inferred activities (arbitrary units) at each dilution rate. TFinfer analysis software (Asif et al., 2010) was used to simultaneously infer the activity of 167 transcription factors across the transcriptomic data set produced in this work. Changes in transcription factor activity were deemed significant if the signal to noise ratio was >2.5. The signal to which each transcription factor responds is also listed. \* indicates that abundance of transcript encoding these transcription factors was also seen to significantly change (Tables 4.3 and 4.4). Bold text indicates where activity of a transcription factor was predicted to be greatest. Transcription factors have been organised into clusters of similar activity trends according to criteria outlined in Section 4.6.

Transcription Factor	Dilution Rate (h <sup>-1</sup> )				Signal
	0.05	0.1	0.2	0.5	
Cluster 1					
ArgR	<b>4.1</b>	2.3	0.3	0.4	L-Arginine
CdaR*	4.0	<b>5.2</b>	0.2	0.2	D-Glycerate
CpxR	<b>3.3</b>	1.9	0.4	0.4	
CRP	2.5	<b>3.2</b>	0.4	0.3	cAMP
DpiA	5.1	<b>6.3</b>	0.2	0.2	Citrate
EvgA*	<b>4.2</b>	2.0	0.3	0.3	
FeaR*	5.4	<b>9.2</b>	0.1	0.2	
FhlA	2.0	<b>3.9</b>	0.4	0.3	Formate
GadE*	3.0	<b>4.5</b>	0.3	0.3	pH Homeostasis
HcaR*	<b>6.4</b>	5.5	0.1	0.3	3-Phenylpropanoate
HyfR	<b>4.4</b>	2.8	0.3	0.3	Formate
MetJ	3.6	<b>3.8</b>	0.2	0.3	S-Adenosylmethionine
NtrC	3.7	<b>5.8</b>	0.2	0.2	Nitrogen limitation
PurR	<b>4.4</b>	3.5	0.2	0.3	Hypoxanthine/guanine
RcsAB	<b>5.8</b>	3.5	0.2	0.3	
RhaR	<b>7.0</b>	5.7	0.2	0.1	L-Rhamnose
TrpR	<b>3.8</b>	2.9	0.3	0.3	L-Tryptophan
XylR	2.9	<b>6.9</b>	0.3	0.2	α-D-xylopyranose
Cluster 2					
AgaR	0.2	0.1	4.9	<b>8.0</b>	N-acetylgalactosamine
AllR	0.8	0.2	2.4	<b>3.6</b>	Allantoin/glyoxolate
AscG	0.2	0.1	4.8	<b>7.5</b>	
CusR	0.2	0.7	2.2	<b>4.5</b>	Copper
DicA	0.3	0.3	2.3	<b>5.8</b>	
FlhDC	0.3	0.5	<b>3.2</b>	2.1	
GalS*	0.4	0.3	<b>3.9</b>	1.9	β-D-Galactose
LsrR*	0.1	0.2	<b>10.5</b>	8.7	Phosphorylated AI-2
PaaX	0.3	0.2	<b>4.7</b>	3.9	Phenylacetyl-CoA
PspF	0.1	1.0	<b>3.7</b>	2.4	
RbsR	0.1	0.2	<b>7.9</b>	6.2	D-Ribose
TreR	1.1	0.0	4.9	<b>5.8</b>	Trehalose-6-phosphate
UlaR	0.5	0.2	<b>3.9</b>	2.8	L-ascorbate-6-phosphate
YiaJ	0.9	0.1	3.4	<b>4.4</b>	
Cluster 3					
BetI	<b>6.1</b>	0.4	1.0	0.4	Choline
CynR	1.1	<b>9.5</b>	1.0	0.1	Cyanate
Fur	0.7	0.5	0.7	<b>4.9</b>	Fe(II)
KdpE	0.2	1.5	<b>15.2</b>	0.2	K <sup>+</sup> Concentration
LldR	0.3	0.2	<b>10.5</b>	1.5	L-Lactate
PrpR	4.9	<b>25.6</b>	0.3	0.0	2-Methylcitrate



**Figure 4.5** Graphs illustrating activity trends of transcription factors **CdaR**, **LsrR** and **CynR** which are representative of their respective clusters. Data for each transcription factor is that displayed in Table 3.4, and was determined by the TFinfer analysis software (Asif et al., 2010). Inferred activity is in arbitrary units. Criteria for inclusion in Cluster 1 (A), Cluster 2 (B) and Cluster 3 (C) are as described in Section 4.6.

high at  $d= 0.5 \text{ h}^{-1}$ , which was coupled to decreased transcription of its target operon. Two of the genes within this operon, *rbsA* and *rbsD*, were seen to be significantly down-regulated in the transcriptomics (Table 4.4).

Conversely, *feaR* encodes a transcriptional activator responsible for the transcription of genes required for oxidation of the aromatic amide phenylethylamine into phenylacetaldehyde (*tynA*) and the subsequent oxidation into phenylacetate (*feaB*), essentially allowing utilisation of phenylethylamine as an alternative nitrogen and carbon source (Parrott *et al.*, 1987; Zeng and Spiro, 2013). In this case FeaR showed a *decrease* in activity on transition from a dilution rate of  $0.1 \text{ h}^{-1}$  to a rate of  $0.2 \text{ h}^{-1}$  (an inferred activity of 9.2 compared to 0.1 respectively) and remained low at  $d= 0.5 \text{ h}^{-1}$ , which is coupled to the decrease in transcription of the target genes *tynA* and *feaB*. Both of these target genes were observed to decrease in transcript abundance at doubling times  $\leq 3.5$  hours in the complete transcriptomic dataset (Supplementary Data 4.1). For example, *tynA* exhibited a 1.08-fold, 0.56-fold and 0.51-fold change in transcript abundance at the dilution rates 0.1, 0.2 and  $0.5 \text{ h}^{-1}$  respectively (values are relative to transcript abundance at  $d= 0.05 \text{ h}^{-1}$ ).

For the transcriptional regulators CdaR, FeaR, GadE, GalS, HcaR and LsrR the predicted changes in activity were accompanied by a change in the transcription of the corresponding gene. This was not apparent for the other regulators suggesting their intracellular concentrations are maintained, at least at the level of transcription. This allows specific responses to levels of signalling molecules. An exception to this was *evgA* which encodes a transcriptional activator that initiates a signalling cascade associated with the activation of genes for acid and multi-drug resistance (Itou *et al.*, 2009). The activity of this transcriptional activator was seen to be highest at dilution rates of 0.05 and  $0.1 \text{ h}^{-1}$  (Cluster 1), and this is accompanied with the decrease in transcript abundance of target genes (e.g. *frc* shows a 0.7-fold decrease in transcript abundance in samples grown at  $d= 0.2$  and  $0.5 \text{ h}^{-1}$ ). However, the gene *evgA* that encodes this regulator showed a significant increase in transcript abundance as growth rate increased (Table 4.3). This is likely to be an example of an increase in transcription being required in order to maintain an intracellular concentration of the regulator protein; however in the increased dilution rate conditions the specific signal to which EvgA responds is no longer present at a sufficient quantity for activation. It is also possible that the increase in acetate detected in the supernatant

of cultures grown at  $d= 0.5 \text{ h}^{-1}$  (Figure 4.1) was perceived as an internal acid stress signal for *evgA* transcription, but was not the specific signal required for subsequent activation of the regulator itself.

The activity of DicA (Cluster 2), increased at  $d= 0.2 \text{ h}^{-1}$  and increased further at  $d= 0.5 \text{ h}^{-1}$ . DicA has been shown to be a repressor of genes involved in cell division, namely *dicB* (Bejar *et al.*, 1986; Cam *et al.*, 1988). DicB is an inhibitor of *E. coli* cell division, therefore its inhibition at higher growth rates by the increased activity of DicA may play an important role in allowing the higher rate of cell doubling required at  $d= 0.2$  and  $0.5 \text{ h}^{-1}$ .

TFinfer also predicted that CRP had increased activity at the lower dilution rates ( $d= 0.05$  and  $0.1 \text{ h}^{-1}$ ), in agreement with the observation that CRP is a positive regulator of genes decreased significantly in transcript abundance at the higher dilution rates ( $d= 0.2$  and  $0.5 \text{ h}^{-1}$ ).

#### **4.7. Quantifying activity of alternative sigma factors FliA and RpoS**

Of the 62 genes of known function that were up-regulated in response to increasing growth rate, 12 genes (covering 7 separate operons) are transcribed from FliA- ( $\sigma^{28}$ ) dependent promoters (Table 4.3). FliA is a minor sigma factor that is known or predicted to be responsible for the transcription of 147 operons in *E. coli* MG1655 (Keseler *et al.*, 2013) many of which are involved in flagella biosynthesis and motility (Koo *et al.*, 2009). The total transcriptomic dataset and TFinfer were utilised to establish the trend of FliA influence across the growth rate range studied here. The output indicates that FliA activity was at its lowest at the lower dilution rates tested. Its inferred activity was 0.0, 0.3, 10.2 and 7.1 at the dilution rates of  $d= 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  respectively (signal to noise ratio = 9.0), again showing the previously observed sharp increase in activity at doubling times  $\leq 3.5$  h. This is reflected in the transcript expression profile of the genes involved in chemotaxis and motility that were significantly up-regulated at increased growth rate.

RpoS encodes another alternative sigma factor ( $\sigma^s$ ) that is predicted or known to control the expression of 226 operons in *E. coli* MG1655 (Keseler *et al.*, 2013). RpoS was responsible for the expression of 6 of the 62 up-regulated genes (6 operons) and 21 of the 108 down-regulated genes (14 operons) as growth rate was



increased (Table 4.3 and Table 4.4). RpoS is considered to direct the transcription of genes involved in stress responses and secondary metabolism (Maciag *et al.*, 2011). Using TFinfer and the total transcriptomic dataset, it was predicted that RpoS was most active at the lower dilution rates, with an inferred activity of 1.5, 1.8, 0.6 and 0.5 at the dilution rates of  $d = 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  respectively (signal to noise ratio = 1.18). This decrease in activity with increasing growth rate correlates with the observed increased dependency on RpoS for gene expression in the down-regulated gene list; RpoS aids in transcription of 9.7% of the up-regulated gene list, and 19.3% of the down-regulated gene list. Thus, at doubling times  $\geq 6.9 \text{ h}$  RpoS is predicted to play a more significant role in gene transcription activation, and as cell doubling time is reduced (at least  $\leq 3.5 \text{ h}$ ) RpoS sees its influence decline, coupled with a down-regulation of the genes it was once targeting.

#### **4.8. Development of a Transposon-Mediated Differential Hybridisation (TMDH) approach in order to test gene fitness over increasing growth rate in *E. coli* MG1655**

The work outlined in this section was carried out in order to develop the use of an *E. coli* MG1655 Tn5-transposon mutant library as a means to identify genes that confer an advantage to cells that progress through the dilution rates already used in the gene-expression microarray analysis.

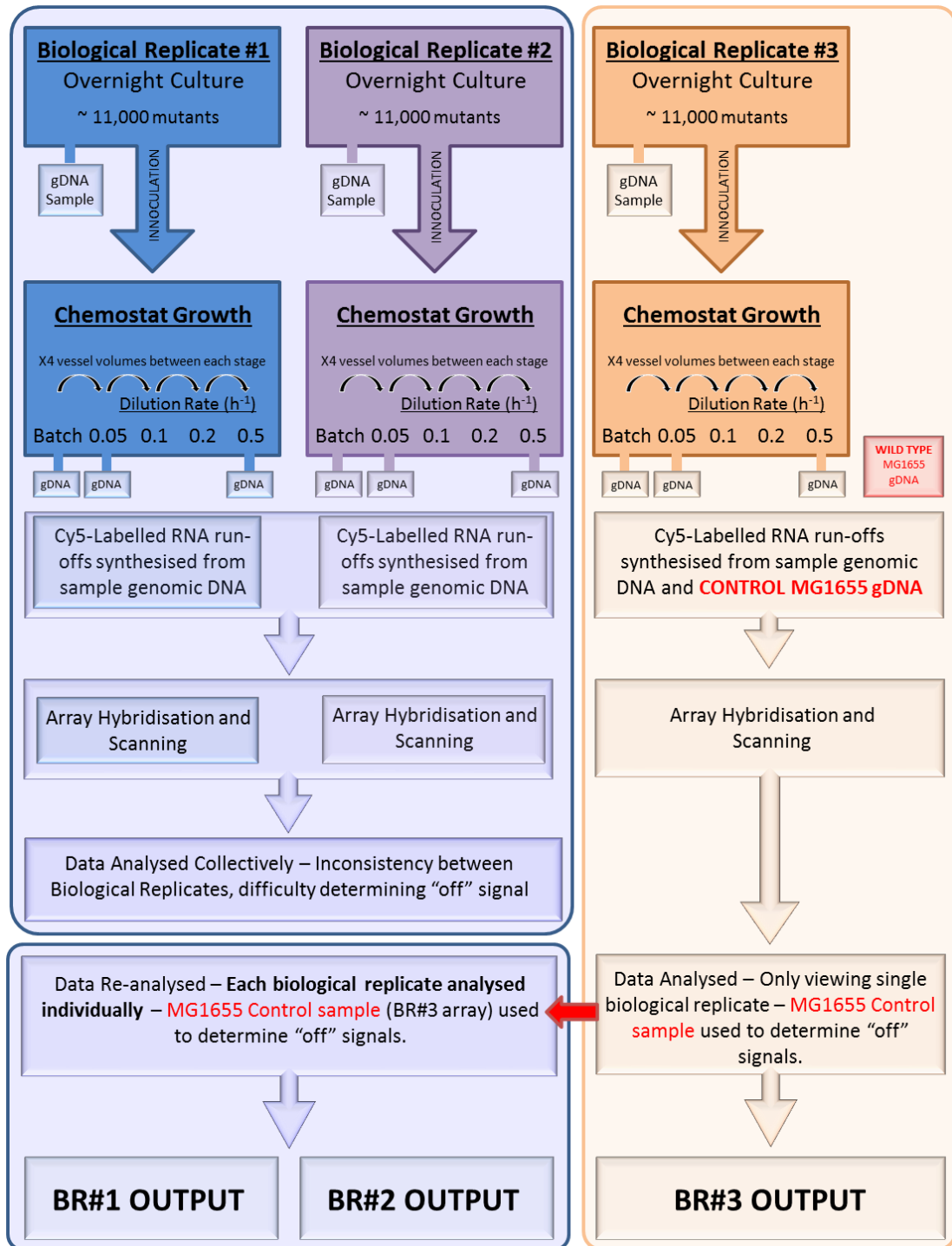
The method is based on the procedure developed by Chaudhuri *et al.* (Chaudhuri *et al.*, 2009a). Transposon-Mediated Differential Hybridisation (TMDH) enables the mapping of transposon insertions, in a culture of many individual transposon mutants grown together, to their location in the genome by synthesis of fluorescently-labelled run-offs from an outward facing T7-promoter engineered within the transposon itself. This procedure has already been utilised to identify genes that are essential for *Staphylococcus aureus* survival (Chaudhuri *et al.*, 2009a) and genes essential for *Salmonella enterica* Serovar Typhimurium infection of mice (Chaudhuri *et al.*, 2009b).

In this work a library of 11,616 *E. coli* MG1655 mutants (made by Mr. I R Kean) each with a single, randomly inserted Tn5-transposon mutation, were grown in a chemostat culture wherein the dilution rate was increased in a step-wise manner from  $d = 0.05 \text{ h}^{-1}$  through  $d = 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$ . A minimum of four vessel volumes

of fresh Evans minimal medium were exchanged before cell pellet samples were taken at each stage. Samples were also obtained from the original inoculum (a 50 ml overnight culture grown in LB) and after initial batch growth in the chemostat. The hypothesis was that after establishment of a culture containing mutants capable of growth at  $d = 0.05 \text{ h}^{-1}$ , as dilution rate was raised, mutants lacking genes advantageous to growth at these increased rates would have diminished residence within the culture as a whole and would eventually be washed out of the chemostat. A work-flow of the procedure that will be discussed is summarised in Figure 4.6, where problems encountered are also mentioned.

Using the TMDH method, a transposon insertion that permitted growth at  $d = 0.05 \text{ h}^{-1}$  and was then disadvantageous to growth at  $d = 0.5 \text{ h}^{-1}$ , could be mapped to the *E. coli* MG1655 genome; and the disrupted gene revealed. Mapping of the transposon insertions involved isolation of the total genomic DNA of the culture sample, digestion with AluI (AG<sup>^</sup>CT) restriction enzyme and then use of this digested genomic DNA as a template for T7 RNA Polymerase production of Cy5-fluorescently labelled run-offs from the T7 promoter within any transposons. The restriction digest with AluI is critical within the TMDH procedure, as it prevents products from T7 promoters in non-essential genomic loci extending through into adjacent essential gene areas. Labelled cRNA samples were then hybridised to microarrays, these arrays represented the entire 4.6 Mb genome of *E. coli* MG1655 (NCBI RefSeq: NC\_000913.2) and were designed by Genotypic Technology Ltd. Each probe was 60 bases in length, spaced approximately 150 bp apart on the genome, on both the sense and anti-sense strands. As a result of this, each array was made up of 61,824 probes (30,912 sense probes and 30,912 antisense probes) representative of the entire *E. coli* MG1655 genome in a tiled format.

After completing two biological replicates of the chemostat culture phase and the subsequent labelling, array hybridisation and scanning phases (Figure 4.6), the development of a method for data filtering highlighted that there appeared to be a low level of reproducibility between the two biological replicates. It also highlighted the difficulty in establishing what was to be classed as an “off” signal. As is stated in Chaudhuri *et al.*, “for TMDH the primary interest is in a discrete binary property” (Chaudhuri *et al.*, 2009a), which is the presence or absence of a transposon producing a signal on the resultant array. The lack of reproducibility and the need for



**Figure 4.6 Schematic work-flow of the stages leading to eventual data output from TMDH study of *E. coli* strain MG1655 gene fitness over increased growth rate.** Of note is the separation of Biological Replicate #3 from the rest of the work carried out. This is due to the fact that initial analysis of data from BR#1 and BR#2 warranted the undertaking of a third biological replicate and scanning of a third array slide including a wild type *E. coli* MG1655 control T7-run-off. Control sample data was then applied to BR#1 and BR#2 datasets retrospectively.

a method of identifying any background signal led to a third biological replicate, and also obtaining a dataset for T7 RNA polymerase products from an *E. coli* MG1655 genomic DNA template with no transposon insertions present.

With the third biological replicate completed, and a dataset for T7 cRNA run-off from a control wild type *E. coli* MG1655 background, a method for data analysis was established. Taking the raw microarray signals for each biological replicate in isolation, a median shift was applied to all samples under test (i.e. (1) LB grown inoculum, (2) minimal medium batch culture, (3) after growth at  $d= 0.05 \text{ h}^{-1}$ , (4) after growth at  $d= 0.5 \text{ h}^{-1}$  and (5) the control *E. coli* MG1655 sample). Probe signals were then expressed as a fold difference compared to the *E. coli* MG1655 background signal. A probe was classed as “on” if it exhibited a signal to background ratio  $\geq 2$ , and “off” if this ratio was  $\leq 1$ . Those with a signal to background ratio between 1 and 2 were classed as inconclusive.

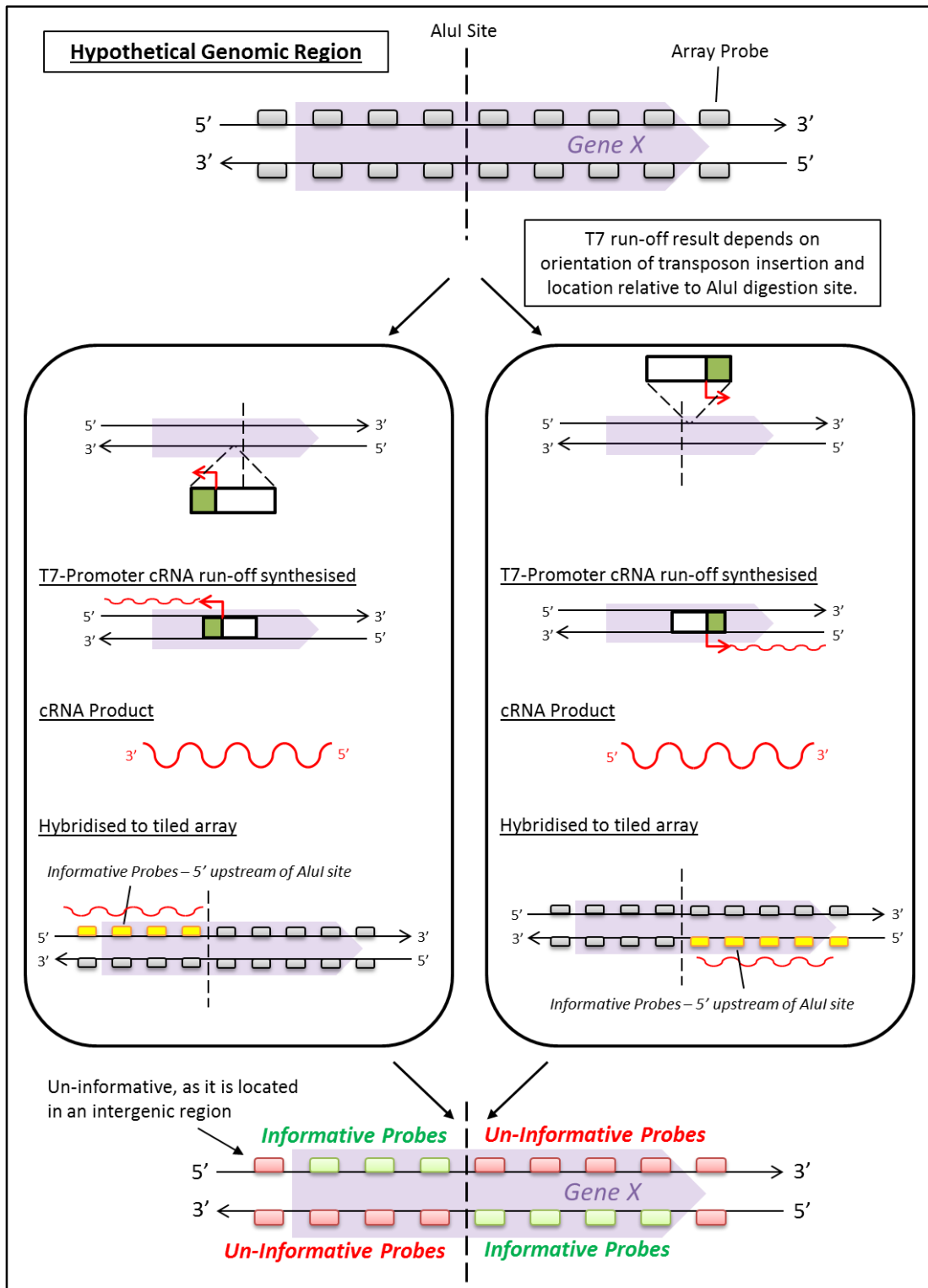
Another filtering step was required in order to remove probes that were non-informative. The various stages of this filtering process are outlined in Table 4.6. A probe was classed as informative if, when taking into account its location within the genome and the position of all AluI digest sites (AG<sup>^</sup>CT), it was representative of only one gene (not in a gene overlap region), was situated entirely *within* a gene, was 5'-downstream of any intragenic AluI sites (more detail on this specific criterion can be found in Figure 4.7) and did not have an AluI site within its sequence. If a probe was situated within a gene that did not contain an intragenic restriction site, it was only classed as informative if it was the lone gene found within that restriction fragment. It must also be noted that since the array was designed based on the *E. coli* MG1655 genome as defined by the NCBI RefSeq NC\_000913.2, this genome definition was updated as of November 2013 to NCBI RefSeq NC\_000913.3. This meant updating the chromosomal locations of all probes and also resulted in the removal of 6 probes due to base changes.

With a full set of data and method to focus on only those probes that were determined to be informative, a simple test was applied; informative probes that were seen to produce a significant signal (i.e.  $\geq 2$ -fold) for the LB-inoculum sample, minimal medium batch sample and the  $d= 0.05 \text{ h}^{-1}$  sample and then have a significant “off” signal (i.e.  $\leq 1$ -fold) when grown in the chemostat at  $d= 0.5 \text{ h}^{-1}$  were

**Table 4.6 Stages in filtering out un-informative probes from subsequent analysis.**

From the total 61,824 probes on the tiled *E. coli* MG1655 genome array, a total of 28,177 probes were removed for any one of the six reasons detailed below. More detail on the criterion that resulted in the removal of 12,990 probes in Stage 4 can be seen in Figure 4.7.

Stage	Reason for probe removal from analysis	Number of Probes Removed	Running Total
	Total number of probes on array.		61,824
1	Removed due to base changes in new <i>E. coli</i> MG1655 genome annotation.	6	61,818
2	Removed as they represented intergenic regions.	9192	52,626
3	Removed as they represented a region where more than one gene overlapped.	168	52,458
4	Probes representing genes that contained an AluI site but were 5'-upstream of intragenic AluI site.	12,990	39,468
5	Probes within genes that did not contain an AluI site, and were not the <u>only</u> gene within its restriction fragment.	3,494	35,974
6	Probe itself contained an AluI site, and < 40 bp of probe was outside that site.	2327	33,647



**Figure 4.7 Rationale behind crucial criterion for classing probes as ‘informative’ or ‘un-informative’.** As transposons can insert in two opposite orientations, and produce labelled cRNA run-off irrespective of gene orientation, only those probes that were 5'-upstream of intragenic AluI digest sites were classed as informative. Those sites that are 5'-downstream can be influenced by run-off from an intergenic transposon outside of the gene and produce a false positive.

interpreted as being influenced by a transposon present within genes that are essential for growth at the increased growth rate. This criterion was applied to each biological replicate dataset individually.

Biological Replicate #1 produced a list of 42 informative probes with “off” signals at  $d= 0.5 \text{ h}^{-1}$ , representative of 36 individual genes. Biological Replicate #2 produced a list of only 2 informative probes representative of 2 individual genes. Finally, Biological Replicate #3 produced a list of 639 probes representative of 408 individual genes. Clearly, there is a large disparity between the three biological replicates, with very little crossover between these experiments. Seven of the genes following the desired trend in Biological Replicate #1 were also observed in the data-set for Biological Replicate #3, and neither of the genes identified in Biological Replicate #2 were present in either of the other replicates. A full list of the informative probes expressing the trend of diminished signal only at  $d= 0.5 \text{ h}^{-1}$  in each Biological Replicate data-set can be seen in the Supplementary Data 4.4.

#### **4.9. Output from a preliminary TMDH study of *E. coli* strain MG1655 comparing lower and higher growth rates**

Despite the lack of reproducibility between the replicates, the list of 408 genes identified in Biological Replicate #3 was probed to determine whether there was any crossover with the full list of those genes that were significantly up-regulated at increased growth rates (detailed in Section 4.4, gene list can be found in Supplementary Data 4.2). The justification for analysing the Biological Replicate #3 data-set was that the array slide on which all the samples of this particular replicate were hybridised also contained the control wild type *E. coli* MG1655 sample that was utilised for normalisation and removal of background signal. The fact all Biological Replicate #3 samples and the control sample were hybridised and scanned within the same technical replicate may go some way towards explaining the shortcomings of the procedure in its current form and lack of consistency between biological replicates. Further discussion as to how the procedure may be improved can be found in Section 4.10.

Genes that were: (1) observed to increase in transcript abundance as growth rate increased in gene expression microarrays and were (2) potentially essential for

growth at  $d = 0.5 \text{ h}^{-1}$  in Biological Replicate #3 of the TMDH study are listed in Table 4.7.

A total of 10 genes that were classed as being essential for growth at increased dilution rate in the TMDH study were also up-regulated at increased rates of growth in the gene-expression data (Section 4.4). Four of these genes (*flgK*, *fliC*, *tar* and *tsr*) are involved in flagellar assembly and regulating the chemotactic response, a functional group previously seen to be over-represented in the gene-expression work described in Sections 4.2 and 4.3. The functional group of lipopolysaccharide biosynthesis genes that was previously seen to be significantly up-regulated at doubling times  $\leq 3.5 \text{ h}$  is also represented in the TMDH analysis by the gene *rfbC* which encodes a dTDP-6-deoxy-D-glucose-3,5 epimerase (Stern *et al.*, 1999). This is an important component for the eventual synthesis of the O-antigen of *E. coli*, though *E. coli* K-12 has long been known to not express the end product of this biosynthetic pathway due to mutations in the *rfb* cluster (Stevenson *et al.*, 1994). Despite the fact that O-antigen synthesis may not be relevant to the specific strain in this study, this finding may still be relevant to those strains that do express the antigen. Also, the disruption of *rfbC* on its own has been shown to diminish resistance to stress from multiple sources in *E. coli* K-12, including UV, mitomycin C and hydrogen peroxide (Han *et al.*, 2010). Another gene in the same operon as *rfbC*, *rfbA*, was identified in the Biological Replicate #1 dataset. This gene encodes dTDP-glucose pyrophosphorylase which is involved in the biosynthesis of dTDP-L-rhamnose a precursor of L-rhamnose which is an essential component of surface antigens including the O-antigen mentioned above (Marolda and Valvano, 1995).

The gene *proP* appeared in both the TMDH work (in Biological Replicate #1 and #3) and the gene expression transcriptomics. This gene encodes an osmoprotectant/proton symporter associated with the uptake of the amino acid proline, and also glycine betaine, that is seen to be upregulated in conditions of both hyperosmolarity and also amino acid starvation (Mellies *et al.*, 1995).

Other genes that are present in both studies *purC* that encodes a protein vital in purine biosynthesis (Zhang *et al.*, 2008), *ompW* which encodes a component of the receptor for colicin S4 (PilsI *et al.*, 1999), *ybbM* (a.k.a. *fetB*) which encodes a component of the FetAB predicted ABC iron export transporter important in



**Table 4.7 Compilation of genes that were observed to be essential in Biological Replicate #3 of the TMDH study, and were up-regulated as growth rate increased in gene-expression microarrays.** A list of genes that contained probes that expressed a transition from significantly “on” at  $d=0.05\text{ h}^{-1}$  to significantly “off” at  $d=0.5\text{ h}^{-1}$  in Biological Replicate #3 of the TMDH study (detailed in Section 3.6) were studied for crossover with those genes observed to be significantly up-regulated in the gene expression microarray approach (detailed in Sections 3.1 and 3.2). Gene *proP* (in bold) was also identified in Biological Replicate #1 of the TMDH study, and *rffC* (\*) is located in an operon with *rffA*, another gene that was identified in Biological Replicate #1 of the TMDH study.

Also shown are (A) the total number of informative probes associated with the indicated gene before filtering for data trends, (B) those probes that express a signal to background ratio  $\geq 2$  in the LB inoculum sample (i.e. are classed as “on”) and (C) those that maintain this “on” state until growth at  $d=0.5\text{ h}^{-1}$  wherein they express a signal to background ratio  $\leq 1$  (i.e. are classed as “off”).

Gene Name	Functional Grouping	A) Total number of Informative Probes	B) Informative probes that are "on" in LB inoculum	C) Probes that show "off" signal on transition from $d=0.05$ to $0.5\text{ h}^{-1}$
<b><i>proP</i></b>	Amino Acid Biosynthesis and Related	16	3	2
<i>rffC</i> *	LPS	3	3	1
<i>flgK</i>	Motility / Chemotaxis	21	2	2
<i>fliC</i>	Motility / Chemotaxis	16	2	2
<i>tar</i>	Motility / Chemotaxis	21	4	3
<i>tsr</i>	Motility / Chemotaxis	20	2	1
<i>purC</i>	Nucleotide Metabolism	9	1	1
<i>ompW</i>	Transport	5	1	1
<i>ybbM</i>	Stress Resistance	8	2	1
<i>ydiY</i>	Unknown	21	4	2

resistance to hydrogen peroxide stress (Nicolaou *et al.*, 2013) and, lastly, *ydiY* which encodes a protein of unknown function.

For all 10 of these genes, a significant “on” signal was never observed for any associated informative probe at  $d= 0.5 \text{ h}^{-1}$  when an “on” signal had been observed for the LB inoculum sample. Also displayed in Table 4.7 are (A) the total number of informative probes associated with the genes mentioned above, (B) the number of probes that were seen to express a significant “on” signal in the LB-grown inoculum sample, and subsequently (C) the number of probes observed to be significantly “off” on the transition from growth at  $d= 0.05 \text{ h}^{-1}$  to growth at  $d= 0.5 \text{ h}^{-1}$  (Table 4.7, columns (A), (B) and (C) respectively). The numbers highlight how the data trend observed is often seen in a low proportion of the total number of informative probes associated with a gene of interest. However, this is not inexplicable, as cRNA fluorescence signal hybridised to a given probe is entirely dependent on where in the gene the transposon has inserted and the orientation of that insertion (and thus which strand the T7 promoter run-off is homologous to). It may be prudent to form a scoring system that takes into account signal trends seen in all other probes associated with a gene, which will streamline the analysis of larger datasets, such as the full list of 408 genes classed as being advantageous to fast growth in the complete Biological Replicate #3 dataset. Before this is developed though, the fundamentals of producing consistent biological replicate data-sets must be addressed.

#### **4.10. Discussion**

A gene expression profile as it pertains to growth rate, and therefore *E. coli* doubling time has been produced for the specific K-12 strain, MG1655. Given the controlled rate of dilution with carbon-limited minimal medium containing 20 mM glucose, and the obtained cell dry weight at each dilution rate, multiple carbon utilisation parameters have been quantified; rate of glucose uptake, rate of biomass synthesis, rate of acetate synthesis and the maximum flux through the TCA cycle were all calculated. The estimated maximal rate of carbon flux through the TCA cycle was in the region of  $28 \text{ mmol C l}^{-1} \text{ h}^{-1}$ , in agreement with previous flux balance studies carried out with *E. coli* (Holms, 1996). Though in-depth quantification of carbon utilisation by *E. coli* grown at different growth rates was not

the primary aim of this work, it has been the focus of many other studies (Holms, 1996; Vemuri *et al.*, 2006; Nahku *et al.*, 2010). However, little in the way of a definitive statement as to the maximum flux of carbon through the TCA cycle has been made. The data provided here is a good basis for such an investigation, though it would be prudent to carry out similar growth of *E. coli* at dilution rates exceeding  $d = 0.5 \text{ h}^{-1}$  to clarify whether the observed maximum flux through the TCA cycle is conserved, and consistently offset by an increase in the proportion of acetate produced.

Though more detail is provided here for those genes deemed to be *significantly* affected by an increase in dilution rate, transcriptomic profiles have been produced for 4,311 genes at dilution rates of 0.05, 0.1, 0.2 and  $0.5 \text{ h}^{-1}$  (Supplementary data 4.1). Also, as has been demonstrated by the arginine biosynthetic pathway (Figure 4.4), just because a gene is not observed to have a significant change in transcript abundance as judged by the statistical filtering used in this work, does not mean that there is no correlation between expression and varying growth rate.

Once a statistical filter was applied, a total of 86 genes were observed to be up-regulated by an increase in dilution rate, and a total of 167 genes were observed to be down-regulated.

Of those genes that displayed an increased abundance at higher dilution rates there was a striking over-representation of genes associated with flagella assembly, motility and chemotaxis. In terms of functional groups, motility and chemotaxis is one that was considered surprising considering the conditions in which the bacteria were cultured (i.e. constant mixing with no gradients of metabolites). Further work into confirming the increased expression of flagella in response to increasing growth rate is described in Chapter 5, though it is postulated here that the reason for the increased abundance in genes across this functional group may be linked to the observed down-regulation in the *lsr* operon genes (of which *lsrA*, *lsrC*, *lsrD*, *lsrF* and *lsrR* have an observed significant down-regulation). This operon encodes the uptake machinery of *E. coli* for the quorum sensing molecule autoinducer-2 (AI-2) (Xavier and Bassler, 2005), a molecule that has been observed to up-regulate flagella synthesis and motility when it is present in high enough concentrations outside the

cell (Sperandio *et al.*, 2002). If the import machinery for this signalling molecule is down-regulated at higher dilution rates, it is logical that there would be a relative build-up of extracellular AI-2 in the increased growth rate conditions. Further analyses of the implications of the down-regulated *lsr* operon are also reported in Chapter 5.

Transcripts associated with the conversion of *E. coli* into a fimbriated state were also observed to increase in abundance as growth rate increased. As mentioned previously this may be a by-product of the increase in cell division rate, and hence overlapping rounds of DNA-replication, causing the inversion of the regulatory *fimS* region to occur at a relatively high frequency (Adiciptaningrum *et al.*, 2009). With the data presented here it cannot be claimed that this triggers a true conversion to fimbriated *E. coli* when growing with doubling times of  $\leq 3.5$  hours, but it can be said that there is a relative up-regulation in some genes associated with this state change (namely the *fimAICDFGH* operon).

Multiple genes associated with various steps in synthesis of the lipopolysaccharide layer of *E. coli* K-12 were also seen to increase in expression as growth rate increased. This prioritisation of resources is likely highlighting its importance in maintaining cell integrity and survival, especially at higher cell doubling rates.

A large proportion of transcripts that exhibited a down-regulation in abundance as growth rate increased encoded proteins involved in either secondary metabolism, transport of alternative metabolites or regulation of genes associated with related functions. Also, within these functional groups a striking proportion was ultimately regulated by the availability of cAMP, an intracellular signal whose build-up within the cell is inversely proportional to the rate of dilution with glucose limited minimal medium (Kuo *et al.*, 2003), and therefore inversely proportional to the rate of glucose consumption. As it has been demonstrated that glucose uptake rate is increased as growth rate is increased in this work, the intracellular concentration of cAMP will decrease accordingly. Hence, in relative terms, cells that are grown at the slower dilution rates utilised in this work are exhibiting a trait of scavenging for alternative sources of metabolite, and though the medium in which the cells are

grown is well defined, and specifically contains little in the way of alternative sources of carbon, scavenging from the remains of dead cells is a possible scenario.

When the total transcriptomic dataset was analysed using TFinfer software, it was possible to infer the activity of 167 transcription factors over the range of growth rates tested. This highlighted a common trend in which the activity of a transcriptional regulator was exhibiting a defined shift in activity between the dilution rates  $d= 0.1$  and  $0.2 \text{ h}^{-1}$ . A trend that was also noted in the general transcriptional profile of genes in the transcriptomic analysis itself. This might indicate that, in terms of steady-state grown cultures, there is a step change in gene expression between those cells that have a doubling time  $\geq 6.9 \text{ h}$  and those that have a doubling time  $\leq 3.5 \text{ h}$ , a shift in which a noticeable proportion of regulation and transcript abundance changes are dedicated to starvation responses and utilisation of secondary metabolites. This observation is coupled with the inferred activity of cAMP-CRP itself decreasing between the dilution rates of  $d= 0.1$  and  $0.2 \text{ h}^{-1}$ . This step change may be occurring at a threshold possibly dictated by the ratio of an individual cell's time spent in a completely glucose starved environment to the rate of glucose being fed into the culture environment. It would thus be intriguing to do similar global gene transcriptomic studies on cultures grown at intermediate dilution rates between  $d= 0.1$  and  $0.2 \text{ h}^{-1}$  in order to extrapolate whether this adaptive shift takes place progressively, or occurs as a distinct step change at a defined dilution rate threshold.

The transcriptomic data and subsequent TFinfer analyses also emphasised, for example in the case of *cspD* expression at slow growth rates, how a gene that is currently stated to be expressed in stationary-phase cells (Yamanaka *et al.*, 2001) may indeed still be expressed in cells undergoing slow steady-state growth. This dataset provides a good source for identifying those genes that may be expressed both in stationary phase and slow growth. It must be stated here that all transcriptomic and inferred regulator activities presented in this work are relative values, and while a gene such as *cspD* may be exhibiting relative up-regulation at slower growth rates it may not be to a physiologically significant level compared to its degree of expression in stationary-phase cells.

The relative activities of the alternative sigma factors FliA and RpoS have also been inferred in this work and are suggested to be important for regulation of genes associated with adaptation to changes in dilution rate in steady-state grown cultures; with FliA ( $\sigma^{28}$ ) showing increased influence at high growth rates, and conversely RpoS ( $\sigma^s$ ) showing increased influence at the slower growth rates. Both sigma factors exhibit the same shift in activity described previously between  $d= 0.1$  and  $d= 0.2 \text{ h}^{-1}$ . RpoS, much like CRP, is associated with the activation of genes associated with secondary metabolism and also stress responses (Maciag *et al.*, 2011). The fact that it activates transcription of a higher proportion of genes at the lower dilution rates is in consonance with CRP's observed increased activity at slower growth rates. The activity of FliA on the other hand is inferred to be higher at the increased dilution rates largely because of the observed increase in genes associated with flagella biosynthesis and motility, reasons why this may be the case at higher rates of growth are explored further in Chapter 5.

In order to complement the data from the growth rate transcriptomic work, steps were made to develop the use of a transposon mutant library to test gene fitness and essentiality amongst a possible total pool of 11,616 uncharacterised random transposon insertion mutants as dilution rate was increased in a step-wise manner. This work was based on the transposon mediated differential hybridisation (TMDH) method established by Chaudhuri *et al.* (Chaudhuri *et al.*, 2009a; Chaudhuri *et al.*, 2009b). Though the data output was ultimately inconsistent between separate biological replicates (i.e. cultures grown) significant steps were made in terms of experimental methodology of data analysis and sample handling. Namely, the establishment and implementation of various criteria necessary for determining informative probes for such an analysis on a tiled array harbouring the entire *E. coli* strain MG1655 genome. The necessity of a dataset for the T7 run-off for a control wild type *E. coli* strain MG1655 background was also established in this work, and it is suggested that in future such a sample is run alongside every sample under investigation. That is to say, as in the gene expression transcriptomics carried out in the work included in this chapter, a control sample labelled with the fluorescent label Cy3 should be hybridised to each array in tandem with each sample being tested. The presence of a consistent background sample will aid in normalisation between biological replicates and account for any differences in separate technical replicates.

It may also be prudent to run a scaled down version of the procedure attempted here. Perhaps fewer uncharacterised mutant strains grown alongside a characterised mutant strain harbouring a mutation for a gene that would be predicted to be necessary for either growth at higher dilution rates, or just in the transfer from growth in the LB inoculum to batch growth in glucose-limited minimal medium (e.g. genes necessary for amino acid biosynthesis or glucose uptake). This would provide definitive evidence as to whether the data output matches a predictable hypothesis.

Though an output was produced for Biological Replicate #3 of the TMDH study carried out in this work, in-depth analysis of the significance of the total 408 genes that are possibly advantageous to growth at higher growth rates is ill-advised until multiple consistent biological replicate datasets are obtained. It must also be stated, if development of this high throughput method of mutant library analysis proves fruitless, the use of a transposon specific primer in order to carry out sequencing of regions flanking insertion sites and subsequent mapping of those sites is always a possibility.

Recently a comparative study of transcriptomic and proteomic changes experienced by *E. coli* MG1655 in response to changes in growth rate was carried out by Valgepea *et al.* (Valgepea *et al.*, 2013). Samples were taken from a glucose-limited (25 mM glucose) accelerostat *E. coli* MG1655 culture at  $\mu$  values of 0.21, 0.31, 0.4 and 0.48 h<sup>-1</sup>. Changes in transcript abundance were calculated relative to those of a steady-state culture with  $\mu = 0.11$  h<sup>-1</sup>. In comparison to the steady-state transcriptional reprogramming reported here, in which 86 genes were up-regulated and 167 genes down-regulated, the accelerostat data showed 1,484 genes up-regulated and 112 genes down-regulated ( $\geq 2$ -fold change at  $\mu = 0.48$  h<sup>-1</sup> relative to  $\mu = 0.11$  h<sup>-1</sup>). These differences could arise from the different culture conditions and the methods of data analysis (for the accelerostat study all samples were obtained from a single culture). Nevertheless, comparison of the two datasets was undertaken Supplementary Data 4.5. Among the 86 genes designated as up-regulated as growth rate is increased here, there were no cases of a reduction in transcript abundance at  $\mu = 0.48$  h<sup>-1</sup> relative to  $\mu = 0.11$  h<sup>-1</sup> in the accelerostat study (51 of the 86 [59.3%] genes exhibited a  $\geq 2$ -fold increase in transcript abundance at  $\mu = 0.48$  h<sup>-1</sup>). The most up-regulated transcript in this work, *azuC* (6.1-fold at  $d = 0.2$  h<sup>-1</sup>), is the product of a small gene (87 bp) of unknown function, consisting of both a sRNA (*isrB*) and a 28

amino acid membrane associated protein (AzuC). However, this gene and other sRNA encoding genes, which were custom built into the array chips used here, were not represented on the standard microarray platform used to determine the transcript profiles from the accelerostat study. Amongst the 95 genes in the down-regulated list produced here, 25 are 2-fold down-regulated (27.5 %), but 9 (9.5 %) are 2-fold up-regulated, at  $\mu = 0.48 \text{ h}^{-1}$  in the accelerostat study. This comparison is compromised because the accelerostat data are often incomplete and only those genes with complete data are included in the comparison. This lack of data for certain genes is a problem highlighted by the genes of the *lsr* operons, that were significantly down regulated as growth rate increased in this work. Measurement values for many of the genes of the *lsrACDBFGtam* operon are absent in the accelerostat study, and in several cases these are contradictory (e.g. *lsrA* has a 7.6-fold increase in transcript abundance at  $\mu = 0.21 \text{ h}^{-1}$  relative to  $\mu = 0.11 \text{ h}^{-1}$ , with no subsequent data at the increased growth rates; whereas *lsrF*, present in the same operon, was down regulated, as observed here). Such inconsistent behaviour of genes within the same operon is a potential concern that was observed in interpreting the accelerostat data.

One of the main conclusions to emerge from the comparison of transcript and protein abundances in the accelerostat study was the apparent dominance of post-transcriptional control in metabolic adaptations in response to increased growth rate (Valgepea *et al.*, 2013). Comparison of the transcriptomic data obtained in this work with the proteomic data presented in Valgepea *et al.* (2013) (Supplementary data 4.6), revealed that 22 of the 86 up-regulated genes and 36 of 167 down-regulated genes have associated proteomic data trends. The proteomic data presented are expressed as fold changes at  $\mu = 0.21, 0.31, 0.4$  and  $0.49 \text{ h}^{-1}$  relative to samples gathered at  $\mu = 0.11 \text{ h}^{-1}$  (Valgepea *et al.*, 2013). Of the 22 up-regulated genes, only three have an incomplete data set (*deaD* has data for  $\mu = 0.11$  and  $0.21 \text{ h}^{-1}$  only, whilst *mmuM* and *yjdM* have no data at  $\mu = 0.4$  and  $0.49 \text{ h}^{-1}$  respectively). The general trend for the 22 up-regulated genes was either no change in protein level or a slight increase at higher growth rates, with exceptions (including *dld*, *mmuM*, *proA* and *yjdM*) representing possible examples of increased protein turnover at the higher growth rates. The proteomic data associated with genes that were classed as being down-regulated at increased growth rate in this work showed strong agreement. Of 167 genes, 36 have



associated proteomic data, all of which showed a downward trend in protein detected as growth rate was increased. Again, there are cases where an incomplete data set was obtained (13 out of the 36 genes), though only in the case of *ytfR* where measurements for only  $\mu = 0.11$  and  $0.21 \text{ h}^{-1}$  were provided does that make it unwise to infer a trend in protein level regulation in relation to growth rate. The gene *gadB* had proteomic data for  $\mu = 0.11$  and  $0.4 \text{ h}^{-1}$  only, but with a significant 0.3-fold decrease in protein level, it can be said that intracellular level of GadB is reduced at increased growth rate. Interestingly, the proteomic data for *lsrA*, *lsrF* and *lsrK* all show a downward trend in relation to growth rate, in accord with the transcript abundance measurements reported here. The relatively strong correlation between transcripts exhibiting decreased abundance that are reported here and decreased abundance of the corresponding proteins in the Valgepea *et al.* (2013) study suggests that the influence of transcription repression in lowering protein abundance/activity might be greater than previously thought (Valgepea *et al.* 2013). Unfortunately, no proteomic data was reported for those proteins associated with flagella biosynthesis that were analyzed here by Western blotting (Chapter 5).

In summary, although there is a degree of overlap between the data reported by Valgepea *et al.* (2010; 2013) it is argued that the transcriptomic data set reported here is; (1) more comprehensive due to the presence of probes representative of sRNAs and lack of gaps in data obtained; and (2) has an increased reliability and physiological relevance due to the greater number of biological replicates examined. Proteomic data obtained by the same group appears to show good correlation with the down-regulated transcripts reported here, although the size of the data set is relatively limited.

## 5. Studying *Escherichia coli* flagellar gene expression and autoinducer-2 production in response to growth rate

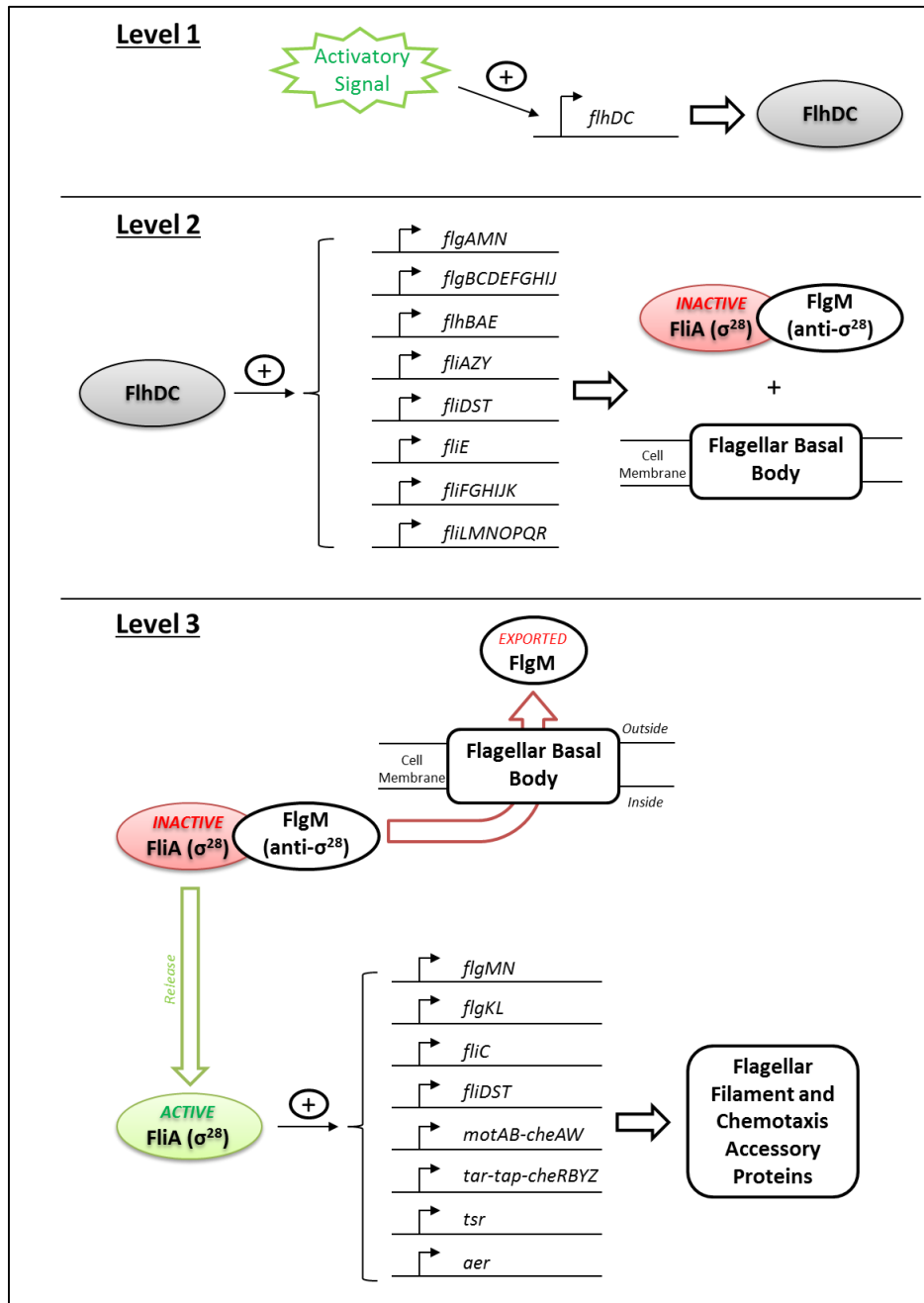
### Main findings

- Increased expression of flagellar complexes was observed as growth rate was increased
- Quantified relative levels of proteins responsible for the regulation of the flagellar regulon over increasing growth rates
- Rate of extracellular autoinducer-2 accumulation increased as growth rate increased, likely a product of both increased LuxS activity and decreased uptake through the machinery encoded by the *lsr* operons
- Increased flagellar expression at increased growth rate deemed as unlikely to be linked to AI-2 build-up

### 5.1 Introduction

Assembly of the macromolecular machinery responsible for flagella motility and chemotaxis has been heavily studied in *S. enterica* Serovar Typhimurium and *E. coli* wherein the genes involved are near identical between the two bacterial species (Macnab, 1992; Chilcott and Hughes, 2000; Chevance and Hughes, 2008). The various components necessary for regulation of flagella assembly and the structural components of the flagella themselves are encoded by over 50 genes spread predominantly across 15 operons, though there are separate operons for various specific chemoreceptors (Macnab, 1996).

Expression of the operons making up the entire flagellar regulon occurs in a hierarchical fashion (Figure 5.1), as shown by the effect of null mutations on the expression of other genes in the regulon situated in separate operons (Kutsukake *et al.*, 1990). The entire flagellar regulon can be divided into three levels, each level representing a different rank in the hierarchy. Firstly, at the apex of this regulation system are the Level 1 genes that encode the flagellar master regulator, FlhDC. This regulator can respond to signals from many sources and ultimately activate transcription of the Level 2 genes involved in construction of the membrane-bound hook-basal body structure of the flagellum (Liu and Matsumura, 1994; Claret and Hughes, 2002). It also activates transcription of two additional genes: (1) *fliA* which



**Figure 5.1 Schematic diagram of hierarchy within flagellar gene expression.** Various activatory signals initiate expression of *flhDC*, coupled with increases in FlhDC protein (all proteins depicted as ovals). FlhDC then initiates transcription of specific operons with Level 2 promoters, which leads to increased production of proteins involved in construction of the flagellar basal body as well as increased amounts of FliA ( $\sigma^{28}$ ) and its associated anti-sigma factor, FlgM. On completion of flagellar basal body structures, FlgM is exported and FliA is released and able to guide RNAP to initiate transcription of operons under the control of Level 3 promoters. This leads to production of filament proteins and completion of the flagellar macromolecular structure. Operons have been organised into the hierarchical levels as based on the review by Chilcott and Hughes (2000) and data from Ecocyc.org (Chilcott and Hughes, 2000; Keseler et al., 2013).

encodes the sigma factor ( $\sigma^{28}$ ) responsible for the activation of Level 3 genes of the flagellar regulon (Liu and Matsumura, 1995) and (2) the gene *flgM*, which encodes the anti-sigma factor for  $\sigma^{28}$  (Chadsey *et al.*, 1998; Chilcott and Hughes, 2000). It has been shown that FlgM is exported from the cell on completion of the flagellar hook-basal body complex (Hughes *et al.*, 1993; Kutsukake, 1994) such that the  $\sigma^{28}$  is then free to activate transcription of those Level 3 genes required for construction of the actual flagella filament only when a complete basal body complex is present.

The aim of this work was to investigate the observed significant up-regulation of a number of Level 3 genes associated with flagellar assembly and motility as growth rate was increased (Section 4.4). Intracellular levels of both FliA ( $\sigma^{28}$ ) and the FlgM anti-sigma factor were analysed over increasing growth rate, and increased expression of structurally complete flagella was also confirmed.

As a link between high extracellular levels of the quorum sensing molecule autoinducer-2 and increased flagellar motility has previously been observed (Sperandio *et al.*, 2002), and it has been shown that expression of the *lsr* operons associated with AI-2 import was significantly down regulated as growth rate increased (Section 4.5), it was thought that there may be a link between extracellular build-up of this quorum sensing molecule at increased growth rate and up-regulation of genes associated with motility and flagellar assembly.

The divergently transcribed *lsrABCDFG-tam* and *lsrRK* operons were significantly down-regulated as growth rate was increased (Section 4.5.) These related operons are associated with the uptake and subsequent processing of the quorum sensing signal, AI-2. Quorum sensing is a process employed by bacteria whereby a small molecule/peptide is synthesised and exported from the cell. If a sufficient population density of cells has been established the extracellular concentration of this signal reaches a minimum threshold, that is detected by the bacteria, which subsequently activate a coordinated adaptive response across the entire population (Miller and Bassler, 2001; Ng and Bassler, 2009). There have been multiple classes of autoinducer molecule reported with *N*-acylated homoserine lactone derivatives being utilised by Gram-negative bacteria, and oligopeptide signals being utilised by Gram-positive bacteria (Antunes and Ferreira, 2009; Ng and Bassler, 2009). Since the initial discovery of these two classes of signalling molecule,

a third class was characterised called autoinducer-2 (AI-2), which is regarded as a universal signal in that its synthesis is carried out by both Gram-negative and Gram-positive bacterial species (Bassler *et al.*, 1997).

In bacterial species that produce AI-2, the intermediate 4,5-dihydroxy-2,3-pentanedione (DPD) is produced by the enzyme LuxS (Schauder *et al.*, 2001). DPD then undergoes further spontaneous reactions in solution to produce various forms of the AI-2 molecule. These various forms are distinct but related, with a range of bacterial species recognising different forms (Miller *et al.*, 2004).

The protein TqsA has been shown to play a role in the export of autoinducer-2 from the *E. coli* cell (Herzberg *et al.*, 2006). Conversely, the *lsrACDBFG-tam* and *lsrRK* operons encode the machinery responsible for the uptake of extracellular AI-2 and its subsequent phosphorylation and degradation (Xavier and Bassler, 2005; Xavier *et al.*, 2007), the transcription of these divergent operons has been shown to be activated by cAMP-CRP (Wang *et al.*, 2005a; Wang *et al.*, 2005b). The functions of the individual genes *lsrA*, *B*, *C* and *D* have been inferred due to sequence similarity to their *S. enterica* homologues wherein LsrB is a periplasmic binding protein that responds to the build-up of extracellular AI-2 which is then internalised by the ATP-binding cassette transporter comprised of LsrA, LsrC and LsrD (Xavier and Bassler, 2005). Once internalised, AI-2 can then be phosphorylated by LsrK, bind to LsrR and de-repress the expression of the *lsr* operons, while LsrF and LsrG have been shown to have an additional role in the breakdown of internalised, phosphorylated AI-2 (Xavier *et al.*, 2007; Marques *et al.*, 2011).

It was the aim of the work in this chapter to clarify the apparent up-regulation of flagellar gene expression and to also quantify the level of extracellular autoinducer-2 accumulation in relation to growth rate.

## **5.2 Analysis of transcriptomic data attributed to genes associated with motility and flagellar assembly at increased growth rates**

The transcriptomic data pertaining to all genes associated with the assembly of the flagellar macromolecular structure are collected in Table 5.1. In this table the genes have been organised into operons, which have subsequently been grouped into Level 1, 2 or 3 based on their status within the flagellar gene expression hierarchy.

**Table 5.1 Fold change in transcript abundance of genes associated with the flagellar regulon, organised by promoter level within the hierarchy.** Table summarises transcriptomic data obtained from work described in Chapter 4 (Supplementary Material 4.1). Numbers represent the appropriate fold change in transcript abundance when cells are grown at  $d=0.5\text{ h}^{-1}$  relative to their abundance at  $d=0.05\text{ h}^{-1}$ . Operons have been organised into the hierarchical levels as based on the review by Chilcott and Hughes (2000) and data from Ecocyc.org (Chilcott and Hughes, 2000; Keseler et al., 2013), with each row representing a different operon. Red highlights those genes observed to be significantly up-regulated at either  $d=0.2$  or  $0.5\text{ h}^{-1}$  (Table 4.3).

<b>Operons with Level 1 Promoters</b>								
<i>flhD</i>	<i>flhC</i>							
1.05	0.83							
<b>Operons with Level 2 Promoters</b>								
<i>flgA</i>	<i>flgM</i>	<i>flgN</i>						
1.00	1.41	1.65						
<i>flgB</i>	<i>flgC</i>	<i>flgD</i>	<i>flgE</i>	<i>flgF</i>	<i>flgG</i>	<i>flgH</i>	<i>flgI</i>	<i>flgJ</i>
1.42	1.06	1.25	1.12	1.55	1.35	1.22	1.20	1.61
<i>flhB</i>	<i>flhA</i>	<i>flhE</i>						
1.03	1.10	1.37						
<i>fliA</i>	<i>fliZ</i>	<i>fliY</i>						
1.23	1.28	1.11						
<i>fliD</i>	<i>fliS</i>	<i>fliT</i>						
1.36	1.58	1.75						
<i>fliE</i>								
1.66								
<i>fliF</i>	<i>fliG</i>	<i>fliH</i>	<i>fliI</i>	<i>fliJ</i>	<i>fliK</i>			
1.06	1.47	1.25	0.83	1.78	1.18			
<i>fliL</i>	<i>fliM</i>	<i>fliN</i>	<i>fliO</i>	<i>fliP</i>	<i>fliQ</i>	<i>fliR</i>		
1.07	1.39	1.43	1.51	1.49	1.20	1.56		
<b>Operons with Level 3 Promoters</b>								
<i>flgM</i>	<i>flgN</i>							
1.41	1.65							
<i>flgK</i>	<i>flgL</i>							
2.25	1.51							
<i>fliC</i>								
1.83								
<i>fliD</i>	<i>fliS</i>	<i>fliT</i>						
1.36	1.58	1.75						
<i>motA</i>	<i>motB</i>	<i>cheA</i>	<i>cheW</i>					
1.89	1.94	1.93	2.10					
<i>tar</i>	<i>tap</i>	<i>cheR</i>	<i>cheB</i>	<i>cheY</i>	<i>cheZ</i>			
1.91	2.36	3.05	2.27	2.37	1.03			
<i>tsr</i>								
2.04								
<i>aer</i>								
1.83								

This organisation of genes within the table is based on data found on Ecocyc.org (Keseler *et al.*, 2013) and the review by Chilcott and Hughes (2000). All transcriptomic data displayed is that of each gene's fold-change in transcript abundance at  $d= 0.5 \text{ h}^{-1}$  relative to the level detected at  $d= 0.05 \text{ h}^{-1}$ , genes that exhibited significant up-regulation at either  $d= 0.2$  or  $0.5 \text{ h}^{-1}$  (Section 3.4) are highlighted.

Viewing the entire list of 52 genes in this way, taking into account the hierarchical nature of flagellar gene expression, it is noted that those genes that are changing in expression most significantly are almost entirely limited to those operons controlled by Level 3 promoters. The only exception to this is the gene *fliJ*, a member of the Level 2 promoter-controlled *fliFGHIJK* operon, which had an observed 2.1-fold up-regulation of gene expression at  $d= 0.2 \text{ h}^{-1}$ .

A significant change in the levels of transcription for the genes encoding the flagellar master regulator, FlhDC, was not observed at high growth rates. The gene *flhD* expressed a fold-change of 1.06, 1.06 and 1.05 at the dilution rates of  $d= 0.1$ ,  $0.2$  and  $0.5 \text{ h}^{-1}$  respectively, with *flhC* expressing a similarly insignificant fold-change of 0.98, 0.87 and 0.83. This suggests that the impact of this regulator on the relative expression of flagellar genes over increasing growth rate is likely controlled at a post-transcriptional level.

The 35 genes associated with the eight operons controlled by FlhDC-regulated Level 2 promoters show, generally, a slightly increased expression at the increased growth rate with an average fold-change across the entire group of 1.33, and 25.7% of those genes expressed a fold-change greater than 1.5 at  $d= 0.5 \text{ h}^{-1}$  relative to  $d= 0.05 \text{ h}^{-1}$ . This is evidence of FlhDC exhibiting at least a small degree of increased activity at the higher growth rate conditions used in this work, and is supported by the TFinfer analysis carried out in Section 4.6 (Table 4.5) implying that FlhDC does indeed exhibit increased activity at  $d= 0.2$  and  $0.5 \text{ h}^{-1}$  relative to  $d= 0.05 \text{ h}^{-1}$ .

As has already been stated, one of the genes under regulation of FlhDC is *fliA*, which encodes the sigma factor ( $\sigma^{28}$ ) responsible for activation of the final set of flagellar genes, regulated by Level 3 promoters (Liu and Matsumura, 1995). Though the transcript level of *fliA* was only increased 1.32- and 1.23-fold at  $d= 0.2$  and  $0.5$

$\text{h}^{-1}$  respectively, the average fold-change across the 20 genes associated with the eight operons controlled by  $\sigma^{28}$ -regulated Level 3 promoters was 1.83 at  $d= 0.5 \text{ h}^{-1}$ , with 85% of those expressing a fold-change greater than 1.5. FliA ( $\sigma^{28}$ ) activity has already been shown, through TFinfer analysis, to increase at higher growth rates (Section 4.7).

It should be noted that the operons *flgMN* and *fliDST* deemed to be regulated by Level 3 promoters can also be activated by FlhDC as they are downstream of Level 2 promoter regions also, and are therefore present in both lists (Table 5.1). The Level 2 operons *fliE*, *fliFGHIJK*, *fliLMNOPQR* are also listed on Ecocyc.org as being regulated by the presence of  $\sigma^{28}$ , as investigated in other work (Kutsukake and Iino, 1994; Liu and Matsumura, 1996; Kalir and Alon, 2004), though it has been shown that in the case of *fliDST* and *fliE* this observed FliA-related activation may be due in some part to increases in *fliZ* transcription, which is located in the same operon as *fliA* (Saini *et al.*, 2008). Also, where  $\sigma^{28}$  activation of Level 2 promoters has been observed, FlhDC is still required (Kutsukake and Iino, 1994). It is clear from these statements that regulation of the flagellar regulon is by no means simple, however it can be said that the Level 2 and Level 3 flagellar operons listed in Table 5.1 absolutely require the presence of *at least* FlhDC or FliA respectively.

### **5.3 Quantifying cellular levels of FlhDC, FliA, FlgM and FliC protein over increasing growth rate**

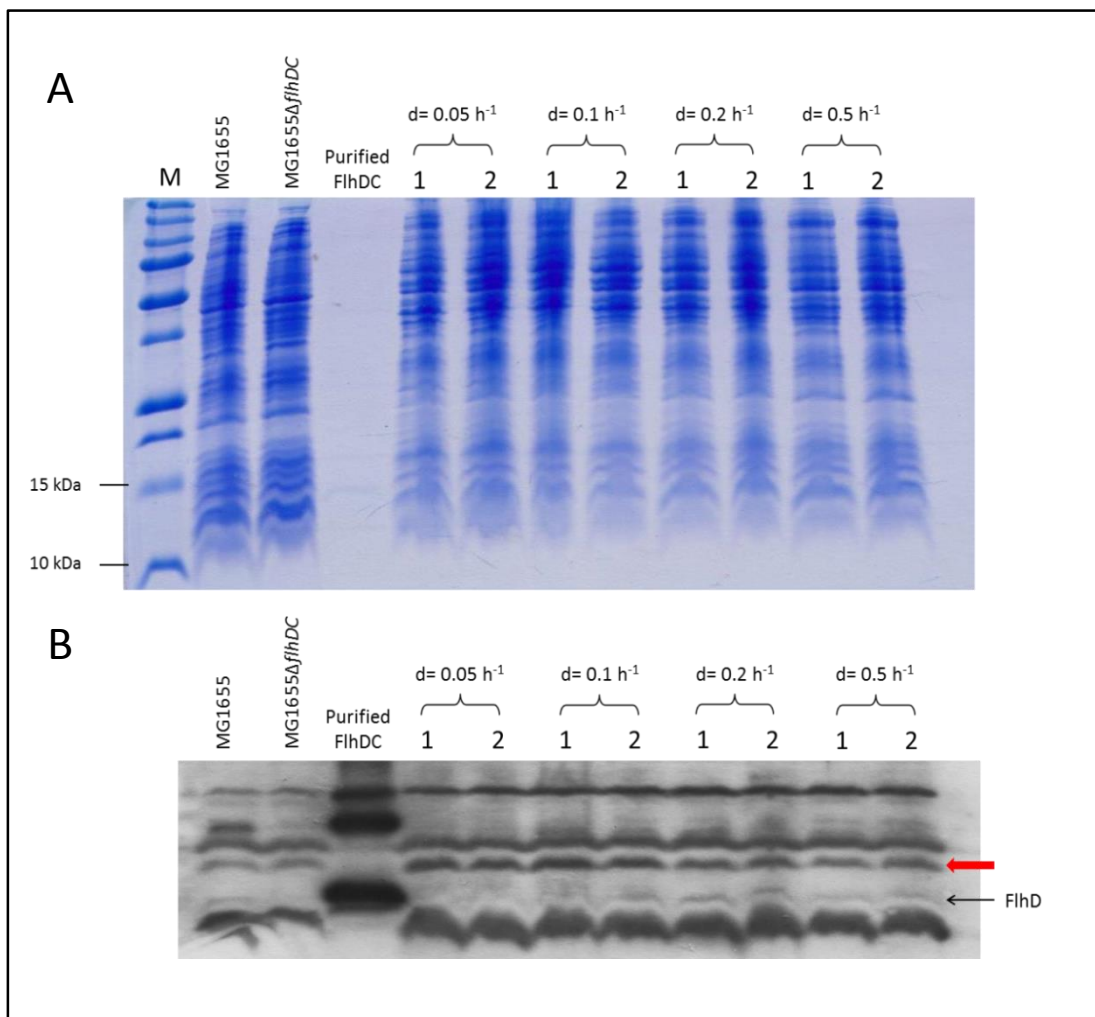
The next step in evaluating the relationship between flagellar gene expression and growth rate was to utilise cell pellet samples obtained from the same cultures grown at  $d= 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  (Section 4.2), from which the transcriptomic data sets were obtained, to quantify the relative level of various key proteins in the flagellar biosynthesis pipeline. These cell pellet samples were used in Western blot analyses that enabled estimation of cellular levels of FlhD, FliA ( $\sigma^{28}$ ), FlgM and FliC. This encompasses the proteins associated with the regulation of both Level 2 (FlhDC) and Level 3 (FliA and its anti-sigma factor FlgM) flagellar regulon promoters and levels of FliC allowed quantification of complete flagellar structures; as it represents the main filament subunit only assembled at the latter stages of flagellar construction (Macnab, 1992; Macnab, 1996).



Equal amounts of cell pellet samples were separated on a 12% SDS-PAGE gel and protein was transferred onto a Hybond-C Extra nitrocellulose membrane as detailed in Section 2.7.6. Each blot was carried out on eight samples, one per biological replicate carried out at each of the four dilution rates. The membrane with bound protein was then probed with the relevant primary antibody at its appropriate dilution (anti-FlhDC 1:5000, anti-FliA 1:1000, anti-FlgM 1:10000 or anti-FliC 1:3000). The result of these Western blots can be seen in Figure 5.1 and Figure 5.2. Samples of MG1655, MC1000 $\Delta$ *flhDC* and purified FlhDC protein (provided by Miss Nicola Whiting) were on the FlhDC Western blots. These additional samples allowed for identification of species representative of specific binding of antibody to FlhDC protein, specifically the smaller (13 kDa) FlhD subunit. The detected level of protein at each dilution rate was then measured using ImageJ software (Schneider *et al.*, 2012), the results of which can be seen in Figure 5.3A, wherein transcriptomic data for *flhD*, *fliA*, *flgM* and *fliC* are displayed alongside for comparison (Figure 5.3B).

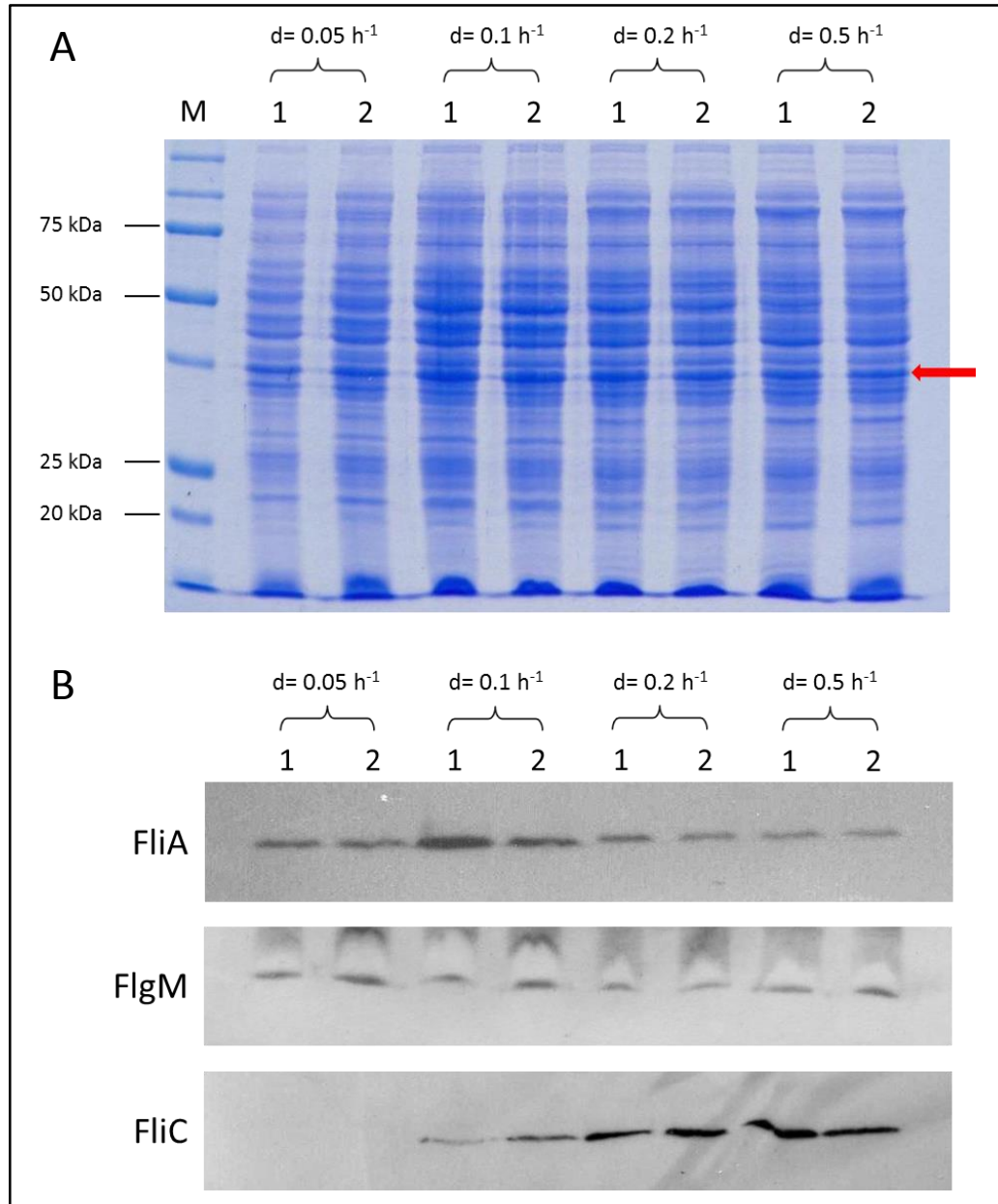
It is simplest to discuss the results of these assays by starting with FliC, a member of the Level 3 promoter controlled group of flagellar genes, and therefore one of the final components synthesised in the flagellar biosynthesis pathway. It is clear from the data in Figure 5.3 that the increase in *fliC* transcription observed at the higher dilution rates was accompanied by an expected significant increase in the FliC protein. This is evidence of increased flagellar macromolecular structures being expressed as growth rate was increased.

Working back through the flagellar regulation hierarchy, one comes to FliA and its associated anti-sigma factor, FlgM. Though the transcriptomic profiles for both of the genes show comparable trends (i.e. an increase in transcript abundance that is most significant at  $d= 0.2$  and  $0.5 \text{ h}^{-1}$ ), little similarity was observed in the relative levels of protein detected across the dilution rate range. FliA protein concentration was observed to be greatest at  $d= 0.1 \text{ h}^{-1}$  with similarly decreased amounts observed at  $d= 0.2$  and  $0.5 \text{ h}^{-1}$ . On the other hand, FlgM appeared to maintain a relatively constant intracellular concentration regardless of growth rate, exhibiting the smallest degree of variation in intracellular concentration. There are two important factors to note that may go some way towards explaining the trends observed here: (1) On completion of the flagellar hook-basal body, FlgM is known



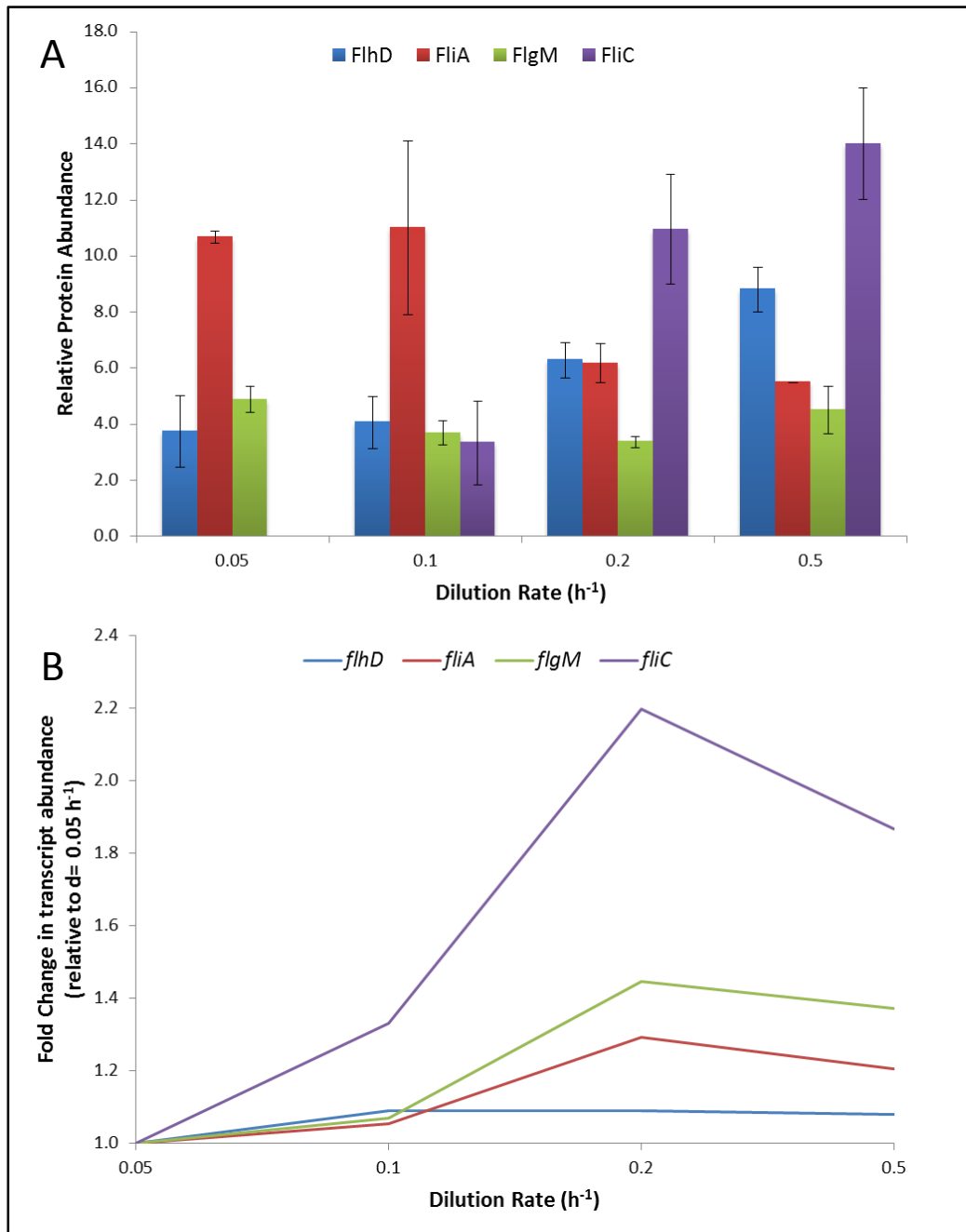
**Figure 5.2 Western blot showing amounts of FlhDC detected in samples taken from steady-state cultures grown at different dilution rates.** Upper gel (A) shows relative amounts of total loaded protein per lane and lower gel (B) shows the Western blot result. Samples were separated on identical 12 % SDS-PAGE gels. The gel shown in (A) was stained with Coomassie Blue. Lane M represents the Precision Plus Protein™ Molecular Weight marker (BioRad) with molecular weights shown. (1) and (2) indicate independent biological replicates. Samples of MG1655, MG1655Δ*flhDC* and purified FlhDC (100 ng) were included, in order to determine which species related to specific anti-FlhDC binding and could be used for quantification with ImageJ. Off-target band used for normalisation, in order to account for any discrepancy in total amount of protein loaded per lane, is highlighted by a red arrow.

In the case of steady-state culture samples, 1 ml of chemostat culture was pelleted and resuspended in at least 50  $\mu$ l (dry-cell weight data was used to ensure equal loading), MG1655 and MG1655Δ*flhDC* samples were grown in LB at 37°C, and diluted to a cell density comparable to steady state-samples. All samples were mixed 1:1 with SDS Loading Buffer. 10  $\mu$ l of each sample was then loaded per lane.



**Figure 5.3** Western blots showing amounts of FliA, FlgM and FliC detected in samples taken from steady-state cultures grown at different dilution rates. Gel (A) shows relative amounts of total loaded protein per lane and gel (B) shows the Western blot result. Samples were separated on identical 12 % SDS-PAGE gels. The gel shown in (A) was stained with Coomassie Blue. Lane M represents the Precision Plus Protein™ Molecular Weight marker (BioRad) with molecular weights shown. (1) and (2) indicate independent biological replicates. Off-target band used for normalisation in ImageJ analyses, in order to account for any discrepancy in total amount of protein loaded per lane, is highlighted by a red arrow.

In the case of culture samples, 1 ml of chemostat culture was pelleted and resuspended in at least 50  $\mu\text{l}$  (dry-cell weight data was used to ensure equal loading), and mixed 1:1 with SDS Loading Buffer. 10  $\mu\text{l}$  of each sample was then loaded per lane.



**Figure 5.4** Graphs displaying (A) relative amounts of protein detected in Western blots and (B) fold change in associated transcript abundance at different dilution rates for the proteins FlhDC, FliA, FlgM and FliC. ImageJ software (Schneider et al., 2012) was used to quantify changes in FlhDC, FliA, FlgM and FliC amounts observed in Western blots of culture samples grown at different dilution rates (Figure 5.1). Results are shown in (A), wherein relative protein amounts are displayed against dilution rate, all values are given in arbitrary units with error bars representing one standard deviation. Graph (B) displays the observed trend in transcript abundance of the relevant genes against dilution rate as determined in gene expression microarray studies (Chapter 4), data is expressed as a fold change in transcript abundance relative to  $d=0.05 h^{-1}$ . In the case of FlhDC, the associated protein amounts and gene expression are represented by FlhD/*flhD* only.

to be exported through that same membrane bound structure in order to liberate FliA ( $\sigma^{28}$ ) and allow transcription of Level 3 flagellar genes only when the foundation of the flagellum is complete (Hughes *et al.*, 1993; Kutsukake, 1994), and (2) FlgM has been observed to protect FliA from proteolysis when the two are bound (Barembuch and Hengge, 2007). Taking these two characteristics into account, it is suggested that at the increased growth rates studied here there was an increase in the manifestation of completed flagellar basal body structures, accompanied by an increased rate of FlgM export from the cell which results in increased instances of liberated  $\sigma^{28}$  which were free to activate transcription at Level 3 promoters, but were also open to proteolytic attack. There is obviously a delicate balance between the amounts of FlgM and FliA present within the cell at any given time and only a snap-shot of this relationship is visualised here, but above all it can be said that this balance was maintained in such a way as to increase Level 3 flagellar gene expression at the higher growth rates studied here.

Finally, at the top of the flagellar regulon hierarchy, the master regulator FlhDC showed no detectable change in *flhD* transcript abundance over the growth rate range studied. However, the detected protein amount did increase slightly with dilution rate, with the maximum observed concentration being achieved at  $d = 0.5 \text{ h}^{-1}$ . The increase in complete flagellar structures at higher growth rates as evidenced by the observed increase in FliC is evidence of FlhDC ultimately exhibiting increased activity. Considering that, alongside the fact there was no detectable change in gene expression, it is suggested here that regulation of FlhDC activity relative to growth rate is primarily carried out at the post-transcriptional level.

Overall, in terms of the transcriptomics summarised in Table 5.1, the three genes related to flagellar gene regulation studied here (*flhD*, *fliA* and *flgM*) all exhibit very minor changes in transcript abundance as growth rate is increased. However this small change in gene expression ultimately leads to a significant increase in the abundance of FliC, a marker for whole-flagellar synthesis.

#### **5.4 Transmission electron microscopy of *E. coli* MG1655 cells cultured at different growth rates**

In order to further substantiate the claim that flagellar biosynthesis is up-regulated as growth rate is increased, samples of cells grown in the same steady-state

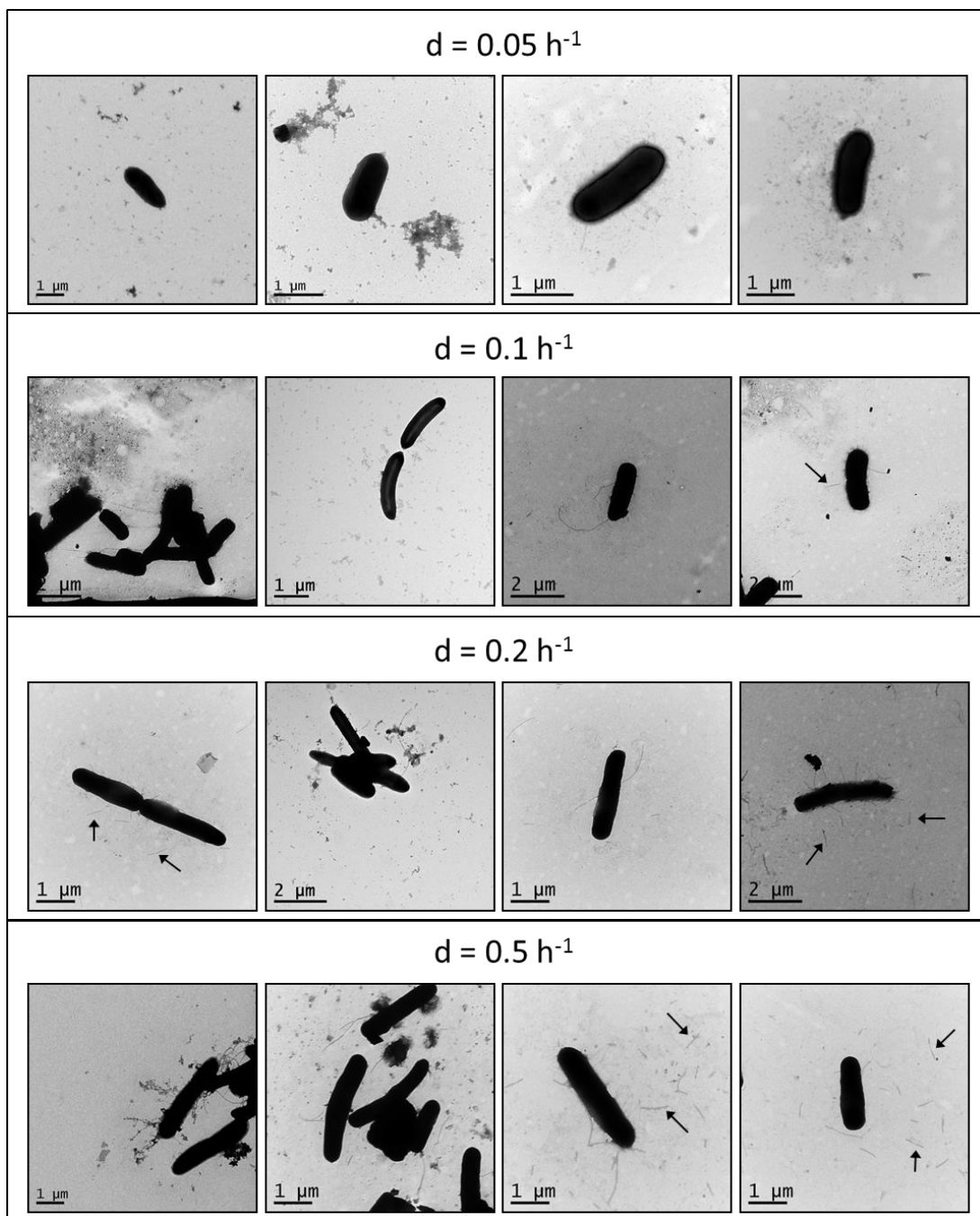
conditions at different dilution rates were directly visualised using Transmission electron microscopy (TEM), the results of which can be seen in Figure 5.4. As is highlighted in the micrographs, in multiple cases only detached flagellar fragments were seen, likely due to the sample handling procedure. However, a case can be made that flagellar expression was visualised with higher frequency at higher dilution rates. In fact, no flagella were visualised at the slowest growth rate of  $d=0.05\text{ h}^{-1}$ , in line with the observed lack of detectable FliC at  $d=0.05\text{ h}^{-1}$  in Western blots (Figure 5.2).

The TEM micrographs were also used to judge average cell size in relation to dilution rate using ImageJ (Schneider *et al.*, 2012). Only a small sample set of eight cells per dilution rate was used, due to limited cell number seen in micrographs especially for samples grown at  $d=0.05\text{ h}^{-1}$ . The results of this analysis can be seen in Table 5.2 and Figure 5.5. An increase in cell size was observed at increased steady-state growth rates and appeared to be maximal at doubling times  $\leq 3.5\text{ h}$ . Variations in cell size appeared to be dictated largely by cell length with little change in cell diameter. These trends are in line with what has previously been observed under similar glucose-limited steady-state conditions with a sample set of over 200 cells per condition (Volkmer and Heinemann, 2011). The high standard deviations in both this study and that of Volkmer and Heinemann are likely due to viewing a pool of cells all at various stages of cell growth which, in *E. coli*, occurs via elongation.

### **5.5 Analysis of transcriptomic data in relation to AI-2 synthesis, export and import as growth rate increases**

All known stages leading to the eventual synthesis of AI-2, its export, import, phosphorylation and subsequent utilisation or degradation are summarised in Figure 5.6, wherein all relevant transcriptomic data is also displayed. The displayed data represent fold changes in transcript abundance at  $d=0.5\text{ h}^{-1}$  relative to  $d=0.05\text{ h}^{-1}$  as determined in Chapter 4 (Section 4.5).

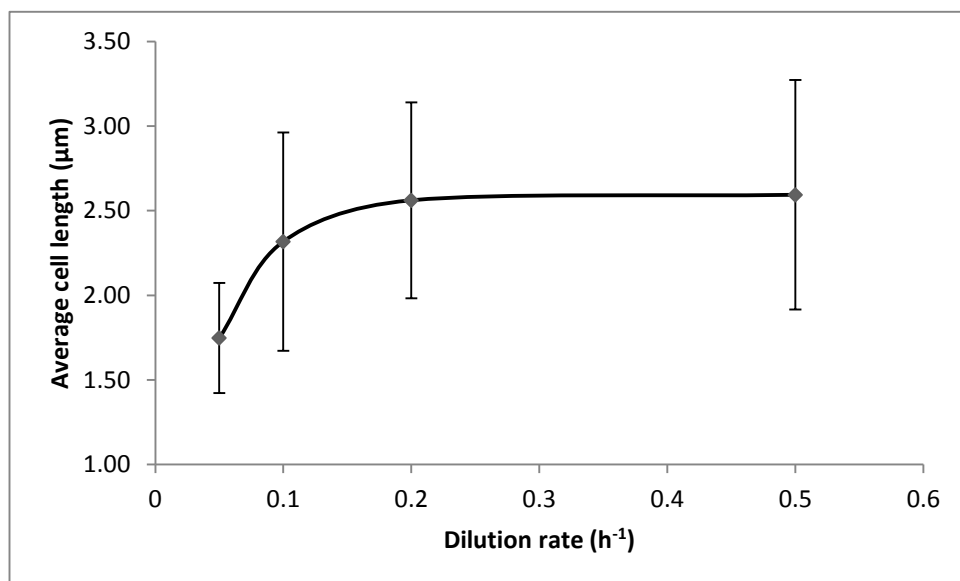
As AI-2 is synthesised by LuxS, with S-ribosyl-L-homocysteine (SRH) as a substrate, the stages leading to SRH synthesis are also summarised, and encompass the stages in methionine synthesis from cysteine. Perhaps unsurprisingly, at increased growth rates all stages in this amino acid biosynthesis network were at least slightly up-regulated, likely to support the increased protein synthesis turnover



**Figure 5.5** Typical transmission electron micrographs of *E. coli* cells grown at dilution rates of 0.05, 0.1, 0.2 and 0.5  $\text{h}^{-1}$ . Samples of *E. coli* grown at various growth rates were taken directly from steady-state cultures grown in the chemostat, diluted to an  $\text{OD}_{600}$  of 0.5 and immediately fixed by mixing in a 1:1 ratio with a solution of 3% glutaraldehyde (v/v) in 1 M cacodylate buffer. Fixed samples were then applied to carbon-coated grids and stained with phosphotungstic acid as described in Section 2.9. Grids were examined, and micrographs taken in a FEI G2 Bio-twin Tecnai 120Kv (Hillsboro, OR, USA). Arrows in micrographs highlight examples of observed detached flagellar fragments; this separation is likely due to shearing forces introduced during sample handling. Scale bars are shown.

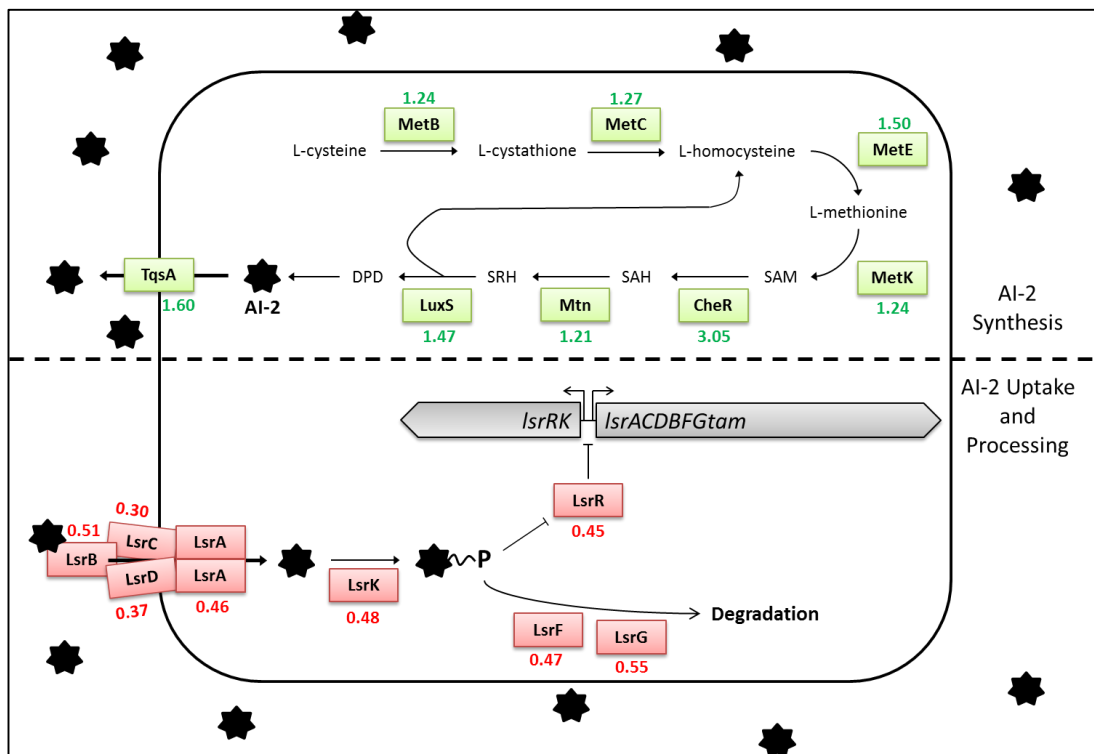
**Table 5.2 Observed average cell length and width in TEM micrographs of *E. coli* cells grown at various dilution rates.** A small sample size of 8 cells per dilution rate was analysed using ImageJ software in order to determine average cell size parameters of samples visualised in TEM micrographs, (examples can be seen in Figure 5.4).

Dilution Rate ( $\text{h}^{-1}$ )	Cell Length ( $\mu\text{m}$ )		Cell Width ( $\mu\text{m}$ )	
	Average	STD	Average	STD
0.05	1.75	0.32	0.64	0.09
0.1	2.32	0.64	0.62	0.12
0.2	2.56	0.58	0.59	0.05
0.5	2.59	0.68	0.59	0.10



**Figure 5.6 Plot of average *E. coli* cell length ( $\mu\text{m}$ ) against dilution rate ( $\text{h}^{-1}$ ).** Here, average cell length data, as shown in Table 5.2, has been plotted against dilution rate. Error bars represent one standard deviation.





**Figure 5.7 Schematic of the stages in AI-2 synthesis, export and import in *E. coli*.** Proteins, intermediates and associated data above the dashed line refer to all stages in AI-2 synthesis and subsequent export from the cell. Conversely, that which is below the dashed line refers to all stages in AI-2 import, phosphorylation and subsequent utilisation or degradation. Numbers displayed represent the corresponding fold change in transcript abundance when cells were grown at  $d=0.5 \text{ h}^{-1}$  relative to their abundance at  $d=0.05 \text{ h}^{-1}$ . Green highlights proteins with a relative up-regulation in associated transcript abundance, whereas red highlights those with a relative down-regulation. SAM, SAH, SRH and DPD refer to *S*-adenosyl-L-methionine, *S*-adenosyl-L-homocysteine, *S*-ribosyl-L-homocysteine and (*S*)-4,5-dihydroxy-2,3-pentandione respectively.

that comes with supporting an increased rate of cell doubling. This would suggest that at increased growth rate there was an increased rate of AI-2 production, in agreement with previous observations that AI-2 build-up is highest in exponential phase rather than stationary phase growth (Surette and Bassler, 1998). The gene *tqsA* which encodes the transporter proposed to export AI-2 into the extracellular environment was also up-regulated at the higher rate of growth.

To the contrary, the genes of the *lsr* operons that are associated with the uptake of AI-2 into the cell were all down-regulated, most of which more than 2-fold. This is in agreement with observations made in Chapter 3, wherein it was suggested that cAMP-CRP plays a significant role in regulating genes in cells growing at a slower rate. As the activation of the *lsr* operons has been attributed to the activity of CRP (Wang *et al.*, 2005a), one would expect that at the higher growth rates studied here there would be lower levels of cAMP and therefore reduced transcription of the *lsr* operons. This too coincides with the previously mentioned correlation between exponential phase growth and increased accumulation of extracellular AI-2. The genes of the *lsrRK* operon were also down-regulated to a similar degree, as the two *lsr* operons share the same regulatory promoter elements.

### **5.6 Quantifying the extracellular build-up of AI-2 with respect to growth rate**

As the transcriptomics suggest that the synthesis and export of AI-2 is increased at higher growth rate, and conversely its import into the cell is decreased, it was presumed that the amounts of detectable AI-2 in culture supernatant samples acquired during chemostat steady-state growth would be higher in those grown at increased dilution rates. In order to test this, the *Vibrio harveyi* strain BB170 was utilised. This mutant strain of *V. harveyi* has been shown to respond exclusively to AI-2 in order to activate the *lux* gene and express bioluminescence (Bassler *et al.*, 1993). A 10% cell-free culture supernatant sample was incubated with prepared *V. harveyi* BB170 cells as described in Section 2.10, and amount of luminescence emitted was recorded. This was carried out with supernatant samples of *E. coli* MG1655 cultures grown at  $d = 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$ , the results of which can be seen in Table 5.3. Values presented are referred to as “AI-2 Activity”, which is the

**Table 5.3 Relative levels of AI-2 detected in steady-state *E. coli* MG1655 culture supernatant samples using the *V. harveyi* strain BB170.** Measurements were obtained using the method outlined in Section 2.10. “Raw AI-2 Activity” refers to the relative amount of AI-2 induced bioluminescence detected in steady-state culture derived supernatant samples, compared to a sterile Evans minimal medium sample; values are expressed as a fold increase. “Weighted AI-2 Activity” refers to the same values after they have been multiplied by the dilution rate, to account for increased rates of supernatant dilution at the higher growth rate culturing conditions. This produces an estimation of the relative rate of extracellular AI-2 accumulation.

Dilution Rate (h <sup>-1</sup> )	Raw AI-2 Activity		Weighted AI-2 Activity	
	Average	STD	Average	STD
<b>0.05</b>	12.80	0.93	0.64	0.04
<b>0.1</b>	10.10	1.32	1.01	0.13
<b>0.2</b>	7.34	2.09	1.47	0.42
<b>0.5</b>	5.35	0.08	2.68	0.04

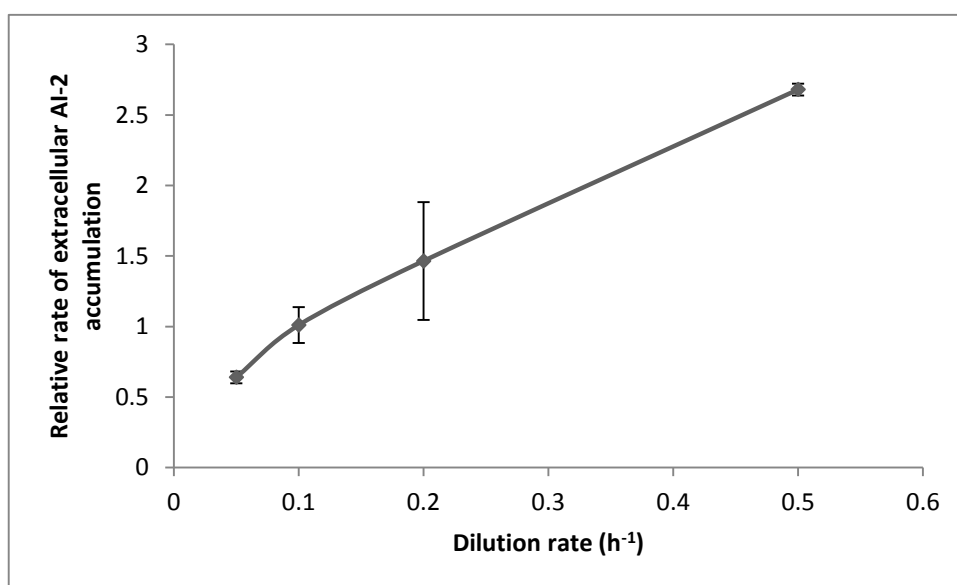
fold increase in AI-2 derived bioluminescence detected in the sample under test relative to the amount detected in a control sample of sterile Evans' minimal medium.

The assay produced a raw value for the relative amount of AI-2 present in the supernatant sample with which the BB170 culture was incubated, however this did not account for the increase in dilution rate impacting the concentration of extracellular AI-2 in a negative manner. Therefore, these raw values were multiplied by the imposed dilution rate in order to estimate the relative rate of AI-2 synthesis in relation to growth rate, the outcome of which is represented in Figure 5.7.

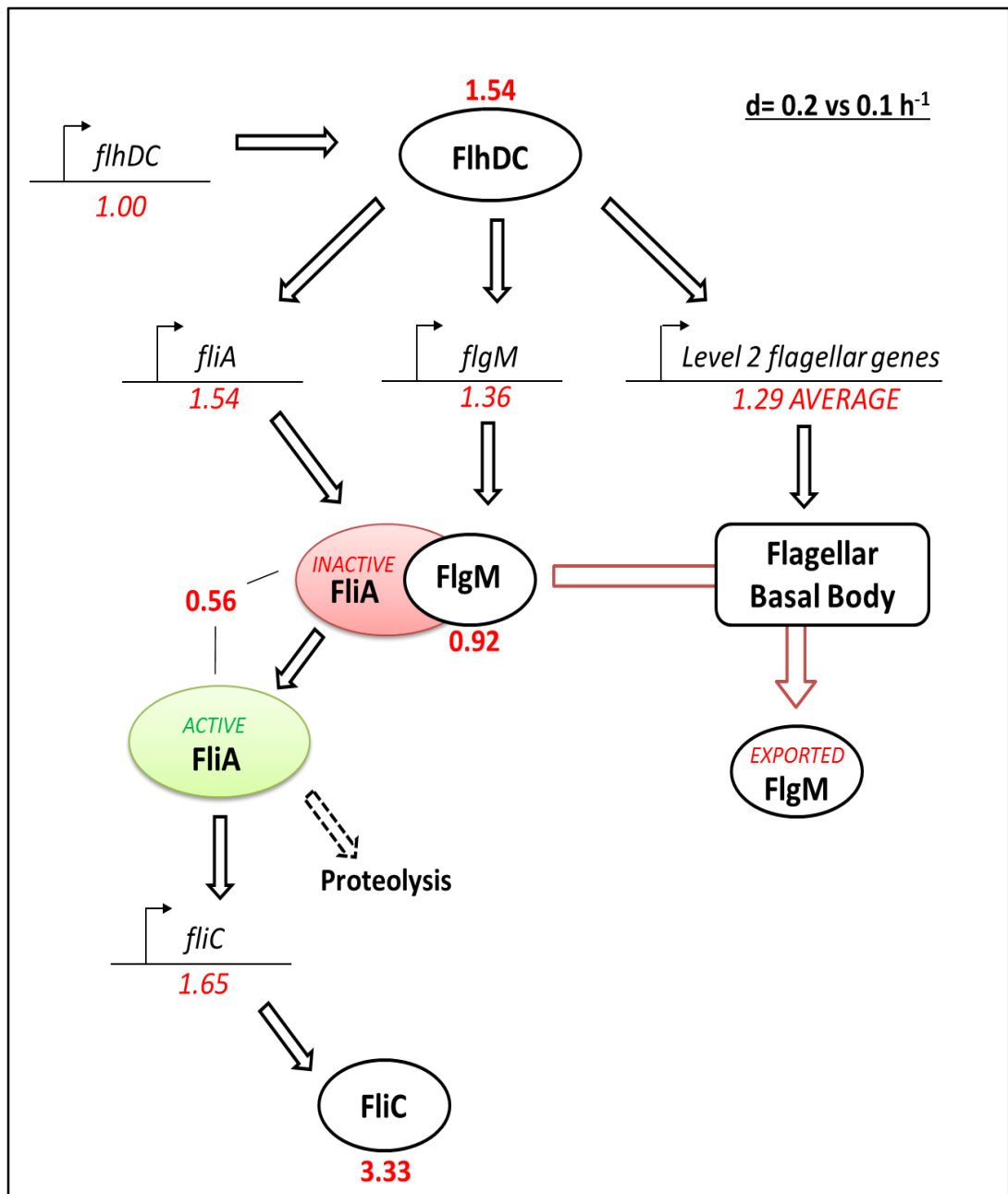
The result of this assay confirms that as growth rate of *E. coli* increased, there was an increase in the rate of extracellular AI-2 accumulation, with a fold increase relative to  $d = 0.05 \text{ h}^{-1}$  of 1.58, 2.29 and 4.19 at  $d = 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  respectively. The change in accumulation rate that was observed is likely due to decreased AI-2 uptake rather than increased AI-2 synthesis and export, as evidenced by the transcriptomic data summarised in Figure 5.6. It is also in direct agreement with the work carried out by DeLisa *et al.* wherein an increase in the rate of extracellular AI-2 synthesis was observed for the *E. coli* strain W3110 grown at incrementally increased rate of growth in Luria-Bertani (LB) medium (DeLisa *et al.*, 2001a).

## **5.7 Discussion**

Building on the broad transcriptomic work discussed in Chapter 3, it has been confirmed here that *E. coli* cells growing at increased rates also express increased number of flagella. This was confirmed both by an increase in the amount of FliC detected by Western blot at increased growth rate, and by direct visualisation of cells by TEM. The amount of FliC detected by Western blot increased 3.33-fold between  $d = 0.1$  and  $0.2 \text{ h}^{-1}$ , with an associated 1.65-fold jump in transcript abundance between those same growth rates. Western blot data was also obtained for the regulatory proteins FlhDC, FliA and FlgM, whose role it is to control the hierarchical expression of the genes of the flagellar regulon. The complexity of this regulatory network has already been well documented, and the data that was obtained through these Western blots coupled with the transcriptomic data trends already observed did not dispute that, these data are summarised diagrammatically in Figure 5.8. Where a general increase in the transcript abundance of FliA and FlgM at higher growth rates was observed, the correlation between their respective



**Figure 5.8 Relative rates of extracellular AI-2 accumulation plotted against dilution rate.** Plotted values represent those presented in Table 5.3 after having been weighted with regards to dilution rate. Error bars represent one standard deviation.



**Figure 5.9 Schematic diagram of flagellar gene regulation with overlaid transcriptomic and Western blot data.** Key stages in the flagellar regulon leading to biosynthesis of FliC and hence complete flagella are shown. Operons involved in this regulon are shown, with associated proteins represented by ovals. Data obtained in this work are shown as red numbers, with transcriptomic data shown in italics and quantified Western blot data shown in bold. All data is given as a fold change in transcript abundance or detected protein at  $d=0.2 \text{ h}^{-1}$  relative to  $d=0.1 \text{ h}^{-1}$ , wherein the largest quantifiable fold change in FliC and its transcript occurred. Transcriptomic and Western blot data pertaining to FlhDC is represented by data for *flhD*/FlhD.

intracellular protein amounts and growth rate was less clear cut. Intracellular FlgM levels remained relatively constant, while FliA protein levels seemed to experience a decrease at  $d = 0.2$  and  $0.5 \text{ h}^{-1}$ . It is suggested that as growth rate increases, particularly at doubling times  $\leq 3.5$  hours, the frequency of flagellar hook-basal body completion reaches a threshold wherein the rate of FlgM export from the cell is sufficient to allow  $\sigma^{28}$  to activate transcription of Level 3 flagellar genes but also be exposed to proteolytic attack. As is illustrated in Figure 5.8, while there is no change in the transcription of the *flhDC* operon on the transition from a growth rate of  $d = 0.1$  to  $0.2 \text{ h}^{-1}$ , there is a detectable 1.54-fold increase in the amount of intracellular FlhDC protein, this ultimately leads to a similar 1.66-fold increase in *fliC* transcription despite the intermediate complexity of the system. This highlights the well-established role of FlhDC as the flagellar master regulator. The observed lack of transcriptional change for the *flhDC* operon is also evidence of FlhDC activity being regulated at the post-transcriptional level. While more data would be required in order to further investigate the complexity of this system, it can be said with some confidence that the outcome of how this regulatory network is controlled in relation to increased growth rate is that there is an increase in the synthesis of fully assembled flagella.

In utilising the *V. harveyi* strain BB170, the quantification of the relative amounts of AI-2 present in supernatant samples of cultures grown at dilution rates of  $d = 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$  was made possible. This revealed that the rate of AI-2 accumulation in the extracellular environment increased with the rate of *E. coli* growth. This is in agreement with the observations that found extracellular AI-2 concentration to be maximal in exponential phase growth (Surette and Bassler, 1998) and increased growth rate (DeLisa *et al.*, 2001a), and also previous studies that have shown that internalisation of AI-2 is activated by a build-up of cAMP (Wang *et al.*, 2005a) which was shown to be modulating CRP activity to a larger extent at the lower growth rates studied here.

As quorum sensing systems exist in order to elicit a coordinated change in gene expression across a bacterial population, there have been multiple studies into the transcriptomic changes that occur upon manipulation of the AI-2 network, whether through the deletion of *luxS* (Sperandio *et al.*, 2001; DeLisa *et al.*, 2001b) or the genes *lsrR* and *lsrK* (Li *et al.*, 2007). The genes linked with AI-2 mediated

quorum sensing regulation, as determined in these studies, span a large functional range. The study by Sperandio *et al.* (2001) carried out a comparison between a wild-type Enterohemorrhagic *E. coli* (EHEC) strain and its *luxS* isogenic mutant. However, they state that *luxS* deletion causes an increase in growth rate, and though both strains were grown to an early stationary phase OD<sub>600</sub> of 1.0, any transcriptomic changes that are reported may be effects of the growth rate differential. The work by DeLisa *et al.* however, utilising an *E. coli* strain W3110 *luxS* mutant, avoids this as they performed a comparison between the same mutant strain with and without the presence of exogenously provided AI-2 (DeLisa *et al.*, 2001b).

Studying the work by DeLisa *et al.* and Li *et al.* exclusively (DeLisa *et al.*, 2001b; Li *et al.*, 2007), DeLisa *et al.* link increased AI-2 concentrations to the regulation of: (1) 22 genes involved in cell division, DNA processing and morphological adaptation; (2) 23 genes involved in processes considered to be quorum sensing regulated, such as virulence, biofilm formation and motility; and (3) 28 genes involved in small-molecule metabolism. On the other hand the work by Li *et al.* focused on the role of AI-2 in the regulation of genes associated with biofilm formation and sRNA expression. Despite studies such as these, the only gene regulator that has been shown a direct response to AI-2 is LsrR, whose own direct binding has only been shown in relation to the promoter region for the divergent *lsr* operons (Wu *et al.*, 2013). This limited evidence of direct AI-2 mediated gene regulation in *E. coli* is the basis of arguments against its role as a quorum sensing molecule, and an argument must be made that the primary role of LuxS is to aid in the conversion of SAH to non-toxic derivatives (Schauder *et al.*, 2001) with AI-2 produced purely as an over-metabolite at an increased rate at higher growth rates, as evidenced by the transcriptomics presented here. As is stated by Xavier and Bassler, “at present, we do not understand the benefit that enteric bacteria derive from producing and releasing AI-2, only to internalize it later” (Xavier and Bassler, 2005), it is suggested here that an increase in the rate of AI-2 synthesis and expulsion from the cell as growth rate is increased is a product of increased flux through the SAH detoxification pathway, and a decrease in *lsr* operon expression is due to the reduction in cAMP-CRP activity at higher rates of growth. Perhaps in cells growing with a longer doubling time, in the conditions studied here, extracellular AI-2 is scavenged back into the cell as an alternative source of carbon.



Despite the lack of clarity with regards to the role of AI-2 in quorum sensing of *E. coli*, a link has previously been made between AI-2 and regulation of genes involved in motility (Sperandio *et al.*, 2001; Herzberg *et al.*, 2006). In a *luxS* mutant strain (i.e. in the absence of AI-2 synthesis) a reduction in the amounts of FlhC was observed, and in a *tqsA* mutant (i.e. in the absence of AI-2 expulsion from the cell), the amount of *fliC* transcription increased. These two studies create a picture wherein the expulsion of AI-2 into the extracellular environment inhibits the expression of flagella. This is contrary to what has been observed in this work; with the increase in growth rate there is an observed increase in the rate of AI-2 expulsion and a decrease in its uptake back into the cell, this is accompanied by an increase in the expression of flagellar structures. It is put forward here that the previously claimed link between extracellular AI-2 accumulation and the expression of flagellar genes is at least partially due to growth rate differences reported in the mutant strains studied; a *luxS* mutant has a noticeable increase in growth rate (Sperandio *et al.*, 2001), though a *tqsA* mutant only shows a slight growth rate differential compared to wild type *E. coli* (Herzberg *et al.*, 2006).

It should also be noted that in the work carried out here, due to the nature of constant dilution with fresh medium in the chemostat-based culturing method, a relative build-up of extracellular AI-2 to a threshold concentration is not possible as dilution rate is increased. For this reason, any transcriptomic changes observed in this work are unlikely to be due to an AI-2 mediated quorum sensing system.

A link between the quorum sensing sensor kinase and its respective response regulator, QseC and QseB respectively, had originally been made wherein transcription of *qseBC* was up-regulated in culture medium pre-conditioned with cells that produce AI-2 and it was shown that expression of this two-component regulatory system positively influences the transcription of the *E. coli* flagellar master regulatory FlhDC (Sperandio *et al.*, 2002). However, it was subsequently demonstrated that QseC activity was in fact modulated by the previously undetected molecule AI-3, another compound which relies on LuxS for its synthesis (Sperandio *et al.*, 2003; Clarke *et al.*, 2006). This discovery casts doubt on the transcriptomic studies mentioned here and how they relate to AI-2 signalling. In the case of the two studies utilising a *luxS* mutant especially, any transcriptomic changes seen could be equally attributed to the presence of AI-3.

These observations highlight the contradictory nature of the data that has been produced in studies focusing on the AI-2 molecule playing a signalling role. Taking into account the observations made in this study and the fact that, in *E. coli*, LsrR is the only proven direct responder to AI-2 and operates solely to regulate the transcription of the genes associated with AI-2 uptake, it is suggested that AI-2 be considered not as a signalling molecule, but an over-metabolite; the synthesis of which increases with growth rate. It is also put forward that a direct link between AI-2 production and flagellar synthesis does not exist, and that they are two independent effects of growth rate. The reason for increased expression of flagella as doubling time is reduced is unclear, though it may prove advantageous for cells that are in a population of high proliferation to move into areas of less competition.

Alternatively, the increased expression of flagella may confer an indirect advantage in that they are proton channels that may aid in redox balance at higher growth rates. In conditions where *E. coli* grows at an increased rate, the rate of glucose uptake increases and subsequent flux through the TCA cycle increases also. However, it is well documented that at the higher limits of growth rate a portion of the glucose metabolised is secreted as acetate (Hollywood and Doelle, 1976; Andersen and Vonmeyenburg, 1980; Vemuri *et al.*, 2006; Nahku *et al.*, 2010). Flux from acetyl-CoA through to acetate produces no NADH, as opposed to the three NADH and one FADH<sub>2</sub> produced from flux of acetyl-CoA through the TCA cycle. This equates to a lower yield of NADH reducing equivalents at higher growth rates, and it has been suggested that limiting flux through the TCA cycle in tandem with acetate production is in order to maintain the NADH/NAD<sup>+</sup> ratio for sustained operation of oxidative phosphorylation in the electron transport chain when growing in the presence of oxygen (van Hoek and Merks, 2012). It is postulated here that the increased presence of flagella when *E. coli* is growing at faster rates may aid in the maintenance of intracellular redox balance by allowing the influx of H<sup>+</sup> through the MotA/B complex, relieving ‘back-pressure’ in the electron transport chain thereby facilitating increased NADH oxidation.

## 6. General discussion

In studying the *E. coli* MarR family gene regulator SlyA through microarray analyses it was determined that it may affect the expression of at least 44 genes, either directly or indirectly. The majority of these affected genes were influenced in a positive way (88.6%) which was in agreement with the already proposed SlyA mechanism of action being anti-repression of H-NS silencing. The likelihood of SlyA predominantly operating via this method of anti-repression was reinforced by the observation that 25% of the affected genes are already thought to be repressed by H-NS and that there was significant crossover with the regulon associated with another gene regulator, LeuO. In fact, of the 39 genes whose expression was positively influenced by the over-expression of SlyA, 64% were also present in a SELEX-chip study determining LeuO binding sites within gene promoter regions. Given that LeuO has been found to activate transcription by antagonising H-NS repression, SlyA operating via a similar mechanism is likely.

In the case of *S. enterica* Serovar Typhimurium SlyA, those genes that are considered to be influenced by the gene regulator are largely attributed with altering the protein profile of the bacterial cell envelope. A similar propensity for controlling genes associated with the inner membrane, outer membrane or periplasmic space was observed for *E. coli* SlyA (44% of the genes with a known or predicted function), with multiple operons that encode fimbrial-like adhesins being affected.

This work has expanded on the list of those genes that are directly influenced by SlyA in *E. coli* K-12 through EMSAs. Direct binding of SlyA to *Pcas*, *Pelf*, *Pgsp*, *Ppaa*, *Pssu*, *Psgc*, *Pfec*, *PleuO* and *PmdtM* has now been shown, with *Pssu*, *Pcas*, *Ppaa*, *Pelf*, *PleuO* and *Pgsp* having previously been demonstrated as targets of LeuO binding. This therefore not only expands the regulon of SlyA in *E. coli* K-12, where *hlyE* and *fimB* were the only targets previously associated with direct regulation, but further reinforces the crossover with the LeuO regulon and the general activatory mechanism being H-NS de-repression.

Further detail into the specificity of SlyA binding was determined in the context of the *mdtM* promoter region. It was found that the binding site motif TTA<6nt>TAA accounted for observed strong binding and, on removal of all likely sites, specific binding to promoter DNA in the presence of heparin was abolished. It

was also found that the presence of at least one complete binding half-site was required for strong SlyA binding to target DNA. These data demonstrate how SlyA may bind DNA with relatively low sequence specificity, but certain sites are bound preferentially. These traits, taken together, are appropriate for a regulator that operates by competing with the widespread, unspecific nucleoid structuring protein: H-NS.

It was demonstrated through microarray analysis of a wild type *E. coli* MG1655 strain compared to an *E. coli* MG1655 $\Delta$ slyA mutant strain grown at different growth rates that slyA expression and its subsequent transcription regulatory effect was not significantly controlled by growth rate. The observed large fold-change in detectable SlyA protein in the overexpression strain relative to the small detected change in transcript level suggests that control of SlyA activity is predominantly carried out at the post-transcriptional level, this may be due to the rare UUG start codon in the slyA open reading frame though this regulation is not responsive to changes in growth rate as was previously suggested (McVicker *et al.*, 2011). Contrary to what was observed for SlyA of *S. enterica* Serovar Typhimurium (Zhao *et al.*, 2008), ppGpp had no effect on the DNA binding properties of *E. coli* SlyA when incubated with the promoter region for the gene *mdtM*, as had been previously observed with the promoter of *fimB* (McVicker *et al.*, 2011). However, it is possible that ppGpp modifies *E. coli* SlyA such that it is targeted to alternative DNA targets. This means that the ligand to which SlyA may respond remains elusive, assuming that SlyA does indeed respond to a ligand. It must also be considered that ligand binding may de-activate the DNA binding potential of SlyA, as is the case for multiple MarR family regulators.

Despite the observation that it had no effect on the expression of SlyA, a full global transcriptomic profile as it pertains to growth rate has been produced for the specific *E. coli* K-12 strain, MG1655, providing a transcriptomic profile for 4,311 genes at dilution rates of 0.05, 0.1, 0.2 and 0.5 h<sup>-1</sup>. Of those genes, 253 were observed to change significantly in transcript abundance at increased growth rates, with 86 being up-regulated and 167 being down-regulated. A comparative study of transcriptomic and proteomic changes experienced by *E. coli* MG1655 in response to changes in growth rate had recently been carried out by Valgepea *et al.* (Valgepea *et al.*, 2013). It has been outlined why the data presented in this work provides a more

comprehensive (including sRNA expression changes) and likely more reliable overview. Interestingly, on comparison of the transcriptomic data produced in this work and the proteomic data presented in the Valgepea study (2013) it was suggested that the influence of transcription repression in lowering protein abundance/activity might be greater than previously thought (Valgepea *et al.*, 2013).

When the total transcriptomic dataset was analysed using TFinfer software, it was possible to infer the activity of 167 transcription factors over the range of growth rates tested, with 38 being predicted to exhibit a significant change in activity. This highlighted a common trend in which the activity of a transcriptional regulator was exhibiting a defined step change in activity between the dilution rates  $d= 0.1$  and  $0.2 \text{ h}^{-1}$ . A trend that was also noted in the general transcriptional profile of genes in the transcriptomic analysis itself. This might indicate that, in terms of steady-state grown cultures, there is a step change in gene expression between those cells that have a doubling time  $\geq 6.9 \text{ h}$  and those that have a doubling time  $\leq 3.5 \text{ h}$ , a shift in which a noticeable proportion of regulation and transcript abundance changes are dedicated to starvation responses and utilisation of secondary metabolites. This observation is coupled with the inferred activity of cAMP-CRP itself decreasing between the dilution rates of  $d= 0.1$  and  $0.2 \text{ h}^{-1}$ , linked with the observed large proportion of down-regulated genes being linked with either secondary metabolism, transport of alternative metabolites or regulation of genes associated with related functions. It has been suggested that cells that are grown at the slower dilution rates utilised in this work are exhibiting a trait of scavenging for alternative sources of metabolite and, though the medium in which the cells are grown is well defined and specifically contains little in the way of alternative sources of carbon, scavenging from the remains of dead cells is a possible scenario. The observed step change in transcription factor activity between  $d= 0.1$  and  $0.2 \text{ h}^{-1}$  activity has been suggested to occur at a threshold possibly dictated by the ratio of an individual cell's time spent in a completely glucose starved environment to the rate of glucose being fed in to the culture.

Genes that expressed an increase in transcript abundance as growth rate increased included those involved with expression of Type I fimbriae and also the synthesis of the *E. coli* lipopolysaccharide layer. However, a striking proportion of up-regulated genes were associated with flagellar assembly, motility and chemotaxis,

a surprising observation given the conditions in which the samples were cultured. The increased expression of flagellar genes at increased rates of growth was confirmed by both Western blot of FliC and observation of increased flagella numbers by transmission electron microscopy. Though data has been obtained that illustrates the complexity in the regulon leading to Level 3 flagellar gene expression, it has also been shown that a 1.4 fold increase in FlhDC protein (Level 1) at increased growth rates resulted in a comparable 1.66 fold change in *fliC* gene expression (Level 3) highlighting the well-established role of FlhDC as the flagellar master regulator.

Reasons as to why increased flagellar expression may be beneficial at increased growth rates have been discussed. A particularly novel explanation has been given; the increased presence of flagella when *E. coli* is growing at faster rates may aid in the maintenance of intracellular redox balance by allowing the influx of additional H<sup>+</sup> through the MotA/B complex, thereby relieving back pressure in the aerobic electron transport chain facilitating NADH reoxidation at the expense of O<sub>2</sub> reduction to H<sub>2</sub>O.

Genes of the *lsr* operon, attributed with the uptake of AI-2, decreased in expression as growth rate increased. This was coupled with an increase in those genes associated with flux through the SAH detoxification pathway, increased expression of *tqsA* associated with AI-2 export from the cell and an observed increase in the amount of detectable AI-2 in culture supernatants of samples grown at increased rates of growth. Autoinducer-2 is widely regarded as a quorum sensing molecule, though its direct influence on the transcription of genes is not well established. On studying the literature pertaining to AI-2 and its effect on gene regulation in *E. coli*, it is suggested that its production and expulsion from the cell is a product of over-metabolism and increased flux through the SAH detoxification pathway with its uptake into the cell at lower growth rates being a product of increased cAMP-CRP mediated expression of the *lsr* operon. This increased uptake at lower growth rates may be an example of scavenging for additional sources of carbon.

Multiple carbon utilisation parameters have been quantified for *E. coli* K-12 grown in glucose-limited minimal medium (20 mM glucose) over the range of

growth rates ( $d= 0.05, 0.1, 0.2$  and  $0.5 \text{ h}^{-1}$ ); rate of glucose uptake, rate of biomass synthesis, rate of acetate synthesis and the maximum flux through the TCA cycle were all calculated. The estimated maximal rate of carbon flux through the TCA cycle was in the region of  $28 \text{ mmol C L}^{-1} \text{ h}^{-1}$ , in agreement with previous flux balance studies carried out with *E. coli* (Holms, 1996). Any excess carbon flux was channelled towards acetate synthesis when *E. coli* was grown at the fastest rate of growth utilised in this study ( $d= 0.5 \text{ h}^{-1}$ ). Little in the way of a definitive statement as to the maximum flux of carbon through the TCA cycle has been made in previous studies of this nature. The data provided here is a good basis for such an investigation, though it would be prudent to carry out similar growth of *E. coli* at dilution rates exceeding  $d= 0.5 \text{ h}^{-1}$  to clarify whether the observed maximum flux through the TCA cycle is conserved, and consistently offset by an increase in the proportion of acetate produced.

In summary, a broad range of data is presented in this work, with the primary findings including: (1) *E. coli* SlyA has been observed to significantly influence the expression of 44 genes, with direct binding to the promoter regions of 9 operons other than *PhlyE*, *PfimB* and *PslyA* having been demonstrated; (2) Though no link between growth rate and SlyA activity was observed, the effect of growth rate on global transcriptomics has been quantified, with 253 genes being significantly influenced and the activity of 38 transcription factors possibly playing a significant role in this adaptation; (3) Multiple genes associated with expression of flagella and motility were observed to be up-regulated as growth rate increased, and this was clarified by Western blot analyses and TEM; (4) Build-up of extracellular AI-2 increased with growth rate which was rationalised by the increase in expression of genes associated with SAH detoxification, AI-2 synthesis and AI-2 export with a coordinated decrease in expression of the *lsr* operon associated with AI-2 uptake; and finally, (5) carbon metabolism parameters with respect to growth rate in a glucose-limited medium were calculated and a maximal carbon flux through the TCA cycle in the region of  $28 \text{ mmol C l}^{-1} \text{ h}^{-1}$  has been proposed.

## 7. Appendices

Supplementary data is provided on the enclosed CD. File names and descriptions are provided below:

<u>File Name</u>	<u>Description</u>
Supplementary data 4.1	Growth rate transcriptomics – total data set without filtering
Supplementary data 4.2	Growth rate transcriptomics – Up-regulated gene list, including those of unknown function.
Supplementary data 4.3	Growth rate transcriptomics – Down-regulated gene list, including those of unknown function.
Supplementary data 4.4	Growth rate TMDH study - full list of the informative probes expressing the trend of diminished signal only at $d= 0.5 \text{ h}^{-1}$ in each Biological Replicate.
Supplementary data 4.5	Growth rate transcriptomics – comparison with the transcriptomic data set presented in the Valgepea <i>et al.</i> (2013) study.
Supplementary data 4.6	Growth rate transcriptomics - comparison with the proteomic data set presented in the Valgepea <i>et al.</i> (2013) study.



## 8. References

- ABE, H., TATSUNO, I., TOBE, T., OKUTANI, A. & SASAKAWA, C. 2002. Bicarbonate ion stimulates the expression of locus of enterocyte effacement-encoded genes in enterohemorrhagic *Escherichia coli* O157 : H7. *Infection and Immunity*, 70, 3500-3509.
- ADICIPTANINGRUM, A. M., BLOMFIELD, I. C. & TANS, S. J. 2009. Direct observation of type 1 fimbrial switching. *Embo Reports*, 10, 527-532.
- ALEKSHUN, M. N., KIM, Y. S. & LEVY, S. B. 2000. Mutational analysis of MarR, the negative regulator of *marRAB* expression in *Escherichia coli*, suggests the presence of two regions required for DNA binding. *Molecular Microbiology*, 35, 1394-1404.
- ALEKSHUN, M. N. & LEVY, S. B. 1999a. Characterization of MarR superrepressor mutants. *Journal of Bacteriology*, 181, 3303-3306.
- ALEKSHUN, M. N. & LEVY, S. B. 1999b. The *mar* regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends in Microbiology*, 7, 410-413.
- ALEKSHUN, M. N., LEVY, S. B., MEALY, T. R., SEATON, B. A. & HEAD, J. F. 2001. The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 angstrom resolution. *Nature Structural Biology*, 8, 710-714.
- ANDERSEN, K. B. & VONMEYENBURG, K. 1980. Are growth-rates of *Escherichia coli* in batch cultures limited by respiration? *Journal of Bacteriology*, 144, 114-123.
- ANTUNES, L. C. M. & FERREIRA, R. B. R. 2009. Intercellular communication in bacteria. *Critical Reviews in Microbiology*, 35, 69-80.
- ASIF, H. M. S., ROLFE, M. D., GREEN, J., LAWRENCE, N. D., RATTRAY, M. & SANGUINETTI, G. 2010. TFInfer: a tool for probabilistic inference of transcription factor activities. *Bioinformatics*, 26, 2635-2636.
- BAREMBRUCH, C. & HENGGE, R. 2007. Cellular levels and activity of the flagellar sigma factor FliA of *Escherichia coli* are controlled by FlgM-modulated proteolysis. *Molecular Microbiology*, 65, 76-89.
- BASSLER, B. L., GREENBERG, E. P. & STEVENS, A. M. 1997. Cross-species induction of luminescence in the quorum-sensing bacterium *Vibrio harveyi*. *Journal of Bacteriology*, 179, 4043-4045.
- BASSLER, B. L., WRIGHT, M., SHOWALTER, R. E. & SILVERMAN, M. R. 1993. Intercellular signaling in *Vibrio harveyi* - sequence and function of genes regulating expression of luminescence. *Molecular Microbiology*, 9, 773-786.

- BATTESTI, A., MAJDALANI, N. & GOTTESMAN, S. 2011. The RpoS-Mediated General Stress Response in *Escherichia coli*. In: GOTTESMAN, S. & HARWOOD, C. S. (eds.) *Annual Review of Microbiology*, Vol 65.
- BEATTY, C. M., BROWNING, D. F., BUSBY, S. J. W. & WOLFE, A. J. 2003. Cyclic AMP receptor protein-dependent activation of the *Escherichia coli acsP2* promoter by a synergistic class III mechanism. *Journal of Bacteriology*, 185, 5148-5157.
- BEJAR, S., CAM, K. & BOUCHE, J. P. 1986. Control of cell-division in *Escherichia coli* - DNA-sequence of *dicA* and of a 2nd gene complementing mutation *dicA1*, *dicC*. *Nucleic Acids Research*, 14, 6821-6833.
- BELTRAMETTI, F., KRESSE, A. U. & GUZMAN, C. A. 1999. Transcriptional regulation of the *esp* genes of enterohemorrhagic *Escherichia coli*. *Journal of Bacteriology*, 181, 3409-3418.
- BELYAEVA, T. A., RHODIUS, V. A., WEBSTER, C. L. & BUSBY, S. J. W. 1998. Transcription activation at promoters carrying tandem DNA sites for the *Escherichia coli* cyclic AMP receptor protein: Organisation of the RNA polymerase alpha subunits. *Journal of Molecular Biology*, 277, 789-804.
- BHENDE, P. M. & EGAN, S. M. 2000. Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with sigma 70. *Journal of Bacteriology*, 182, 4959-4969.
- BLATTER, E. E., ROSS, W., TANG, H., GOURSE, R. L. & EBRIGHT, R. H. 1994. Domain organization of RNA-polymerase alpha-subunit - C-terminal 85 amino acids constitute a domain capable of dimerization and DNA-binding. *Cell*, 78, 889-896.
- BRADFORD, M. M. 1976. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.
- BRAUN, V. 1995. Energy-coupled transport and signal-transduction through the gram-negative outer-membrane via TonB-ExbB-ExbD-dependent receptor proteins. *Fems Microbiology Reviews*, 16, 295-307.
- BRAUN, V., MAHREN, S. & SAUTER, A. 2006. Gene regulation by transmembrane signaling. *Biometals*, 19, 103-113.
- BROWN, N. L., STOYANOV, J. V., KIDD, S. P. & HOBMAN, J. L. 2003. The MerR family of transcriptional regulators. *Fems Microbiology Reviews*, 27, 145-163.

- BROWNING, D. F., COLE, J. A. & BUSBY, S. J. W. 2000. Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Molecular Microbiology*, 37, 1258-1269.
- BROWNING, D. F., COLE, J. A. & BUSBY, S. J. W. 2004. Transcription activation by remodelling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli nir* promoter. *Molecular Microbiology*, 53, 203-215.
- BUCHMEIER, N., BOSSIE, S., CHEN, C. Y., FANG, F. C., GUINEY, D. G. & LIBBY, S. J. 1997. SlyA, a transcriptional regulator of *Salmonella typhimurium*, is required for resistance to oxidative stress and is expressed in the intracellular environment of macrophages. *Infection and Immunity*, 65, 3725-3730.
- BURGESS, R. R. 1969. Separation and characterization of subunits of ribonucleic acid polymerase. *Journal of Biological Chemistry*, 244, 6168-&.
- BURGESS, R. R., TRAVERS, A. A., DUNN, J. J. & BAUTZ, E. K. F. 1969. Factor stimulating transcription by RNA polymerase. *Nature*, 221, 43-&.
- BUSBY, S. & EBRIGHT, R. H. 1997. Transcription activation at class II CAP-dependent promoters. *Molecular Microbiology*, 23, 853-859.
- CAM, K., BEJAR, S., GIL, D. & BOUCHE, J. P. 1988. Identification and sequence of gene *dicB* - translation of the division inhibitor from an in-phase internal start. *Nucleic Acids Research*, 16, 6327-6338.
- CARAMEL, A. & SCHNETZ, K. 1998. Lac and lambda repressors relieve silencing of the *Escherichia coli bgl* promoter. Activation by alteration of a repressing nucleoprotein complex. *Journal of Molecular Biology*, 284, 875-883.
- CASTANIE-CORNET, M.-P., CAM, K., BASTIAT, B., CROS, A., BORDES, P. & GUTIERREZ, C. 2010. Acid stress response in *Escherichia coli*: mechanism of regulation of *gadA* transcription by RcsB and GadE. *Nucleic Acids Research*, 38, 3546-3554.
- CATHELYN, J. S., ELLISON, D. W., HINCHLIFFE, S. J., WREN, B. W. & MILLER, V. L. 2007. The RovA regulons of *Yersinia enterocolitica* and *Yersinia pestis* are distinct: evidence that many RovA-regulated genes were acquired more recently than the core genome. *Molecular Microbiology*, 66, 189-205.
- CHADSEY, M. S., KARLINSEY, J. E. & HUGHES, K. T. 1998. The flagellar anti-sigma factor FlgM actively dissociates *Salmonella typhimurium* sigma 28 RNA polymerase holoenzyme. *Genes & Development*, 12, 3123-3136.

- CHAUDHURI, R. R., ALLEN, A. G., OWEN, P. J., SHALOM, G., STONE, K., HARRISON, M., BURGIS, T. A., LOCKYER, M., GARCIA-LARA, J., FOSTER, S. J., PLEASANCE, S. J., PETERS, S. E., MASKELL, D. J. & CHARLES, I. G. 2009a. Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *Bmc Genomics*, 10.
- CHAUDHURI, R. R., PETERS, S. E., PLEASANCE, S. J., NORTHEN, H., WILLERS, C., PATERSON, G. K., CONE, D. B., ALLEN, A. G., OWEN, P. J., SHALOM, G., STEKEL, D. J., CHARLES, I. G. & MASKELL, D. J. 2009b. Comprehensive Identification of *Salmonella enterica* Serovar Typhimurium Genes Required for Infection of BALB/c Mice. *Plos Pathogens*, 5.
- CHEKABAB, S. M., JUBELIN, G., DOZOIS, C. M. & HAREL, J. 2014. PhoB Activates *Escherichia coli* O157: H7 Virulence Factors in Response to Inorganic Phosphate Limitation. *PLoS One*, 9.
- CHEN, C. C., CHOU, M. Y., HUANG, C. H., MAJUMDER, A. & WU, H. Y. 2005. A cis-spreading nucleoprotein filament is responsible for the gene silencing activity found in the promoter relay mechanism. *Journal of Biological Chemistry*, 280, 5101-5112.
- CHEVANCE, F. F. V. & HUGHES, K. T. 2008. Coordinating assembly of a bacterial macromolecular machine. *Nature Reviews Microbiology*, 6, 455-465.
- CHILCOTT, G. S. & HUGHES, K. T. 2000. Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* Serovar Typhimurium and *Escherichia coli*. *Microbiology and Molecular Biology Reviews*, 64, 694-+.
- CLARET, L. & HUGHES, C. 2002. Interaction of the atypical prokaryotic transcription activator FlhD(2)C(2) with early promoters of the flagellar gene hierarchy. *Journal of Molecular Biology*, 321, 185-199.
- CLARKE, M. B., HUGHES, D. T., ZHU, C., BOEDEKER, E. C. & SPERANDIO, V. 2006. The QseC sensor kinase: A bacterial adrenergic receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 10420-10425.
- COOPER, S. & HELMSTETTER, C. E. 1968. Chromosome Replication and Division Cycle of *Escherichia coli* B/r. *Journal of Molecular Biology*, 31, 519-540.
- CORBETT, D., BENNETT, H. J., ASKAR, H., GREEN, J. & ROBERTS, I. S. 2007. SlyA and H-NS regulate transcription of the *Escherichia coli* K5 capsule gene cluster, and expression of slyA in *Escherichia coli* is temperature-dependent, positively autoregulated, and independent of H-NS. *Journal of Biological Chemistry*, 282, 33326-33335.
- DALEBROUX, Z. D. & SWANSON, M. S. 2012. ppGpp: magic beyond RNA polymerase. *Nature Reviews Microbiology*, 10, 203-212.

- DANGI, B., GRONENBORN, A. M., ROSNER, J. L. & MARTIN, R. G. 2004. Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Molecular Microbiology*, 54, 45-59.
- DATSENKO, K. A. & 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*, 97, 6640-6645.
- DELISA, M. P., VALDES, J. J. & BENTLEY, W. E. 2001a. Mapping stress-induced changes in autoinducer AI-2 production in chemostat-cultivated *Escherichia coli* K-12. *Journal of Bacteriology*, 183, 2918-2928.
- DELISA, M. P., WU, C. F., WANG, L., VALDES, J. J. & BENTLEY, W. E. 2001b. DNA microarray-based identification of genes controlled by autoinducer 2-stimulated quorum sensing in *Escherichia coli*. *Journal of Bacteriology*, 183, 5239-5247.
- DILLON, S. C. & DORMAN, C. J. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology*, 8, 185-195.
- DOLAN, K. T., DUGUID, E. M. & HE, C. 2011. Crystal Structures of SlyA Protein, a Master Virulence Regulator of *Salmonella*, in Free and DNA-bound States. *Journal of Biological Chemistry*, 286, 22178-22185.
- DOVE, S. L., DARST, S. A. & HOCHSCHILD, A. 2003. Region 4 of sigma as a target for transcription regulation. *Molecular Microbiology*, 48, 863-874.
- EBRIGHT, R. H. 1993. Transcription Activation at Class-I Cap-Dependent Promoters. *Molecular Microbiology*, 8, 797-802.
- EICHHORN, E., VAN DER PLOEG, J. R. & LEISINGER, T. 1999. Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry*, 274, 26639-26646.
- EICHHORN, E., VAN DER PLOEG, J. R. & LEISINGER, T. 2000. Deletion analysis of the *Escherichia coli* taurine and alkanesulfonate transport systems. *Journal of Bacteriology*, 182, 2687-2695.
- ELLISON, D. W. & MILLER, V. L. 2006. H-NS represses *inv* transcription in *Yersinia enterocolitica* through competition with RovA and interaction with YmoA. *Journal of Bacteriology*, 188, 5101-5112.
- ESQUERRE, T., LAGUERRE, S., TURLAN, C., CARPOUSIS, A. J., GIRBAL, L. & COCAIGN-BOUSQUET, M. 2014. Dual role of transcription and transcript stability

in the regulation of gene expression in *Escherichia coli* cells cultured on glucose at different growth rates. *Nucleic Acids Research*, 42, 2460-2472.

- FANG, M., MAJUMDER, A., TSAI, K. J. & WU, H. Y. 2000. ppGpp-dependent *leuO* expression in bacteria under stress. *Biochemical and Biophysical Research Communications*, 276, 64-70.
- FERRANDEZ, A., GARCIA, J. L. & DIAZ, E. 2000. Transcriptional regulation of the divergent *paa* catabolic operons for phenylacetic acid degradation in *Escherichia coli*. *Journal of Biological Chemistry*, 275, 12214-12222.
- FERRANDEZ, A., MINAMBRES, B., GARCIA, B., OLIVERA, E. R., LUENGO, J. M., GARCIA, J. L. & DIAZ, E. 1998. Catabolism of phenylacetic acid in *Escherichia coli* - Characterization of a new aerobic hybrid pathway. *Journal of Biological Chemistry*, 273, 25974-25986.
- FIC, E., BONAREK, P., GORECKI, A., KEDRACKA-KROK, S., MIKOLAJCZAK, J., POLIT, A., TWORZYDLO, M., DZIEDZICKA-WASYLEWSKA, M. & WASYLEWSKI, Z. 2009. cAMP Receptor Protein from *Escherichia coli* as a Model of Signal Transduction in Proteins - A Review. *Journal of Molecular Microbiology and Biotechnology*, 17, 1-11.
- FINN, R. D., BATEMAN, A., CLEMENTS, J., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., HEGER, A., HETHERINGTON, K., HOLM, L., MISTRY, J., SONNHAMMER, E. L. L., TATE, J. & PUNTA, M. 2014. Pfam: the protein families database. *Nucleic Acids Research*, 42, D222-D230.
- FRANCETIC, O., BELIN, D., BADAUT, C. & PUGSLEY, A. P. 2000. Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. *Embo Journal*, 19, 6697-6703.
- FRANCETIC, O. & PUGSLEY, A. P. 1996. The cryptic general secretory pathway (*gsp*) operon of *Escherichia coli* K-12 encodes functional proteins. *Journal of Bacteriology*, 178, 3544-3549.
- FU, J. H., GNATT, A. L., BUSHNELL, D. A., JENSEN, G. J., THOMPSON, N. E., BURGESS, R. R., DAVID, P. R. & KORNBERG, R. D. 1999. Yeast RNA polymerase II at 5 angstrom resolution. *Cell*, 98, 799-810.
- GAJIWALA, K. S., CHEN, H., CORNILLE, F., ROQUES, B. P., REITH, W., MACH, B. & BURLEY, S. K. 2000. Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature*, 403, 916-921.
- GALLY, D. L., BOGAN, J. A., EISENSTEIN, B. I. & BLOMFIELD, I. C. 1993. Environmental-regulation of the fim switch controlling type-1 fimbrial phase variation in *Escherichia coli* K-12 - effects of temperature and media. *Journal of Bacteriology*, 175, 6186-6193.

- GEERTZ, M., TRAVERS, A., MEHANDZISKA, S., SOBETZKO, P., JANGA, S. C., SHIMAMOTO, N. & MUSKHELISHVILI, G. 2011. Structural Coupling between RNA Polymerase Composition and DNA Supercoiling in Coordinating Transcription: a Global Role for the Omega Subunit? *Mbio*, 2.
- GORKE, B. & STULKE, J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature reviews. Microbiology*, 6, 613-24.
- GOURSE, R. L., ROSS, W. & GAAL, T. 2000. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Molecular Microbiology*, 37, 687-695.
- GRAINGER, D. C., WEBSTER, C. L., BELYAEVA, T. A., HYDE, E. I. & BUSBY, S. J. W. 2004. Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with its DNA target site and with domain 4 of the RNA polymerase sigma subunit. *Molecular Microbiology*, 51, 1297-1309.
- GRIGOROVA, I. L., PHLEGER, N. J., MUTALIK, V. K. & GROSS, C. A. 2006. Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 5332-5337.
- GRISHIN, A. M., AJAMIAN, E., TAO, L. M., ZHANG, L. H., MENARD, R. & CYGLER, M. 2011. Structural and Functional Studies of the *Escherichia coli* Phenylacetyl-CoA Monooxygenase Complex. *Journal of Biological Chemistry*, 286, 10735-10743.
- GROSS, C. A., CHAN, C., DOMBROSKI, A., GRUBER, T., SHARP, M., TUPY, J. & YOUNG, B. 1998. The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor Symposia on Quantitative Biology*, 63, 141-155.
- GRUBER, T. M. & GROSS, C. A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology*, 57, 441-466.
- HAIDER, F., LITHGOW, J. K., STAPLETON, M. R., NORTE, V. A., ROBERTS, R. E. & GREEN, J. 2008. DNA recognition by the *Salmonella enterica* Serovar Typhimurium transcription factor SlyA. *International microbiology : the official journal of the Spanish Society for Microbiology*, 11, 245-50.
- HAN, X., DORSEY-ORESTO, A., MALIK, M., WANG, J.-Y., DRLICA, K., ZHAO, X. & LU, T. 2010. *Escherichia coli* genes that reduce the lethal effects of stress. *Bmc Microbiology*, 10.
- HARLE, C., KIM, I., ANGERER, A. & BRAUN, V. 1995. Signal transfer through 3 compartments - transcription initiation of the *Escherichia coli* ferric citrate transport system from the cell surface. *Embo Journal*, 14, 1430-1438.

- HARLEY, C. B. & REYNOLDS, R. P. 1987. Analysis of *Escherichia coli* promoter sequences. *Nucleic Acids Research*, 15, 2343-2361.
- HERZBERG, M., KAYE, I. K., PETI, W. & WOOD, T. K. 2006. YdgG (TqsA) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. *Journal of Bacteriology*, 188, 587-598.
- HINTON, D. M., PANDE, S., WAIS, N., JOHNSON, X. B., VUTHOORI, M., MAKELA, A. & HOOK-BARNARD, I. 2005. Transcriptional takeover by sigma appropriation: remodelling of the sigma(70) subunit of *Escherichia coli* RNA polymerase by the bacteriophage T4 activator MotA and co-activator AsiA. *Microbiology-Sgm*, 151, 1729-1740.
- HOLDSWORTH, S. R. & LAW, C. J. 2012. Functional and biochemical characterisation of the *Escherichia coli* major facilitator superfamily multidrug transporter MdtM. *Biochimie*, 94, 1334-1346.
- HOLDSWORTH, S. R. & LAW, C. J. 2013. Multidrug resistance protein MdtM adds to the repertoire of antiporters involved in alkaline pH homeostasis in *Escherichia coli*. *Bmc Microbiology*, 13.
- HOLLYWOOD, N. & DOELLE, H. W. 1976. Effect of specific growth-rate and glucose concentration on growth and glucose-metabolism of *Escherichia coli* K-12. *Microbios*, 17, 23-33.
- HOLMS, H. 1996. Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *Fems Microbiology Reviews*, 19, 85-116.
- HOMMAIS, F., KRIN, E., LAURENT-WINTER, C., SOUTOURINA, O., MALPERTUY, A., LE CAER, J. P., DANCHIN, A. & BERTIN, P. 2001. Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular Microbiology*, 40, 20-36.
- HOOK-BARNARD, I. G. & HINTON, D. M. 2007. Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene regulation and systems biology*, 1, 275-93.
- HORVATH, P. & BARRANGOU, R. 2010. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327, 167-170.
- HUDSON, B. P., QUISPE, J., LARA-GONZALEZ, S., KIM, Y., BERMAN, H. M., ARNOLD, E., EBRIGHT, R. H. & LAWSON, C. L. 2009. Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 19830-19835.



- HUGHES, K. T., GILLEN, K. L., SEMON, M. J. & KARLINSEY, J. E. 1993. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science*, 262, 1277-1280.
- IGARASHI, K., HANAMURA, A., MAKINO, K., AIBA, H., MIZUNO, T., NAKATA, A. & ISHIHAMA, A. 1991. Functional map of the alpha-subunit of *Escherichia coli* RNA-polymerase - 2 modes of transcription activation by positive factors. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 8958-8962.
- ISHIHAMA, A. 2000. Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol*, 54, 499-518.
- ISLAM, M. S., BINGLE, L. E. H., PALLEN, M. J. & BUSBY, S. J. W. 2011. Organization of the LEE1 operon regulatory region of enterohaemorrhagic *Escherichia coli* O157:H7 and activation by GrlA. *Molecular Microbiology*, 79, 468-483.
- ISMAIL, W., MOHAMED, M. E., WANNER, B. L., DATSENKO, K. A., EISENREICH, W., ROHDICH, F., BACHER, A. & FUCHS, G. 2003. Functional genomics by NMR spectroscopy - Phenylacetate catabolism in *Escherichia coli*. *European Journal of Biochemistry*, 270, 3047-3054.
- ITOU, J., EGUCHI, Y. & UTSUMI, R. 2009. Molecular Mechanism of Transcriptional Cascade Initiated by the EvgS/EvgA System in *Escherichia coli* K-12. *Bioscience Biotechnology and Biochemistry*, 73, 870-878.
- JAIN, D., NICKELS, B. E., SUN, L., HOCHSCHILD, A. & DARST, S. A. 2004. Structure of a ternary transcription activation complex. *Molecular Cell*, 13, 45-53.
- KALIR, S. & ALON, U. 2004. Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell*, 117, 713-720.
- KESELER, I. M., MACKIE, A., PERALTA-GIL, M., SANTOS-ZAVALA, A., GAMACASTRO, S., BONAVIDES-MARTINEZ, C., FULCHER, C., HUERTA, A. M., KOTHARI, A., KRUMMENACKER, M., LATENDRESSE, M., MUNIZ-RASCADO, L., ONG, Q., PALEY, S., SCHROEDER, I., SHEARER, A. G., SUBHRAVETI, P., TRAVERS, M., WEERASINGHE, D., WEISS, V., COLLADO-VIDES, J., GUNSALUS, R. P., PAULSEN, I. & KARP, P. D. 2013. EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Research*, 41, D605-D612.
- KLUMPP, S., ZHANG, Z. G. & HWA, T. 2009. Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139, 1366-1375.

- KOLB, A., BUSBY, S., BUC, H., GARGES, S. & ADHYA, S. 1993. Transcriptional regulation by cAMP and its receptor protein. *Annual Review of Biochemistry*, 62, 749-795.
- KOO, B.-M., RHODIUS, V. A., CAMPBELL, E. A. & GROSS, C. A. 2009. Mutational analysis of *Escherichia coli* sigma(28) and its target promoters reveals recognition of a composite-10 region, comprised of an 'extended-10' motif and a core-10 element. *Molecular Microbiology*, 72, 830-843.
- KOREA, C. G., BADOURALY, R., PREVOST, M. C., GHIGO, J. M. & BELOIN, C. 2010. *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environmental Microbiology*, 12, 1957-1977.
- KORZHEVA, N., MUSTAEV, A., KOZLOV, M., MALHOTRA, A., NIKIFOROV, V., GOLDFARB, A. & DARST, S. A. 2000. A structural model of transcription elongation. *Science*, 289, 619-625.
- KUBITSCHKEK, H. E. & FREEDMAN, M. L. 1971. Chromosome Replication and the Division Cycle of *Escherichia coli* B/r. *Journal of Bacteriology*, 107, 95-99.
- KUO, J. T., CHANG, Y. J. & TSENG, C. P. 2003. Growth rate regulation of *lac* operon expression in *Escherichia coli* is cyclic AMP dependent. *Febs Letters*, 553, 397-402.
- KUTSUKAKE, K. 1994. Excretion of the anti-sigma factor through a flagellar substructure couples flagellar gene-expression with flagellar assembly in *Salmonella typhimurium*. *Molecular & General Genetics*, 243, 605-612.
- KUTSUKAKE, K. & IINO, T. 1994. Role of the FliA-FlgM regulatory system on the transcriptional control of the flagellar regulon and flagellar formation in *Salmonella typhimurium*. *Journal of Bacteriology*, 176, 3598-3605.
- KUTSUKAKE, K., OHYA, Y. & IINO, T. 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *Journal of Bacteriology*, 172, 741-747.
- LAEMMLI, U. K. 1970. Cleavage of structural proteins during assembly of head of bacteriophage T4. *Nature*, 227, 680-&.
- LEE, S. J., XIE, A. G., JIANG, W. N., ETCHEGARAY, J. P., JONES, P. G. & INOUE, M. 1994. Family of the major cold-shock protein, CspA (CS7.4), of *Escherichia coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding-proteins. *Molecular Microbiology*, 11, 833-839.
- LEWIS, M. 2005. The *lac* repressor. *Comptes Rendus Biologies*, 328, 521-548.

- LI, J., ATTLA, C., WANG, L., WOOD, T. K., VALDES, J. J. & BENTLEY, W. E. 2007. Quorum sensing in *Escherichia coli* is signaled by AI-2/LsrR: Effects on small RNA and Biofilm architecture. *Journal of Bacteriology*, 189, 6011-6020.
- LIANG, S. T., XU, Y. C., DENNIS, P. & BREMER, H. 2000. mRNA composition and control of bacterial gene expression. *Journal of Bacteriology*, 182, 3037-3044.
- LIBBY, S. J., GOEBEL, W., LUDWIG, A., BUCHMEIER, N., BOWE, N., FANG, F. C., GUINEY, D. G., SONGER, J. G. & HEFFRON, F. 1994. A Cytolysin Encoded By *Salmonella* Is Required For Survival Within Macrophages. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 489-493.
- LIU, X. Y. & MATSUMURA, P. 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class-II operons. *Journal of Bacteriology*, 176, 7345-7351.
- LIU, X. Y. & MATSUMURA, P. 1995. An alternative sigma factor controls transcription of flagellar class-III operons in *Escherichia coli* - gene sequence, overproduction, purification and characterization. *Gene*, 164, 81-84.
- LIU, X. Y. & MATSUMURA, P. 1996. Differential regulation of multiple overlapping promoters in flagellar class II operons in *Escherichia coli*. *Molecular Microbiology*, 21, 613-620.
- LLOYD, G. S., NIU, W., TEBBUTT, J., EBRIGHT, R. H. & BUSBY, S. J. W. 2002. Requirement for two copies of RNA polymerase alpha subunit C-terminal domain for synergistic transcription activation at complex bacterial promoters. *Genes & Development*, 16, 2557-2565.
- LOPILATO, J. E., GARWIN, J. L., EMR, S. D., SILHAVY, T. J. & BECKWITH, J. R. 1984. D-ribose metabolism in *Escherichia coli* K-12 - genetics, regulation, and transport. *Journal of Bacteriology*, 158, 665-673.
- MACIAG, A., PEANO, C., PIETRELLI, A., EGLI, T., DE BELLIS, G. & LANDINI, P. 2011. *In vitro* transcription profiling of the Sigma(S) subunit of bacterial RNA polymerase: re-definition of the Sigma(S) regulon and identification of Sigma(S)-specific promoter sequence elements. *Nucleic Acids Research*, 39, 5338-5355.
- MACNAB, R. M. 1992. Genetics and biogenesis of bacterial flagella. *Annual Review of Genetics*, 26, 131-158.
- MACNAB, R. M. 1996. In F. C. Neidhart et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C., 123-145.

- MAEDA, H., FUJITA, N. & ISHIHAMA, A. 2000. Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Research*, 28, 3497-3503.
- MAROLDA, C. L. & VALVANO, M. A. 1995. Genetic Analysis of the dTDP-Rhamnose Biosynthesis Region of the *Escherichia coli* VW187 (O7:K1) *rfb* Gene Cluster: Identification of Functional Homologs of *rfbB* and *rfbA* in the *rff* Cluster and Correct Location of the *rffE* Gene. *Journal of Bacteriology*, 177, 5539-5546.
- MARQUES, J. C., LAMOSAS, P., RUSSELL, C., VENTURA, R., MAYCOCK, C., SEMMELHACK, M. F., MILLER, S. T. & XAVIER, K. B. 2011. Processing the Interspecies Quorum-sensing Signal Autoinducer-2 (AI-2): Characterization of phospho-(s)-4,5-dihydroxy-2,3-pentanedione isomerization by LsrG protein. *Journal of Biological Chemistry*, 286, 18331-18343.
- MARTIN, R. G., BARTLETT, E. S., ROSNER, J. L. & WALL, M. E. 2008. Activation of the *Escherichia coli* *marA/soxS/rob* regulon in response to transcriptional activator concentration. *Journal of Molecular Biology*, 380, 278-284.
- MARTIN, R. G. & ROSNER, J. L. 1995. Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 5456-5460.
- MARTINEZ-ANTONIO, A. & COLLADO-VIDES, J. 2003. Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology*, 6, 482-489.
- MAUZY, C. A. & HERMODSON, M. A. 1992. Structural homology between *rbs* repressor and ribose binding-protein implies functional similarity. *Protein Science*, 1, 843-849.
- MCLEOD, S. M. & JOHNSON, R. C. 2001. Control of transcription by nucleoid proteins. *Current Opinion in Microbiology*, 4, 152-159.
- MCVICKER, G., SUN, L., SOHANPAL, B. K., GASHI, K., WILLIAMSON, R. A., PLUMBRIDGE, J. & BLOMFIELD, I. C. 2011. SlyA Protein Activates *fimB* Gene Expression and Type 1 Fimbriation in *Escherichia coli* K-12. *Journal of Biological Chemistry*, 286, 32026-32035.
- MEIBOM, K. L., KALLIPOLITIS, B. H., EBRIGHT, R. H. & VALENTIN-HANSEN, P. 2000. Identification of the subunit of cAMP receptor protein (CRP) that functionally interacts with CytR in CRP-CytR-mediated transcriptional repression. *Journal of Biological Chemistry*, 275, 11951-11956.
- MELLIES, J., WISE, A. & VILLAREJO, M. 1995. Two Different *Escherichia coli* *proP* Promoters Respond to Osmotic and Growth Phase Signals. *Journal of Bacteriology*, 177, 144-151.

- MERRICK, M. J. 1993. In a class of its own - the RNA-polymerase sigma-factor sigma(54) (sigma(n)). *Molecular Microbiology*, 10, 903-909.
- MILLER, M. B. & BASSLER, B. L. 2001. Quorum sensing in bacteria. *Annual Review of Microbiology*, 55, 165-199.
- MILLER, S. T., XAVIER, K. B., CAMPAGNA, S. R., TAGA, M. E., SEMMELHACK, M. F., BASSLER, B. L. & HUGHSON, F. M. 2004. *Salmonella typhimurium* recognizes a chemically distinct form of the bacterial quorum-sensing signal AI-2. *Molecular Cell*, 15, 677-687.
- MOORES, A., CHIPPER-KEATING, S., SUN, L., MCVICKER, G., WALES, L., GASHI, K. & BLOMFIELD, I. C. 2014. RfaH Suppresses Small RNA MicA Inhibition of *fimB* Expression in *Escherichia coli* K-12. *Journal of Bacteriology*, 196, 148-156.
- MUKHERJEE, K. & CHATTERJI, D. 1997. Studies on the omega subunit of *Escherichia coli* RNA polymerase - Its role in the recovery of denatured enzyme activity. *European Journal of Biochemistry*, 247, 884-889.
- MULLER-HILL, B. 1996. The *lac* operon: A short history of a genetic paradigm. *Walter de Gruyter, Berlin*.
- MURAKAMI, K. S., MASUDA, S., CAMPBELL, E. A., MUZZIN, O. & DARST, S. A. 2002a. Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science*, 296, 1285-1290.
- MURAKAMI, K. S., MASUDA, S. & DARST, S. A. 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 angstrom resolution. *Science*, 296, 1280-1284.
- MURAKAMI, K. S., MASUDA, S. & DARST, S. A. 2002c. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 angstrom resolution. *Science*, 296, 1280-1284.
- NAGEL, G., LAHRZ, A. & DERSCH, P. 2001. Environmental control of invasins expression in *Yersinia pseudotuberculosis* is mediated by regulation of RovA, a transcriptional activator of the SlyA/Hor family. *Molecular Microbiology*, 41, 1249-1269.
- NAHKU, R., VALGEPEA, K., LAHTVEE, P. J., ERM, S., ABNER, K., ADAMBERG, K. & VILU, R. 2010. Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures. *Journal of Biotechnology*, 145, 60-65.
- NAKANISHI, N., ABE, H., OGURA, Y., HAYASHI, T., TASHIRO, K., KUHARA, S., SUGIMOTO, N. & TOBE, T. 2006. ppGpp with DksA controls gene expression in

the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic *Escherichia coli* through activation of two virulence regulatory genes. *Molecular Microbiology*, 61, 194-205.

NASSER, W., SHEVCHIK, V. E. & HUGOUVIEUX-COTTE-PATTAT, N. 1999. Analysis of three clustered polygalacturonase genes in *Erwinia chrysanthemi* 3937 revealed an anti-repressor function for the PecS regulator. *Molecular Microbiology*, 34, 641-650.

NAVARRE, W. W., HALSEY, T. A., WALTHERS, D., FRYE, J., MCCLELLAND, M., POTTER, J. L., KENNEY, L. J., GUNN, J. S., FANG, F. C. & LIBBY, S. J. 2005. Co-regulation of *Salmonella enterica* genes required for virulence and resistance to antimicrobial peptides by SlyA and PhoP/PhoQ. *Molecular Microbiology*, 56, 492-508.

NEIDHARDT, F. C., INGRAHAM, J. L. & SCHAECHTER, M. 1990. *Physiology of the Bacterial Cell: A Molecular Approach*, Sinauer Associates, Inc., Sunderland, Mass.

NEWBERRY, K. J. & BRENNAN, R. G. 2004. The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *Journal of Biological Chemistry*, 279, 20356-20362.

NG, W.-L. & BASSLER, B. L. 2009. Bacterial Quorum-Sensing Network Architectures. *Annual Review of Genetics*.

NICKELS, B. E., DOVE, S. L., MURAKAMI, K. S., DARST, S. A. & HOCHSCHILD, A. 2002. Protein-protein and protein-DNA interactions of sigma(70) region 4 involved in transcription activation by lambda cl. *Journal of Molecular Biology*, 324, 17-34.

NICOLAOU, S. A., FAST, A. G., NAKAMARU-OGISO, E. & PAPOUTSAKIS, E. T. 2013. Overexpression of *fetA* (*ybbL*) and *fetB* (*ybbM*), Encoding an Iron Exporter, Enhances Resistance to Oxidative Stress in *Escherichia coli*. *Applied and Environmental Microbiology*, 79, 7210-7219.

NORTE, V. A., STAPLETON, M. R. & GREEN, J. 2003. PhoP-responsive expression of the *Salmonella enterica* Serovar Typhimurium *slyA* gene. *Journal of Bacteriology*, 185, 3508-3514.

OH, S.-Y., SHIN, J.-H. & ROE, J.-H. 2007. Dual role of OhrR as a repressor and an activator in response to organic hydroperoxides in *Streptomyces coelicolor*. *Journal of Bacteriology*, 189, 6284-6292.

OKADA, N., OI, Y., TAKEDA-SHITAKA, M., KANOU, K., UMEYAMA, H., HANEDA, T., MIKI, T., HOSOYA, S. & DANBARA, H. 2007. Identification of amino acid residues of *Salmonella* SlyA that are critical for transcriptional regulation. *Microbiology-Sgm*, 153, 548-560.

- OSCARSSON, J., MIZUNOE, Y., UHLIN, B. E. & HAYDON, D. J. 1996. Induction of haemolytic activity in *Escherichia coli* by the *slyA* gene product. *Molecular Microbiology*, 20, 191-199.
- PARROTT, S., JONES, S. & COOPER, R. A. 1987. 2-Phenylethylamine Catabolism by *Escherichia coli* K12. *Journal of General Microbiology*, 133, 347-351.
- PEDERSEN, H., SOGAARDANDERSEN, L., HOLST, B. & VALENTINHANSEN, P. 1991. Heterologous cooperativity in *Escherichia coli* - the CytR repressor both contacts DNA and the cAMP receptor protein when binding to the *deoP2* promoter. *Journal of Biological Chemistry*, 266, 17804-17808.
- PERERA, I. C. & GROVE, A. 2010. Molecular Mechanisms of Ligand-Mediated Attenuation of DNA Binding by MarR Family Transcriptional Regulators. *Journal of Molecular Cell Biology*, 2, 243-254.
- PEREZ, J. C., LATIFI, T. & GROISMAN, E. A. 2008. Overcoming H-NS-mediated transcriptional silencing of horizontally acquired genes by the PhoP and SlyA proteins in *Salmonella enterica*. *Journal of Biological Chemistry*, 283, 10773-10783.
- PETERSEN, C., MOLLER, L. B. & VALENTIN-HANSEN, P. 2002. The cryptic adenine deaminase gene of *Escherichia coli* - Silencing by the nucleoid-associated DNA-binding protein, H-NS, and activation by insertion elements. *Journal of Biological Chemistry*, 277, 31373-31380.
- PILSL, H., SMAJS, D. & BRAUN, V. 1999. Characterization of colicin S4 and its receptor, OmpW, a minor protein of the *Escherichia coli* outer membrane. *Journal of Bacteriology*, 181, 3578-3581.
- POTRYKUS, K. & CASHEL, M. 2008. (p)ppGpp: Still Magical? *Annual Review of Microbiology*. Palo Alto: Annual Reviews.
- PRESSLER, U., STAUDENMAIER, H., ZIMMERMANN, L. & BRAUN, V. 1988. Genetics of the iron dicitrate transport-system of *Escherichia coli*. *Journal of Bacteriology*, 170, 2716-2724.
- RAETZ, C. R. H. & WHITFIELD, C. 2002. Lipopolysaccharide endotoxins. *Annual Review of Biochemistry*, 71, 635-700.
- REIZER, J., MICHOTEY, V., REIZER, A. & SAIER, M. H. 1994. Novel phosphotransferase system genes revealed by bacterial genome analysis - unique, putative fructose-specific and glucoside-specific systems. *Protein Science*, 3, 440-450.

- REIZER, J., REIZER, A. & SAIER, M. H. 1997. Is the ribulose monophosphate pathway widely distributed in bacteria? *Microbiology-Uk*, 143, 2519-2520.
- REVELL, P. A. & MILLER, V. L. 2000. A chromosomally encoded regulator is required for expression of the *Yersinia enterocolitica inv* gene and for virulence. *Molecular Microbiology*, 35, 677-685.
- REVERCHON, S., ROUANET, C., EXPERT, D. & NASSER, W. 2002. Characterization of indigoidine biosynthetic genes in *Erwinia chrysanthemi* and role of this blue pigment in pathogenicity. *Journal of Bacteriology*, 184, 654-665.
- RHODIUS, V. A. & BUSBY, S. J. W. 2000. Interactions between activating region 3 of the *Escherichia coli* cyclic AMP receptor protein and region 4 of the RNA polymerase sigma(70) subunit: Application of suppression genetics. *Journal of Molecular Biology*, 299, 311-324.
- RICHET, E., VIDALINGIARDI, D. & RAIBAUD, O. 1991. A New Mechanism for Coactivation of Transcription Initiation - Repositioning of an Activator Triggered by the Binding of a 2nd Activator. *Cell*, 66, 1185-1195.
- ROSS, W., ERNST, A. & GOURSE, R. L. 2001. Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes & Development*, 15, 491-506.
- ROSS, W., SCHNEIDER, D. A., PAUL, B. J., MERTENS, A. & GOURSE, R. L. 2003. An intersubunit contact stimulating transcription initiation by *E. coli* RNA polymerase: interaction of the alpha C-terminal domain and sigma region 4. *Genes & Development*, 17, 1293-1307.
- ROSSITER, A. E., BROWNING, D. F., LEYTON, D. L., JOHNSON, M. D., GODFREY, R. E., WARDIUS, C. A., DESVAUX, M., CUNNINGHAM, A. F., RUIZ-PEREZ, F., NATARO, J. P., BUSBY, S. J. W. & HENDERSON, I. R. 2011. Transcription of the plasmid-encoded toxin gene from Enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Molecular Microbiology*, 81, 179-191.
- ROUANET, C., NOMURA, K., TSUYUMU, S. & NASSER, W. 1999. Regulation of *pelD* and *pelE*, encoding major alkaline pectate lyases in *Erwinia chrysanthemi*: Involvement of the main transcriptional factors. *Journal of Bacteriology*, 181, 5948-5957.
- ROUANET, C., REVERCHON, S., RODIONOV, D. A. & NASSER, W. 2004. Definition of a consensus DNA-binding site for PecS, a global regulator of virulence gene expression in *Erwinia chrysanthemi* and identification of new members of the PecS regulon. *Journal of Biological Chemistry*, 279, 30158-30167.



- SAINI, S., BROWN, J. D., ALDRIDGE, P. D. & RAO, C. V. 2008. FlhZ is a posttranslational activator of FlhD(4)C(2)-dependent flagellar gene expression. *Journal of Bacteriology*, 190, 4979-4988.
- SANCHEZ, A., OSBORNE, M. L., FRIEDMAN, L. J., KONDEV, J. & GELLES, J. 2011. Mechanism of transcriptional repression at a bacterial promoter by analysis of single molecules. *Embo Journal*, 30, 3940-3946.
- SANDERSON, A., MITCHELL, J. E., MINCHIN, S. D. & BUSBY, S. J. W. 2003. Substitutions in the *Escherichia coli* RNA polymerase sigma(70) factor that affect recognition of extended-10 elements at promoters. *Febs Letters*, 544, 199-205.
- SAWERS, G., KAISER, M., SIRKO, A. & FREUNDLICH, M. 1997. Transcriptional activation by FNR and CRP: Reciprocity of binding-site recognition. *Molecular Microbiology*, 23, 835-845.
- SCHAECHTER, M., MAALOE, O. & KJELDGAARD, N. O. 1958. Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Journal of General Microbiology*, 19, 592-606.
- SCHAUDER, S., SHOKAT, K., SURETTE, M. G. & BASSLER, B. L. 2001. The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Molecular Microbiology*, 41, 463-476.
- SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671-675.
- SCOTT, S., BUSBY, S. & BEACHAM, I. 1995. Transcriptional co-activation at the *ansB* promoters - involvement of the activating regions of CRP and FNR when bound in tandem. *Molecular Microbiology*, 18, 521-531.
- SEMSEY, S., GEANACOPOULOS, M., LEWIS, D. E. A. & ADHYA, S. 2002. Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *Embo Journal*, 21, 4349-4356.
- SEMSEY, S., KRISHNA, S., SNEPPEN, K. & ADHYA, S. 2007. Signal integration in the galactose network of *Escherichia coli*. *Molecular Microbiology*, 65, 465-476.
- SEMSEY, S., TOLSTORUKOV, M. Y., VIRNIK, K., ZHURKIN, V. B. & ADHYA, S. 2004. DNA trajectory in the Gal repressosome. *Genes & Development*, 18, 1898-1907.
- SHAH, I. M. & WOLF, R. E. 2004. Novel protein-protein interaction between *Escherichia coli* SoxS and the DNA binding determinant of the RNA polymerase alpha subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-

- element-containing promoters to SoxS-dependent promoters during oxidative stress. *Journal of Molecular Biology*, 343, 513-532.
- SHAW, D. J., RICE, D. W. & GUEST, J. R. 1983. Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli*. *Journal of Molecular Biology*, 166, 241-247.
- SHI, Y. X., LATIFI, T., CROMIE, M. J. & GROISMAN, E. A. 2004. Transcriptional control of the antimicrobial peptide resistance *ugtL* gene by the *Salmonella* PhoP and SlyA regulatory proteins. *Journal of Biological Chemistry*, 279, 38618-38625.
- SHIMADA, T., BRIDIER, A., BRIANDET, R. & ISHIHAMA, A. 2011. Novel roles of LeuO in transcription regulation of *E. coli* genome: antagonistic interplay with the universal silencer H-NS. *Molecular Microbiology*, 82, 378-397.
- SHIN, M., KANG, S., HYUN, S. J., FUJITA, N., ISHIHAMA, A., VALENTIN-HANSEN, P. & CHOY, H. E. 2001. Repression of *deoP2* in *Escherichia coli* by CytR: conversion of a transcription activator into a repressor. *Embo Journal*, 20, 5392-5399.
- SPERANDIO, V., TORRES, A. G., GIRON, J. A. & KAPER, J. B. 2001. Quorum sensing is a global regulatory mechanism in enterohemorrhagic *Escherichia coli* O157 : H7. *Journal of Bacteriology*, 183, 5187-5197.
- SPERANDIO, V., TORRES, A. G., JARVIS, B., NATARO, J. P. & KAPER, J. B. 2003. Bacteria-host communication: The language of hormones. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 8951-8956.
- SPERANDIO, V., TORRES, A. G. & KAPER, J. B. 2002. Quorum sensing *Escherichia coli* regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Molecular Microbiology*, 43, 809-821.
- SPORY, A., BOSSERHOFF, A., VON RHEIN, C., GOEBEL, W. & LUDWIG, A. 2002. Differential regulation of multiple proteins of *Escherichia coli* and *Salmonella enterica* Serovar Typhimurium by the transcriptional regulator SlyA. *Journal of Bacteriology*, 184, 3549-3559.
- STAPLETON, M. R., NORTE, V. A., READ, R. C. & GREEN, J. 2002. Interaction of the *Salmonella typhimurium* transcription and virulence factor SlyA with target DNA and identification of members of the SlyA regulon. *Journal of Biological Chemistry*, 277, 17630-17637.
- STERN, R. J., LEE, T. Y., LEE, T. J., YAN, W. X., SCHERMAN, M. S., VISSA, V. D., KIM, S. K., WANNER, B. L. & MCNEIL, M. R. 1999. Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the *rmlC* gene products of *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology-Uk*, 145, 663-671.

- STEVENSON, G., NEAL, B., LIU, D., HOBBS, M., PACKER, N. H., BATLEY, M., REDMOND, J. W., LINDQUIST, L. & REEVES, P. 1994. Structure of the O Antigen of *Escherichia coli* K-12 and the Sequence of Its *rfb* Gene Cluster. *Journal of Bacteriology*, 176, 4144-4156.
- STOCK, A. M., ROBINSON, V. L. & GOUDREAU, P. N. 2000. Two-component signal transduction. *Annual Review of Biochemistry*, 69, 183-215.
- STRANEY, D. C., STRANEY, S. B. & CROTHERS, D. M. 1989. Synergy between *Escherichia coli* CAP protein and RNA polymerase in the *lac* promoter open complex. *Journal of Molecular Biology*, 206, 41-57.
- STRATMANN, T., MADHUSUDAN, S. & SCHNETZ, K. 2008. Regulation of the *yjjQ-bglJ* operon, encoding LuxR-type transcription factors, and the divergent *yjjP* gene by H-NS and LeuO. *Journal of Bacteriology*, 190, 926-935.
- STUDIER, F. W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41, 207-234.
- SUNDARARAJ, S., GUO, A., HABIBI-NAZHAD, B., ROUANI, M., STOTHARD, P., ELLISON, M. & WISHART, D. S. 2004. The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of *Escherichia coli*. *Nucleic Acids Research*, 32, D293-D295.
- SURETTE, M. G. & BASSLER, B. L. 1998. Quorum sensing in *Escherichia coli* and *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 7046-7050.
- TEUFEL, R., MASCARAQUE, V., ISMAIL, W., VOSS, M., PERERA, J., EISENREICH, W., HAEHNEL, W. & FUCHS, G. 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 14390-14395.
- TSUJIKAWA, L., TSODIKOV, O. V. & DEHASETH, P. L. 2002. Interaction of RNA polymerase with forked DNA: evidence for two kinetically significant intermediates on the pathway to the final complex. *Proc Natl Acad Sci U S A*, 99, 3493-8.
- USHIDA, C. & AIBA, H. 1990. Helical phase dependent action of CRP - effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic Acids Research*, 18, 6325-6330.
- VALGEPEA, K., ADAMBERG, K., SEIMAN, A. & VILU, R. 2013. *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Molecular Biosystems*, 9, 2344-2358.

- VAN DER PLOEG, J. R., IWANICKA-NOWICKA, R., BYKOWSKI, T., HRYNIEWICZ, M. M. & LEISINGER, T. 1999. The *Escherichia coli* *ssuEADCB* gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. *Journal of Biological Chemistry*, 274, 29358-29365.
- VAN HOEK, M. J. A. & MERKS, R. M. H. 2012. Redox balance is key to explaining full vs. partial switching to low-yield metabolism. *Bmc Systems Biology*, 6.
- VEMURI, G. N., ALTMAN, E., SANGURDEKAR, D. P., KHODURSKY, A. B. & EITEMAN, M. A. 2006. Overflow metabolism in *Escherichia coli* during steady-state growth: Transcriptional regulation and effect of the redox ratio. *Applied and Environmental Microbiology*, 72, 3653-3661.
- VOLKMER, B. & HEINEMANN, M. 2011. Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data Conversion for Systems Biology Modeling. *PLoS One*, 6.
- VRENTAS, C. E., GAAL, T., ROSS, W., EBRIGHT, R. H. & GOURSE, R. L. 2005. Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA. *Genes & Development*, 19, 2378-2387.
- WADA, A., IGARASHI, K., YOSHIMURA, S., AIMOTO, S. & ISHIHAMA, A. 1995. Ribosome modulation factor - stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 214, 410-417.
- WADE, J. T., BELYAEVA, T. A., HYDE, E. I. & BUSBY, S. J. W. 2001. A simple mechanism for co-dependence on two activators at an *Escherichia coli* promoter. *Embo Journal*, 20, 7160-7167.
- WADE, J. T., ROA, D. C., GRAINGER, D. C., HURD, D., BUSBY, S. J. W., STRUHL, K. & NUDLER, E. 2006. Extensive functional overlap between sigma factors in *Escherichia coli*. *Nature Structural & Molecular Biology*, 13, 806-814.
- WAGNER, N. J., LIN, C. P., BORST, L. B. & MILLER, V. L. 2013. YaxAB, a *Yersinia enterocolitica* Pore-Forming Toxin Regulated by RovA. *Infection and Immunity*, 81, 4208-4219.
- WANG, L., HASHIMOTO, Y., TSAO, C. Y., VALDES, J. J. & BENTLEY, W. E. 2005a. Cyclic AMP (cAMP) and cAMP receptor protein influence both synthesis and uptake of extracellular autoinducer 2 in *Escherichia coli*. *Journal of Bacteriology*, 187, 2066-2076.
- WANG, L., LI, J., MARCH, J. C., VALDES, J. J. & BENTLEY, W. E. 2005b. *luxS*-Dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling. *Journal of Bacteriology*, 187, 8350-8360.

- WEICKERT, M. J. & ADHYA, S. 1992. A family of bacterial regulators homologous to Gal and Lac repressors. *Journal of Biological Chemistry*, 267, 15869-15874.
- WESTRA, E. R., PUL, U., HEIDRICH, N., JORE, M. M., LUNDGREN, M., STRATMANN, T., WURM, R., RAINE, A., MESCHER, M., VAN HEEREVELD, L., MASTOP, M., WAGNER, E. G. H., SCHNETZ, K., VAN DER OOST, J., WAGNER, R. & BROUNS, S. J. J. 2010. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Molecular Microbiology*, 77, 1380-1393.
- WILKINSON, S. P. & GROVE, A. 2006. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Current Issues in Molecular Biology*, 8, 51-62.
- WOSTEN, M. 1998. Eubacterial sigma-factors. *Fems Microbiology Reviews*, 22, 127-150.
- WRIGHT, K. J., SEED, P. C. & HULTGREN, S. J. 2007. Development of intracellular bacterial communities of uropathogenic *Escherichia coli* depends on type 1 pili. *Cellular Microbiology*, 9, 2230-2241.
- WU, M., TAO, Y., LIU, X. & ZANG, J. 2013. Structural Basis for Phosphorylated Autoinducer-2 Modulation of the Oligomerization State of the Global Transcription Regulator LsrR from *Escherichia coli*. *Journal of Biological Chemistry*, 288, 15878-15887.
- WYBORN, N. R., STAPLETON, M. R., NORTE, V. A., ROBERTS, R. E., GRAFTON, J. & GREEN, J. 2004. Regulation of *Escherichia coli* Hemolysin E expression by H-NS and *Salmonella* SlyA. *Journal of Bacteriology*, 186, 1620-1628.
- XAVIER, K. B. & BASSLER, B. L. 2005. Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in *Escherichia coli*. *Journal of Bacteriology*, 187, 238-248.
- XAVIER, K. B., MILLER, S. T., LU, W., KIM, J. H., RABINOWITZ, J., PELCZER, I., SEMMELHACK, M. F. & BASSLER, B. L. 2007. Phosphorylation and processing of the quorum-sensing molecule autoinducer-2 in enteric bacteria. *Acs Chemical Biology*, 2, 128-136.
- XU, J. M. & JOHNSON, R. C. 1995. Fis activates the RpoS-dependent stationary-phase expression of proP in *Escherichia coli*. *Journal of Bacteriology*, 177, 5222-5231.
- XUE, P., CORBETT, D., GOLDRICK, M., NAYLOR, C. & ROBERTS, I. S. 2009. Regulation of Expression of the Region 3 Promoter of the *Escherichia coli* K5 Capsule Gene Cluster Involves H-NS, SlyA, and a Large 5' Untranslated Region. *Journal of Bacteriology*, 191, 1838-1846.

- YAMAGISHI, M., MATSUSHIMA, H., WADA, A., SAKAGAMI, M., FUJITA, N. & ISHIHAMA, A. 1993. Regulation of the *Escherichia coli* *rmf* gene encoding the ribosome modulation factor - growth phase-dependent and growth rate-dependent control. *Embo Journal*, 12, 625-630.
- YAMANAKA, K., ZHENG, W. D., CROOKE, E., WANG, Y. H. & INOUE, M. 2001. CspD, a novel DNA replication inhibitor induced during the stationary phase in *Escherichia coli*. *Molecular Microbiology*, 39, 1572-1584.
- YANG, Y. H. & SPEED, T. 2002. Design issues for cDNA microarray experiments. *Nature Reviews Genetics*, 3, 579-588.
- ZENG, J. & SPIRO, S. 2013. Finely Tuned Regulation of the Aromatic Amine Degradation Pathway in *Escherichia coli*. *Journal of Bacteriology*, 195, 5141-5150.
- ZHANG, G. Y., CAMPBELL, E. A., MINAKHIN, L., RICHTER, C., SEVERINOV, K. & DARST, S. A. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 angstrom resolution. *Cell*, 98, 811-824.
- ZHANG, G. Y. & DARST, S. A. 1998. Structure of the *Escherichia coli* RNA polymerase alpha subunit amino-terminal domain. *Science*, 281, 262-266.
- ZHANG, Y., MORAR, M. & EALICK, S. E. 2008. Structural biology of the purine biosynthetic pathway. *Cellular and Molecular Life Sciences*, 65, 3699-3724.
- ZHAO, G., WEATHERSPOON, N., KONG, W., CURTISS, R. & SHIA, Y. X. 2008. A dual-signal regulatory circuit activates transcription of a set of divergent operons in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 20924-20929.
- ZHENG, D. L., CONSTANTINIDOU, C., HOBMAN, J. L. & MINCHIN, S. D. 2004. Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research*, 32, 5874-5893.
- ZHOU, Y. H., BUSBY, S. & EBRIGHT, R. H. 1993. Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell*, 73, 375-379.