The University Of Sheffield.

Studying the adaptation of *Escherichia coli* K-12 in response to trimethylamine-*N*-oxide

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Department of Molecular Biology and Biotechnology,

The University of Sheffield

Katie Jane Denby

MBiolSci (Hons) University of Sheffield

September 2016

Abstract

Escherichia coli is a Gram-negative, metabolically versatile facultative anaerobe, which utilises either fermentation, anaerobic respiration or aerobic respiration for energy generation and growth. Trimethylamine-*N*-oxide (TMAO) is used by *E. coli* as an alternative terminal electron acceptor, being reduced to trimethylamine (TMA). Although the regulation and operation of the *E. coli* TMAO respiratory chain (TorCAD) are well characterised, there is no understanding of the dynamic adaptive processes that occur in *E. coli* during transition from fermentative to TMAO-respiratory growth. Here, glucose-limited chemostat cultures were used to study these adaptive processes. Analyses of the transcriptional and metabolic changes occurring when *E. coli* K-12 responds to TMAO revealed a number of unexpected components of the adaptive process. Firstly, it was found that growth on a sub-optimal concentration of TMAO resulted in mixed metabolism, with two distinct sub-populations of cells that either activate or do not activate the transcription of the *torCAD* operon in response to TMAO. DNA methylation was found to contribute to this regulation in response to low concentrations of TMAO. Secondly, it was found that *E. coli* possesses TMAO demethylase activity, as both the products of this activity, dimethylamine and formaldehyde and the induction of the *frmRAB* operon was detected when cells were grown in the presence of TMAO. Work was carried out to characterise the regulator protein, FrmR, of the *frmRAB* operon. Here, it is shown that the FrmR transcriptional repressor specifically reacts with formaldehyde, resulting in the formation of two inter-molecular methylene bridges between the *N*-terminal Pro2 and Cys35 residues of adjacent FrmR subunits, leading to alleviation of repression. It was also shown that growth yield was higher in the presence of TMAO under aerobic conditions eventhough TMAO reduction to TMA was not detected. Futhermore, the nitrate response regulator NarL does not directly regulate the *torCAD* operon.

Acknowledgements

Firstly, I would like to thank Professor Jeff Green for all his advice, support and patience throughout the last four years. I would also like to thank Professor Robert Poole, whose advice was greatly valued.

Thank you to everyone in F10 (both past and present), for their help and making the workplace a lovely place to be. Special thanks must of course go to Dr Matt Rolfe, for his endless patience whilst teaching me all he knows; a lot of the work done here would not have been possible without his help. Thanks also go to Sue Clark for her assistance with flow cytometry and Dr Christa Walther for her advice regarding fluorescence microscopy. I also have to thank our collaborators at Durham University, for their insights into working with FrmR.

Thanks also to all of my friends in Sheffield, within MBB, who over the past eight years have made it a truly excellent place to both work and live.

Many thanks go to my family. Mum and Dad, without the encouragement and support you have given me over the years, none of this would have been possible. Laura and Sam, thanks for being genuinely awesome siblings who always provide much laughter and fun. Thanks to you all for showing interest in my work and listening to my incessant science-related ramblings. You are truly the best family anyone could have, for that I am privileged.

Lastly, and most importantly, is my husband Rich. You have seen more than anyone else, witnessed the tears; restored the lost documents; provided endless support and at times much needed perspective. Thank you for trying to learn and understand biological terms and principles, despite not having a background in science (being a muggle, as you say), in order to perhaps become able to understand my project. You have picked me up and encouraged me to carry on through the tough times and moments of self-doubt. It goes without saying that completing this thesis would not have been possible without you.

Publications and Presentations

Publications

Denby, K. J., Rolfe, M. D., Crick, E., Sanguinetti, G., Poole, R. K., and Green, J. **(2015)** Adaptation of anaerobic cultures of *Escherichia coli* K-12 in response to environmental trimethylamine-*N*-oxide. *Environmental Microbiology,* **17**: 2477-2491.

Denby, K. J., Iwig, J., Bisson, C., Westwood, J., Rolfe, M. D., Sedelnikova, S. E., Higgins, K., Maroney, M. J., Baker, P. J., Chivers, P. T., and Green, J. **(2016)**. The mechanism of a formaldehyde-sensing transcriptional regulator. *Scientific Reports*. *Submitted.*

Presentations

Society of General Microbiology (SGM) Autumn Conference - University of Sussex (2-4th September 2013)

Poster presented: Transcriptomic and biochemical changes in *Escherichia coli* during a shift from fermentation to Trimethylamine-*N*-oxide (TMAO) respiration

Bacterial Electron Transfer Processes and Their Regulation Conference - Vimeiro, Portugal (15-18th March 2015)

Poster presented: Adaptation of anaerobic cultures of *Escherichia coli* K-12 in response to environmental trimethylamine-*N*-oxide

Table of contents

List of Figures

List of Tables

Abbreviations

Ap	Ampicillin
APS	Ammonium per sulfate
ATP	Adenosine triphosphate
BLItz	Biolayer interferometry
bp	Base pair
cDNA	Complementary DNA
cdw	Cell dry weight
CFE	Cell-free extract
Cm	Chloramphenicol
CRP	cAMP receptor protein
C_T	Cycle threshold
CTP	Cytidine triphosphate
CVD	Cardiovascular disease
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DIC	Differential interference contrast
DMA	Dimethylamine
DMSO	Dimethyl sulfoxide
DNA	Deoxynucleic acid
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
ESI	Electrospray ionisation
FITC	Fluorescein isothiocyanate
FMO	Flavin monoxygenase
FRT	FLP (flippase) recognition target
GTP	Guanosine triphosphate
ICP-MS	Inductively coupled plasma-mass spectrometry

IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kn	Kanamycin
LB medium	Luria-Bertani medium
LC-MS	Liquid chromatography–mass spectrometry
MBP	Maltose binding protein
mRNA	Messenger ribonucleic acid
MWCO	Molecular weight cut off
NAD+/NADH	Nicotinamide adenine dinucleotide oxidised/reduced
NMR	Nuclear magnetic resonance
NO_2^-	Nitrite
NO_3^-	Nitrate
OD	Optical density
ORF	Open reading frame
poly [d(I-dC)]	Poly-deoxy-inosinic-deoxy-cytidylic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PTS	Phosphotransferase system
qRT-PCR	Quantitative reverse transcription-PCR
RNA	Ribonucleic acid
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
*Se*FrmR	*Salmonella enterica* serovar Typhimurium FrmR
TCA	Tricarboxylic acid cycle
TEMED	Tetramethylethylenediamine
TF	Transcription factor
T_m	Melting temperature
TMA	Trimethylamine
TMAO	Trimethylamine *N*-oxide
TSP	Trimethylsilylpropanoic acid
TTP	Thymidine triphosphate
UPEC	Uropathogenic *E. coli*
UTP	Uridine triphosphate

1. Introduction

1.1. The metabolic modes of *Escherichia coli*

Escherichia coli is a Gram-negative, metabolically versatile facultative anaerobe, which is able to utilise different metabolic pathways to generate energy and promote growth. These pathways differ fundamentally in terms of the mechanism of ATP synthesis and redox considerations. The metabolic versatility of *E. coli* means that the bacterium is well adapted to ensure survival and reproduction in the gastrointestinal tract of a suitable host (Ingledew and Poole, 1984). Certain pathogenic strains, such as uropathogenic *E. coli* (UPECs), are able to survive in other tissues and locations (Slonczewski and Foster, 2009). *Escherichia coli* can thrive on low concentrations of a relatively limited number of low molecular mass compounds, sometimes only available transiently. Through adaptation, there are three metabolic modes available to *E. coli*, aerobic respiration, anaerobic respiration and fermentation (summarised in Figure 1.1).

The redox energy provided by substrate oxidation is conserved by oxido-reduction systems that require one or more terminal electron acceptors of higher redox potential than the donor. When molecular oxygen is available, pre-eminent amongst these electron acceptor reactions is the four-electron reduction of oxygen to water. Maximal energy is generated under aerobic conditions when, via the citric acid cycle, pyruvate is oxidised to carbon dioxide. The reducing equivalents generated enter an electron transport chain, terminating in the reduction of oxygen to water. This mode of metabolism gives the highest yield of energy possible (up to 38 ATPs per mole of glucose) (Slonczewski and Foster, 2009) (Figure 1.1a) .

Oxygen can readily diffuse into the cytoplasm, under normal aerobic conditions (~20 mbar). However, any lower (under 2-5 mbar) and oxygen availability becomes limiting and pathway, which utilise alternative electron acceptors to oxygen, are employed. There are a number of alternative electron acceptors used by *E. coli* including fumarate, nitrate and trimethylamine-*N*-oxide (TMAO)(Patschkowski *et al*., 2000; Unden *et al*., 1995; Unden and Bongaerts, 1997). Under anaerobic respiratory conditions, the energy yield is much less (4 ATP + 8[H] per mole of glucose), due the citric acid cycle becoming non-cyclic and thus leading to the incomplete oxidation of the substrate (Guest, 1996) (Figure 1.1b). In order to utilise different alternative electron acceptors, adaptation must occur, which involves the synthesis of a number of co-enzymes and co-factors that are essential for anaerobic

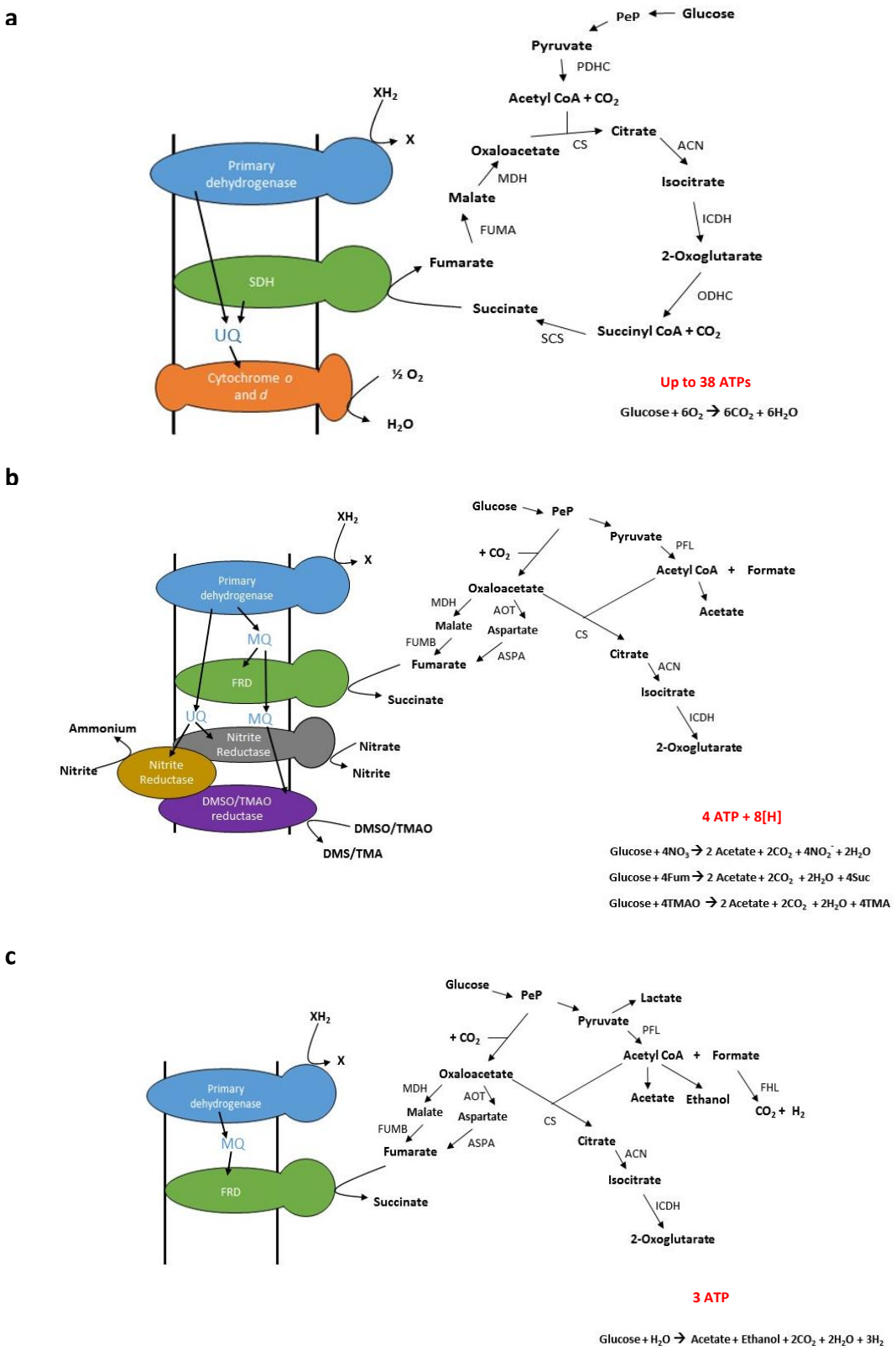


Figure 1.1a, b and c. <u>The metabolic modes of *Escherichia coli*</u>

The main pathways of energy and carbon metabolism are illustrated for: (a) aerobic respiration, (b) anaerobic respiration and (c) fermentation.

Enzymes shown are aspartase (ASP); aspartate-oxaloacetate aminotransferase (AOT); formate hydrogen-lyase (FHL); fumarate reductase (FRD); isocitrate dehydrogenase (ICDH); malate dehydrogenase (MDH); menaquinone (MK); succinyl-CoA synthase (SCS); succinate dehydrogenase (SDH); 2-oxoglutarate dehydrogenase complex (ODHC); pyruvate dehydrogenase complex (PDHC); pyruvate formate-lyase (PFL) and ubiquinone (UQ). Adapted from Guest (1996).

respiration and repression of any genes involved in oxygen metabolism (Kiley and Beinert, 1998; Lynch and Lin, 1996).

Fermentation is the metabolic mode which is adopted by *E. coli w*hen neither oxygen nor any exogenous alternative electron acceptors are available. This metabolic mode utilises carbon and energy sources to produce more reduced products (Figure 1.1c). Substrate level phosphorylation is used to conserve energy, generating much less energy than other metabolic modes (3 ATP per mole of glucose) (Guest, 1996). *Escherichia coli* produces a number of different end products during fermentation and thus is classed as a mixed acid fermenter. Redox is balanced during fermentation by producing varying amounts of end products; these include lactate, acetate, ethanol, CO_2, H_2 and succinate. The relative amount of each is dependent upon both on the environmental conditions and the carbohydrate being fermented (Guest, 1996; Iuchi and Lin, 1993).

1.2. The hierarchy of metabolism in *E. coli* and its regulation

Escherichia coli operates a metabolic hierarchy in which aerobic respiration is preferred to anaerobic respiration, which in turn is preferred to fermentation (Gunsalus, 1992; Lin and Iuchi, 1991; Spiro and Guest, 1991; Unden *et al*., 1994). This hierarchy is maintained by the coordinated action of transcription factors (TFs) and reflects the relative amounts of energy that can be conserved by each metabolic mode (Unden *et al*., 1994). The development of these pathways and sensitive regulatory mechanisms have conferred on *E. coli* the ability to maximise ATP production under many redox conditions. The ability of *E. coli* to adopt different metabolic modes is based on a number of redox-responsive regulatory mechanisms that coordinate this successful adaptation (Unden and Bongaerts, 1997). Although *E. coli* is viable across changing redox environments there is a cost to this adaptability as the bacterium must provide genetic and enzymatic "equipment" for each metabolic system. The metabolic pathways used by *E. coli* are dependent on oxygen availability, the nature of the growth substrate and also the availability of any alternative electron acceptors. *Escherichia coli* can respire aerobically with oxygen and anaerobically by using a variety of terminal electron acceptors for electron transport-coupled oxidative phosphorylation (Table 1.1) (Unden and Bongaerts, 1997). These compounds include, in order of decreasing potential energy: nitrate, dimethyl sulfoxide (DMSO), trimethylamine-*N*-oxide (TMAO) and fumarate (Unden and Bongaerts, 1997). It has been suggested that *E. coli* regulates the utilisation of these compounds on the basis of the amount of energy potentially available in order to generate maximum cellular energy. For example, in the presence of the preferred electron

Redox Couple	Reduction potential (E_0') (V)	Genes
O_2 / H_2O	+ 0.82	*cyoABCDE, cydAB, appBC*
NO_3^- / NO_2^-	+ 0.42	*narGHJI, narZYWV*
NO_2^- / NH_4^+	+ 0.36	*nrfABCDEFG*
DMSO / DMS	+ 0.16	*dmsABC*
TMAO / TMA	+ 0.13	*torCAD, torZY, dmsABC*
Fumarate / Succinate	+ 0.03	*frdABCD*

Table 1.1. Electron acceptors utilised by *E. coli*

The different electron acceptors utilised by *Escherichia coli* and what these acceptors are reduced to during catabolism are listed. The redox couples are listed in order of reduction potential, with the most energetically favourable redox couples towards to the top of the table. The TMAO/TMA redox couple is the second to least energetically favourable redox couples utilised by *E. coli*. Adapted from Unden and Bongaerts (1997).

acceptor oxygen, production of nitrate reductase and fumarate reductase enzymes is repressed. Furthermore, TMAO respiration has been shown to be repressed approximately two-fold in the presence of nitrate, the preferred anaerobic electron acceptor (Takagi *et al*., 1981), with this repression not occurring in a *narL* mutant (Iuchi and Lin, 1987).

Thus, this hierarchy reflects the relative amounts of energy that can be conserved by these metabolic modes, with oxygen availability as the major signal that governs which metabolic mode is adopted. Ultimately, in *E. coli,* this metabolic hierarchy is rooted in and maintained by altered patterns of gene expression controlled by the main regulators involved in the aerobic/anaerobic response (Unden and Bongaerts, 1997).

1.2.1. FNR

In *E. coli*, the transcriptional regulator FNR (**f**umarate **n**itrate **r**eduction regulator) controls the expression of ≥200 genes, activating those involved in anaerobic respiration and the reduction of alternative electron acceptors (Constantinidou *et al*., 2006; Covert *et al*., 2004; Gonzalez *et al*., 2003; Myers *et al*., 2013; Salmon *et al*., 2003; Woltjer *et al*., 2007). FNR belongs to the CRP/FNR superfamily of transcription factors whose members are widely distributed in bacteria (reviewed in (Green *et al*., 2001). Other regulators in addition to FNR are often involved in the fine tuning of gene expression in response to availability of oxygen. FNR is the primary transcriptional regulator, which mediates the transition from aerobic to anaerobic growth by repressing genes involved in aerobic metabolism and activating genes involved in anaerobic metabolism and (Guest, 1996). More recent CHIP-seq analysis suggested that FNR occupancy at many target sites within the *E. coli* genome is strongly influenced by nucleoid-associated proteins (NAPs) that restrict access to many FNR binding sites (Myers *et al*., 2013). It was also found that genome-wide FNR occupancy did not correlate with the match to consensus a binding sites, which may be explained by genome accessibility. However, FNR was found bound some promoters without regulating expression, highlighting a more complex system for FNR regulation than previously known (Myers *et al*., 2013).

It is known that under anaerobiosis, FNR acquires a $[4Fe\text{-}4S]^{2+}$ cluster (Khoroshilova *et al*., 1997; Lazazzera *et al*., 1996) that causes a conformational change and dimerization of the protein, this conformation has increased affinity for target DNA, which in turn promotes site-specific DNA-binding, most likely through a reorganisation of the dimer interface (Crack *et al*., 2004; Green *et al*., 1996; Jordan *et al*., 1997; Khoroshilova *et al*., 1997; Moore and Kiley, 2001; Scott and Green, 2002). Recently, the crystal structure of FNR from *Aliivibrio fischeri*

was solved, this is the same length as *E. coli* FNR and shares 84% amino acid sequence identity (Volbeda *et al*., 2015). The crystal structure suggests that there is a tuneable monomer-dimer equilibrium that is mainly determined by two interactions: the imperfect interface of the two α-C helices and the top salt bridges. In light of the crystal structure, the authors proposed that FNR monomerisation, involves a dimer "unzipping" process, which starts with the dissociation of a salt bridge near the *N*-terminal domain (Volbeda *et al*., 2015). This process propagates toward the *C*-terminal end. Activated FNR (under anaerobic conditions) is able to bind a specific a 14 bp imperfect palindromic sequence of DNA (**TTGAT**NNNN**ATCAA**; bold residues highlight palindromic regions; N, any base) (Eiglmeier et al., 1989; Gerasimova et al., 2001; Scott and Green, 2002). Dimeric FNR binds symmetrically to this region, with each subunit recognising one half of the binding site, allowing it to either repress or activate transcription through interactions and/or recruitment of RNA polymerase (Spiro, 1994).

Upon exposure to oxygen, the FNR $[4Fe\text{-}4S]^{2+}$ cluster is rapidly disassembled, possibly via a $[3Fe\text{-}4S]^{+}$ state to a $[2Fe\text{-}2S]^{2+}$ cluster, causing FNR monomerisation (Crack *et al*., 2007; Popescu *et al*., 1998). It has been suggested that oxygen-mediated [4Fe-4S] to [2Fe-2S] cluster degradation indirectly initiates monomerisation through a conformational change, which leads to the disruption of the salt-bridge in the *N*-terminal domain (Volbeda *et al*., 2015). Monomeric FNR cannot bind DNA and thus cannot interact with RNA polyermase (Scott and Green, 2002). This conversion in response to oxygen is fast *in vitro* and *in vivo*, occurring within minutes of exposure allowing for a fast regulatory response. Initial cluster disassembly *in vitro* is biphasic, thus ensuring that FNR reacts optimally at low oxygen concentrations (Jordan *et al*., 1997). After prolonged oxygen exposure, the [2Fe-2S] cluster is destroyed, and apo-FNR, which lacks an Fe-S cluster, is the primary form of FNR under aerobiosis (Popescu *et al*., 1998). Under aerobiosis, the apo-FNR monomer is exposed and can be degraded by the ClpXP protease (Mettert and Kiley, 2005).

Thus, FNR acts as the master switch, as its activation under anaerobic conditions leads to activation of genes involved in anaerobic metabolism and repression of genes involved in aerobic metabolism.

1.2.2. The ArcBA system

The ArcBA system (**a**erobic **r**espiration **c**ontrol) is responsible for the regulation of numerous genes and operons across aerobic and microaerobic environments (5-10 mbar oxygen tension). It is a two-component system consisting of a membrane bound sensor (ArcB) and

cytoplasmic regulator (ArcA)(Bauer *et al*., 1999; Iuchi *et al*., 1989; Iuchi and Lin, 1988; Sawers, 1999). The ArcB sensor is part of a complex class of sensor kinases that contain a primary transmitter domain, central receiver domain and secondary terminal histidine phosphotransfer domain that possesses autophosphorylation activity (Bauer *et al*., 1999; Sawers, 1999). The actual signal detected by ArcB has been shown not to be oxygen but rather a number of other factors related to the metabolic and redox state of the cell (Georgellis *et al*., 2001). It has been established that oxidised quinones, ubiquinone and menadione, were able to inhibit the autophosphorylation of ArcB unlike the reduced forms which were unable to inhibit the process (Georgellis *et al*., 2001). Under anaerobic or microaerobic conditions, quinones accumulate in their reduced form thus stimulating ArcB autophosphorylation and the subsequent activation of ArcA (ArcA~P). It has also been shown that the autophosphorylation of ArcB is dependent on upon the formation of a number of disulfide bonds between ArcB monomers so it is possible that quinones may play a further role in the direct oxidation of the cysteines involved in this bond formation (Malpica *et al*., 2004). The activity of ArcBA is also sensitive to cellular NADH/NAD^+ ratios. *In vitro* studies have also shown ArcB responds to metabolites such as, lactate, acetate and pyruvate that accumulate under anaerobic or microaerobic conditions and are indicative of lower oxygen tensions. High concentrations of these anaerobic or microaerobic intermediates are able to inhibit the intrinsic phosphatase activity of ArcB resulting in a prolonged signal duration (Georgellis *et al*., 1999; Rodriguez *et al*., 2004). Thus, ArcB responds to electron flux in the aerobic respiratory chain rather than sensing oxygen directly (Iuchi and Lin, 1995). The ArcA protein is a typical response regulator with an amino terminal receiver domain and carboxy terminal electron acceptor domain which binds a proposed consensus [A/T]GTTAATTA[A/T] determined from alignment of a number of ArcA target promoters and DNase I protection assays (Drapal and Sawers, 1995). Generally, ArcA~P acts as a repressor preventing the synthesis of aerobically required enzymes during anaerobic conditions, although it is responsible for the activation of a number of genes notably the *cydAB* operon, which encodes the high affinity terminal cytochrome *d* oxidase. Regulation by the ArcBA system in coordination with another transcription factor is not uncommon, most notably FNR to coordinate the aerobic/anaerobic interface response with numerous examples being recorded (Table 1.2)(Govantes *et al*., 2000). More recently, genome-wide analysis using CHIP-seq has revealed that ArcA reprograms metabolism under anaerobic conditions such that carbon oxidation pathways, which recycle redox carriers via respiration are transcriptionally repressed by ArcA (Park *et al*., 2013). Also, this ChIP-seq analysis revealed

Enzyme	Gene	ArcA Effect	FNR Effect
1. <u>Citric acid cycle enzymes</u>			
Aconitase A	*acnA*	-	-
Fumarase A	*fumA*	-	-
Fumerase B	*fumB*	+	+
Succinate dehydrogenase	*sdhCDAB*	-	-
α-Ketoglutarate dehydrogenase	*sucAB*	-	-
Succinyl coenzyme A synthetase	*sucCD*	-	-
2. <u>Respiratory chain dehydrogenases</u>			
Formate dehydrogenase	*fdhGHI*	-	+
NADH dehydrogenase I	*nuo*	-	-
3. <u>Terminal reductases</u>			
Cytochrome *o* oxidase	*cyoABCDE*	-	-
Cytochrome *d* oxidase	*cydAB*	+	-
4. <u>Other enzymes</u>			
Formate transport and pyruvate-formate lyase	*focA-pfl*	+	+
Pyruvate dehydrogenase complex and regulator	*pdhR-aceEF-lpd*	-	-

Table 1.2. <u>Regulation by the ArcBA system in coordination with FNR</u>

This table shows the operons that are regulated by ArcBA and FNR and the effect that each regulator has on the transcription. (-) represents inhibition and (+) represents activation of transcription by the transcription factor.

that most ArcA binding sites contain additional direct repeat elements beyond the two required for binding an ArcA dimer (Park *et al*., 2013).

Thus, the ArcBA regulon is vital to the metabolic versatility of *E. coli,* consisting of more than thirty loci with products coding for cytochrome oxidases, enzymes of the citric acid cycle, glyoxylate shunt and fatty acid degradation pathways (Bauer *et al*., 1999; Liu and De Wulf, 2004).

1.2.3. The Nar system

The Nar system regulates the expression of numerous genes involved in both anaerobic respiration and fermentation pathways. It comprises the homologous dual response regulator proteins NarL and NarP and their corresponding sensors NarX and NarQ. These sensors respond to the presence of nitrate and nitrite within the environment (Darwin *et al*., 1996; Rabin and Stewart, 1993; Stewart, 1993) and activate NarL and NarP via a phosphotransfer relay allowing them to bind DNA and control transcription (Lee *et al*., 1999). Once active, these proteins are able to act directly as conventional regulators, both recognising and binding to a consensus sequence known as a NarL heptamer (TACYYMT; Y, C or T residue; M, A or C residue) (Darwin *et al*., 1997) although, like ArcA, the Nar system often works in concert with the anaerobic response regulator FNR. The FNR protein often binds near position -41, activating transcription whilst site(s) for NarL and/or NarP are situated further upstream (Bell *et al*., 1990; Darwin *et al*., 1993). These upstream sites contribute to activation, for example the *nrf* gene encoding a formate-dependent nitrite reductase (Hussain *et al*., 1994) where the Nar binding site is located at -74.5 (Tyson *et al*., 1994), and the *nir* gene which encodes a NADH-dependent nitrite reductase with an upstream Nar binding site situated at -69.5 (Tyson *et al*., 1993). Through a series of deletions in the *nir* promoter, it was determined that removal of the Nar binding site (and preceding regions) allowed for full activation by FNR alone, rather than requiring the contributory effects of Nar binding for full expression (Wu *et al*., 1998). This raised the possibility that co-regulation may not be a simple case of independent contacts with the RNAP between the FNR and Nar proteins but that Nar may contribute by interfering with other upstream factors that in the absence of Nar inhibit FNR-dependent activation. This was confirmed by Browning and co-workers when Fis and IHF were identified as these factors whose negatory effects are nullified by Nar binding (Browning *et al*., 2000).

Thus, the Nar system is integrated with other regulatory systems playing a crucial role in maintaining the metabolic hierarchy of *E. coli* ensuring that under anaerobic conditions

preferred alternative electron acceptors such as nitrate and nitrite are used to maximise energy production.

1.3. Reprogramming of the metabolic mode of *E. coli* under stable (steady-state) and dynamic conditions (transitions)

Escherichia coli has long been a model organism for studying aerobic respiration, anaerobic respiration and fermentation (Ingledew and Poole, 1984). Many bacteria undergo transitions between environments with differing oxygen availabilities as part of their natural lifestyles. However, the dynamics of adaptation when bacteria experience changes in oxygen and alternative electron acceptor availability are much less well characterised (Bettenbrock *et al*., 2014).

1.3.1. Gene expression and transcription factor (TF) activities under steady-state conditions in response to oxygen availability

Widespread changes in gene expression have been associated with the reprogramming of *E. coli* metabolism in response to oxygen availability (summarised in Bettenbrock *et al*., 2014). Comparison of gene expression data from both aerobic and anaerobic cultures of *E. coli* grown on various different media have been previously reported (Constantinidou *et al*., 2006; Kang *et al*., 2005; Salmon *et al*., 2005). It has been suggested, based on these studies, that many as 1450 genes have altered expression under anaerobic conditions when compared to those expressed under aerobic conditions. However, only 63 genes were significantly regulated when steady-state cultures grown with excess oxygen were compared with anaerobic steady-state cultures (Rolfe *et al*., 2011). These data are representative of the minimal set of differentially expressed genes that distinguish aerobic and anaerobic growth, without the fluctuations in batch culture affecting the gene expression (Hoskisson and Hobbs, 2005).

Under aerobic conditions, the most important adaptive responses appeared to be enhanced expression of TCA cycle and glyoxylate shunt genes (Rolfe *et al*., 2011). The cyclical form of the TCA cycle was restored as oxygen availability increased by the expression of the operon encoding succinate dehydrogenase (SDH)-succinyl CoA synthetase. It was shown that initially the bacteria might not fully oxidize the substrate glucose by the induction of the glyoxylate shunt under aerobic conditions, but continue to excrete a proportion of the carbon supplied as acetate (Rolfe *et al*., 2011). Acetate was undetectable in the culture medium because it was presumably 'immediately' taken up via the acetate transporter ActP (induced >10-fold under aerobic conditions) for metabolism via the glyoxylate cycle (Rolfe *et al*., 2011). The switch to aerobic respiratory growth was completed by the down-regulation

of genes associated with anaerobic fermentation and the upregulation of the relatively low-oxygen affinity terminal oxidase cytochrome *bo'* (Rolfe *et al*., 2011). The transcriptional changes observed were consistent with the inactivation of FNR and relief of ArcA-mediated repression.

A total of 89 genes exhibited expression changes when the complete set of transcript profiles obtained in the microaerobic–aerobic range was compared with that of the anaerobic steady-state (Rolfe *et al*., 2011). Transcripts with altered abundance were associated with central metabolic functions (39%), unknown function (11%), transcription regulation (7%), outer membrane proteins and adhesions (5%), and the oxidative stress response (4%) (Rolfe *et al*., 2011). The transcriptomic changes observed at each steady state suggested a stepwise switch to aerobic respiratory metabolism. For example, the *cyoABCDE* operon (encoding cytochrome *bo'*) was maximally expressed under fully aerobic conditions, whereas the *cydAB* operon (encoding cytochrome *bd-I*) was maximally expressed at 56% aerobiosis, and *appCB* (encoding cytochrome *bd-II*) was maximally expressed at 0% aerobiosis (fully anaerobic conditions). Cytochrome *bo*' has a relatively low affinity for oxygen (K_m 16–350 nM) (D'Mello *et al*., 1995) whilst cytochrome *bd-I* has an extremely high-oxygen affinity (K_m 3–8 nM) (D'mello *et al*., 1996). Thus, the expression profiles match the oxygen reactivities of these oxidases. Furthermore, changes in abundance of the *cydAB* and *cyoA-E* transcripts correlated well with the changes in the amounts of cytochrome *bd-I* and cytochrome *bo'* in the cell as shown by spectroscopic measurements (Rolfe *et al*., 2011). Thus, these shifts in oxidase expression were consistent with a switch from fermentation to microaerobic respiration (with cytochrome *bd-I* acting as the terminal oxidase), and then to aerobic respiration (with the terminal oxidase cytochrome *bo'*), as oxygen availability increased.

Changes to the metabolic signals that are sensitive to oxygen availability, or in the activities of regulators that respond to oxygen itself, are responsible for the observed transcript and protein profiles. Simultaneous inference (using TFInfer; Asif *et al*., 2010) of the activities of 134 TFs in *E. coli* revealed that 23 responded in at least one of the steady states examined. All seven of the *E. coli* global regulators (ArcA, CRP, Fis, FNR, H-NS, IHF and Lrp; Martínez-Antonio *et al*., 2003) were amongst these 23 TFs. The predicted FNR activities from the TFInfer model were confirmed experimentally and showed that the activity of FNR decreased rapidly at aerobiosis values greater than 56% (Rolfe *et al*., 2012). In contrast, the activity of the indirect oxygen-sensor ArcBA was predicted to decrease linearly with increasing aerobiosis (Georgellis *et al*., 2001), but not with the redox state of the ubiquinone pool. Measurement of the amount of phosphorylated ArcA revealed a good correlation with the

predicted ArcA activities and with aerobiosis suggesting that, under those conditions, fermentation product-mediated inhibition of ArcB phosphatase activity was the dominant mechanism for regulating ArcA activity (Rolfe *et al*., 2012).

Thus, the switch between anaerobic to microaerobic to aerobic and *vice versa* is predominately mediated by the TFs ArcBA and FNR. Regulation by these TF allows maintenance of the regulatory hierarchy of respiration in *E. coli.*

1.3.2. Gene expression and transcription factor (TF) activities during transitions between oxygen-replete and oxygen-deficient environments

Transitions between oxygen-replete and oxygen-deficient environments (Partridge *et al*., 2007; Partridge *et al*., 2006; Rolfe *et al*., 2012; Trotter *et al*., 2011) are able to mimic real-life environmental conditions, allowing investigation into the adaptation to new oxygen concentrations. Transcriptomic analysis of the transitions showed that 590 genes (12.1% of the genome) in an aerobic to anaerobic switch and 396 genes (8.1% of the genome) exhibited expression in the anaerobic to aerobic switch and during the acute phase of adaptation (first 20 min) (Partridge *et al*., 2007; Rolfe *et al*., 2012). Thus, many more transcripts change during transitions than were observed in simple comparison of aerobic and anaerobic steady-state cultures (Section 1.3.1). This presumably reflects the dynamic nature of the adaptive processes and the many stresses encountered by *E. coli*, whilst oxygen availability is changing. Once metabolism has been reprogrammed, these acute stresses no longer exert their influence on the transcriptome and are not apparent from analysis of steady-state cultures.

Of the 396 and 590 genes that exhibited altered expression during the acute phases of the anaerobic to aerobic and aerobic to anaerobic transitions, respectively, 195 and 284 could be classified into 23 functional groups (Rolfe *et al*., 2012). The major changes in expression were observed for transcripts encoding proteins with roles in central metabolic pathways. Presumably, endogenous oxidative stress is experienced in the presence of oxygen as indicated by the altered expression of genes associated with the oxidative stress response. Introduction of oxygen alters the redox state of copper as the responses of systems involved in copper ion homeostasis suggest (Partridge *et al*., 2007; Rolfe *et al*., 2012). Although a large proportion of the transcriptional reprogramming that occurred during the aerobic to anaerobic transition was essentially the reverse of that of the anaerobic to aerobic transition, transcript changes in the aminosugars, ascorbate, C5-branched dibasic acids and galactose metabolism functional classes, as well as the translation category, were differentially

affected (Rolfe *et al*., 2012). Altered patterns of abundance in the transitions were exhibited by the transcripts of 204 genes of unknown function. Of these genes of unknown function, 119 responded in the aerobic to anaerobic transition, 51 in the anaerobic to aerobic transition, and 34 in both (Rolfe *et al*., 2012). This observation emphasises the new insights that can be obtained by perturbation of systems by helping to assign function to some of these genes.

More information about the complex regulatory events that occur during adaptation can be gained by inference of TF activities. Inference of TF activities (Asif *et al*., 2010) when aerobic or anaerobic steady states were perturbed revealed the presence of a core set of 16 TFs (including FNR and ArcA) that exhibited expression changes (Rolfe *et al*., 2012). This core network contains TFs that respond to oxygen availability and regulate central metabolism, oxidative stress responses, metal-ion homeostasis and amino acid biosynthesis. Predicted activity from transcriptomic data suggested that FNR remained active, in the the DNA binding form of the protein, under microaerobic conditions (Rolfe *et al*., 2012). This could be because the bacterial cytoplasm remains essentially anaerobic under these conditions, perhaps due to the consumption of oxygen by high-affinity cytochrome *bd-I* (Cyd). A *lacZ* reporter assay confirmed the predicted FNR expression under anaerobic and microaerobic conditions (Rolfe *et al*., 2012). Thus, spatial constraints have significant implications, even on the scale of the bacterial cell, for example here, allowing 'hybrid' metabolism to proceed under microaerobic conditions, where anaerobic processes can be supported in the cytoplasm and aerobic respiration proceeds close to the cell membrane. Furthermore, this could account for the need for two TFs consisting of a membrane-associated sensor interacting with the respiratory chain (ArcBA) and a sensor (FNR) responding to cytoplasmic oxygen, both of which work in concert to coordinate global gene expression in response to oxygen availability.

Other TFs exhibited expression changes only during a specific transition (11 TFs for aerobic to anaerobic and 4 for anaerobic to aerobic)(Rolfe *et al*., 2012). Furthermore, of the 44 TFs predicted to respond under one or more of the conditions investigated, 23 were specifically associated with the transitions, whereas only five were specifically linked to the different steady states. This is perhaps indicative of the additional stresses encountered during the environmental changes occurring during transitions compared to the consistent environmental conditions of the steady-state cultures.

Thus, there is a large body of information about the reprogramming of *E. coli* that occurs during stable (steady-state) and dynamic conditions (transitions) in response to oxygen availability. However, the reprogramming of *E. coli* that occurs in response to alternative electron acceptor availability has not been characterised.

1.4. Trimethylamine-*N*-oxide (TMAO)

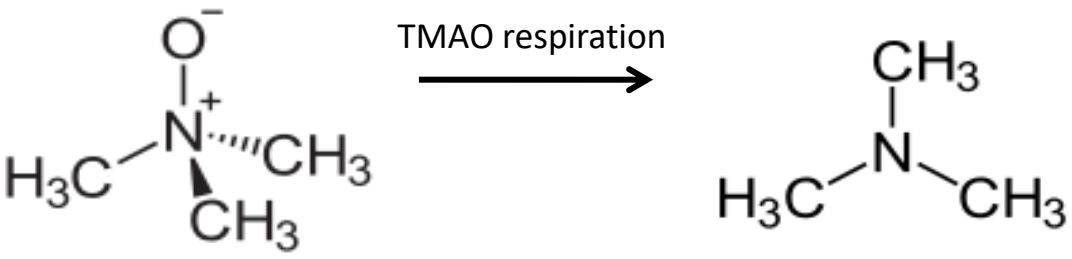
1.4.1. Bacterial TMAO respiration

Trimethylamine-*N*-oxide (TMAO) belongs to the class of amine oxides and has the formula $(CH_3)_3NO$. It is a common metabolite in many living organisms and for various bacteria serves as an alternative electron acceptor in anaerobic metabolism. The various bacteria that can reduce TMAO inhabit three main environments: animal intestines (Enterobacteriaceae), the marine environment (e.g., Alteromonas and Vibrio) and the brackish pond (non-sulfur photosynthetic bacteria) (Barrett and Kwan, 1985). In the Enterobacteriaceae, which includes *Escherichia coli*, a normal part of the human gut flora, reduction of TMAO has been shown to support oxidative phosphorylation (Barrett and Kwan, 1985). TMAO also represents a significant nutrient for marine bacteria (Lidbury *et al*., 2014).

The biochemistry and genetics of TMAO utilisation as an alternative electron acceptor in respiration by *E. coli* have been characterised over the last two decades. During TMAO respiration, TMAO is reduced to the malodorous trimethylamine (TMA) (Figure 1.2). The redox potential for the TMA/TMAO couple is +130 mV (Gon *et al.* 2001a), making TMAO a relatively poor electron acceptor when compared to other electron acceptors utilised by *E. coli* (Table 1.1). A large body of research investigating marine-fish spoilage organisms, which form TMA by the reduction of TMAO that is naturally present in fish, began last century as reviewed elsewhere (Barrett and Kwan, 1985). Since this time, characterisation of TMAO reductases and other components of the TMAO respiratory chain has been conducted.

1.4.1.1. The *E. coli* TMAO Respiratory Chain, TorCAD

In the early 1980s, the observation of TMAO reduction-linked proton translocation confirmed a role in energy conservation processes (Takagi *et al*., 1981). Research on the *E. coli* TMAO respiratory chain has shown that it utilises either a menaquinol or demethylmenaquinol electron donor (Wissenbach *et al*., 1990), consists of a membrane-associated *c*-type cytochrome (TorC) and a molybdenum-containing terminal reductase (TorA) (Mejean *et al*., 1994)(Figure 1.3). TorA is typical of the DMSO reductase family of molybdoenzymes and is located in the periplasm of *E. coli* (Czjzek *et al*., 1998). TorC, which belongs to the NirT/NapC family of *c*-type cytochromes, is a 46 kDa penta-haem *c*-type



Trimethylamine - *N* – oxide
(TMAO)

Trimethylamine
(TMA)

Figure 1.2. <u>The chemical structures of TMAO and TMA</u>

This figure shows the chemical structures of TMAO and TMA, the product of TMAO respiration.

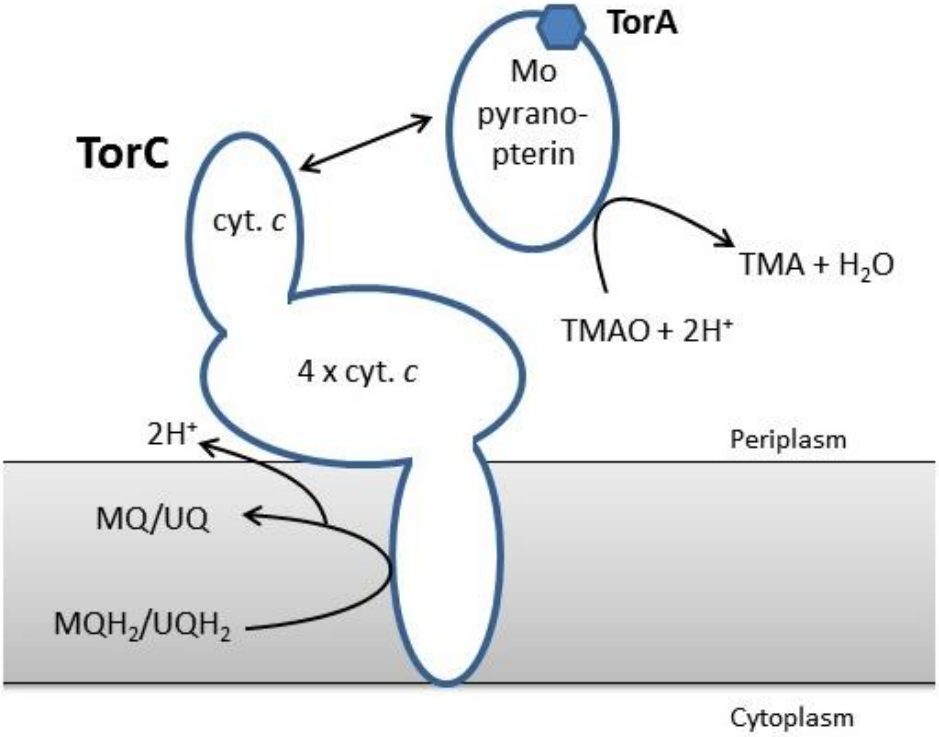


Figure 1.3. The *E. coli* TorCA respiratory chain

The TMAO respiratory chain in *E. coli* is composed of *c*-type cytochrome (TorC), which is anchored to the cytoplasmic membranes and takes electrons from the menaquinone pool and donates them to the molybdenum-containing TMAO reductase (TorA). The TMAO reductase is located in the periplasm and reduces TMAO and $2H^+$ to TMA and H_2O.

cytochrome (Mejean *et al*., 1994). Approximately 20 hydrophobic amino acid residues of the *N*-terminal region of TorC ($TorC_N$) comprise a single membrane-spanning helix, which anchors the protein to the cytoplasmic membrane (Gon *et al.,* 2001a). A large portion of the TorC protein exists in a globular form facing the periplasm (Gon *et al.,* 2001a). The $TorC_N$, which is homologous to NirT and NapC (membrane-anchored multihaem *c*-type cytochromes), possesses four haem-binding sites (C-X-X-C-H motifs) and the *C*-terminal region ($TorC_C$), which is specific for TMAO respiration, contains the fifth haem-binding sites (Mejean *et al*., 1994). *In vitro* reconstitution of electron transfer experiments with the purified *N*- ($TorC_N$) and *C*-terminal domains of TorC ($TorC_C$) and TorA, showed that only full TorC and $TorC_C$ were capable of electron transfer to TorA (Gon *et al.,* 2001a). Protein interaction studies have shown that it is $TorC_N$ domain that binds to TorA (Gon *et al.,* 2001a).

Thus, the tetra haem $TorC_N$ oxidises the menaquinone and transfers electrons to $TorC_C$, and eventually to TorA. $TorC_N$ binding to TorA could ensure a confirmation suitable for electron transfer (Gon *et al.,* 2001a).

1.4.1.1.1. Molecular properties of TMAO reductases

1.4.1.1.1.1. The Molybdenum cofactor

TorA catalyses the reduction of TMAO to TMA and the active site contains a molybendenum cofactor (Méjean *et al*., 1994; Zhang *et al*., 2008). Molybdenum-containing enzymes all contain a Molybdenum (Mo) ion coordinated by one or two modified pterin derivatives, known as molybdopterin or pyranoterin. The molybdenum cofactor in TorA is considered to be a bis(guanylyl molybdopterin)(Genest *et al*., 2008). TMAO reductases belong to the DMSO reductase family of molybdenum enzymes, found exclusively in prokaryotes (Hille *et al*., 1998; McCrindle *et al*., 2005).

1.4.1.1.1.2. Structure of TMAO reductases

The full structure of the *E. coli* TorA has not yet been reported. However, structural characterisation of the molybdenum site of recombinant *E. coli* TorA using X-ray absorption spectroscopy has been carried out (Zhang *et al*., 2008). This analysis suggested that the active site of TMAO reductase can exist in three different forms, as isolated; reduced and redox-cycled, with a considerable resemblance to DMSO reductase (Zhang *et al*., 2008). Although the crystal structure of *E. coli* TorA has not been resolved, many TMAO reductases from other organisms have. The periplasmic *Shewanella massilia* TMAO reductase, which is homologous to the *E. coli* TorA has been isolated and characterised (Figure 1.4). The overall fold of the *S. massilia* TMAO reductase was found to be essentially the same as the DMSO

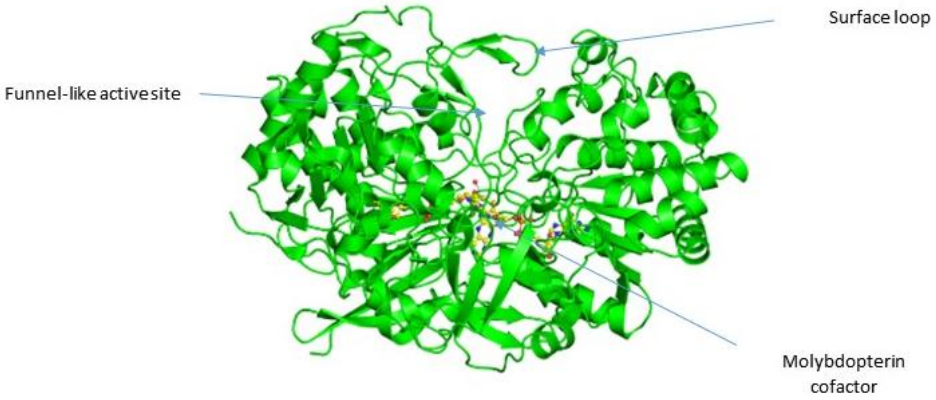


Figure 1.4. <u>The crystal structure of the TMAO reductase from *S. massilia*</u>

Shown is the crystal structure of the TMAO reductase from *S. massilia* in the oxidised state at a resolution of 2.65 Å. The view shown is from the side. Indicated on the structure are the locations of the active site, containing the molybdenum cofactor, the active site funnel and the surface loop and the entrance of the active site tunnel. This structure was obtained from the Protein Data Bank (PDB) (PDB code: 1tmo).

reductase structures, and is organised into four domains (Czjzek *et al*., 1998). However, there were structural differences found in the active site of the protein, which could correlate to the differences in the substrate specificities of these enzymes. In the DMSO reductase structure it was shown that there is a tyrosine residue closely located to the molybdenum ion, which is missing in the TMAO reductase structure, leaving a wider accessible space perhaps to accommodate the bulkier TMAO molecule (Czjzek *et al*., 1998). There are also differences in the arrangement and number of positively charged residues lining the inner surface of the funnel-like entrance to the active site, which could also account for the differences in substrate specificity. A surface loop at the entrance of the active site funnel was observed in the oxidised form of TMAO reductase, and is believed to form a lid on the active site in the substrate bound form (Czjzek *et al*., 1998).

1.4.1.1.1.3. Substrate Specificity of TMAO reductases

TMAO and DMSO reductases can be classified according to their substrate specificity, with TMAO reductases having high substrate specificity towards TMAO and DMSO reductases having broad specificity for both *S*- and *N*-oxides (Iobbi-Nivol *et al*., 1996). TorA is able to catalyse the reduction of several different *N*-oxides, including TMAO, but does not reduceany of the *S*-oxides (Iobbi-Nivol *et al*., 1996). Early experiments carried out on purified *E. coli* TorA highlighted high enzymatic activity towards TMAO, adenosine *N*-oxide, as well as lower activities towards *N*-oxides of nicotinamide, nicotinic acid and various picolines (Shimokawa and Ishimoto, 1979; Yamamoto *et al*., 1986). More detailed kinetic analyses have more recently been conducted (Iobbi-Nivol *et al*., 1996). By using kinetic analysis, this group showed that out of a wide variety of *N*- and *S*-oxides, only two *N*-oxides, TMAO and 4-methylmorpholine-*N*-oxide could be considered good substrates for TorA on the basis of their k_{cat}/K_m ratios (Iobbi-Nivol *et al*., 1996). This suggests that TMAO would be the only substrate for TorA in nature, as 4-methylmorpholine-*N*-oxide is a synthetic compound. This demonstrates the high substrate specificity for TMAO exerted by *E. coli* TorA TMAO reductase.

1.4.1.1.2. Genetic organisation and regulation of operons encoding TorCAD

There is a common pattern of gene arrangement in the TMAO reductase systems. Analysis of the operon showed that the structural genes for the reductase and the cognate membrane-bound reductase are encoded in the same operon. Located in close proximity to the structural genes is a cluster of regulatory genes (sensor kinase/response regulator), which is transcribed in the opposite direction.

1.4.1.1.2.1. The *torCAD* operon of *E. coli*

Escherichia coli has two TMAO reductase operons, *torCAD* (Figure 1.5a) and *torZY* (discussed later). In the 1980s, mutagenesis and screening experiments identified the *E. coli torCAD* operon (Takagi, 1983). Further experiments mapped the *torA* gene to 28 min on the *E. coli* chromosome (Pascal *et al*., 1984) and a *tor* regulatory gene, *torR,* was mapped to 22 min (Pascal *et al*., 1991). The *E. coli torCAD* operon is known to consist of three contiguous open reading frames (ORFs) (Mejean *et al*., 1994).

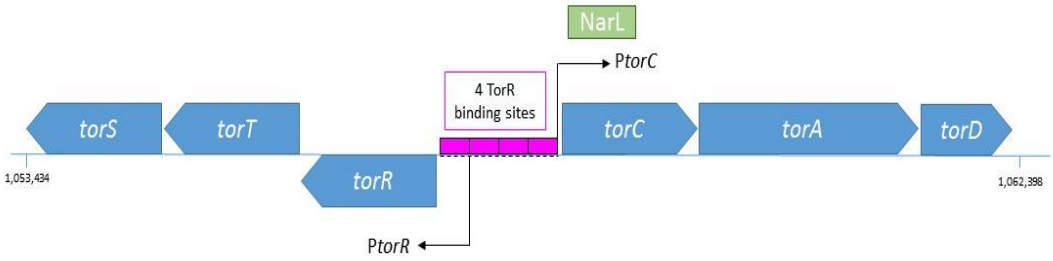
1.4.1.1.2.2. The regulatory two-component system, TorRS

Components of the TorCAD respiratory chain are strictly under the control of TMAO. Thus, TMAO reductase is only expressed when TMAO is sensed in the environment.

The *tor* locus also encodes the associated regulatory system, which involves three genes, *torS, torT* and *torR* (Jourlin *et al*., 1996a; Jourlin *et al*., 1996b; Simon *et al*., 1994). The *torR-torCAD* intergenic region contains four copies of a 10 nucleotide sequence motif (5'-CTGTTCATAT-3') called the "*tor* box" (Simon *et al*., 1994). TorR binds to these *tor* boxes, one of which is contained within the *torR* transcribed region, thus inducing the transcription of both *torC* and *torR* (Simon *et al*., 1994). Expression of the *torCAD* operon is predominately under the control of the TorS/TorR two-component system. This system is composed of the TorS complex sensor histidine kinase, which induces expression of the *tor* operon through a phosphoryl transfer pathway involving the response regulator TorR (McCrindle *et al*., 2005) (Figure 1.5b). The TorS transmitter domain contains three histidine phosphorylation sites and an aspartate phosphorylation sites in the receiver domain of TorS (Jourlin *et al*., 1997).

All three phosphorylation sites proved to be essential for the induction of *tor* operon *in vivo* (Jourlin *et al*., 1997). It was also shown that expression of the second transmitter (*C*-terminal) TorS transmitter domain as a cytoplasmic protein is sufficient to cause constitutive expression of the *tor* operon (Jourlin *et al*., 1997). This shows that it is the second transmitter domain that is involved in phosphoryl transfer to TorR. TorS is involved in both the phosphorylation and dephosphorylation of TorR (Jourlin *et al*., 1997). The *torT* gene encodes another regulatory protein essential for *torCAD* expression (Jourlin *et al*., 1996b). The product of this gene, TorT, binds TMAO promoting a conformational change in the protein (Baraquet *et al*., 2006). This conformational change in TorT is then transduced to TorS, leading to TorS activation (Baraquet *et al*., 2006), which can go on to induce the expression of the *torCAD* operon through TorR. Using crystal structures, it was shown that TorS and apoTorT form an asymmetric 2:2 complex, which binds TMAO with negative cooperativity to

a



b

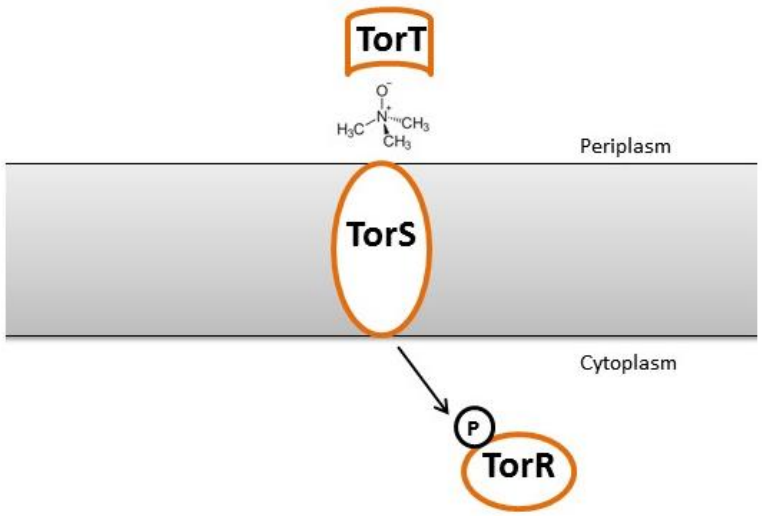


Figure 1.5a and b. Regulation of the *E. coli torCAD* operon

a. The *torCAD and torRST* genes are encoded for by two divergently transcribed operons on *E. coli* MG1655 chromosome, one coding the structural genes of the respiratory chain (*torCAD*) and the other operon codes for the associated regulatory two-component system (*torRST).* Pink boxes show binding sites for 4 TorR TFs (*tor* boxes), which activate transcription of *torCAD* and *torRST*. The *tor* boxes 1,2 and 4 contain the consensus sequence (5'-CTGTTCATAT-3'), whereas box 3 contains a degenerate consensus sequence (5'-C**C**GTTCAT**CC**-3'; bold letters indicate different nucleotides to consensus) (Simon *et al.*, 1995). NarL (shown in green) is hypothesised to inhibit the expression of the *torCAD* operon.

b. The two-component system (TorRS) is located in the cytoplasmic membrane of *E. coli* and is the main system for sensing TMAO. TorT is a periplasmic binding protein (PBP) that is involved in signal transduction and is essential for TMAO sensing.

form an active symmetric kinase (Moore and Hendrickson, 2012). The expression of the *tor* operon is negatively regulated by apocytochrome TorC via an interaction of the *C*-terminal domain of apocytochrome *c* with the periplasmic detector region of sensor TorS (Gon *et al.,* 2001b). Another level of *tor* regulation comes from a novel negative regulator, *torI*, which is of phage origin and encoded for by small ORF (66 amino acids) and is believed to decrease *tor* expression by preventing RNA polymerase recruitment to the *torC* promoter (Ansaldi *et al*., 2004). The TorS/TorR system also activates alkalinity-resistance genes and represses acid-resistance genes in *E. coli*, which is proposed to be due to the anticipation of alkalinisation due to the production of alkaline TMA during TMAO respiration (Bordi *et al*., 2003).

1.4.1.1.3. Expression and Assembly of TMAO reductases

1.4.1.1.3.1. Protein transport to the periplasm

In prokaryotes, a number of pathways are used for protein translocation, including the general secretory (Sec) system (Pugsley, 1993) and the twin-arginine translocation (Tat) pathway (Sec-independent) (Berks, 1996). The membrane-associated TorC, like other cytochromes *c*, is translocated across the plasma membrane by the Sec system, whilst the required haem groups are transported by specific export complexes to the periplasm. The Sec pathway transports the unfolded protein across the membrane by a threading mechanism, which utilises both ATP and proton motive force (Pugsley, 1993). The *E. coli* TMAO reductases are translocated by the Tat export system (Figure 1.6). The *E. coli* TorA is often used as a tool to study the Tat system (Barrett *et al*., 2003; Thomas *et al*., 2001). The presence of conserved twin (adjacent) arginine residues in the signal sequences of certain periplasmic proteins, is the reason why the Tat system is so named (Berks, 1996). Unlike the Sec transport system, proteins that are transported across the membrane by the Tat apparatus depend solely on the proton motive force and do not require ATP (Yahr and Wickner, 2001). Tat substrates are typically transported in the folded state, after cytoplasmic folding of the protein, with cofactors already inserted. Tat signal peptides are found in all of the well-characterised bacterial TMAO reductases (Gon *et al*., 2000).

1.4.1.1.3.2. The role of the TorD chaperone

TMAO reductases are assembled in the cytoplasm and translocated across the membrane to the periplasm, fully assembled with the molybdenum cofactor already inserted as previously detailed. TMAO reductases require protein-specific chaperones for cofactor insertion and export. The first of these specific chaperones to be discovered was *torD,* a chaperone for *E. coli* TorA (Genest *et al*., 2009). TorD is a cytoplasmic protein, which before the insertion of

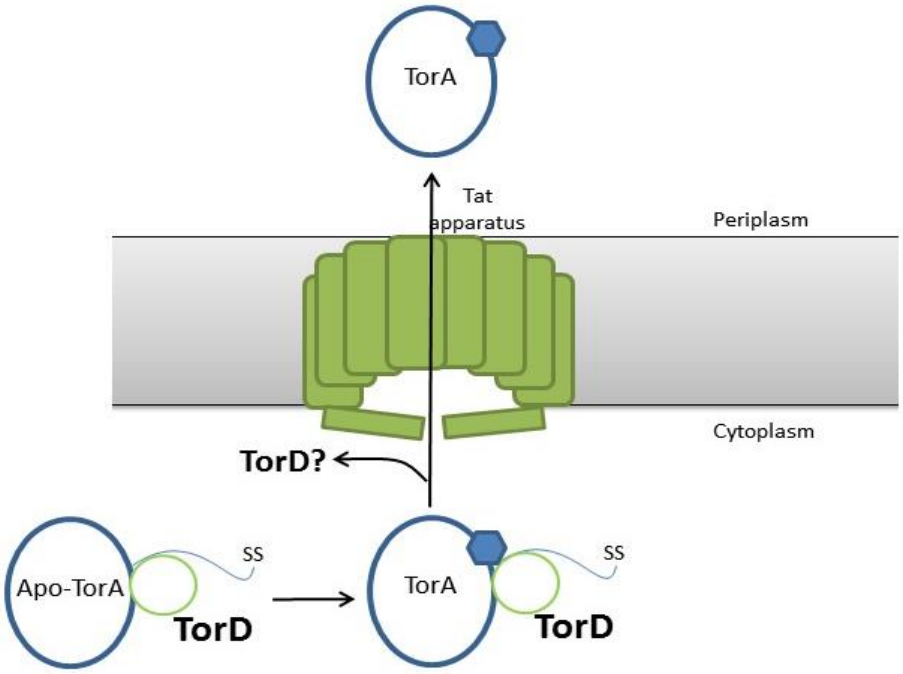


Figure 1.6. <u>Translocation of the *E. coli* TMAO reductase to the periplasm by the Tat export system</u>

Apo-TorA has the molybdenum cofactor inserted and becomes fully folded in the cytoplasm, prior to being bound by the chaperone TorD and translocated across the cytoplasmic membrane into the periplasm via the Tat apparatus. SS represents the Tat signal sequence; purple hexagon, the molybdenum cofactor; and green rectangles, the Tat apparatus.

the molybdenum cofactor interacts with the unfolded TorA (Pommier *et al*., 1998). Characterisation of TorD highlighted a specific interaction between TorA and TorD (Pommier *et al*., 1998). More recently, *in vitro* work showed that low levels of TorA were assembled when the molybdenum cofactor was added to the TorA apoprotein, however, when TorD was also added a 4-fold increase of apo-TorA activation was seen (Ilbert *et al*., 2003). Incubation of apoTorA with TorD before the addition of molybdenum, led to an increase in apoTorA activation (Ilbert *et al*., 2003). This data suggests that TorD is involved in the first step of TorA maturation, causing a conformational change, which allows TorA to receive the cofactor. More recent work suggests that as TorA must be correctly folded for the TorD chaperone to bind and thus the chaperone acts as "proofreading" step in the process (Hatzixanthis *et al*., 2005).

1.4.1.2. The alternative TMAO reductases of *E. coli*

Escherichia coli also utilises two alternative TMAO reductases in addition to the TorCAD system. A second TMAO reductase is encoded for by the *torZY* operon. TorZY has been shown to have a low level of constitutive expression and is not induced by TMAO or DMSO (Gon *et al*., 2000). Somewhat different to TorA is the second *E. coli* TMAO reductase enzyme, TorZ. Like TorA, only 4-methylmorpholine-*N*-oxide and TMAO are considered good substrates for TorZ (McCrindle *et al*., 2005). However, TorZ can also catalyse the reduction of some *S*-oxides. DMSO is also known to act as a competitive inhibitor of TorZ-catalysed of TMAO reduction, unlike TorA. DMSO is able to bind or block the active site of TorZ, without being reduced itself (Gon *et al*., 2000). The *torZ* gene was previously known as *bisZ* (Biotin sulfoxide (BSO) reductase), however, it was renamed TorZ due to its ability to reduce TMAO more efficiently than BSO and also due to its periplasmic location; even though TorZ is more homologous to BisC than TorA (63 and 42% identity, respectively). T*orZ* belongs to the *torYZ* operon. A pentahemic *c*-type cytochrome homologous to the TorC cytochrome of the TorCAD respiratory system is encoded for by the first gene, *torY* (formerly *yecK*). Surprisingly, in the chromosomal region around *torYZ,* no *torD* homologue is found.

The DMSO reductase (DmsABC) of *E. coli* has a broad substrate specificity, which includes TMAO (Bilous and Weiner, 1988; Weiner *et al*., 1988). Although the DMSO reductase reduces TMAO *in vitro,* it is not induced by TMAO (Cotter and Gunsalus, 1989). The turnover number of this enzyme with DMSO and TMAO were 3,000 and 9,100, respectively (Weiner *et al*., 1988). The K_m values for both these substrates were reported as 0.18 mM for DMSO and 0.60 mM for TMAO (Weiner *et al*., 1988) In light of the turnover number and K_m values,

it was suggested that both TMAO and DMSO would serve as the true physiological substrates for DMSO reductase in the natural environment (Weiner *et al*., 1988).

The DmsABC reductase consists of three subunits (DmsA, B and C). The DmsA subunit contains a molybdo-bis(pyranopterin guanine dinucleotide) (Mo-bisPGD) cofactor and a [4Fe-4S] cluster (Tang *et al*., 2011). Although DmsA does not appear to be exported to the periplasm, it contains a twin-arginine leader peptide, which targets the protein to the membrane (Winstone *et al*., 2013). The leader peptide is also essential for expression of DmsA and stability of the DmsAB dimer and interacts with the redox enzyme maturation protein DmsD (Winstone *et al*., 2013). The DmsB subunit of DMSO reductase subunit contains four [4Fe-4S] clusters (Bilous and Weiner, 1988). DmsC is the integral membrane subunit of the DMSO reductase complex, which is predicted to have 8 transmembrane helices with the *N*- and *C*-termini located in the periplasm (Sambasivarao *et al*., 1990). This subunit anchors and stabilizes the catalytic subunits DmsA and DmsB (Sambasivarao *et al*., 1990).

1.4.1.3. TMAO and the regulatory hierarchy

The prevailing regulatory rule leads to the induction of the most energetically-efficient respiratory system by *E. coli*, whereby pathways with high ATP or growth yields are favoured as discussed in Section 1.2. There is no evidence to suggest that the *E. coli* TorCAD respiratory system is regulated by FNR or by ArcBA and thus differs from the other anaerobic respiratory systems (Pascal *et al*., 1984; Simon *et al*., 1994). Unlike other alternative electron acceptors that are reduced by *E. coli,* previous work has demonstrated that both the expression of *tor* and TMAO respiration occurred in the presence of the preferred electron acceptor oxygen (Ansaldi *et al*., 2007). It was proposed that an advantage of TMAO reduction during aerobiosis is the production of alkaline TMA, which could enhance growth by maintaining the pH of the culture.

However, TMAO respiration has been shown to be repressed ~two-fold in the presence of nitrate, the preferred anaerobic electron acceptor (Takagi *et al*., 1981), with this repression not occurring in a *narL* mutant (Iuchi and Lin, 1987). A more recent microarray analysis suggested that *torCAD* transcript levels are decreased in the presence of nitrate (Constantinidou *et al*., 2006). Nitrate-repression of *tor* would be consistent with the current dogma that the most energetically efficient mode is adopted because the redox potential of the NO_3^-/NO_2^- pair is +420 mV compared to +130 mV generated by the TMAO/TMA pair

(Simon *et al*., 1994). However, the mechanism of nitrate-mediated regulation of *tor* is unknown.

Thus, it appears that reduction of TMAO may sit outside of the prevailing regulatory hierarchy due to the reduction of TMAO taking place in the presence of oxygen. However, the regulatory rule appears to apply to TMAO reduction in the presence of nitrate.

1.4.2. TMAO and its role in human disease

1.4.2.1. The gut microbiome and disease

The term gut microbiome describes the composite microbial genome found in mammalian gastrointestinal tract. Around 95% of the bacterial genes in the human body, which equates to ~3.6 million genes are located in the large intestine (Qin *et al*., 2010). Over 1,000 different species colonise the human gut (Consortium, 2012). There is a high inter-individual variability in the gut microbiota composition, which is believed to have a role in both maintaining health and the development of disease (D'Argenio and Salvatore, 2015). A number of functions have been assigned to the microbiome, including immune system development, fat storage, synthesis of vitamins, polysaccharide digestion and angiogenesis regulation as summarised in D'Argenio and Salvatore (2015). Over the last decade, next-generation sequencing has developed rapidly, which has had a major impact this field. For example, a specific microbiome can be characterised in depth using these methods. This allows analysis and comparison between the gut microbiomes of both healthy subjects and individuals with specific diseases. It is known that the microbiome can be influenced by a number of factors, including diet (David *et al*., 2014). Thus, the modulation of the gut microbiota via diet to gain a healthy status is a major research area.

1.4.2.2. Metabolism of TMAO

Although some marine organsims endogenously synthesise TMAO, in mammals it mostly comes from exogenous sources (Goldstein and Funkhouser, 1972; Koeth *et al*., 2013; Yancey *et al*., 2001; Zeisel *et al*., 1989). The role of TMAO in human disease and its metabolism has recently been reviewed by Ufnal *et al.* (2015). The gut microbiome plays a crucial role in the metabolism of TMAO and its dietary precursors in the human body, as summarised in Figure 1.7. It has been shown that the concentration of TMAO in blood increases after ingestion of dietary L-carnitine and choline (Miller *et al*., 2014; Zeisel *et al*., 1983). Choline is endogenously synthesised, although this synthesis does not meet human requirements, thus, the majority of choline is acquired from diet (Ufnal *et al*., 2015). There are a number of dietary sources rich in choline including red meat, poultry and eggs (Ufnal *et al*., 2015).

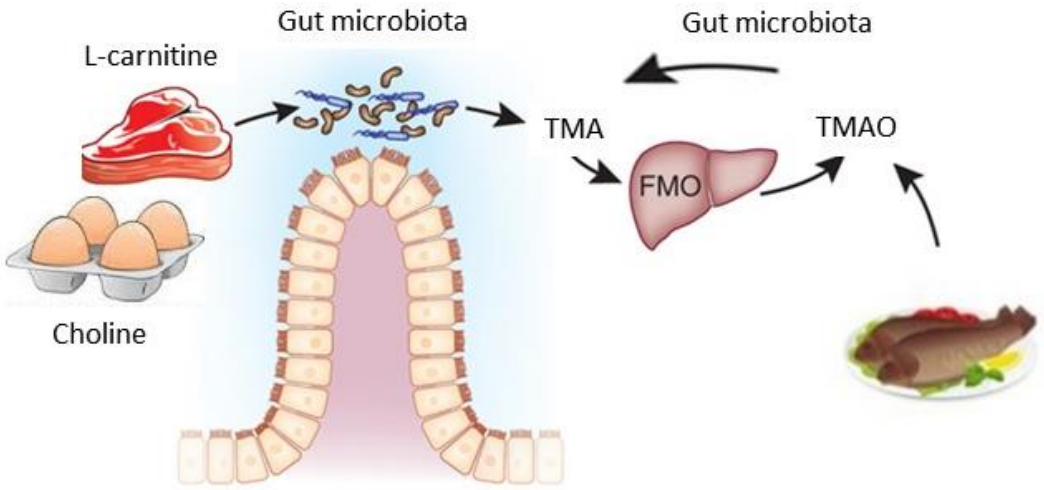


Figure 1.7. <u>Metabolism of TMAO in the human body and its dietary sources</u>

The metabolism of TMAO and its dietary sources are shown. Choline and L-carnitine are converted to TMA by the gut microbiota, which is then oxidised to TMAO by FMOs in the liver. TMAO can also be directly ingested trough the intake of fish. This TMAO can be reduced to TMA by enteric bacteria in the gut. Adapted from Backhed (2013).

There is a threshold concentration of choline that must be ingested for it to be converted to TMAO (Miller *et al*., 2014; West *et al*., 2014; Zeisel *et al*., 1989). Mediated transport in the small intestine is responsible for choline absorption, which is half saturated at choline concentrations in the intestines of 200 to 300 mM (Kuczler *et al*., 1977; Sanford and Smyth,1971; Sheard and Zeisel, 1986; Ufnal *et al*., 2015). When the concentration of choline in the small intestine exceeds transport capacity, normally when choline-rich food is ingested, it is able to reach the large bowel, where intestinal bacteria metabolise it, producing TMA and dimethylamine (DMA) (Wang *et al*., 2011; Zeisel *et al*., 1985). In humans, L-carnitine, which is both acquired from diet and biosynthesis, is another important source of TMAO (Rigault *et al*., 2006; Ussher *et al*., 2013). It is found at high levels in poultry, red meat and some dairy products (Rigault *et al*., 2008). However, little L-carnitine is present in eggs and fish. In a similar way to choline, TMA is produced from L-carnitine by bacterial reduction in the large bowel. Flavin-containing monooxygenase-3 (FMO3), a hepatic enzyme rapidly oxidises TMA derived from bacterial metabolism to TMAO in the bloodstream (Wang *et al*., 2011; Zeisel, 1981; Zeisel *et al*., 1983). Moreover, a high-protein diet has been associated with high urine TMAO concentration (Rasmussen *et al*., 2012). Modification of the gut microbiota can be caused by a high-protein diet, which may affect the absorption of choline and L-carnitine in the small intestines. In humans, DMA, TMA, and TMAO are excreted mainly in urine, as well as in sweat and exhaled air (Ufnal *et al*., 2015). In healthy individuals, the concentration of TMAO in human blood plasma is in the range of 3 mmol/L, this increases to 40 mmol/L in patients with renal failure (Bain *et al*., 2006; Tang *et al*., 2013).

There is also evidence that dietary TMAO and TMA may also be direct sources of TMAO in humans. TMAO, TMA, and DMA is present at high concentrations in seafood (Bell *et al*., 1991; Svensson *et al*., 1994). Compared to the British population, increased urinary excretion of TMAO in the Swedish population has been reported and was linked with the higher fish consumption of the Swedish population (Lenz *et al*., 2004). This is also seen when the Japanese population are compared with that of North America (Dumas *et al*., 2006). After treatment with broad-spectrum antibiotics, a reduction in the TMAO concentration in the blood of humans is seen (Tang *et al*., 2013; Wang *et al*., 2011). This confirms that the key source of TMAO in the blood is gut bacteria and reduction of dietary TMAO by enteric bacteria, such as *E. coli,* is a likely contributing factor, as the TMA produced enters the bloodstream and is oxidised to TMAO in the liver in the gut producing TMA (see above) (Barrett and Kwan, 1985).

1.4.2.3. TMAO in cardiovascular disease

Recent evidence shows that in humans, an elevated concentration of plasma TMAO is marker of an increased cardiovascular risk in humans (Koeth *et al*., 2013; Tang *et al*., 2013; Trøseid *et al*., 2015; Wang *et al*., 2014). Foods that are abundant in phosphatidylcholine and L-carnitine, such as red meat, egg yolks, and other products are considered to be atherogenic, suggesting a link between diet and CVD (Koeth *et al*., 2013). A positive correlation between an increased cardiovascular risk and high blood plasma TMAO levels and has been determined (Koeth *et al*., 2013; Wang *et al*., 2014). TMAO has been shown to affect lipid and hormonal homeostasis. One study found that the development of atherosclerosis arise through TMAO as it may regulate sterol and cholesterol metabolism (Koeth *et al*., 2013). This study showed that suppression of reverse cholesterol transport through gut microbiota-dependent mechanisms is mediated by TMAO (Koeth *et al*., 2013). Angiotensin II, a hormone involved in homeostasis of the circulatory system, is also affected by TMAO (Ufnal *et al*., 2014). Finally, TMAO is also know to affect protein folding (Bennion and Daggett, 2004; Ma *et al*., 2014; Mondal *et al*., 2013; Sarma and Paul, 2013), which may affect the activity of the enzymes and hormones involved in the control of the circulatory system homeostasis.

However, TMAO may reduce the effects of oxidative stress and thus play a protective role in cardiovascular system. For example, a significant increase in plasma TMA and TMAO levels after oral L-carnitine supplementation, decreases intracellular adhesion molecule-1, vascular cell adhesion molecule-1 and malondialdehyde, which are markers of vascular injury and oxidative stress (Fukami *et al*., 2015).

1.4.2.4. Trimethylaminuria

Trimethylaminuria, or fish odour syndrome, is an autosomal recessive disease (Hernandez *et al*., 2003). Trimethylaminuria is characterised by a FMO3 deficiency leading to both the accumulation of TMA in the body and its excretion in urine, breath, and sweat (Hernandez *et al*., 2003). However, a number of healthy individuals have increased excretion of TMA in the absence of the classical features of trimethylaminuria (D'Angelo *et al*., 2013).

1.4.2.5. Kidney disease

End-stage renal disease results in low TMAO plasma clearance, which leads to the accumulation of TMA and TMAO (Bell *et al*., 1991). Kidney ischemic damage causes the release of TMAO from the renal medulla due, which may be responsible for increased urine and plasma TMAO levels (Doucet *et al*., 2004; Hauet *et al*., 2000).

1.4.2.6. Diabetes

High plasma TMAO levels may affect glucose metabolism and thus can be associated with diabetes mellitus (Gao *et al*., 2014; Li *et al*., 2012). Compared to healthy individuals, patients with diabetes have a significantly higher median plasma TMAO concentration (Lever *et al*., 2014). Interestingly, TMAO was found to promote normal protein folding in diabetic rats, counteracting endoplasmic reticulum stress (Lupachyk *et al*., 2013).

1.4.2.7. Cancer

A recent study has positively linked plasma TMAO levels with colorectal cancer (CRC) (Bae *et al*., 2014). It is known that a high consumption of red meat is associated with an increased risk for CRC (Norat *et al*., 2005). TMA and DMA can become nitrosated (Lijinsky *et al*., 1972), which may act as potent carcinogens (Bartsch and Montesano, 1984). Interestingly, whereas production of TMAO requires bacterial processing, the production of promutagenic nitrosated amines does not require the presence of bacteria to occur (Lunn *et al*., 2007; Tang *et al*., 2013; Ufnal *et al*., 2015).

1.4.3. TMAO and fish spoilage

In humans, most TMAO seems to come from the oxidation of TMA that is, produced by gut flora from TMAO (Koeth *et al*., 2013). However, may be synthesised endogenously by marine animals, although the mechanism of this is still unknown (Goldstein and Funkhouser, 1972). There is a higher concentration of TMAO in marine animals than in any other organism (Yancey, 2005). Some contaminating bacteria, which may be present on marine organisms, are able to reduce TMAO to TMA. It is the TMA that is responsible for the characteristic smell of decaying seafood. This led to the usage of the TMA-to-TMAO ratio to indicate freshness in the fishing industry (Lundstrom and Racicot, 1983). In fact, TMAO was first discovered in 1909 in the muscle tissue of dogfish (*Acanthias vulgaris)* and bacterial reduction of TMAO was discovered through studying fish spoilage (Suwa, 1909). The relationship between TMA production and increases in bacterial counts during the storage of marine fish was confirmed (Beatty and Gibbons, 1937). Assays of the TMAO content in fish tissue showed that the concentration of this compound in all marine fish and molluscs was frequently more than 1% of the wet weight (Barrett and Kwan, 1985). However, the concentration of TMAO in freshwater fish was shown to be negligible (Barrett and Kwan, 1985). In the fishery industry, the demethylation of TMAO to DMA and formaldehyde catalysed by endogenous enzymes in the fish muscle (TMAO demethylases) is now a bigger problem than the bacterial reduction of TMAO to TMA (Sotelo *et al*., 1995).

1.5. Mechanisms of transcriptional repression

Bacteria use different mechanisms to sense and respond to conditions within and outside the cell membrane. These responses most frequently lead to the activation and/or repression of a number of genes to adapt cell physiology or metabolism to the new conditions. Bacteria contain several sigma (σ) factors, each one directing RNA polyermase to a specific set of promoters (Gross *et al*., 1992), a strategy that is in itself the first level of regulation of transcription initiation. In principle, any factor inhibiting the access of RNA polyermase to the promoter can be considered a repressor. This definition includes not only the classical repressors but the anti-sigma factors as well.

Regulatory proteins play key roles in responding to the different conditions that bacteria are exposed to. Regulatory proteins come in two forms: activators or repressors. Many repressors prevent transcription by binding DNA at positions that interfere directly with RNA polymerase binding (Gralla and Collado-Vides, 1996). Thus, at many promoters subject to repression, operator sequences for a repressor are found to overlap or be immediately adjacent to the transcription start site. Repressor proteins bind to operator sequences and directly prevent the transcription of target genes. The process of repression can happen in one of two ways, depending on the repressor. Some repressors bind operator DNA themselves and impede RNA polymerase access and thus transcription. Expression of the target gene requires that a specific ligand, called an inducer, bind to the repressor, causing it to release from the operator DNA. Because a small inducer molecule is required, the increased expression of the target gene is called induction. The *E. coli* LacI when binding at the O_1 operator of the *lac* promoter works in this manner (Straney and Crothers, 1987). LacI and RNA polymerase can bind simultaneously to the *lac* promoter, forming a nonproductive complex that can be rendered productive upon addition of the inducer isopropyl-b-D-thiogalactopyranoside (Straney and Crothers, 1987). To exclude RNA polymerase from the promoter, a repressor does not necessarily need to bind to a site overlapping the RNA polymerase binding region. For example, the CytR repressor binds to position -70 at the E. coli *deo* promoter, helped by two flanking cyclic AMP receptor protein (CRP) dimers bound to positions -40 and -93 (as reviewed in Valentin-Hansen *et al*., 1996).

Other repressor proteins do not bind well to operator DNA unless they first bind a small ligand called a co-repressor. When the repressor protein binds to the co-repressor ligand, the complex binds to operator DNA strongly (Slonczewski and Foster, 2009). As the co-repressor ligand disappears, it is no longer available to bind to the repressor. Thus, the repressor releases from the operator and the target gene is transcribed. This process is

called derepression. The *E. coli* tryptophan biosynthesis operon (*trp*) is an example of a derepression system. The expression of the *trp* operon is repressed 300-fold by the repressor TrpR (Klig *et al*., 1988). TrpR-tryptophan bind the *trpL* promoter, possibly directly blocking binding by RNA polymerase and thus repressing the expression of the tryptophan biosynthesis machinery (Klig *et al.,* 1988).

Thus, a number of different molecular mechanisms are responsible for the repression of transcription in bacteria.

1.6. Aims of this work

The information provided above shows that TMAO is present in many environments that *E. coli* and other enteric bacteria inhabit. Thus, enteric bacteria have evolved to utilise TMAO as an alternative electron acceptor. This process has been well characterised. However, respiration of TMAO is hypothesised to sit outside the conventional hierarchy of regulation of respiration in *E. coli*. TMAO respiration is not regulated by FNR or ArcBA, as other modes of the metabolism are to ensure the most energetically efficient electron acceptor present is used. Further evidence for this mode of respiration sitting out of the regulatory hierarchy is shown because TMAO is respired by *E. coli* in the presence of oxygen, the preferred electron acceptor. Recently, TMAO has been the focus of a number of publications, linking the molecule to various disease processes in humans. Based on current evidence, the hypothesis is that respiration of TMAO by *E. coli* is not regulated in the same way as other modes of the respiration, i.e. no involvement of ArcBA or FNR, and the presence of TMAO simply results in the induction of the *E. coli* TMAO reductases. The experiments described in this thesis were carried out to test this hypothesis using a range of molecular techniques.

A range of different state of the art molecular technqiues are planned to be used in order to study this hypothesis and see if using these newer techniques can highlight any additional information, adding further to the body of previous knowledge. One of the techniques planned to be used is transcriptional profiling using microarrays. Using this method should enable the study of the change in transcript levels in *E. coli* in response to TMAO. This will highlight any changes that occur in dynamic environmental conditions in real-time. Metabolite profiling is also planned to be used, in order to determine the different metabolites produced by *E. coli* in the presence of TMAO. NMR is used to carry out metabolite profiling and enables accurate detection of different metabolites in complex cultures. Transcript and metabolite profiling can then be used to garner information on the expression levels and activities of various TFs involved in the initiation of the *E. coli* response to TMAO. All of this will be carried out on chemostat cultures, to ensure reproducibility and

accuracy. It is hoped that using such molecular techniques may highlight new findings, which could add to the body of knowledge around TMAO respiration by *E. coli.*

2. Materials and Methods

2.1. Strains and plasmids

The strains of *Escherichia coli* used in this study are listed in Table 2.1. Strains were maintained on LB agar plates (containing antibiotics where appropriate) at 4°C and subcultured every 6 – 8 weeks. Longer term storage was in LB medium (containing antibiotics where appropriate) containing 7% (v/v) dimethylsulfoxide (DMSO) at -70°C.

Plasmids used in this study are described in Table 2.2. Plasmids were stored in distilled H_2O (dH_2O) at -20°C.

2.2. Media

All heat stable solutions were sterilised by autoclaving at a pressure of 1.1 kg cm^{-2} at 121°C for 20 min. Heat labile solutions were sterilised by filtration through a 0.2 µm membrane.

2.2.1. Rich media

Luria Bertani (LB) medium and agar	(L^{-1})
Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g
Agar (when added)	15 g
	pH 7.0

Rich medium	(L^{-1})
Tryptone	32 g
Yeast extract	20 g
$MgSO_4.7H_2O$	0.16 g
K_2HPO_4	6 g
$NH_4NaHPO_4.4H_2O$	2.4 g
	pH 7.0

2.2.2. Minimal media

Evans minimal medium	(L^{-1})
2 M $NaH_2PO_4.2H_2O$	5 ml
2 M KCl	5 ml
0.25 M $MgCl_2.6H_2O$	5 ml
4 M NH_4Cl	25 ml

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

Strains	Relevant characteristics	Source or reference
K-12		
MG1655	F^- λ^- *ilvG*$^-$ *rfb*-50, *rph*-1	Lab collection
JRG6703	MG1655 Δ*frmRAB* (Kn^R)	This work
JRG6705	MG1655 $P_{torC\text{-}gfp}$, *ΔtorC* (Kn^R)	This work
JRG6726	JRG6703/ pGS2486 (Kn^R, Ap^R)	This work
JRG6825	JRG6703/ pGS2547 (Kn^R, Ap^R)	This work
JRG6826	JRG6703/ pGS2548 (Kn^R, Ap^R)	This work
JRG6437	MG1655/ pKD4 (Kn^R)	Lab collection
ER2925	*ara*-14 *leuB*6 *fhuA*31 *lacY*1 *tsx*78 *glnV*44 *galK*2 *galT*22 *mcrA dcm*-6 *hisG*4 *rfbD*1 R(*zgb*210::Tn10)TetS *endA*1 *rpsL*136 *dam*13::Tn9 *xylA*-5 *mtl*-1 *thi*-1 *mcrB*1 *hsdR*2 (Cm^R)	New England BioLabs, Inc.
JRG6842	ER2925 $P_{torC\text{-}gfp}$, *ΔtorC* (Cm^R, KnR)	This work
B-strain		
BL21 λ(DE3)	*Escherichia coli* BL21 λ(DE3) lysogen carrying a copy of the T7 RNA polymerase under the control of the IPTG-inducible *lacUV5* promoter	Lab collection
JRG6782	BL21/ pGS2516 (Ap^R)	This work
JRG6783	BL21/ pGS2517 (Ap^R)	This work
JRG6784	BL21/ pGS2518 (Ap^R)	This work
JRG6437	BL21/ pVJS1043 (Ap^R)	Lab collection

Ap^R – ampicillin resistant, Kn^R- kanamycin resistant, Cm^R – chloramphenicol resistant strain.

Table 2.2. Plasmids used in this study

Plasmid	Relevant characteristics	Source or reference
pGS2486	Ap^R, plasmid encoding the MG1655 *frmRAB* operon plus 300 bp of the upstream native promoter region.	This work
pGS2547	Ap^R, plasmid encoding the MG1655 *frmRAB* operon plus 300 bp of the upstream native promoter region, with the *frmR* gene carrying a C35A mutation.	This work
pGS2548	Ap^R, plasmid encoding the MG1655 *frmRAB* operon plus 300 bp of the upstream native promoter region, with the *frmR* gene carrying a P2A mutation.	This work
pVJS1043	Ap^R, pMAL plasmid containing *narL* gene for overexpression of MBP-NarL fusion protein.	Lab collection
pKD4	Km^R, plasmid encoding a kanamycin resistance cassette for the use in λred recombination method of gene deletion.	Lab collection
pSIM18	Hyg^R, deriviative of pSC101 carrying 3 λred genes (g*am*, b*et* and *exo*) under the control of the λ pL promoter, which is regulated by the temperature sensitive Cl857 repressor.	Lab collection
pDOC-G	Km^R, plasmid allowing *C*-terminal tagging of chromosomal genes or promoter regions with a GFP tag or deletion of chromosomal regions.	Lab collection
pGS2516	Ap^R, pET22a plasmid carrying *frmR* for overexpression of untagged FrmR	Dr Peter Chivers, University of Durham
pGS2517	Ap^R, pET22a plasmid carrying *frmR,* with a C35A, mutation for overexpression of untagged FrmR C35A	Dr Peter Chivers, University of Durham
pGS2518	Ap^R, pET22a plasmid carrying *frmR,* with a P2A, mutation for overexpression for overexpression of untagged FrmR P2A	Dr Peter Chivers, University of Durham

Ap^R – ampicillin resistance, Kn^R - kanamycin resistance, Hyg^R – hygromycin resistance genes on the plasmids listed.

0.4 M Na_2SO_4	5 ml
Titriplex (nitrilotriacetic acid)	0.38 g
0.004 M $CaCl_2.2H_2O$	5 ml
Trace elements	5 ml
3 mg/ml sodium selenite	10 µl
	pH 6.9

Evans medium buffered with 50 mM phosphate buffer	(L^{-1})
As above except 2 M $NaH_2PO_4.2H_2O$ (5 ml) was replaced with:	
1 M Na_2HPO_4	33 ml
1 M $NaH_2PO_4.2H_2O$	17 ml

Trace elements	(L^{-1})
ZnO	0.41 g
$FeCl_3.6H_2O$	5.40 g
$MnCl_2.4H_2O$	2.00 g
$CuCl_2.2\ H_2O$	0.17 g
$CoCl_2.6\ H_2O$	0.48 g
H_3BO_3	0.06 g
$Na_2MoO_4.2\ H_2O$	4.00 mg
HCl (37%)	8.00 ml

2.2.3. Media supplements

Bacteria carrying antibiotic resistance plasmids or cassettes were grown in the presence of the appropriate antibiotics. Antibiotics were added to the autoclaved media (cooled to 50°C) at the following concentrations:

Ampicillin	100 µg ml^{-1}
Kanamycin	25 µg ml^{-1}
Hygromycin	100 µg ml^{-1}
Chloramphenicol	25 µg ml^{-1}

Other supplements were used at the following concentrations

Glucose	20 mM

2.3. Culture techniques

2.3.1. Chemostat culture of *Escherichia coli* MG1655

Escherichia coli K-12 sub-strains MG1655 and JRG6705 were used to establish steady-state continuous cultures in a 2 L Labfors chemostat (Infors-HT, Bottmingen, Switzerland) at 37°C,

a 0.2 h^{-1} dilution rate, a 950 ml working volume, a stirring rate of 400 rpm and a gas flow rate of 1 L/min. The pH was maintained at 6.9 by titration with 1 M NaOH. Continuous cultures were grown in 20 mM glucose-limited Evans minimal medium plus antibiotic when required. Initially, 5 ml of overnight culture was added to 1 L Evans medium in the chemostat. This was batch cultured aerobically overnight, before starting the medium feed (0.2 h^{-1}) and setting the gas mix to anaerobic (N_2 and 5 % CO_2 gas, (BOC)) to establish a fermentative steady-state over the next 24 h. After 24 h the fermentative steady-state samples were taken from the chemostat using SureSEAL universal tubes. Then the culture was perturbed by adding 45.76 ± 1.48 mM TMAO to both the Evans feed and directly into the chemostat. Transcriptomic and metabolite samples (Section 2.7) were taken at 2, 5, 10, 15, 20 and 60 min after the addition of TMAO (~40 mM). The chemostat was then left to establish a TMAO respiratory steady-state over 24 h, before final samples were taken.

2.3.2. Aerobic batch culture of *Escherichia coli* MG1655

Escherichia coli strains were grown in liquid batch cultures in either LB, Evans medium pH 7.0 or in Evans medium pH 7.0 buffered with 50 mM phosphate, with supplements as indicated. Cultures were grown in conical flasks with vigorous shaking (250 rpm) at 37°C, unless otherwise indicated. Cell growth was monitored by measuring optical density at 600 nm. Cultures on solid medium were grown on LB plates at 37°C unless otherwise indicated.

2.3.3. Anaerobic batch culture of *Escherichia coli* MG1655

Liquid cultures were grown in in either LB, Evans medium pH 7.0 or in Evans medium pH 7.0 buffered with 50 mM phosphate, with supplements as indicated. Cultures were grown in air-tight bottles filled to the top. Anaerobic cultures were grown at 37°C without shaking.

2.4. Cell viability assays

Anaerobic cultures were grown overnight at 37°C in Evans medium, buffered with 50 mM phosphate buffer, pH 7.0, and supplemented with 20 mM glucose in 7 ml bijoux filled to the top. After overnight growth, a 10% inoculum of the cultures was added to 7 ml fresh Evans medium in an anaerobic workstation. Pre-treatment viable counts were measured from this culture, by serial dilution in Evans medium. The 7 ml culture was then split in half (2 x 3.5 ml), and 40 mM TMAO was added to one of the two subcultures. The cultures were incubated in parallel anaerobically for 1 h at 37°C, before post-treatment viable counts for both cultures were measured. Plated cells were grown overnight at 37°C on LB agar, supplemented with antibiotics when appropriate, and CFUs ml^{-1} calculated. Viability assays were carried out with three technical repeats and three biological replicates for each strain.

2.5. Formaldehyde-sensitivity of FrmR mutants *in vivo*

A 1% (v/v) inoculum of an overnight culture of *E. coli* JRG6703, harbouring either pGS2497 (*frmR* wild type), pGS2547 (*frmR* P2A) or pGS2548 (*frmR* C35A) was grown in LB aerobically until A_{600} 0.5 was reached. Cultures were then diluted to A_{600} 0.05 in LB, in LB medium in a sterile 96-well plate. Formaldehyde was added to the cultures, with an increasing concentration gradient across the plate (0-1650 µM), with total volume adjusted to 200 µl. The growth of the cultures aerobically at 37°C with shaking (250 rpm) was monitored using a Sunrise absorbance reader (Tecan), with A_{595} recorded for 25 kinetic cycles at intervals of 20 min. Biological replicates were carried out in triplicate. The maximum growth rate (µmax h^{-1}) for each strain was determined using the steepest part of the growth curve.

2.6. Transformation

2.6.1. Preparation of electrocompetent cells

Strains were grown overnight from a single colony in 5 ml LB. An aliquot (2.5 ml) of this culture was then used to inoculate 250 ml LB and grown at 37°C with 250 rpm shaking until an A_{600} of 0.6 was reached. Cells were chilled on ice for ~20 min before being harvested by centrifugation at 6,000 *g* for 10 min at 4°C. The cells were then kept on ice throughout the rest of the protocol. Cells were resuspended in 50 ml ice-cold 10% (v/v) glycerol. Centrifugation and washing was repeated for 3 washes. Finally, the cells were resuspended in 1-2 ml of ice cold 10% (v/v) glycerol to form a thick cell suspension. Competent cells were kept on ice and used with 2 h.

2.6.2. Bacterial transformation

Electrocompetent cells (50 µl), were mixed with up to 1 µg of purified plasmid DNA. The cells were left on ice for 10 min and then transformed by electroporation using a Hybaid Cell Shock unit. Luria-Bertani medium (1 ml) was immediately added to the cells, before being incubated for 1 h with shaking at 37°C. The cells were then plated on selective agar plates and grown overnight at 37°C. Any colonies that grew overnight on selective media were analysed.

2.7. Sampling of chemostat cultures

2.7.1. RNA

At each time-point, culture samples (6 ml) were removed from the chemostat vessel (anaerobic) using SUREseal tubes (SLS select), directly into 12 ml RNAprotect Bacteria Reagent (QIAGEN). The samples were vortexed immediately for 5 s, before being centrifuged at 4,020 *g* for 10 min. Pellets were resuspended in the residual supernatant, which was then

centrifuged at 8,000 *g* for 3 min. Excess supernatant was removed by gentle tapping of the inverted tubes.

2.7.2. Cell pellets

Chloramphenicol (20 µl of 25 mg ml^{-1}) was added to 6 ml of culture to prevent protein synthesis after removal from the chemostat. The culture was pelleted via centrifugation for 3 min at 10,000 *g*. Cell pellets were then stored at -80°C.

2.7.3. Supernatant

Samples of culture (6 ml) were pelleted by centrifugation at 3,400 *g* for 10 min. The supernatant was then filtered using a 0.2 µm syringe filter (Sartorius Stedim biotech). Samples were at -80°C until required.

2.7.4. Stabilisation of phospho-Asp residues

Chemostat culture samples (6 ml) were collected directly into chloramphenicol (50 mg ml^{-1} final concentration) and formic acid (1 M final concentration) as previously described (Rolfe *et al*., 2011). Aliquots of the suspension (500 µl) were pelleted by centrifugation (8,000 *g* for 3 min) in 1.5 ml Eppendorf tubes and resuspended in 50 µl of 1 M formic acid, before storing at -80°C.

2.7.5. Cell dry weight

Falcon tubes (15 ml) were weighted to a milligram degree of accuracy. Culture (10 ml) was added to half of the tubes, before centrifugation at 3,400 *g* for 10 min. The supernatant was poured off, and the pellet resuspended in molecular grade water (Sigma), before further centrifugation at 3,400 *g* for 10 min. All tubes were dried in an oven overnight, after the removal of the supernatant. The tubes were re-weighed, the following day; weight change of tubes containing sample after addition of observed weight loss due to heating of the tubes lacking cells was equivalent to cell dry weight.

2.8. Nucleic acid methods

2.8.1. PCR amplification

2.8.1.1. Oligonucleotide primers

Oligonucleotide primers (20 – 60 bases) were designed to anneal to DNA sequences flanking those to be amplified. Generally, the G and C content was 50-60%, the T_m was 55-80°C with comparable annealing temperature for primer pairs. Primer design was carried out using Primer3 software. All oligonucleotides were synthesised by Eurofins Genomics (Ebersberg). The synthetic oligonucleotides used in this work are listed in Appendix 1.

2.8.1.2. Polymerase chain reaction (PCR)

All PCR reactions were carried out using a Techne TC-3000 thermocycler. Typical conditions were as follows:

Forward primer	0.5 µM
Reverse primer	0.5 µM
Template DNA	~250 ng
DMSO	3% (v/v)
Phusion High-Fidelity PCR Master Mix[1] (2X)	25 µl
dH_2O	to 50 µl

[1] The Phusion High-Fidelity PCR Master Mix with HF buffer (New England BioLabs Inc.) was composed of Phusion DNA polymerase (1 U), 1.5 mM $MgCl_2$, dNTPs (200 µM each) in HF buffer.

The DNA polymerase was added last to the mixtures, immediately prior to placing the reactions in the thermocycler. The following basic programme was used:

98°C, 30 s	1 cycle
98°C, 10 s; 55 – 70°C[2], 30 s; 72°C, 30 – 90 s[3]	30 cycles
72°C, 8 min	1 cycle
4°C	

[2]The annealing temperature depended on the T_m of the primers.

[3]The extension time was dependent on the size of the fragment being amplified, with 30 s per kb being used.

The PCR products were purified by PCR purification kit (QIAGEN) (Section 2.8.2) or by gel extraction kit (QIAGEN) (Section 2.8.3) after analysis by gel electrophoresis (Section 2.8.8).

2.8.2. PCR purification

PCR products were purified using QIAquick PCR purification kit (QIAGEN) following manufacturer's instructions, with DNA being eluted into 50 µl of molecular grade water (Sigma). Purified PCR products were stored at -20°C.

2.8.3. Gel extraction

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

2.8.4. Purification of plasmid DNA

Plasmid DNA purification from 5 ml (high-copy number plasmid DNA) or 10 ml (low-copy plasmid DNA) of aerobic overnight *E. coli* cultures in LB broth was carried out using the QIAprep Spin Miniprep Kit (QIAGEN), according to the manufacturer's instructions.

2.8.5. Purification of genomic DNA

Genomic DNA was purified using DNeasy kits (QIAGEN) from 1 ml of an overnight aerobic culture. This protocol yields up to 40 µg of genomic DNA. The kit was used to purify genomic DNA using the protocol outlined in the manufacturer's instructions.

2.8.6. Isolation of RNA

RNA was isolated from sample pellets using a RNeasy Mini kit (50) (QIAGEN) according to the manufacturer's instructions. Briefly, RNeasy-stabilised cell pellets (Section 2.7.1) were resuspended in 200 µl TE buffer containing lysozyme (15 mg ml^{-1}), mixed by vortexing for 10 s and incubated at room temperature for 5 min. Samples were vortexed for 10 s every minute during the incubation time. Buffer RTL, 700 µl, containing β-mercaptoethanol (10 µl ml^{-1}) was added, followed by vigorous vortexing and centrifugation at 20, 000 *g* for 2 min. The resulting supernatant was transferred to a clean tube and 500 µl of ethanol (96% (v/v)) was added. The lysate (including any precipitation) was transferred to an RNeasy Mini spin column, placed in a 2 ml collection tube, and centrifuged for 15 s at 8,000 *g*, with the flow through discarded. Buffer RW1 (350 µl) was added to the column and centrifuged for 15 s at 8,000 *g* to wash the membrane. After the flow through was discarded, a mixture of 10 µl DNase I (27 Kunitz units) and 70 µl Buffer RDD was transferred directly onto the RNeasy column gel membrane and incubated at room temperature for 15 min. Buffer RW1, 350 µl, was then added to the spin column and centrifuged for 15 s at 8,000 *g* following 5 min incubation at room temperature. The column was then transferred to a new 2 ml collection tube and washed twice with buffer RPE (500 µl). To elute RNA, 30 µl RNase-free H_2O was pipetted directly on the membrane of the column and centrifuged for 1 min at 8,000 *g*. The concentration and purity of isolated RNA was determined (Section 2.8.7).

2.8.7. Quantification of nucleic acid concentration and purity

Determination of DNA and RNA purity and concentration was by means of a Nanodrop ND-1000 UV-VIS spectrophotometer, using 2 µl of sample. The pre-configured methods for measuring dsDNA and RNA concentrations at 260 nm wavelength were used. The purity of samples was assessed using the A_{260}/A_{280} ratio, whereby a ratio between 1.7 and 2.0 shows high-purity as defined in the Two-Color Microarray-Based Prokaryote Analysis (FairPlay III Labeling) Protocol (Agilent, 252009). The purity of RNA samples was also determined by running on an agarose gel (Section 2.8.8) and looking for the presence of two distinct species corresponding to ribosomal RNA (rRNA) to show a lack of RNA degradation in the samples.

2.8.8. Agarose gel electrophoresis

Analysis, recovery and purification of target DNA fragments were carried out using agarose gel electrophoresis using a Flowgen electrophoresis apparatus and concentrations of 1-3% (w/v) agarose gels (depending on size of fragment) buffered with 1X TAE buffer supplemented with 1:10,000 GelRed (Biotium). DNA loading buffer, 5X, (Bioline) was added to DNA samples with 10 to 100 ng of DNA loaded per lane. HyperLadder I (Bioline) (200 ng) was run as a marker alongside the samples to estimate the sizes of DNA fragments. DNA was visualized by exposure to a source of ultraviolet light (254 nm).

50x TAE buffer	
Tris	242 g
Glacial acetic acid	57.1 ml
EDTA, pH 8.0 (0.5 M)	100 ml
dH_2O	to 1 L

2.8.9. DNA digestion with restriction endonucleases

Restriction endonucleases were used with compatible 10X buffers. In the case of double enzyme digestion, a 10X buffer was chosen to maintain conditions for both enzyme activities. Digestions were usually performed in a 20 µl reaction mixture containing 0.5 – 50 µg of DNA and 5 units of enzyme per 1 µg of DNA. Reactions were incubated at 37^{o}C for 1 – 3 h. Reactions were loaded directly onto agarose gels for analysis and purification of the separated fragments.

2.8.10. Site-Directed Mutagenesis (SDM)

Site-Directed Mutagenesis was carried out using QuickChange II XL Site-Directed Mutagenesis kit (Agilent). The manufacturer's instructions were followed, using the following oligonucleotide primers to introduce the required mutations to pGS2486: C35A, pGS2516 C35A SDM F and pGS2516 C35A SDM R; P2A, pGS2516 P2A SDM F and pGS2516 P2A SDM R (Appendix 1). Briefly, two complementary primers containing the desired mutation flanked by the unmodified nucleotide sequence, were designed and synthesised. Thermo cycling was used to synthesise the mutant strand, using the reaction:

10X reaction buffer	5 µl
pGS2486 DNA	10 ng
Forward primer	125 ng
Reverse primer	125 ng
dNTP mix	1 µl

QuikSolution reagent	3 µl
PfuUltra HF DNA polymerase (2.5 U/µl)	1 µl
dH2O	to 50 µl

Each reaction was cycled following the profile outlined below:

95°C, 1 min	1 cycle
95°C, 50 s	18 cycles
60°C, 50 s	18 cycles
68°C, 7 min	18 cycles
68°C, 7 min	1 cycle

The reactions were incubated on ice for 2 min after thermo cycling. The reactions were then digested with *Dpn*1 (10 U) for 1 h at 37°C. Digested DNA products (2 µl) were used to transform *E. coli* XL10-Gold ultracompetent cells (Agilent) via heat shock at 42°C as per manufacturer's instructions. Each reaction was plated on agar plates with the appropriate antibiotic and incubated at 37°C for 24 h. Plasmids were then isolated and sequenced to confirm base changes, before being used to transform relevant host strains.

2.8.11. Gene deletion and creating a reporter fusion using the λred system

All mutant and reporter fusion strains used in this study were constructed using λred recombineering (Datsenko and Wanner, 2000). In this procedure, recombination requires the phage λred recombinase, which is synthesised under the control of an inducible promoter. Firstly, PCR products were generated using primers with 40 nucleotide extensions (Appendix 1), which were homologous to regions adjacent to the gene to be modified and the template plasmids (pKD3, pKD4 or p-DOC-G) (Table 2.2) carrying the appropriate antibiotic resistance genes/reporter genes that are flanked by FRT (FLP recognition target) sites. PCR products were purified using a PCR purification kit (Section 2.8.2) or by gel extraction (Section 2.8.3) to select for DNA of the correct size, if multiple PCR products were present. Mutants of the required gene/reporter fusions were isolated as antibiotic resistant colonies after transformation of bacteria carrying the λred expression plasmid, pSIM18 (Table 2.2), by the linear PCR product (3 – 5 µg). Correct insertion of the antibiotic resistance cassette was confirmed by extracting total genomic DNA, carrying out PCR using upstream and downstream sequencing primers (Appendix 1) and then sequencing the PCR product. Once a mutant had been successfully created, the cassette was transferred to a genetically clean strain, using generalised P1 transduction (Section 2.9).

2.8.12. In-Fusion HD cloning system

The In-Fusion HD cloning system (Clonetech) was used to create plasmid pGS2486. A purified PCR fragment encoding the *frmRAB* operon and its own promoter region with appropriate 15 bp overhangs (synthesised with primers Frm In-Fusion F and Frm In-Fusion R) (Appendix 1) was incubated with linearised pBR322 vector (digested with *Eco*R1 as described in Section 2.8.11) according to manufacturer's instructions. The reaction was as follows:

5X In-Fusion HD Enzyme Premix	2 µl
Linearised pBR322 vector	50 ng
frmRAB PCR product with 15 bp overhangs	80 ng
Nuclease-free water	To 10 µl final volume

Part of the reaction mix (2.5 µl) was then transferred into Stellar Competent Cells, which were then plated onto ampicillin containing LB agar plates. Ampicillin resistant colonies were then selected, re-streaked onto fresh plates, and the insertion of *frmRAB* was confirmed by isolating the plasmid and carrying out PCR (with primers Frm In-Fusion seq F and Frm In-Fusion seq R). The DNA sequence of the *frmRAB* insertion was verified by sequencing and the plasmid was used to transform *E. coli* JRG6703 (Table 2.1).

2.8.13. Transcriptomics: Microarray analyses

An indirect method of microarray analysis was used, with cDNA samples labelled with the Cy3 fluorophore and *E. coli* MG1655 genomic DNA labelled with Cy5 fluorophore as a reference. The indirect method has advantages over dye-swap experiments, such as multiple datasets, for which genomic DNA has been used as a reference can be compared (Yang and Speed, 2002).

2.8.13.1. Direct labelling of genomic DNA

Labelling of genomic DNA was carried out using a protocol based on the BioPrime DNA Labeling Kit (Invitrogen, 18094-011). *Escherichia coli* MG1655 chromosomal DNA (2.0 µg) (Section 2.8.5), was added to a sterile Eppendorf tube and the volume made up to 21 µl. To this was added 20 µl of 2.5X Random Primer/Buffer Mix (Invitrogen) before boiling for 5 min, followed by incubation on ice for 5 min. On ice 5 µl of 10X dNTP mix (1.2 mM each dATP, dGTP, dTTP; 0.6 mM dCTP, 10 mM Tris pH 8.0; 1 mM EDTA), 3 µl Cy5-dCTP (1 mM stock, Amersham) and 1 µl of Klenow (40 U) (Invitrogen) was added, before being incubated overnight at 37°C. Following the incubation, the labelling reaction was cleaned up using the QIAquick PCR purification kit (QIAGEN) (Section 2.8.2), making sure to wash with buffer PE

twice and eluting the labelled DNA into 50 µl RNase-free H_2O. The concentration of DNA and labelling efficiency of the Cy5 dye were determined using the Nanodrop spectrophotometer.

2.8.13.2. cDNA synthesis and labelling

RNA (16 µg) (Section 2.7.1 and 2.8.6) for each sample was incubated with 5 µg random primers (Invitrogen) in a PCR block for 10 min at 72°C, then chilled on ice for 10 min. cDNA synthesis was initiated by adding a reaction mixture consisting of 6 µl 5X First Strand (FS) buffer (Invitrogen), 3 µl 0.1 M DTT (Invitrogen), 0.6 µl 50X dNTPs (25 mM dATP, dGTP, dTTP and 10 mM dCTP) and 2.9 µl nuclease-free H_2O. Cy3-dCTP (1 mM stock, GE Healthcare), 2 µl, and 1.5 µl SuperScript III (200 U ml^{-1}) were added to each sample followed by a 5 min incubation at 25°C, then incubation overnight at 50°C. Samples were hydrolysed by addition of 15 µl of freshly prepared 0.1 M NaOH and incubation at 72°C for 10 min, followed by neutralisation with 15 µl 0.1 M HCl. Labelling reactions were purified using the QIAquick PCR purification kit (QIAGEN) (Section 2.8.2), washing with buffer PE twice and eluting the labelled DNA into 50 µl RNase-free H_2O. Labelled Cy3-cDNA was quantified using the Nanodrop spectrophotometer. The following equations were used to calculate the yield of cDNA and its specific activity:

Yield of cDNA:

cDNA (ng) = A_{260} x 330 ng μl^{-1} x 50 µl (sample volume)

Specific activity:

(*Concentration of Cy3*) / (*Concentration of cDNA*) * 1000 = pmol Cy3 or per µg cDNA

The samples were hybridised if the yield was >825 ng and the specific activity was >8 pmol Cy3 per µg cDNA.

2.8.13.3. Hybridisation and washing of hybridised array slides

To a 1.5 ml nuclease-free microfuge tube was added an appropriate volume of Cy3-labelled cDNA and Cy5-labelled gDNA to give a final concentration of 400 ng of each. Nuclease-free water was added to make up the final volume to 25 µl, before boiling samples at 100°C in a PCR block for 2 min followed by chilling on ice for 2 min. Blocking agent (5 µl of 10X stock) was then added to each reaction, followed by 25 µl 2X GEx Hybridization Buffer HI-RPM before being immediately loaded onto the array. The Agilent SureHyb hybridisation chamber was assembled and 40 µl of sample was slowly dispensed onto a gasket well in "drag and dispense" manner. An array slide (Agilent *E.coli* Custom Gene Expression Microarray) was placed "active side" down onto the SureHyb gasket slide. The slide chamber was placed in a rotisserie in a hybridisation oven set to 65°C with gentle rotation at 10 rpm for 17 h. The

washing of hybridised slides was carried out according to the Two-Color Microarray-Based Prokaryote Analysis (FairPlay III Labeling) Protocol. Before use, the equipment was washed with MilliQ H_2O at least 4 times. Gene Expression (GE) wash buffers were supplemented with 0.005% Triton X-102 (Agilent).

2.8.13.4. Scanning of slides

Slides were scanned immediately after washing using an Agilent DNA microarray scanner (Agilent Technologies, G2505) using the Agilent Scan Control software (v8.5), according to the scanning instructions in the Two-Color Microarray-Based Prokaryote Analysis (FairPlay III Labeling) Protocol (Agilent, 252009). This resulted in two colour .tiff image files being produced for further analysis.

2.8.13.5. Analysis of data

Data acquisition from the .tiff images was carried out using Agilent Feature Extraction software (v6.5), which allows quantification of the Cy3 fluorescence of each feature from the scanned array image. Data were normalised and interpreted using GeneSpring 7.3.1. The LOWESS normalisation was applied and identification of statistically significant gene expression changes was achieved by applying a t-test with a 3-fold cut off and $p < 0.05$. Three biological replicates were carried out for microarrays, using three different chemostat cultures. The relative activities of TFs were inferred using the software package TFInfer as previously described (Asif *et al*., 2010).

2.8.14. Quantitative real time-PCR (qRT-PCR)

The Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent) was used as a one-step for quantitative PCR amplifications. In this system, direct detection of the PCR product was monitored by measuring the increase in fluorescence, as an indicator of amplicon production, caused by binding of SYBR green dye to double-stranded DNA. Fluorescence data were acquired when the PCR amplification was in the exponential phase. The cycle threshold (C_T) at which the SYBR green dye emission intensity rose above the background noise was determined. C_T is inversely proportional to the copy number of the target template and therefore the greater the template concentration, the lower the C_T value.

2.8.14.1. RT-PCR oligonucleotide primer design

Primers used for RT-PCR were designed using Primer3 software to give PCR products within the target genes of approximately 100 bp in size. The primers were approximately 20 nucleotides in length, had T_m ~60^{o}C and primer pairs were designed to have similar annealing

temperatures. All oligonucleotides were synthesised by Eurofins Genomics (Ebersberg). The synthetic oligonucleotides used for RT-PCR are listed in Appendix 1.

2.8.14.2. Amplification of target and control genes

Anaerobic cultures were grown overnight at 37°C in Evans medium buffered with 50 mM phosphate buffer, pH 7.0 and supplemented with 20 mM glucose. The amounts of stabilised *frmR* mRNA in samples of total RNA (100 ng; stabilised as in Section 2.7.1 and isolated as in Section 2.8.6) after exposure to different concentrations of TMAO (0 – 40 mM) for 30 min were determined on an RT-PCR plate in a Mx3005P Thermocycler (Agilent Technologies). In order to calculate expression levels of the target gene, *frmR*, a control housekeeping gene, *gyrA*, was also amplified to allow a relative comparison to baseline expression levels. A single reagent mixture for replicate experimental reactions was prepared, so that each individual reaction was composed of the following:

SYBR Green QRT-PCR Master Mix[4] (2X)	10 µl
Upstream primer	500 nM
Downstream primer	500 nM
DTT (100 mM)	0.2 µl
RT/RNase block[5]	1 µl
Nuclease-free dH_2O	up to 20 µl
Isolated RNA (20 ng μl^{-1})	5µl

[4] SYBR Green QRT-PCR Master Mix was composed of a mutated *Taq* DNA polymerase, 2.5 mM $MnCl_2$, dNTPs, DNA-binding dye SYBR Green I and a buffer formulated for fast cycling.

[5] RT/RNase block provided in the SYBR Green QRT-PCR Master Mix

As a negative control, no-template control reactions were produced for each primer pair, with the addition of 5 µl nuclease-free water instead of RNA. The reactions were gently mixed without creating bubbles and then briefly centrifuged. The following programme was used on the Mx3005P Thermocycler:

50°C, 10 min	1 cycle
95°C, 3 min	1 cycle
95°C, 15 s; 60°C, 20 s	40 cycles

Fluorescence data was collected for each cycle at the 60°C step and analysed by the machine. Each primer pair was tested in duplicate and biological repeats (e.g. RNA from cells with and without TMAO treatment) were carried out in triplicate.

2.8.14.3. Standard curve method

For a comparison of expression levels to be valid, the rate of amplification of the target and control genes must be similar. This was determined by measuring the C_T values for each primer pair using genomic DNA as the template. The genomic DNA was serially diluted and the C_T values measured and plotted against the log concentration of DNA. A best fit line was applied and the gradient and intercept derived from the equation of the straight line.

Each reaction was carried out in 25 µl volumes, as in Section 2.8.14.2, but the DNA template used here was genomic DNA (Section 2.8.5). Serial dilutions of the genomic DNA were carried out across the plate in a range of 100 ng to 0.2 ng, for each primer pair. As a negative control, no-template control reactions were added for each primer pair, with the addition of 5 µl ultra-pure water instead of the genomic DNA. The data was collected and analysed by the machine. The mean C_T value of the duplicates was calculated and plotted against the $\log_{10}$ of the genomic DNA concentration.

The relative expression levels of the genes of interest in the TMAO-treated samples compared housekeeping gene (*gyrA*) were calculated according to the protocol for the standard curve method in the *User Bulletin #2* (*ABI PRISM 7700 Sequence Detection System*, Subject: Relative quantification of gene expression) supplied by Applied Biosystems. To summarise, from each C_T value, the intercept of the corresponding genomic DNA plot was subtracted and the resulting value divided by the gradient of that plot. The mean of the primer duplicates was calculated and the antilog of the value taken. This value was divided by the value for the housekeeping gene in the corresponding sample to normalise the data. Averages of the replicate samples were then calculated to determine the fold change for each target gene in the experimental samples compared to the control samples (no TMAO).

2.9. Generalised transduction with bacteriophage P1 vir

2.9.1. Preparation of lysates

Lysates were prepared according to Miller (1972). Donor cells were grown overnight at 37°C in TY medium supplemented with $CaCl_2$ (5 mM). P1 *vir* stock was diluted from 10^8 – 10^3 PFU ml^{-1}; each was mixed with a 100 µl culture and incubated at 37°C for 20 min. Pre-warmed (37°C) TB (1 ml) and 1.5 ml warm (55°C) TB soft agar were mixed and added to the phage/cell mix and poured onto phage lysate plates. The plates were incubated at 37°C in a "wet box" (with a humid atmosphere) until plaques had taken on a "lacey" appearance, after which the plates were chilled at 4°C for 30 min and an overlay of cold phage dilution buffer added. Plates were stored overnight at 4°C and the overlaying liquid harvested using a Pasteur

pipette. Lysates were stored at 4°C and for long term storage a few drops of chloroform were added.

TY broth	(L^{-1})
Tryptone	8.0 g
NaCl	5.0 g
Yeast extract	5.0 g

TB broth	($1\ L^{-1}$)
Tryptone	4.0 g
NaCl	5 g
Agar (when required for soft agar)	7 g

Phage lysate plates	($0.5\ L^{-1}$)
Tryptone	4.0 g
NaCl	2.5 g
Yeast extract	2.5 g
Glucose	1 g
Agar	6.0 g

After autoclaving, the medium was cooled to 50°C and the following was added:

0.5 M $CaCl_2$	5 ml
1 M $MgSO_4$	5 ml
10 mM $FeCl_3$	0.5 ml

Phage dilution buffer	
Trisma base	0.60 g
$MgSO_4.7H_2O$	1.23 g
$CaCl_2.2H_2O$	0.37 g
NaCl	1.46 g
dH_2O	up to 500 ml
	pH 7.5

2.9.2. Generalised transduction

Recipient cells were grown overnight at 37°C in 2.5 ml TY medium supplemented with $CaCl_2$ (5 mM). P1 *vir* stock from a lysate preparation (100 µl) were mixed with the same volume of recipient cells and incubated at 37°C for 20 min. The mixture was spread onto P1 plates containing appropriate antibiotics plus $Na_4P_2O_7$ (0.125 mM) and incubated overnight at 37°C

(Miller, 1972). After overnight incubation, potential transductants were re-plated onto selective medium and the phenotype verified.

2.10. Protein methods

2.10.1. Overproduction of proteins

2.10.1.1. Overproduction of FrmR and related variants

For isolation of FrmR and the C35A and P2A variants, cultures of the *E. coli* expression strains (JRG6782, 6783 and 6784, respectively) were grown at 37°C in LB containing ampicillin, to an A_{600} of ~0.6, at which point 1 mM IPTG was added and the cultures incubated for a further 3 h. Typically, 2 x 500 ml cultures were grown. Bacteria were pooled, collected by centrifugation and stored at -20°C after incubation with IPTG.

2.10.1.2. Overproduction of MBP-NarL

NarL was amplified as a MBP-NarL fusion in *E. coli* JRG6437. Generally, 2 x 500 ml cultures were grown aerobically at 37°C in rich medium containing both ampicillin (100 μg ml^{-1}) and glucose (2%) to A_{600} ~0.6, at which point expression was induced by the addition of 300 μM IPTG. Incubation was continued for a further 3 h when bacteria were pooled, collected by centrifugation and stored at -20°C.

2.10.2. Production of cell-free extracts (CFE)

2.10.2.1. Production of FrmR containing CFE

Cells were lysed after re-suspension in Buffer A (50 mM Tris containing 0.1 M NaCl, 2 mM EDTA, 10 mM dithiothreitol (pH 8.0)) in a ratio of 1:8 (w/v) (pellet:buffer) by sonication (Soniprep150 ultrasonic disintegrator) at ~16 microns for 2 cycles of 20 s, followed by centrifugation.

2.10.2.2. Production of Maltose binding protein (MBP) fused NarL containing CFE

Clarified cell-free extracts were produced by re-suspending the bacteria from 1 L culture in 15 ml of extraction buffer (20 mM Tris-HCl containing 1 mM EDTA, 100 mM NaCl and 10 mM β-mercaptoethanol), and lysing the cells by two passages through a French pressure cell, followed by centrifugation. The fraction containing the soluble proteins was then extracted for analysis.

2.10.3. Protein purification

2.10.3.1. Purification of FrmR and its variants

Purification of FrmR and its variants was carried out by Dr Svetlana E. Sedelnikova (Department of Molecular Biology and Biotechnology, University of Sheffield, UK). Cell-free extract was applied to a Heparin-HP column (GE Healthcare) and FrmR was eluted using a NaCl gradient (0 to 0.5 M) in Buffer A (Section 2.10.2.1). Fractions containing FrmR protein, as detected by SDS-polyacrylamide gel electrophoresis, were combined and concentrated

using a VivaSpin column 500 MWCO 30,000 (GE Healthcare). The concentrated fractions were applied to a HiLoad Superdex 200 column (GE Healthcare) equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl. FrmR-containing fractions were pooled and the purity of the samples was assessed by SDS-polyacrylamide gel electrophoresis. The authenticity of FrmR was confirmed by total amino acid analysis (Section 2.10.9) allowing calibration of FrmR protein concentrations measured by the A_{280} (Section 2.10.4.2). FrmR protein was stored for up to 2 weeks at room temperature until required.

2.10.3.2. Purification of MBP-NarL

The MBP-NarL fusion protein (~5 ml crude extract) was absorbed onto a 15 ml amylose resin column (BioLabs Inc.). Prior to absorption of the protein, the column was equilibrated with 50 ml column buffer (20 mM Tris-HCl (pH 7.4) containing 1 mM EDTA (pH 8.0), 10 mM NaCl and 10 mM β-mercaptoethanol). The crude extract was passed through the column, which was then washed with 50 ml column buffer. MBP-NarL was eluted from the column in ~ 2 ml column buffer containing maltose (10 mM). Eluate was collected in 200 µl fractions and analysed using SDS-PAGE. Fractions containing MBP-NarL were pooled and concentrated using a VivaSpin column 500 MWCO 30,000 (GE Healthcare). The concentration of the purified protein was determined using Bio-Rad protein reagent and was stored at – 20°C until required. To release NarL from the fusion protein, NarL (~160 µM) was incubated with 3 units of enterokinase (BioLabs Inc.) at room temperature for ~ 18 h.

2.10.4. Estimation of protein concentration

2.10.4.1. Bradford assay

Protein concentration was estimated by the dye-binding method of Bradford (1976) with bovine serum albumin as the standard. The protein sample (10 µl) was mixed with 1 ml of the 1:5 diluted reagent (Bio-Rad), according to manufacturer's instructions. Absorbance was then measured at 595 nm.

2.10.4.2. Measuring protein concentration using A_{280}

To measure the protein concentration of purified FrmR, the scan function on a Varian Cary 50 Bio UV-Visible Spectrophotometer was used to record the protein spectrum between 200-400 nm, with absorbance at 280 and 320 nm being recorded. A total volume of 1 ml was used with a 1:10 ratio of protein to protein buffer (50 mM Tris pH 8.0 containing 0.5 M NaCl) in a Quartz cuvette. A blank containing protein buffer was used. To convert absorbance to protein concentration the equation below was used.

$$\text{Protein concetration (M)} = \frac{(\text{A}_{280\text{ nm}} - \text{A}_{320\text{ nm}})}{\text{Extinction coefficient}}$$

Molar extinction coefficient of FrmR (2980 $M^{-1}cm^{-1}$)

The extinction coefficient was determined using the ProtParam tool on ExPASy (Gasteiger *et al.*, 2005). For FrmR, the first methionine residue is cleaved upon translation, thus, the amino acid sequence entered excluded this.

2.10.5. *In vitro* protein phosphorylation

Purified proteins (NarL and ArcA (provided by Dr Matthew Rolfe, Department of Molecular Biology and Biotechnology, University of Sheffield, UK)) were phosphorylated *in vitro* as described by Bekker *et al*. (2010). Briefly, proteins were incubated in TEGD buffer (50 mM Tris (pH 7.5) containing 0.5 mM EDTA, 10% glycerol), to which 5 mM $MgCl_2$ and 100 mM (final concentration) carbamoyl phosphate were added at 30°C for 1 h. It was assumed that this procedure yielded 100% phosphorylation of the protein. In parallel, similar incubations with mixtures lacking carbamoyl phosphate were conducted to prepare unphosphorylated protein.

2.10.6. SDS-PAGE

SDS-PAGE was carried out according to Laemmli (1970), using Bio-Rad Mini Protean gel kits. A 15% resolving gel and 6% stacking gel were routinely used. The gel mixtures were made up as below, and a Tris-glycine SDS running buffer was used. Protein samples were diluted with dH_2O, where appropriate, to equalise protein concentrations. SDS sample buffer was added to a final concentration of 1X to samples, which were then denatured by boiling for 10 min. The gels were electrophoresed at 200 V per gel, until the dye front was judged to have travelled a sufficient distance. All Blue standards (Bio-Rad) were run alongside the samples.

SDS-PAGE gels

	Resolving gel (15 %)	Stacking gel (6%)
30% (w/v) acrylamide/bisacrylamide (37.5:1)	4.00 ml	0.75 ml
Tris-HCl, pH 8.8 (1.5 M)	2.00 ml	
Tris- HCl, pH 6.8 (1 M)		0.47 ml
SDS (10% w/v)	80 µl	37.5 µl
dH_2O	8.00 ml	2.46 ml
APS (10% w/v)	80 µl	37.5 µl
TEMED	8 µl	3.75 µl

SDS sample buffer

Tris-HCl, pH 6.8 (1 M)	1.88 ml

SDS	600 mg
Glycerol	3.00 ml
Bromophenol-blue	A tiny amount of powder
dH_2O	10.00 ml
β-mercaptoethanol	1.50 ml
Tris-glycine SDS running buffer	(L^{-1})
Tris	3.0 g
Glycine	14.4 g
SDS	1.0 g

Gels were stained using Coomassie brilliant blue. Staining was carried out for 30 min to 16 h, followed by thorough de-staining.

	Staining solution	Destaining solution
Coomassie brilliant blue R250	0.575 g	
Methanol	200 ml	400 ml
Acetic acid	50 ml	100 ml
dH_2O	250 ml	500 ml

2.10.7. Total amino acid analysis

Amino acid analysis was carried out by Alta Biosciences (University of Birmingham, UK) by complete acid hydrolysis for 24 h at 110°C of aliquots of FrmR that had previously been assayed for protein concentration (Section 2.10.5.2).

2.10.8. Estimation of the number of reactive sulphydryl groups

The number of reactive sulphydryl groups in purified protein was estimated using the protocol outlined by Thelander (1973). In a Quartz cuvette, a total reaction volume of 1 ml was prepared consisting of purified protein (100 µl), 700 µl 50 mM Tris containing 0.5 M NaCl (pH 8.0) and 200 µl DTNB solution (0.2 ml 0.4% (w/v) 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) in ethanol and 1.8 ml 1 M Tris-HCl (pH 8.0)). Absorbance at 412 nm was immediately measured using a Cary 50 Bio UV Vis Spectrophotometer that had been blanked with a sample of buffer and DTNB solution. The NTB concentration in the sample was calculated using the equations below to work out the number of reactive sulphydryl groups:

$$\text{NTB concetration (M)} = \frac{A_{412\ \text{nm}}}{\varepsilon l}$$

ε = Molar extinction coefficient of NTB (13,600 $M^{-1}cm^{-1}$)

l = Path length (cm)

$$\text{Reactive sulphydryl groups} = \frac{\text{NTB concetration (M)}}{\text{FrmR concetration (M)}}$$

2.10.9. Intrinsic Fluorescence Spectroscopy

In a fluorescence cuvette (1 ml), FrmR protein was diluted in protein elution buffer (50 mM Tris-HCl, pH 8.0 containing 0.5 M NaCl) to give a peak fluorescence intensity of at least 100 a.u. on a Cary Eclipse Fluorescence Spectrophotometer. Protein samples were maintained at 25°C for emission spectra (290-400 nm; excitation at 278 nm). FrmR was titrated with increasing zinc and manganese concentrations, typically to a 10-fold molar excess of [Zn(II)/Mn(II)]:[tetrameric FrmR]. The peak intensity of the fluorescence emission at 304 nm was measured and the value was corrected to account for dilution of the protein.

2.10.10. Liquid chromatography Mass Spectrometry (LC-MS)

Protein samples (~12 µM) were incubated with either an 8-fold or 40-fold molar excess of formaldehyde relative to FrmR tetramer concentration at room temperature. After 3 min, reactions were quenched with 10 mM glycine. Samples were loaded onto an Agilent 1260 Infinity liquid chromatograph fitted with an Agilent Extended C18 column (2.1 mm x 50 mm) and eluted with a gradient of 5-95% (v/v) acetonitrile in 0.1% (v/v) formic acid at 400 µl min^{-1} over 8 min. The eluent was directly coupled to an Agilent 6530 Q-ToF mass spectrometer fitted with an electrospray ionisation (ESI) source for determination of the mass of each peak.

2.10.11. Inductively couple plasma mass spectrometry (ICP-MS)

ICP-MS was carried out by Dr Jamie Young (Biomolecular Sciences Research Centre, Sheffield Hallam University, UK). The total metal ion content of FrmR was estimated by ICP-MS on samples of purified FrmR protein (200 µM) in 50 mM Tris (pH 8.0) containing 0.5 M NaCl buffer. The ion content of the buffer was also determined. FrmR or buffer was incubated with concentrated nitric acid (1:1 ratio) at 60°C for 1 h. Samples were cooled, diluted with dH_2O (up to 10 ml) and filtered. Samples were analysed on a Perkin Elmer Nexlon ICP-Ms system. Ions were full quantified in both the buffer and protein samples using an external calibration method. A dilution series of certified multi element reference standard (Sigma-Aldrich) was created and analysed via ICP-MS. Counts per second values for elements in the buffer and protein samples were then compared with the calibration curve to determine actual concentrations.

2.10.12. Western Immunoblotting

Pellets of chemostat culture (Section 2.7.2) were stored at -80°C until required. Defrosted samples were standardised with respect to their OD_{600} value by dilution with water and

separated on an SDS-PAGE gel. Wet transfers were performed using mini-transblot apparatus (Bio-Rad). The cassette was assembled as described in the manufacturer's instructions. All apparatus was pre-soaked with transfer buffer (for at least 15 min) and the blotting process was carried out in transfer buffer. Blotting was carried out at 100 V for 1 h, allowing transfer of the protein onto Hybond-C Extra membrane (Amersham). After blotting, the membrane was placed into 10 ml blocking solution (5% (w/v) Marvel in PBST) and incubated for 1 h with gentle shaking at room temperature. Membranes were then washed with 50 ml PBST for 1 x 15 min and 3 x 5 min washes. Primary anti-body was either diluted 1:10,000 (anti-ArcA antisera) or 1:1,000 (anti-TorA antisera; supplied by Dr Chantal Lobbi-Nivol, National Centre for Scientific Research, France) in 20 ml PBST, added to membrane and incubated at room temp for 1 h. Washes were repeated as above and followed by 1 h incubation at room temperature with a 1:100,000 secondary antibody (Amersham ECL™ Rabbit IgG, HRP-linked antibody) in 20 ml PBST. Secondary antibody containing solution was removed and washed as above, but using PBS for the last wash. ECL Plus™ Western blot detection reagents (GE Healthcare) were then placed over the membrane and incubated for 5 min at room temperature in the dark. The membrane was then and exposed to Hyperfilm-ECL for 5 – 30 s in a dark room. Hyperfilm was then developed and fixed in Kodak developing and fixing solutions and left to dry.

Transfer buffer	(L^{-1})
Tris	5.8 g
Glycine	2.7 g
SDS (10% w/v)	3.7 ml
Methanol	200 ml

PBST	(L^{-1})
NaCl	8.00 g
KCl	0.20 g
Na_2HPO_4	1.44 g
KH_2PO_4	0.24 g
Tween 20	0.05%
	pH 7.4

2.10.13. Phos-tag™-acrylamide gel electrophoresis

The degree of ArcA phosphorylation was measured using Phos-tag™-acrylamide gel electrophoresis and subsequent Western immunoblotting. The samples (Section 2.7.4) were

defrosted and diluted to 2.5 OD_{600} units ml^{-1} for running on a 37.5 µM Phos-tag™ (NARD Institute, Ltd., Amagasaki City, Japan) 10% acrylamide resolving gel in the following proportions:

Volume of cells in 1 M formic acid	56.4%
Volume of 3S SDS loading buffer	33.3%
Volume of 5 M NaOH	10.3%

Samples were then immediately resolved on a Phos-tag™ gel, with a 6% stacking gel, for 1 h at 200 V in SDS running buffer. Purified ArcA (provided by Dr Matthew Rolfe, Department of Molecular Biology and Biotechnology, University of Sheffield, UK), which had been phosphorylated with 100 mM carbamoyl phosphate (Section 2.10.5) and unphosphorylated ArcA were run as controls alongside. The Phos-tag™ gel was soaked in transfer buffer containing 1 mM EDTA for 15 min, and then pre-equilibrated for 15 min in transfer buffer (no EDTA). Standard wet transfer to Hybond-C membrane (Amersham) and immunodetection using rabbit ArcA antiserum and the ECL detection system (GE Healthcare) were carried out.

Phos-Tag™ resolving gel (10%)	
dH_2O	1.44 ml
30% (w/v) acrylamide (29:1)	1.36 ml
Tris-HCl, pH 8.8 (1.5 M)	1.08 ml
SDS (10% w/v)	40 µl
Phos-Tag™ acrylamide (5 mM)	20 µl
$MnCl_2$ (10 mM)	20 µl
APS (10% w/v)	50 µl
TEMED	20 µl

2.11. Protein-DNA interaction techniques

2.11.1. Bio-Layer Interferometry (BLItz)

Biotinylated-promoter DNA for *frm* (P_{frm}) and *ydhY* (P_{ydhY}) was amplified from *E. coli* MG1655 genomic DNA by PCR using oligonucleotides described in Appendix 1. Where indicated, FrmR tetramers were incubated with formaldehyde (up to an 800-fold molar excess) for 3 min (before quenching with 10 mM glycine) or Zn(II) (16-fold molar excess) at room temperature, prior to taking association measurements. Analysis of the interaction between purified FrmR and biotinylated-promoter DNA was carried out using the BLItz system (FortéBio), with all measurements taken at room temperature. Streptavidin (SA) biosensors (FortéBio) were hydrated by soaking the tips in 250 µl protein elution buffer (50 mM Tris, 0.5M NaCl (pH 8.0))

for at least 10 min. The measurements were made using the Advanced Kinetics function of the BLItz Pro 1.1.0.31 software. The sequence for each run was as follows:

1. Baseline step with protein elution buffer for 60 s
2. Loading step with 100 nM biotinylated P_{frm} DNA in buffer for 240 s
3. Baseline step with buffer for 60 s
4. Association step with various concentrations of purified FrmR in buffer for 240 s
5. Dissociation step with buffer (or formaldehyde-containing buffer) for 120 s

Baseline and dissociation steps were carried out by placing the biosensor in a black 500 µl Eppendorf tube filled with 250 µl buffer. Loading and association steps were carried out by placing the biosensor in the drop holder containing either 5 µl of P_{frm} DNA or FrmR protein depending on the step. For each subsequent run, a new biosensor was used. Analysis of the data was carried out using the BLItz Pro 1.1.0.31 software using global fitting and correcting both association and dissociation curves.

2.11.2. *In vitro* transcription (IVT) assays

The promoter and part of the coding regions of *frmR* and *ndh* (~200 bp upstream of the start codon to ~190 bp (*frmR*) or ~100 bp (*ndh*) into the gene) were amplified from *E. coli* MG1655 genomic DNA using appropriate oligonucleotides (Appendix 1). These DNA fragments (~0.1 pmol) were incubated for 30 min at 37°C in a 10.5 µl reaction volume containing 40 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 1 mM dithiothreitol, 75 mM KCl, 0.1 mM EDTA, 5% (v/v) glycerol, 250 µg ml^{-1} bovine serum albumin, 20 units of RiboLoc RNase inhibitor (Fermentas), 1 pmol *E. coli* RNA polymerase holoenzyme (New England BioLabs, Inc.) and 0 or 1 nM *Ec*FrmR tetramer. FrmR was reduced with 10 mM dithiothreitol and when required treated with 200-fold molar excess of formaldehyde for 5 min at room temperature, before quenching with 10 mM glycine. Transcription was initiated by the addition of 2 µl solution containing UTP at 50 µM; ATP, CTP and GTP at 1 mM; and 2.5 µCi of [α-^{32}P]UTP (800 Ci $mmol^{-1}$, PerkinElmer Life Sciences), followed by incubation for 15 min at 37°C. Reactions were terminated by the addition of 12.5 µl Stop/Loading dye solution (95% (v/v) formamide, 20 mM EDTA, pH 8.0, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol). Samples (10 µl) of each reaction were loaded onto a 6% (w/v) acrylamide, 1X TBE, 8 M urea gel, run at 20 mA buffered with 1X TBE and analysed using phosphorimaging (Section 2.11.2.1). Markers (0.1- 1 kb) were prepared using Perfect RNA Marker template mix (Novagen). A 20 µl reaction containing 0.75 µg of RNA template mix, 80 mM HEPES, pH 7.5, 12 mM $MgCl_2$, 10 mM NaCl, 10 mM dithiothreitol, 2 mM ATP, 2 mM GTP, 2 mM CTP, 0.1 mM UTP, 5 µCi of [α-^{32}P]UTP (800 Ci $mmol^{-1}$, PerkinElmer Life Sciences), 20 units of RiboLoc RNase inhibitor (Fermentas) and 50

units of T7 RNAP (Novagen), was incubated for 1 h at 37°C, before storing at -20°C. Markers from ~20 ng template were used for gel calibration.

6% acrylamide resolving gel containing 1X TBE, 8 M urea	
40% (w/v) Acrylamide	3 ml
TBE (10X)	1 ml
Urea	9.6 g
10% (w/v) APS	0.20 ml
TEMED	0.04 ml
dH_2O	up to final volume of 20 ml

10X TBE	(L^{-1})
Tris	108 g
Boric acid	55 g
0.5 M EDTA (pH 8.8)	40 ml

2.11.2.1. Phosphorimaging

In vitro transcription assays were visualized using phosphorimaging. The gel was exposed to a phosphor screen in a Hypercassette (Amersham) for ~16 h at 4°C. The phosphor screen was scanned using a Typhoon FLA 700 laser scanner (GE Healthcare) using the IP filter type (band pass of 390 nm). Scanned images were analysed and converted to .tiff images using ImageQuant TL 8.1 Software (GE Healthcare).

2.11.3. Electrophoretic Mobility Shift Assays (EMSAs)

LightShift Chemiluminescent EMSA Kit (Thermo Scientific) was carried out to investigate protein binding to biotin-labelled promoter DNA, following manufacturer's instructions. All reactions were carried out in a 20 µl volume, incubated on ice for 20 min, and contained; 20 fmol Biotin-labelled promoter DNA; 10X Binding buffer (10 mM Tris containing 50 mM KCl, 1mM DTT, pH 7.5); Poly-deoxy-inosinic-deoxy-cytidylic acid (poly [d(I-dC)]) (1 µg), where indicated; purified NarL (0 to 20, 000 nM). The biotin-labelled promoter DNA being studied and NarL were added in amounts described in Section 6.3). Samples were separated on a 6% (w/v) native-PAGE gel (below) for 1 h in 0.5X TBE (Section 2.11.2), followed by transfer onto Hybond-N+ nylon membrane (GE Healthcare) and crosslinking 120 mJ $cm^{2\ -1}$ using a commercial UV-light crosslinking instrument (Amersham, 80-6222-31) for 60 s. Visualisation of the binding reactions was carried out using the Nucleic Acid Detection Module (Thermo Scientific) as described in the manufacturer's instructions. The X-ray film was then developed and fixed in Kodak developing and fixing solutions and left to dry.

6% Native PAGE resolving gel	
30% (w/v) Acrylamide	4 ml
TBE (10X)	1 ml
dH_2O	14.85 ml
10% (w/v) APS	0.20 ml
TEMED	0.04 ml

2.12. Nuclear magnetic resonance (NMR) spectroscopy

Metabolite measurements using NMR were carried out by Dr Matthew Rolfe (Department of Molecular Biology and Biotechnology, University of Sheffield, UK) as previously described (Trotter *et al*., 2011). Cell-free supernatant fractions (Section 2.7.3) were analysed in a volume of 500 µl containing 450 µl supernatant sample, 50 µl deuterium oxide (D_2O) and 1 mM trimethylsilyl-propionate (TSP) as a standard using a Bruker DRX-500 spectrometer operating at 500 MHz, with samples analysed in 5 mm diameter tubes at 298 K. In order to reduce the H_2O signal, a pre-saturation for 2 s was applied during the recycle time. Carbon decoupling was applied during acquisition to suppress ^{13}C satellites. Topspin was used to process the spectra and quantify the peaks. Metabolite concentrations were established by reference to TSP. When required, peak identities were verified using spiking experiments with the corresponding metabolite at a known concentration. To measure ethanol loss from the chemostat, ethanol was added directly to the chemostat and Evans medium feed and then the remaining ethanol concentration in the chemostat after 20 h was measured using NMR.

2.13. Biochemical assays

2.13.1. Methyl viologen linked – TMAO reductase assay

Whole cell samples were taken from the chemostat in a final concentration of 50 mg ml^{-1} chloramphenicol in 50 ml aliquots and were pelleted by centrifugation at 4,020 *g* for 10 min. Cell pellets were defrosted on ice, washed in 25 ml sodium phosphate (25 mM) buffer pH 7.0, and resuspended in a final volume of 10 ml buffer. OD_{600} measurements were made for each sample. An Agilent CARY 50 UV-Visible, pre-warmed to 37°C, was used for measurements. To a screw septa capped cuvette, 960 µl sodium phosphate buffer, 20 µl methyl viologen (10 mM) and 10 µl cell suspension were added. The spectrophotometer was blanked with the cuvette, before the assay solution was sparged with argon for 6 min. Around 3 µl of dithionite solution (made by adding one small spatula of sodium dithionite per 5 ml dH_2O) was added to the cuvette though the septa with a Hamilton syringe and inverted. Once the absorbance was constant for about 60 s, a final concentration of 5 mM

TMAO was added to the cuvette, which was quickly inverted twice before methyl viologen oxidation was measured at 585 nm for 300 s. Assays were carried out in duplicate for each sample.

100 mM sodium phosphate buffer (pH 7.0)	(L^{-1})
Na_2HPO_4 (1 M)	57.7 ml
NaH_2PO_4 (1 M)	42.3 ml

To make 25 mM sodium phosphate buffer, the 100 mM sodium phosphate buffer stock was diluted 1 in 4 with dH_2O and pH adjusted to 7.0.

2.13.2. Copper (II) reduction assays

Cultures of *E. coli* MG1655 were grown anaerobically overnight at 37°C in Evans medium (pH 7.0), supplemented with 20 mM glucose and 40 mM TMAO. Cells were pelleted and washed three times in anaerobic Evans medium, pH 7.0. Cells were resuspended in Evans medium plus 20 mM glucose to an OD_{600} ~25. The assays were performed as previously described (Volentini *et al*., 2011) in an anaerobic workstation. Briefly, 100 µl of washed cells were incubated with 0.05 mM $CuSO_4$, 40 mM TMAO (when appropriate) in a total volume of 1 ml citrate-phosphate buffer, pH 7.0, at 37°C. Control assays with no cells and no $CuSO_4$ were carried out in parallel. At the indicated times, the assays were stopped by centrifuging samples for 3 min at 8,000 *g*. The resulting supernatant had 0.1 mM bathocuproine disulfonate (Sigma-Aldrich) added. The absorbance at 412 nm for the supernatant was measured and the molar coefficient for the Cu (I) complex (12,220 M^{-1} cm^{-1}) used to determine Cu (II) reduction by samples per A_{600} unit.

Citrate-phosphate buffer	
Citric Acid (0.1 M)	13 ml
Na_2HPO_4 (0.2 M)	87 ml
Adjust volume to 200 ml with dH_2O, pH to 7.0	

2.13.3. Measurement of formaldehyde production

Formaldehyde was measured using acetylacetone reagent (Nash, 1953). Anaerobic cultures of *E. coli* MG1655 were grown overnight with stirring at 37°C in LB supplemented with 0.2% (w/v) glucose. Preparation of cell samples and setting up of assays were performed in an anaerobic workstation. Cells were pelleted and washed in 25 mM sodium phosphate buffer pH 7.0 and finally resuspended in a final volume of 2 ml anaerobic cell breakage buffer. An oxygen scavenger, which would convert any oxygen to water, 1 ml, was added to the resuspended cells before sonication (6 x 15 s, 15 micron amplitude). Any unbroken cells

were removed by centrifugation at 4,020 *g* for 1 min. Cell-free extracts were added to a 1 ml total volume of 25 mM sodium phosphate buffer and 40 mM TMAO (in TMAO-containing samples only). Assays were incubated anaerobically at 37°C and at the indicated times were stopped by precipitation using 5% (w/v) trichloroacetic acid and centrifugation at 20,000 *g* for 2 min. Resulting supernatants were treated as previously described (Fu *et al.*, 2006). Briefly, the stopped reaction was combined with the acetylacetone reagent in a 1:1 ratio and incubated at 60°C for 15 min. The OD_{412} of the reaction was recorded. Formaldehyde content was calculated using the extinction coefficient at 412 nm (8,000 M^{-1} cm^{-1}). Protein content of the cell-free extracts was determined using the Bio-Rad assay.

Cell breakage buffer	
Tris	20 mM
NaCl	500 mM
Glycerol	5% (w/v)

Adjust to required volume with dH_2O, pH to 7.5.

Oxygen scavenger	
Glucose	10% (w/v)
Catalase	10 µg ml^{-1}
Glucose oxidase	5 µg ml^{-1}

Acetylacetone reagent (Nash reagent)	(L^{-1})
Ammonium acetate	150 g
Acetic acid	3 ml
Acetylacetone	2 ml
	Filter sterilise

2.14. Fluorescence Microscopy and Flow Cytometry

2.14.1. Cell preparation for microscopy and flow cytometry

Escherichia coli JRG6705 cells were grown anaerobically at 37°C in Evans medium buffered with phosphate buffer. After overnight growth, TMAO (2 – 40 mM) was then added to the culture, in an anaerobic workstation. After the addition of TMAO, samples were incubated for 1 h at 37°C. After incubation, cells were harvested by centrifugation at 4,020 *g* for 10 min and washed in sterile dH_2O, with the resulting cell pellet being fixed. Cells were also grown in continuous culture in a 1 L culture vessel under anaerobic conditions as previously described (Section 2.1.5.1) and washed as described above.

2.14.2. Fixing of cells with paraformaldehyde

Cell pellets were resuspended in fixative for 30 min at room temperature. Fixed cells were washed at least twice in sterile PBS by centrifugation at 4,020 *g* for 3 min at room temperature. Samples were resuspended in PBS to give a suitable concentration of cells and were immediately fixed to microscope slides for imaging. Alternatively, for flow cytometry analysis, fixed samples were stored at 4°C in the dark for a maximum of 24 h before analysis.

Preparation of 16% (w/v) paraformaldehyde:

100 mM sodium phosphate buffer (pH 7)	50 ml
Paraformaldehyde	8.0 g

To prepare the paraformaldehyde, the solution was heated to 60°C in a water bath and mixed vigorously. NaOH (≥5 M) solution was added drop wise, until the solution cleared and was then stored at 4°C.

Fixative

Paraformaldehyde, 16% (w/v)	0.5 ml
100 mM sodium phosphate buffer (pH 7)	2.00 ml

Phosphate Buffered Saline (PBS)

NaCl	8.00 g
KCl	0.20 g
Na_2HPO_4	1.44 g
KH_2PO_4	0.24 g

The volume was adjusted to 1 L with dH_2O and pH adjusted to 7.4.

2.14.3. Slide preparation for light microscopy

Poly-Prep Slides (Sigma) were wiped with 70% (v/v) ethanol and left to air dry. An appropriate sample dilution was prepared in dH_2O, and 5 µl spread gently onto the poly-L-lysine coated slide and left to air-dry. A coverslip was then mounted with 5 µl dH_2O, and sealed by spotting colourless nail varnish at each corner.

2.14.4. Fluorescence microscopy

Fluorescence microscopy was carried out using a DeltaVision deconvolution microscope (Applied precision) equipped appropriate filters for imaging various fluorophores. To image GFP the FITC (Fluorescein isothiocyanate) filter set was used (details shown below). The objective lens used was 100X (with immersion oil). Images were deconvolved using SoftWoRx suite v.3.5.1. Deconvolved fluorescence images were processed using ImageJ version 1.48v.

Filter set	Excitation filter/bandpass (mm)	Emission filter/bandpass(nm)
FITC	490/20	528/38

2.14.5. Flow cytometry

Flow cytometric analysis was performed using an Attune Autosampler (Life Technologies Ltd., Paisley, UK). Typically, the fixed culture (Section 2.4.2) was diluted 10-fold with PBS in order to obtain the correct concentration of cells to analyse. The instrument settings were chosen using a GFP-negative *E. coli* JRG6705 control sample, which had no TMAO added. Instrument settings were kept the same for each sample. A GFP-negative sample was run for each flow cytometry experiment in order to correctly gate for GFP-positive cells. Routinely, 10,000 events were recorded for each analysis. The BL1 detector was used for detecting GFP (details shown below). Flow cytometry data was analysed using FlowJo V10 (Tree Star Inc., USA).

Detector	Filter (center/bandpass) (mm)	Filter range
BL1	530/30	515-545

3. Adaptation of anaerobic cultures of *Escherichia coli* K-12 in response to trimethylamine-*N*-oxide

3.1. Introduction

Trimethylamine-*N*-oxide (TMAO), a common metabolite and an osmolyte in marine organisms, is often found in many anaerobic environments. Consequently several enteric bacteria, including *E. coli* have evolved to use TMAO as an alternative electron acceptor (Barrett and Kwan, 1985). As a metabolically versatile enteric bacterium, *E. coli* is able to grow in both the presence and absence of oxygen. It has three major metabolic modes, including aerobic respiration, anaerobic respiration and fermentation (Bock and Sawers, 1996; Gennis and Stewart, 1996; Guest *et al*., 1996). Anaerobic respiration with TMAO as the terminal electron acceptor is more energetically efficient than fermentation, but less than aerobic respiration (Guest *et al*., 1996).

As explained in Chapter 1, there is a large body of information concerning both the regulation and operation of the *E. coli* K-12 respiratory chain. However, there is little or no understanding of the dynamic adaptive processes that occur during the transition from anaerobic respiration to TMAO respiratory growth. In order to address this issue, glucose-limited chemostat cultures have been used to study the effects of perturbing anaerobic fermentative *E. coli* K-12 cultures by addition of the terminal electron acceptor TMAO. Transcript profiling, metabolite and biochemical measurements, combined with probabilistic modelling of transcription factor activities, have all been undertaken in order to obtain a deeper understanding of the adaptive process.

3.2. Analysis of over-metabolite production by anaerobic cultures of *E. coli* K-12 in the absence and presence of TMAO

In order to study the process of *E. coli* K-12 adaptation to the presence of TMAO, glucose-limited (20 mM) anaerobic fermentative steady state chemostat cultures of *E. coli* K-12 MG1655 were established (Section 2.3.1). Analysis of over-metabolite production by anaerobic chemostat cultures (dilution rate 0.2 h^{-1}) of *E. coli* MG1655 in the absence and presence of TMAO were carried out using NMR (Section 2.12).

Glucose-limited anaerobic fermentative growth of *E. coli* under these conditions can be described using the equations (Equations **1** and **2**) below:

$C_6H_{12}O_6 + H_2O \rightarrow CH_3COOH + CH_3CH_2OH + 2\ HCOOH$ **(1)**

$HCOOH \rightarrow CO_2 + H_2$ **(2)**

In addition to this, other minor fermentation products (over-metabolites) such as succinate and lactate are produced (Table 3.1). Measurement of over-metabolites present in the chemostat culture medium during the fermentative steady state showed the presence of acetate (13.7 ± 0.6 mM), ethanol (10.9 ± 0.2 mM), formate (13.6 ± 2.7 mM), succinate (3.9 ± 0.2 mM), fumarate (0.03 ± 0.01 mM), lactate (0.08 ± 0.01 mM) and orotate (0.16 ± 0.01 mM), with the concentrations measured being consistent with anaerobic fermentative growth, where ~40% of the formate produced was converted to CO_2 and H_2 by formate hydrogen-lyase (Table 3.1, Equations **1** and **2**).

The anaerobic fermentative steady-state was perturbed by the introduction of 45.8 + 1.1 mM TMAO both directly to the chemostat culture and the growth medium feed. Equation **3** describes anaerobic respiratory growth on glucose with TMAO as the electron acceptor:

$C_6H_{12}O_6 + 4\ (CH_3)_3NO \rightarrow 2\ CH_3COOH + 4\ (CH_3)_3NH_2 + 2\ CO_2 + 2\ H_2O$ **(3)**

However, as expected, introduction of 45.8 ± 1.1 mM TMAO to the culture, resulted in a mixed anaerobic TMAO respiratory and fermentative metabolism (TMAO-respiratory/fermentative growth) (Equation **4**).

$C_6H_{12}O_6 + 2.4\ (CH_3)_3NO \rightarrow 1.6\ CH_3COOH + 0.4\ CH_3CH_2OH + 2.4\ (CH_3)_3NH_2 + 0.5\ HCOOH + 1.5\ CO_2 + 0.76\ H_2O$ **(4)**

A remodelling of metabolism was seen during the first 60 min after TMAO was added to the chemostat. The concentration of fumarate increased (0.08 ± 0.03 mM), whereas succinate (3.10 ± 0.20 mM) and formate (11.00 ± 3.00 mM) decreased (Table 3.1). During this period, the concentration of TMAO concomitantly decreased from 45.8 ± 1.1 mM to 40.4 ± 2.30 mM.

The new TMAO-respiratory/fermentative steady-state was established after 1440 min, where TMAO was undetectable and TMA, the product of TMAO respiration, reached a concentration of 45.8 ± 1.4 mM. Higher production of acetate (24.5 ± 1.2 mM) and orotate (0.37 ± 0.02 mM), but less succinate (1.5 ± 0.6 mM) and ethanol (2.3 ± 0.4 mM) was detected in the external milieu of the new steady-state.

During the transition phase, physiological changes accompanied the metabolic changes described above, with an increase in biomass being observed. The biomass supported by the TMAO-respiratory/fermentative steady state was ~2-fold greater (0.61 g L^{-1}) than that under fermentative conditions (0.29 g L^{-1}). The Y_{ATP} value for *E. coli* growing under anaerobic conditions at a dilution rate of 0.2 h is ~5 g cell dry weight (cdw) per mol of ATP (Hempfling and Mainzer, 1975). The cdw of the initial fermentative steady state is consistent with this

Table 3.1. Concentrations of extra-cellular metabolites produced by *E. coli* MG1655 during transition from anaerobic fermentative growth to TMAO anaerobic respiratory growth measured using NMR

Time after addition of TMAO (min)	TMAO (mM)	TMA (mM)	DMA (mM)	Acetate (mM)	Formate (mM)	Ethanol (mM)[a]	Succinate (mM)	Lactate (mM)	Fumarate (mM)	Orotate (mM)	Biomass (mg cell dry weight mL^{-1})
0	ND	ND	ND	13.7±0.6	13.6±2.7	10.9±0.2	3.9±0.2	0.08±0.01	0.03±0.01	0.16±0.01	0.29±0.07
2	45.8±1.1	0.4±0.1	0.08±0.01	13.6±0.7	11.0±2.8	10.7±0.1	3.7±0.2	0.05±0.01	0.02±0.01	0.13±0.02	
5	46.0.5±1.2	0.5±0.1	0.09±0.01	13.7±0.7	10.9±2.8	10.7±0.1	3.7±0.2	0.06±0.01	0.03±0.01	0.11±0.01	
10	45.0±1.2	0.8±0.1	0.10±0.01	13.5±0.8	10.6±2.8	10.4±0.1	3.6±0.2	0.05±0.01	0.03±0.01	0.12±0.01	
15	44.5±1.3	1.1±0.1	0.12±0.02	13.4±0.9	10.6±3.0	10.2±0.1	3.6±0.2	0.05±0.01	0.04±0.01	0.13±0.02	
20	43.2±2.0	1.4±0.1	0.12±0.02	13.2±1.0	10.5±3.0	10.0±0.2	3.4±0.2	0.06±0.01	0.05±0.01	0.13±0.02	
60	40.4±2.3	4.2±0.5	0.23±0.01	13.6±1.0	11.0±3.0	10.7±1.1	3.1±0.2	0.07±0.01	0.08±0.03	0.12±0.04	0.44±0.08
1440	ND	45.8±1.4	0.15±0.03	24.5±1.2	6.7±1.4	2.3±0.4	1.5±0.6	ND	0.05±0.01	0.37±0.02	0.61±0.04

a. The values reported for ethanol are corrected to account for losses to the gas phase

Data are the mean values ± standard deviation (*n* = 3). ND = not detected.

Y_{ATP}, assuming that glucose fermentation yields 3 ATP per glucose molecule (60 mmol ATP from 20 mmol glucose yielding 0.29 g cdw, which is equivalent to Y_{ATP} = 4.8 g cdw per mol of ATP) (Table 3.1). Addition of TMAO to the chemostat, results in an increase in biomass (an additional 0.32 g L^{-1}). The dry weight of an *E. coli* cell is 0.28 pg (Neidhardt *et al*., 1990). Thus, the increase in biomass was equivalent to an additional ~1.2 x 10^{12} bacteria in the TMAO-respiratory/fermentative steady state. A contribution of 6 H^+ per oxidised NADH molecule to the proton motive force is made from an electron transport chain, consisting of NADH dehydrogenase I and TMAO reductase (Keseler *et al*., 2013). Per TMAO molecule, 1.8 ATP molecules are produced with a H^+/ATP ratio of 3.3 (Futai *et al*., 2012), which equates to the production of 5 x 10^{22} molecules of ATP in the TMAO-respiratory/fermentative steady state. The synthesis of an *E. coli* cell requires ~40 x 10^9 ATP molecules; thus, the additional ATP gained from respiration with TMAO as the terminal electron acceptor accounts for the synthesis of an additional ~1.3 x10^{12} cells (0.364 g). Therefore, the metabolic map shown in Figure 3.1 accurately captures the physiological state of the TMAO-respiratory/fermentative steady state.

3.3. The transcriptional response of *E. coli* K-12 during adaptation to TMAO respiratory conditions

Transcript profiles were obtained from the *E.coli* K-12 chemostat cultures under anaerobic fermentative conditions and then 2, 5, 10, 15, 20 and 60 min after the addition of the electron acceptor TMAO; then again at 1440 min after the new TMAO-respiratory/fermentative steady state was established (Section 2.7.1). If at one or more of the time-points, the abundance of a transcript was ≥3-fold greater or smaller after the addition of TMAO when compared to the initial fermentative steady state, then an operon was deemed to be significantly regulated (Table 3.2). Using the ≥3-fold cut-off, 34 operons were judged to be significantly regulated in response to TMAO, with 29 operons being upregulated and 5 being downregulated.

The transcripts that were upregulated in response to the presence of TMAO could be classified into six groups based on function (Table 3.2). The transcription of the *torCAD* operon (encodes TMAO reductase) was induced within 2 min of the addition of TMAO to the culture (Table 3.2, group A). This upregulation suggests that the presence of the electron acceptor had been detected and a transcriptional response had been initiated within 2 min. Concomitantly, within 2 min TMA (0.4 $\pm$ 0.1 mM) was detected in the medium by NMR (Table 3.1), which also indicates that TMAO reduction had been initiated. The abundance of the *torCAD* transcript was maximal (11.6-fold greater than in the fermentative steady state) 10

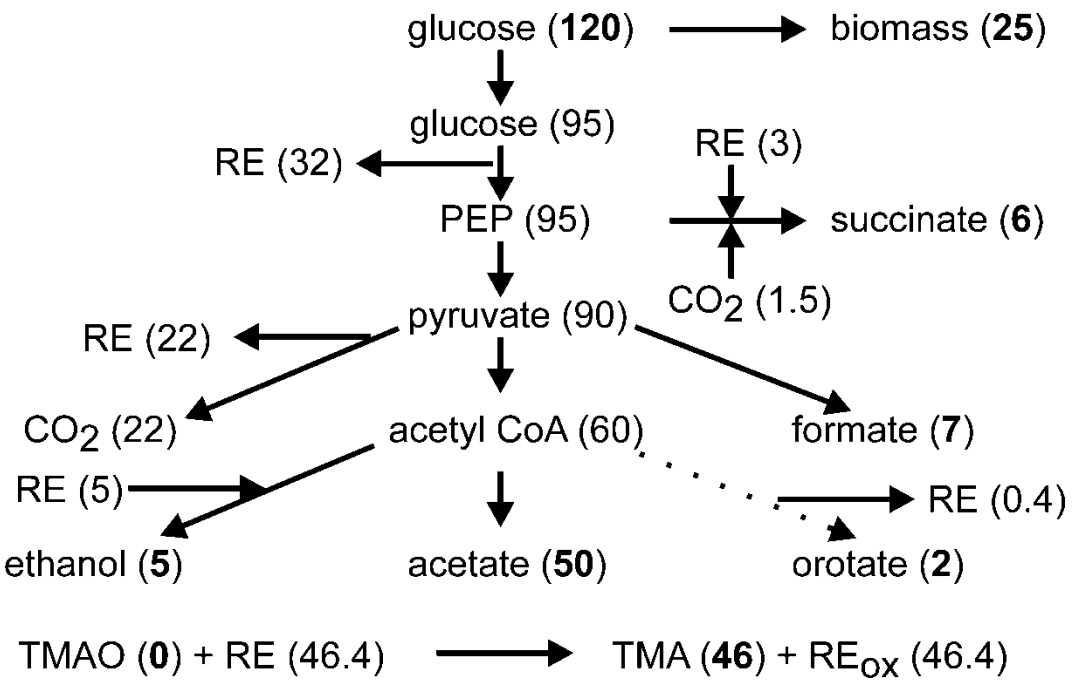


Figure 3.1. **Metabolic map of the TMAO-respiratory/fermentative steady state**

Glucose provides **120** atoms of carbon to the system, **25** of which are converted into biomass based on ~50% of cdw is carbon (Heldal *et al*., 1985). Therefore, 95 carbon atoms are able to enter glycolysis, producing phosphoenolpyruvate (PEP) and in the process of doing so generating 32 reducing equivalents (RE). CO_2 reacts with a small amount of PEP and is converted to succinate (**6** carbon atoms) (Table 3.1), concomitantly re-oxidizing three RE (Table 3.1). Pyruvate formate-lyase then converts the 26% of the PEP to pyruvate and the remainder (74%) of the PEP (90 carbon atoms in total) is converted to pyruvate also, generating 22 RE. Acetyl CoA (60 carbon atoms) is produced and this is converted to acetate (50 carbon atoms) along with small amounts of ethanol (**5** carbon atoms; re-oxidizing five RE) and orotate (**2** carbon atoms; generating 0.4 RE) (Table 3.1). In this scheme, reduction of TMAO to **46** molecular equivalents of TMA results in a total of 46.4 RE being created and re-oxidized (REox) (Table 3.1). Bold numbers in parentheses are measured quantities as shown in Table 3.1 and those shown in normal font are estimates.

Table 3.2. Transcripts that are altered in abundance by ≥3-fold in response to the addition of TMAO to anaerobic fermentative steady-state cultures of *E. coli* K-12 MG1655

Operon[a]	Product	Fold change in abundance relative to the anaerobic fermentation steady state[b]							Regulatory proteins[c]
		2 min	5 min	10 min	15 min	20 min	60 min	TMAO respiratory steady state	
	Predominantly up-regulated operons								
	(A) Respiration								
cyoA-E	Cytochrome *bo*' terminal oxidase	2.5	3.8	6.4	3.3	3.2	2.2	1.1	ArcA (-), Cra (-), CRP (+), FNR (-), Fur (-), GadE (+), PdhR (-)
torCAD	Trimethylamine-N-oxide reductase	4.4	6.1	11.6	7.9	7.6	7.6	3.5	NarL (-) TorR (+)
	(B) Carbon metabolism								
acs-yjcH-actP	Acetyl-CoA synthetase, inner membrane protein, acetate transporter	1.7	3.3	6.1	3.1	1.8	1.0	1.2	CRP (+), Fis (-), IHF (-)
frmRAB	Transcription regulator, formaldehyde dehydrogenase, *S*-formylglutathione hydrolase	2.4	4.6	4.9	6.1	10.9	9.5	1.3	FrmR (-)

glpABC	Anaerobic glycerol-3-phosphate dehydrogenase	2.1	3.5	2.1	1.3	-1.2	-1.4	1.0	ArcA (-), CRP (+), Fis (+), FlhDC (+), FNR (+), GlpR (-)
glpD	Aerobic glycerol-3-phosphate dehydrogenase	2.9	4.6	2.8	1.6	1.2	1.1	1.2	ArcA (-), CRP (+), GlpR (-)
glpFK	Glycerol channel and glycerol kinase	2.2	3.8	3.4	1.2	-1.3	-1.3	1.3	CRP (+), GlpR (-)
pdhR-aceEF-lpd	PdhR regulator, pyruvate dehydrogense complex	1.6	1.7	2.3	3.1	2.7	1.6	1.0	ArcA (-), Cra (-), CRP (+), Fis (+), FNR (+/-), Fur (+/-), NsrR (-), PdhR (-)
sdhCDAB-sucA-D	Succinate dehydrogenase, 2-oxo-glutarate dehydrogenase	2.3	3.6	3.6	2.4	1.6	1.3	1.2	ArcA (+/-), CRP (+), FNR (-), Fur (+)
	(C) Methionine metabolism								
metBL	*O*-Succinylhomoserine lyase, aspartate kinase	1.2	-1.3	1.2	2.4	5.6	1.2	1.2	MetJ (-), PhoP (+), YjiE (+)

metF	5, 10-Methylenetetrahydrofolate reductase	-1.1	-1.5	-1.2	1.5	3.8	1.1	1.1	MetJ (-)
metR	Methionine regulator	1.0	-1.3	-1.2	1.4	3.2	1.1	1.0	MetR (-)
mmuPM	*S*-Methyl-L-methionine transport	1.1	1.1	1.1	1.9	3.5	1.1	1.1	
sbp	Sulfate-binding proten	-1.1	1.0	1.0	1.8	3.6	1.7	-1.2	
ybdL	Methionine aminotransferase	1.1	-1.6	1.0	1.8	4.0	1.0	1.1	
	(D) Transport proteins								
cstA	Peptide transport	2.2	4.0	3.6	1.4	1.1	1.1	1.6	CRP (+)
dctA	C4-dicarboxylate/orotate:proton symport	2.1	3.8	4.3	1.8	1.6	1.3	1.6	ArcA (-), CRP (+), DcuR (+)
mglBAC	Galactose transport	2.1	2.5	3.7	1.1	1.0	-1.2	2.0	Crp (+), GalR (-), GalS (-)

ugpBAECQ	Glycerol-3-phosphate transport	1.8	3.3	2.4	1.5	1.2	1.2	1.2	PhoB (+)
	(E) Motility and chemotaxis								
*flgA**M**N*	Flagella biosynthesis	1.0	1.1	-1.1	1.1	1.1	-1.1	3.0	FlhDC (+)
flgB-J	Flagella proteins	-1.1	-1.2	-1.1	-1.2	-1.1	1.0	3.3	FlhDC (+)
fliAZY	Flagella regulation and biosynthesis	-1.1	-1.1	-1.3	-1.1	-1.1	1.1	4.0	FlhDC (+), H-NS (+), MarA (-), NsrR (-)
*motB-che**A**W*	Flagella and chemotaxis proteins	1.1	-1.1	1.0	1.1	1.0	1.0	3.2	CpxR (-)
tar-tap-cheRBYZ	Chemotaxis proteins	1.0	1.3	1.5	1.1	-1.1	1.0	3.1	FNR (+)
	(F) Others								
betIBA	Betaine aldehyde dehydrogenase and regulatory protein	1.8	3.3	2.5	1.9	1.6	1.4	1.0	ArcA (-), BetI (-)
bsmA	Conserved protein	1.7	3.0	1.4	1.0	-1.2	1.0	1.0	
glcC	Glycolate regulator	1.8	3.9	2.3	1.4	1.1	1.0	1.0	Cra (-), CRP (+), Fis (-), GlcC, (-)

putA	Fused transcriptional repressor-proline dehydrogenase	1.6	2.8	5.5	2.2	2.2	1.7	1.2	BasR (-), MarA (+), PutA (-)
ybdH	Predicted oxidoreductase	1.3	1.2	1.3	1.9	3.0	1.1	1.2	
	Predominantly down-regulated operons								
borD	Predicted lipoprotein	-1.8	-3.0	-1.4	1.5	1.8	1.2	1.1	PhoP (+)
cusCFBA	Copper efflux	-1.1	-2.2	-3.3	-2.5	-3.0	-2.9	1.0	CusR (+)
manXYZ	Mannose PTS permease	-1.8	-1.4	-3.7	-5.0	-3.0	-2.3	-1.1	CRP (+), Cra (-), DgsA (-), NagC (-)
ompF	Outer membrane protein	1.0	-2.4	-3.1	-2.4	-1.3	1.1	1.7	CpxR (-), CRP (+), EnvY (+), Fur (+), IHF (+/-), OmpR (+/-), RstA (-)
thrABC	Threonine and homoserine biosynthesis	-1.6	-1.6	-2.4	-3.2	-2.3	1.0	-1.5	DksA (+)

[a]Unless indicated in bold type the data shown are for the first gene in the operon.

[b]Numbers are fold increase or decrease (by at least 3-fold, $p \leq 0.05$, at one or more sampling points) in transcript abundance after introduction of TMAO.

[c]Regulatory proteins are indicated (-) negative regulation, (+) positive regulation, (+/-) dual regulation.

PTS, phosphotransferase system

min after initial exposure of the culture to TMAO. The increase in abundance of the transcript was maintained across the transition, but declined to be only 3.5-fold greater than that of the initial fermentative steady state in the new TMAO respiratory/fermentative steady state. With the exception of the motility and chemotaxis transcripts (Table 3.2, group E), the *torCAD* operon was the only transcript to exhibit >3-fold increased abundance in the final TMAO-respiratory/ fermentative steady state. Thus, during the transition there is a high expression phase of capacity building for TMAO reduction during initial adaptation, this was followed by decreased expression in the new TMAO-respiratory/fermentative steady state to a level commensurate with the growth rate of the culture. Surprisingly, the transcript encoding a low oxygen affinity terminal oxidase (cytochrome *bo'*; Cyo) was strongly induced during the first 10 min of the transition before returning to its initial level of abundance, despite anaerobic conditions being maintained throughout (Table 3.2, group A). However, there was an absence of the broad transcriptional response that is observed upon transfer of *E. coli* cultures to aerobic growth conditions (Partridge *et al*., 2006; Rolfe *et al*., 2012; Trotter *et al*., 2011), which indicates that unintended introduction of oxygen to the chemostat was not the cause of the increased abundance of the *cyo* transcript.

There were seven upregulated operons that were involved in carbon metabolism (Table 3.2, group B). These operons returned to their initial abundances in the TMAO respiratory/fermentative steady state, but, were all transiently expressed during the transition to TMAO-respiratory growth, like the *cyoA-E* operon (Table 3.2, group B). This group of operons includes transcripts encoding key components of central metabolism: pyruvate dehydrogenase, succinate dehydrogenase and 2-oxo-glutarate dehydrogenase indicating that the presence of the electron acceptor TMAO invokes a shift towards enhanced substrate oxidation. Also upregulated were transcripts associated with glycerol and acetate acquisition and metabolism, including both the anaerobic and aerobic glycerol-3-phosphate dehydrogenase operons and the glycerol-3-phosphate transport system (*ugp* operon, assigned to group D). The establishment of respiratory chains with succinate and glycerol-3-phosphate as the electron donors in the presence of TMAO is suggested by the activation of the anaerobic and aerobic glycerol-3-phosphate dehydrogenases and succinate dehydrogenase (complex II). The decrease in extracellular succinate concentration and increase in fumarate concentration after addition of TMAO was consistent with the induction of complex II. However, extracellular glycerol was not detected by NMR (Table 3.1). Osmotic re-balancing in the 5 min period after addition of the osmolyte TMAO to the culture was suggested by the activation of the *betIBA* operon (glycine-betaine synthesis from choline).

Transcripts which encode a regulator and enzymes involved in formaldehyde detoxification (*frmRAB* operon) exhibited a more sustained (>60 min after TMAO addition) expression profile across the transition (Table 3.2, group B).

Group C is formed of transcripts involved in methionine metabolism, which were transiently upregulated 15-20 min after the addition of TMAO, before returning to pre-TMAO addition level (Table 3.2, group C). MetJ is the regulator of three of the six transiently induced operons in group C, which suggests that the supply of *S*-adenosylmethionine, a molecule used in methyl transfer reactions, is restricted during this period of enhanced biomass production. Group D contained peptide, C_4-dicarboxylic acid/orotate, galactose and glycerol-3-phosphate transport systems (Table 3.2, group D).

In the new TMAO respiratory/fermentative steady state, enhanced abundance of the flagella and chemotaxis transcripts were observed and these form group E. The increased abundance of these transcripts was only seen in the final TMAO-respiratory/ fermentative steady state (Table 3.2, group E).

Group F consisted of transcripts that were transiently induced, including transcripts encoding proteins involved in osmotic homeostasis (*betIBA* operon), regulatory proteins and proteins of unknown function.

In response to addition of TMAO, five operons were transiently downregulated ≥3-fold, these operons encode proteins of various biological functions. The downregulated operons have roles in copper resistance, threonine biosynthesis, mannose transport and outer membrane function (Table 3.2, group G). During the initial phases of adaptation, a decrease in the *cus* operon transcripts was observed. This operon encodes an anaerobic copper detoxification system, which suggests that any copper stress experienced by the culture during this phase of the transition was diminished.

3.4. Using TFInfer to simultaneous infer the activities of transcription factors during the switch of metabolic mode

Complex interactions between signalling molecules and transcription factors (TFs) have resulted in changes in gene expression, substrate utilisation and over-metabolite production (Sections 3.2 and 3.3), during the adaptation of *E. coli* K-12 to the presence of TMAO. During the transition, there were widespread changes that are components of the transcriptional response that sit below the statistical filter used (≥3-fold (t-test, p=0.05)) to deem an operon significantly regulated. Therefore, in order to capture the full breadth of the transcriptional

response when the fermentative culture was perturbed by the addition of TMAO, open access TFInfer software (Asif *et al*., 2010) was used to simultaneously infer the activities of 167 TFs that are recorded as able to directly regulate target genes in the EcoCyc database (Keseler *et al*., 2013) from the transcript profiles. TFInfer is based on a state space model that employs a linear approximation (in log space) to the dynamics of transcription and treats noise in a way such that the estimated TF activities are associated with confidence limits (Sanguinetti *et al*., 2006). The TFInfer model has been previously used in several studies (Partridge *et al*., 2007; Rolfe *et al*., 2012). During this analysis here, TFs that exhibited a signal to noise ratio of >2 were deemed to have activities significantly altered following the introduction of TMAO (Figure 3.2). From the analysis, the inferred activities of only 20 TFs could account for many of the changes in transcript abundance during the transition (Table 3.2), as one or more TFs exhibiting altered activity are known to regulate transcription from the corresponding promoters, thus potentially regulating the expression of a number of operons (Table 3.2). A visual representation of the regulatory events (as inferred by TFInfer) during the time course is shown in Figure 3.3.

When the TMAO-respiratory steady state was compared to the initial fermentative steady state, there were only two TFs (FlhDC and TorR) that displayed >2-fold enhanced activity (Figure 3.2). The inferred activity of TorR in the TMAO respiratory/fermentative steady state which was ~6-fold lower (0.51) than the maximal (3.16), which occurred 10 min after introduction of TMAO, is sufficient to maintain an appropriate concentration of TMAO reductase in the new steady state. It is unclear why there is enhanced FlhDC and thus increased expression of motility and chemotaxis genes.

During the first 60 min of the transition, four TFs (ArcA, CusR, CsgD and PutA) showed lower activity before returning to their original levels in the TMAO-respiratory/fermentative steady state (Figure 3.2). In order to validate the inferred activities of ArcA, the amounts of phosphorylated ArcA (ArcA~P) present in the bacteria were measured (Section 2.10.13) (Figure 3.4). In order to do this, ArcA and ArcA~P were separated by Phos-tag gel electrophoresis (Wako Pure Chemical Industries Ltd.) and using anti-ArcA serum were detected via Western blotting (Section 2.10.12). The proportion of ArcA~P in the samples was then determined using quantitative densitometry. This analysis showed that the ArcA~P decreased to 0% in the 2–20 min period after TMAO addition from the expected ~50% level in the fermentative steady state (Rolfe *et al*., 2011), before slowly increasing to ~30% after 60 min and returning to ~50% in the TMAO-respiratory/fermentative steady state (Figure 3.4). Active ArcA is believed to exist in an octameric form, with 50% of the ArcA monomers

TF	Inferred TF activity at the indicated time (min) after addition of TMAO							
	0	2	5	10	15	20	60	1440
ArcA	2.42	0.80	0.33	0.26	0.68	1.20	1.63	2.25
CusR	2.96	2.39	0.61	0.31	0.46	0.36	0.39	2.64
CsgD	2.13	0.74	0.78	1.23	0.51	0.61	0.61	2.98
PutA	2.83	1.20	0.36	0.13	0.67	0.88	1.38	1.97
PurR	0.78	2.03	5.80	2.30	0.57	0.33	0.37	0.85
ArgR	1.15	1.68	4.48	2.38	0.36	0.15	1.08	0.82
CRP	0.65	1.51	2.86	2.49	0.72	0.52	0.48	1.03
MalT	0.76	2.21	1.12	4.53	1.76	0.37	0.20	1.33
Lrp	0.70	1.24	1.71	4.17	1.73	0.53	0.44	0.57
DgsA	0.33	0.92	0.92	3.28	4.54	2.10	1.08	0.40
IlvY	0.86	1.18	1.30	3.20	2.11	0.59	0.34	0.70
TorR	0.19	0.88	1.83	3.16	3.04	2.85	2.97	0.51
MetJ	1.36	1.41	2.22	2.40	0.72	0.09	0.93	1.31
GalS	0.74	0.62	0.61	0.88	6.17	1.15	1.01	0.60
OxyR	0.81	0.71	0.62	0.92	3.64	1.02	1.14	0.75
CysB	0.79	0.62	0.77	0.41	0.99	5.22	2.08	0.72
RcsAB	1.37	0.80	1.61	0.85	2.07	2.70	0.77	0.16
GlpR	2.04	0.36	0.13	0.25	1.17	2.65	3.30	1.56
PhoP	1.07	0.59	0.24	0.71	2.01	2.47	1.51	1.16
FlhDC	0.87	0.57	0.72	0.58	0.69	0.64	0.85	11.51

Figure 3.2. Changes in TF activities in response to perturbation of anaerobic fermentative cultures of *E. coli* with the electron acceptor TMAO

At the indicated times (min) after TMAO addition to anaerobic fermentative steady state cultures, the output from the TFInfer software (Asif *et al*., 2010) of TF activities with the signal to noise ratios ≥2 are shown. The colour of each cell provides a visual representation of the TF activities inferred from the complete transcript profile dataset. TF activity 0–1, blue; 1–2, green; 2–3, yellow; 3–4, orange; >4, red.

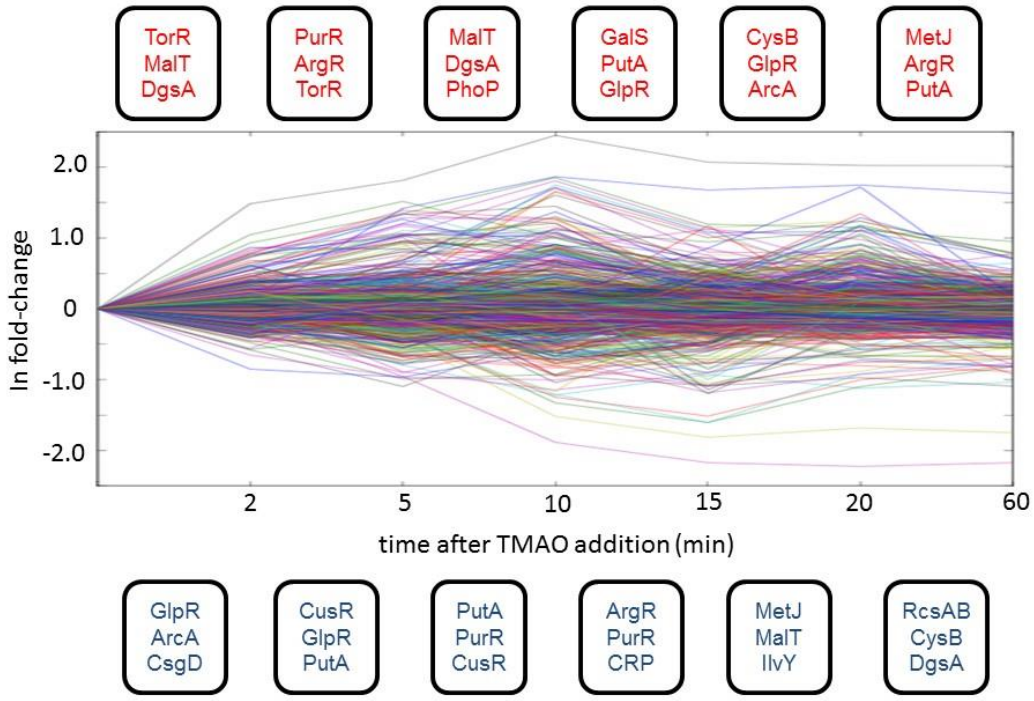


Figure 3.3. **<u>The regulatory dynamics that occurred after the addition of TMAO to anaerobic fermentative steady state cultures</u>**

Graphical representation of regulatory dynamics. The figure shows the time course of all 1,381 genes used in the TFInfer analysis; time is re-scaled to have equal spacing between time points. The panels above and below the time course show the three TFs exhibiting the greatest fold increase (red) or decrease (blue) in activity compared with previous time point as inferred by TFInfer, providing a depiction of the main regulatory events that occurred in response to the addition of TMAO.

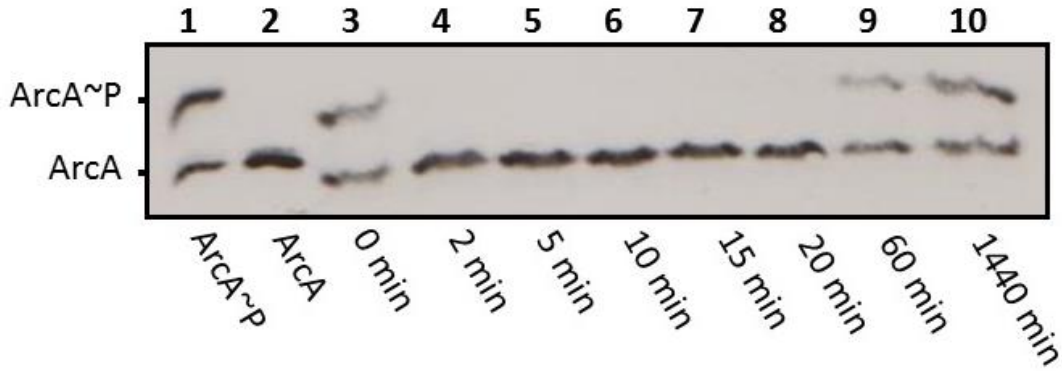


Figure 3.4. <u>Measurement of ArcA~P validate the inferred activity of ArcA</u>

A typical Phos-tag™ gel developed with anti-ArcA serum. Lane 1 contains purified His-tagged ArcA protein (1.2 µg) phosphorylated *in vitro* by incubation with carbamoyl phosphate (50 mM); lane 2, purified His-tagged ArcA dephosphorylated; lane 3-10, whole cell samples taken at time = 0 (anaerobic fermentative steady state), 2, 5, 10, 15, 20, 60 and 1,440 min (TMAO respiratory/fermentative steady state) after addition of TMAO to the anaerobic fermentative steady state culture. The locations of both ArcA and ArcA~P are indicated.

being phosphorylated (1:1 ratio ArcA:ArcA~P). However, inactive ArcA is in a dimeric form and does not contain any phosphorylated monomers (Jeon *et al*., 2001). Thus, direct measurement of the ArcA phosphorylation profile correlated well with the predicted ArcA activities inferred from transcript profiles, as a decrease in ArcA activity is suggested by these measurements. At least three signals have been suggested to affect the activtity of the cognate histidine kinase for ArcA, ArcB: oxidised ubiquinone acts negatively; reduced menaquinone acts positively; and fermentation products, e.g. acetate, promote ArcB kinase activity and ArcA phosphorylation (Alvarez *et al*., 2013; Bekker *et al*., 2010; Georgellis *et al*., 1999, 2001; Rodriguez *et al*., 2004; Rolfe *et al*., 2012; Rolfe *et al*., 2011). The dynamics of ArcA activity here suggests that ArcA was initially dephosphorylated after the introduction of TMAO to the fermenting culture, presumably because the addition of TMAO caused net oxidation of the menaquinone pool. Then, perhaps mediated by the enhanced production of acetate (Table 3.1), ArcA activity was restored in the final TMAO-respiratory/fermentative steady state. The changes in ArcA activity account for 7 of the 29 transcripts that were upregulated ≥3-fold in response to the presence of TMAO, as these transcripts are repressed by ArcA, including the unexpected transient upregulation of the cytochrome *bo*' operon, as well as genes encoding the pyruvate, 2-oxoglutarate and succinate dehydrogenases (Table 3.2).

Maximal activity during the adaptive phase (2-60 min) of the transition was seen for the majority (15 out of 20) of the responsive TFs (signal to noise ratio ≥2), including TorR, before their activities returned to values close to those in the initial fermentative steady state, in the final steady state. The greatest activities for ArgR, CRP and PurR were 5 min after the introduction of TMAO; IlvY, Lrp, MalT, MetJ and TorR after 10 min; CysB, RcsAB and PhoP after 20 min; and GlpR after 60 min (Figure 3.2).

3.5. Regulation of the TorA protein is at the level of transcription as suggested by measurement of TorA protein and TMAO reductase activity

After the introduction of TMAO measurements of the TorA protein content (Section 2.10.12) (Figure 3.5a) and TMAO reductase activities (Section 2.13.1) (Figure 3.5b) were made. Western blotting was used to estimate the TorA content of the bacteria using TorA anti-sera (Dr Iobbi-Nivol, CNRS, Marseille) (Figure 3.5a). This showed that the abundance of TorA increased during the first 60 min of the transition and was maintained in the final steady state. The TMAO reductase activity of bacteria grown under fermentative conditions was low, but increased ~2-fold 60 min after the addition of TMAO (Figure 3.5b). The changes in

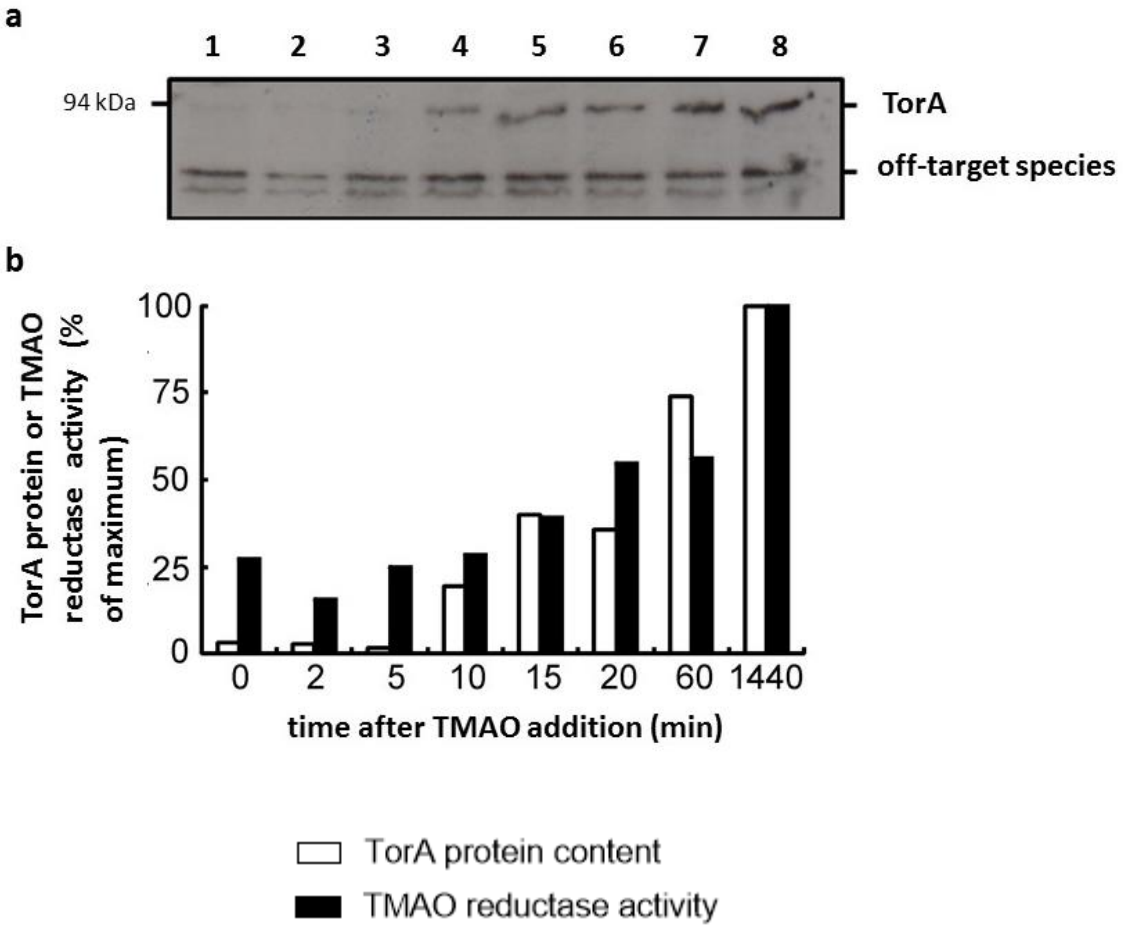


Figure 3.5a and b. **<u>The amount of TorA protein and TMAO reductase activity during adaptation to TMAO-respiratory growth correlate with the changes in *tor* operon transcription</u>**

- **a.** Typical Western blot, developed with anti-TorA serum, of whole cell samples taken from anaerobic fermentative chemostat cultures 0, 2, 5, 10, 15, 20, 60 and 1,440 min after the addition of TMAO.
- **b.** Quantitative densitometric analysis of Western blot (shown in Figure 5a) was used to measure the increase in the amount of TorA protein present in the bacteria after TMAO addition (open bars). The amounts of TorA shown are relative to the level in the 1,440 min sample (100%). TMAO reductase activities (filled bar) were measured for duplicate samples obtained at the same time points used for quantification of TorA protein (Figure 5a). The TMAO reductase activity data are also shown relative to the activity in the 1,440 min sample (100%).

TMAO reductase enzyme activity in the new TMAO-respiratory/fermentative steady state compared to the initial fermentative sample (~4-fold higher), correlated with the 3.5-fold increase in *tor* operon transcript abundance (Table 3.2). This suggests an absence of any post-transcriptional regulation of TMAO reductase activity. Western blotting of TorA protein during the transition suggested a greater overall increased in abundance than the TMAO reductase activity increase. However, there was a good positive correlation between enzyme activity and TorA protein content in the samples obtained in the 10 – 1,440 min period after TMAO addition (Figure 3.5a and b).

3.6. Formaldehyde dehydrogenase FrmAB induction is likely to be the result of the presence of TMAO demethylase activity

Unexpectedly, there was a transient increase in abundance of the *frmRAB* transcript upon exposure of the cultures to TMAO (Table 3.2). This was surprising because the *frmRAB* operon is known to be induced by formaldehyde (Herring and Blattner, 2004). The expression of the operon is negatively regulated by the transcriptional regulator FrmR, which is encoded by the first gene of the operon. The FrmA and FrmB proteins catalyse the glutathione-dependent two-step conversion of toxic formaldehyde to non-toxic formate (Gutheil *et al*., 1992) (Equation **5**).

$$\text{Formaldehyde} + \text{Glutathione} \xrightarrow{\text{FrmA}} \textit{S}\text{-formylglutathione} \xrightarrow{\text{FrmB}} \text{Formate} + \text{Glutathione} + H^+ \quad (\mathbf{5})$$

As this upregulation was unexpected, confirmation of *frmRAB* induction in the presence of TMAO was first sought by quantitative reverse transcription polymerase chain reaction (qRT-PCR). Analyses of RNA isolated from batch cultures of *E. coli* MG1655 after 60 min exposure to a range of different TMAO concentrations (0–40 mM) were carried out (Section 2.8.14.1). The abundance of the *frmRAB* transcript was significantly increased when the bacterial cultures were exposed to ≥5 mM TMAO (Figure 3.6). Therefore, these experiments confirmed the transcriptional response detected by the transcript profiling. They also show the sensitivity of the induction, as upregulation of *frmRAB* occurred at concentrations of TMAO ≥5 mM (Figure 3.6) and that this adaptation occurred both in continuous and batch cultures (Table 3.2).

In separate experiments, the Nash reagent (Nash, 1953) was used to detect formaldehyde accumulation by *E. coli* K-12 cell-free extracts incubated with TMAO (40 mM). No formaldehyde was accumulated in reactions lacking cell-free extract or TMAO (Section

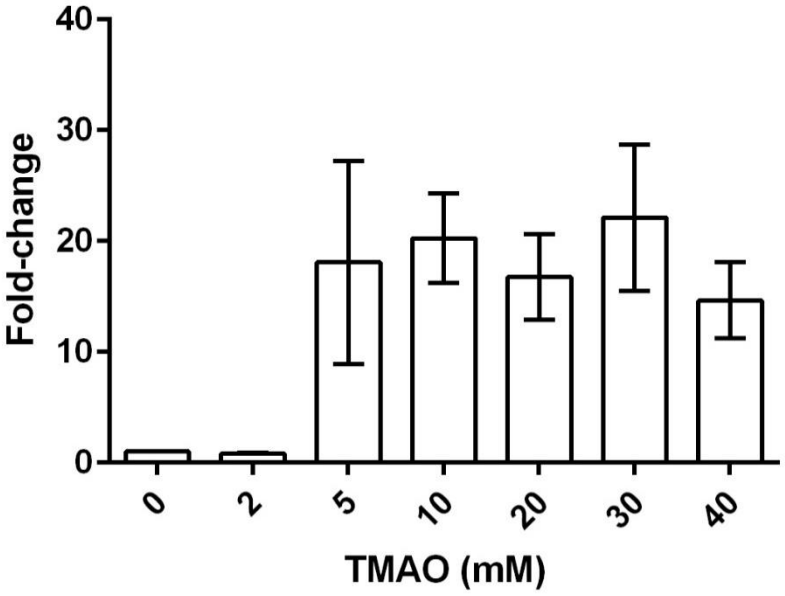


Figure 3.6. The *frmRAB* operon is induced after exposure of *E. coli* MG1655 to concentrations of TMAO ≥5 mM

Anerobic batch cultures were exposed to different concentrations (0-40 mM) of TMAO. After 30 min incubation with TMAO, total RNA was isolated from the cultures for qRT-PCR analysis of *frmR* mRNA. The data shown are the mean and standard deviation for the fold increase relative to the 0 mM TMAO culture from three technical replicates and are representative of three indepent biologial experiments.

2.13.3). However, formaldehyde was detected in the complete assays producing 20 nmole mg^{-1} of cell-free extract after 40 min incubation (Figure 3.7). Furthermore, the presence of 0.23 mM dimethylamine (DMA) was found in the extracellular metabolite profile of the chemostat cultures after the addition of TMAO (Table 3.1). These observations suggest that *E. coli* K-12 possesses TMAO demethylase activity (EC4.1.2.32), which results in the accumulation of DMA and formaldehyde, with the production of the latter inducing *frmRAB* expression. The *frmRAB* operon has previously been shown to be induced when cultures of *E.coli* are exposed to 0.25 mM formaldehyde (Herring and Blattner, 2004), which is a concentration very similar to the amount of DMA detected here (Table 3.1). The induction of the *frmRAB* operon during transition to TMAO respiratory growth; the detection of formaldehyde production by cell-free extracts incubated with TMAO and the detection of DMA in the culture supernatant suggested that a *frmRAB* mutant would be unable to detoxify formaldehyde and therefore would have compromised viability in the presence of TMAO. To test this, an isogenic *E. coli* MG1655 *frmRAB* mutant (JRG6703) was created using λred recombineering (Section 2.8.11) with the kanamycin resistance cassette from the pKD4 plasmid, along with a plasmid (pGS2486) to complement the mutant by expressing *frmRAB* under the control of its own promoter. The complementation plasmid was created using the In-Fusion Cloning Kit (Clonetech) (Section 2.8.12). The numbers of colony forming units (cfu) present in anaerobic cultures of the parent, mutant and complemented mutant before and after incubation with TMAO were measured (Section 2.4). The cfu values of the parent and complemented mutant cultures increased after 60 min exposure to TMAO (40 mM), reflecting the utilisation of the more efficient metabolic mode. However, the cfu value of the *frmRAB* mutant did not increase (Figure 3.8). This result shows that induction of a regulatory response in order to detoxify the production of potentially toxic concentrations of formaldehyde is necessary to prevent inhibition of growth during transition to anaerobic respiration with TMAO as the terminal electron acceptor.

3.7. Copper(II) is in competition with TMAO/TMAO reductase for electrons from the anaerobic electron transport chain

The decrease in *cus* transcripts in the transcriptional profiling and the inferred decrease in activity of the TF CusR implied that, after introduction of TMAO, the exposure of the *E. coli* cultures to Cu(I) was diminished. This is because the cognate sensor for CusR is CusS, for which Cu(I) is the signal. Under aerobic conditions, *E. coli* can use electrons transferred to ubiquinone by NADH dehydrogenase II (E_m NAD^+/NADH, -320mV; ubiquinone/ubiquinol,

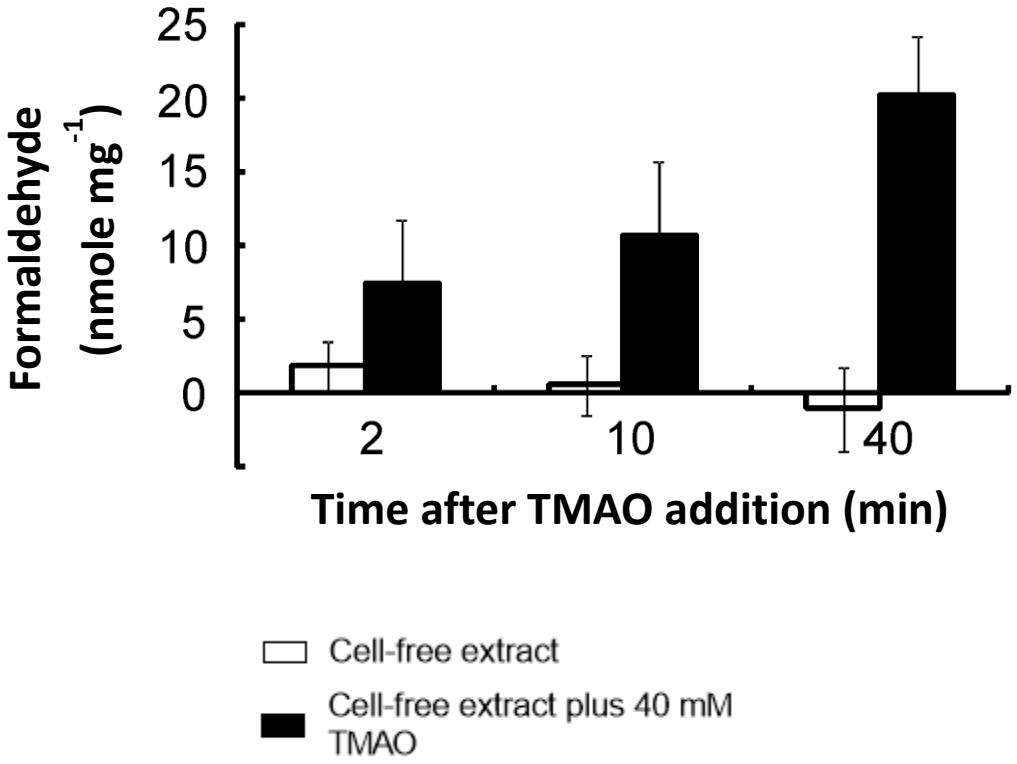


Figure 3.7. Detectable concentrations of formaldehyde are produced upon incubation of *E. coli* MG1655 cell-free extracts with TMAO

Cell-free extracts of *E. coli* K-12 were incubated with TMAO (40 mM) at 37°C. At the indicated times formaldehyde production was measured using the Nash reagent. Open bars, cell-free extract only; closed bars, cell-free extract plus TMAO. The data shown are the mean and standard deviation values obtained from three independent experiments.

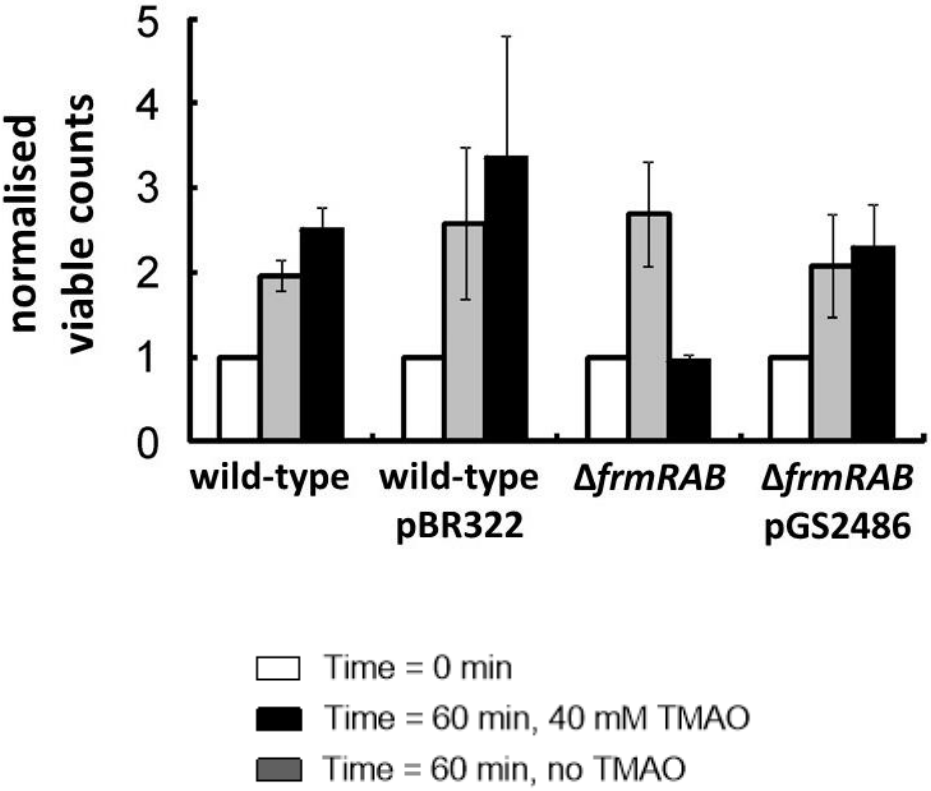


Figure 3.8. The *E. coli* MG1655 *frmRAB* mutant is attenuated when exposed to TMAO

Anaerobic cultures of *E. coli* MG1655 (wild type), wild type transformed with the vector pBR322, the *frmRAB* mutant, and the mutant complemented with the *frmRAB* expression plasmid (pGS2486) were grown on glucose minimal medium buffered with 50 mM phosphate buffer, pH 7.0 at 37°C. In an anaerobic cabinet, each culture was divided into two separate cultures and samples for viable count measurements were taken (t = 0 min; white bars). TMAO (40 mM) was added to one (black bars), but not the other (grey bars), of each pair of cultures and anaerobic incubation was continued for a further 60 min, at which point samples for viable count measurements were taken. The data shown are the mean and standard deviation values, normalised to those obtained at t=0 min, from three independent biological replicates.

+110 mV) to reduce the trace element Cu(II) to Cu(I) (E_m + 340 mV) (Thauer *et al*., 1977; Volentini *et al*., 2011). Menaquinone (E_m menaquinone/menaquinol, -74 mV) is the electron donor for the TMAO reductase under anaerobic conditions, (Wissenbach *et al*., 1992) and has the potential to reduce Cu(II), like ubiquinone. Therefore, under anaerobic conditions in both the presence and absence of TMAO, the capacity of *E. coli* K-12 MG1655 whole cells to reduce Cu(II) to Cu(I) was tested (Section 2.13.2). Reduction of Cu(II) was detected in both the presence and absence of TMAO, but was diminished ~1.6-fold in the presence of TMAO (Figure 3.9). This was interpreted as follows: under anaerobic fermentative conditions there is a relatively low flux of electrons from the menaquinone pool to fumarate reductase to reduce fumarate to succinate, as suggested by the analysis of over-metabolite production (Table 3.1). This allows electrons to leak directly or indirectly to periplasmic Cu(II) from reduced quinones, resulting in the production of Cu(I) and activation of the CusSR transcriptional response (Table 3.2, Figure 3.10a). An additional outlet for electrons in the menaquinone pool is provided when TMAO is supplied, resulting in the decreased reduction of Cu(II) to Cu(I) (Figure 3.9) and decreased reduction of fumarate to succinate (Table 3.1) during the initial adaptive phase (Figure 3.10b). Also, the presence of TMAO invokes a shift towards a more oxidised menaquinone pool, which would contribute to the inactivation of ArcA and hence de-repression of *cyoA-E*, *phdR-aceEF-lpd* and *sdhCDAB-sucA-D* operons (Figure 3.4, Table 3.2).

3.8. Discussion

The experiments described in this chapter provide new insights about the adaptive processes, which occur when *E. coli* K-12 switches metabolic mode from fermentative growth to anaerobic growth with TMAO as the terminal electron acceptor (Figure 3.10a and b).

Several of the measured changes in transcript and metabolite profiles were predictable, such as the reduction of TMAO to TMA via the induction of the *torCAD* operon, the main TMAO reductase of *E. coli* K-12. However, some unexpected components of the adaptive process were highlighted. For example, the detection of DMA and formaldehyde suggested that *E. coli* possesses TMAO demethylase activity (EC4.1.2.32). This activity produced detectable and physiologically significant amounts and of formaldehyde when TMAO was supplied in excess. The amount of formaldehyde generated by this activity was of sufficient quantity to induce the expression of the formaldehyde detoxifying enzymes (FrmAB) when TMAO was abundant. TMAO demethylase activity, although never reported before in *E. coli*, has been found in a number of different marine wild life species, including the Jumbo Squid (*Dosidicus*

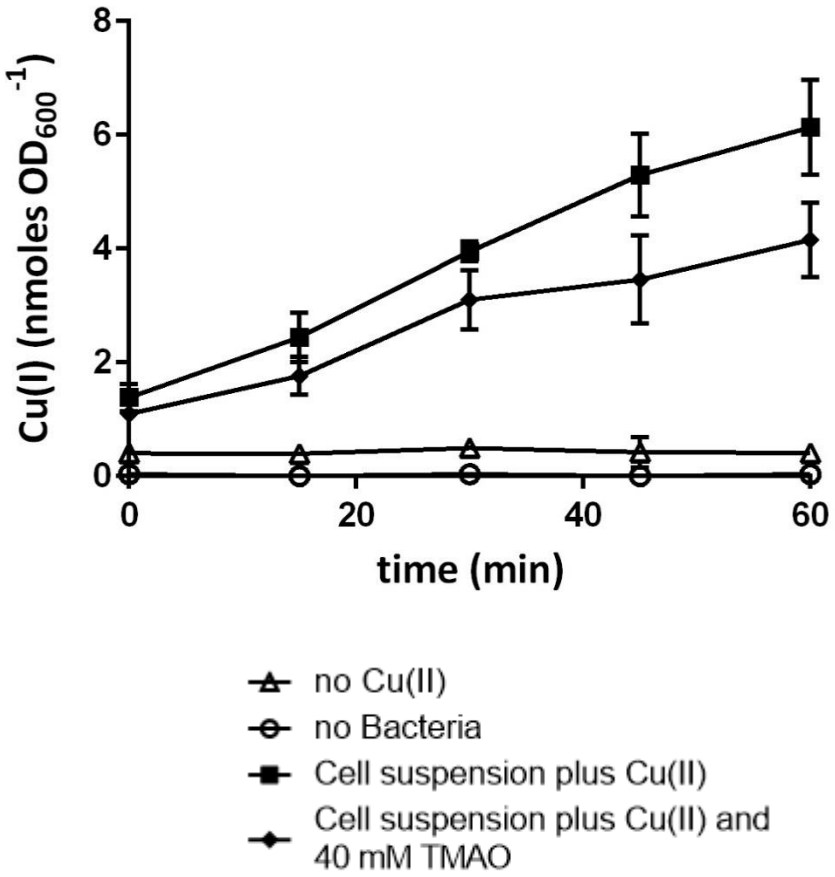


Figure 3.9. <u>The effect of TMAO on Cu(II) reduction by *E. coli* MG1655</u>

Bacterial cell suspensions (final OD_{600} ~2.5) in Evans medium were incubated at 37°C for up to 60 min in the presence and absence of $CuSO_4$ (0.05 mM) and/or TMAO (40 mM). At the indicated time points, aliquots were removed and the amount of Cu(I) present was measured. Open circles, no bacteria; open triangles, no Cu(II); closed diamonds, bacteria plus Cu(II) plus TMAO; closed squares, bacteria plus Cu(II). The data are the mean values ± standard deviation obtained from three independent experiments.

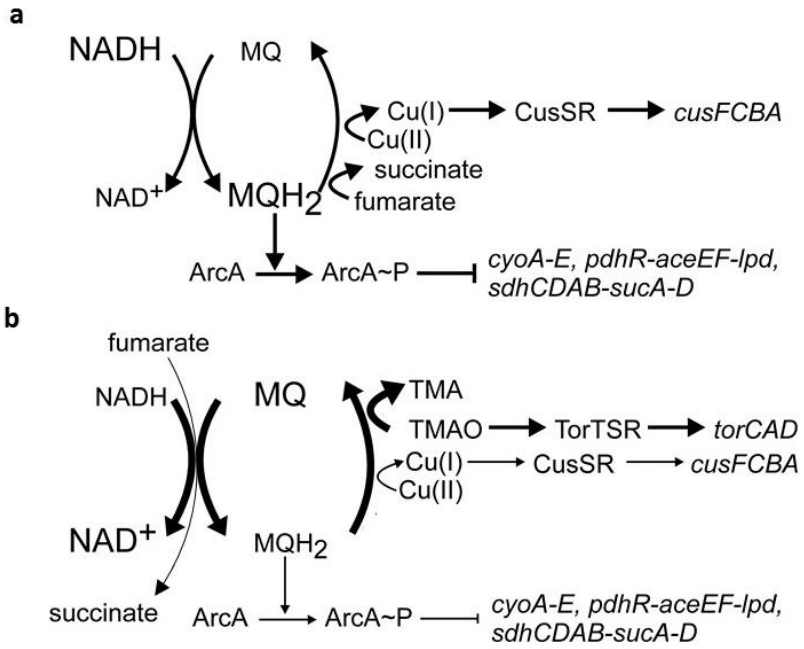


Figure 3.10a and b. <u>Excess TMAO activates TorTSR, inhibits Cu(II) reduction by the electron transport chain and permits activation of ArcA</u>

- **a.** During anaerobic fermentation, at the expense of menaquinol (MQH_2), fumarate is reduced to succinate and the trace element Cu(II) is reduced to Cu(I). The Cu(I) generated by this process is sensed by the CusSR two-component system and the copper efflux system CusFCBA is activated. There is a high MQH_2:MQ ratio and flux through the electron transport chain is relatively low, resulting in the phosphorylation of ArcA (ArcA~P) and repression of the *cyoA-E*, *pdhR-aceEF-lpd* and *sdhCDAB-sucA-D* operons.
- **b.** Addition of TMAO to the fermenting cultures is sensed by the TorTSR sensor-regulator system, activating the expression TMAO reductase (*torCAD* operon). The MQH_2:MQ ratio becomes low as TMAO reductase is able to utilise for the electrons of the menaquinone pool. This results in a decreased phosphorylation of ArcA resulting in de-repression of the *cyoA-E*, *pdhR-aceEF-lpd* and *sdhCDAB-sucA-D* operons. A respiratory chain with succinate as the electron donor and TMAO as the electron acceptor, could also be permitted and is consistent with the observed transcriptional profile and the concentrations of succinate and fumarate in the culture medium. The capacity of the bacteria to reduce Cu(II) to Cu(I) is lower as the flux of electrons to TMAO increased, thus the TF CusR is switched off. The width of the regulatory arrows is an indication of the relative rate of each step and font size is used to indicate the $NADH:NAD^+$ and MQH_2:MQ ratios.

gigas) and Walleye Pollock (*Gadus chalcogrammus*). In Walleye Pollock, the protein Aspolin is responsible for iron-mediated TMAO demethylation and is an extremely aspartic acid-rich protein (179 Asp residues) composed of 186 amino acids (Takeuchi *et al*., 2003). The TMAO demethylase enzyme from Jumbo Squid has also been isolated and characterised (Fu *et al*., 2006). The molecular mass of the TMAO demethylase was 17.5 kDa, had an optimum temperature of 55^{o}C and converted TMAO to DMA and formaldehyde stoichiometrically with a K_m of 26.2 mM (Fu *et al*., 2006). Work on TMAO demethylase activity in fish is of commercial significance, as upon death of the fish, production of formaldehyde by the demethylase activity could result in food spoilage. But, TMAO demethylase activity is not restricted to marine lifeforms, as recently the activity has been discovered in prokaryotes including *Pseudomonas putida* (Liffourrena *et al*., 2010) and *Methylocella silvestris* (Zhu *et al*., 2014). Studies on the recombinant TMAO demethylase of *M. silvestris* suggest that it is hexameric, has a high affinity for TMAO and has narrow substrate specificity (Zhu *et al*., 2014). The enzyme also does not share any sequence homology or similar characteristics to eukaryotic TMAO demethylases (Zhu *et al*., 2014). Sequence analysis of the *E. coli* K-12 genome was carried out and showed that there was no similar gene, with sequence homology present. In light of these recent findings, work was attempted to identify the source of the TMAO demethylase activity in *E. coli* K-12. However, technical difficulties with using the Nash reagent to detect low concentrations of formaldehyde produced from the TMAO demethylase activity prevented this work from being followed through. The failure to purify the TMAO demethylase activity could be simply due to low enzymatic activity and thus difficulty in obtaining the pure enzyme responsible. However, there is also the possibility that the TMAO demethylase activity detected here is a side reaction of another cellular demethylase enzyme present in *E. coli* K-12, and that TMAO is demethylated by this enzyme when TMAO is present in excess. It should still be feasible to identify and purify the source of the activity regardless of where the source is specifically a TMAO demethylase or another demethylase enzyme present in *E. coli,* which recognises and is able to demethylate TMAO and work needs to continue here.

Despite maintaining anaerobic culture conditions through the experiments, there was a surprising induction of the *cyoA-E* transcript, which encodes a low oxygen-affinity terminal oxidase. This was explained by the transient inactivation of ArcA, which was predicted by probabilistic modelling of transcription factor activities and validated by the direct measurement of ArcA~P. A number of other ArcA~P-repressed operons, as well as *cyoA-E*, were induced and this could be accounted for by the temporary inactivation of ArcA. The

two-component system, ArcBA, acts as an indirect sensor of oxygen, as ArcB monitors the redox state of the electron transport chain and the production of fermentation products, such as acetate. Therefore, the activity of ArcBA is sensitive to cellular NADH/NAD^+ ratios (Holm *et al*, 2010). Using glucose-limited chemostat cultures, it was shown that *E. coli* are affected by the availability of electrons acceptors fumarate, nitrate and oxygen, which act as effective NADH sinks. During anaerobic fermentative growth, the greatest NADH/NAD^+ ratios were observed. Lower NADH/NAD^+ ratios were measured when an external electron acceptor (~0.3 with 7.5 mM nitrate; ~0.4 with 70 mM fumarate; and ~0.08 under fully aerobic conditions) was present (de Graef *et al*., 1999; Alexeeva, 2000). Adding excess TMAO to an anaerobic fermentative culture would be expected to result in a decreased NADH/NAD^+ ratio, as an additional outlet for electrons in the menaquinone pool is being provided. The result of the additional outlet for electrons results in the inactivation of ArcBA, which is then activated again by increased production of acetate as TMAO-respiratory metabolism is established. This work highlights further the complexity of ArcBA regulation and the number of different factors, which effect the two-component system's activity.

The downregulation of copper resistance genes in the presence of excess TMAO can be accounted for by competition for electrons from the menaquinone pool. Work previously published showed that electrons from the *E. coli* respiratory chain can reduce Cu(II) to Cu(I) (Volentini *et al*., 2011). Here, it was shown that the presence of the electron acceptor TMAO was able to inhibit reduction of Cu(II) by *E. coli*, which suggests that TMAO reductase/TMAO are effective competitors for electrons that might otherwise have contributed to the reduction of Cu(II) to Cu(I). The inactivation of CusR and downregulation of copper resistance genes are a result of the lower Cu(I) production.

Performing similar experiments with mutants lacking some of the key regulators (e.g. ArcA, FlhDC, FrmR and MetJ) identified here could further characterise the roles played by their respective regulons in adaptation to TMAO exposure. It would also be interesting to use other alternative electron acceptors (e.g. nitrate, nitrite, DMSO) in these experiments to characterise the different dynamic adaptive processes that *E. coli* K-12 utilises in order to respond to other electron acceptors. However, the systematic analysis of *E. coli* K-12 cultures carried out in this chapter has provided new insights into the dynamics of transcriptional and physiological changes required to successfully adapt to anaerobic environments containing TMAO.

4. Analysis of the transcriptional initiation of the *torCAD* operon in individual *E. coli* cells after exposure to TMAO

4.1. Introduction

The major pathway for TMAO respiration in *E. coli* requires the expression of the *torCAD* operon, which encodes the TorC, TorA and TorD proteins (Gon *et al*., 2001a; Méjean *et al*., 1994). TMAO respiration through the TorCAD pathway is strictly regulated by the presence of TMAO in the environment. This strict regulation is exerted by the TorT-TorS-TorR signal transduction system and means that components of the *E. coli* TMAO respiratory chain are only expressed in the presence of TMAO (Jourlin *et al*., 1997; Jourlin *et al*., 1996b).

As shown in Chapter 3, *E. coli* MG1655 grown under anaerobic conditions with a sub-optimal concentration of TMAO as the terminal electron acceptor and glucose as the sole carbon source, resulted in mixed TMAO respiratory and fermentative metabolism (TMAO-respiratory/fermentative growth) after 24 h. Both the products of TMAO respiration and fermentation were present in culture supernatants that had been exposed to TMAO. This behaviour might arise from a single population of cells that simultaneously ferment and respire, or sub-populations that exclusively ferment or respire (Figure 4.1).

In order to test these hypotheses, single cell experiments were carried out to determine whether heterogeneity in *torCAD* transcription was responsible for the mixed metabolism reported in Chapter 3, i.e. did mixed metabolism arise from all cells utilising both respiration and fermentation or from a mixed population of TMAO-respiring cells and cells carrying out fermentation. To carry out single cell studies, a $P_{torC\text{-}gfp}$ reporter strain of *E. coli* K-12 was created (*E. coli* JRG6705), to determine whether all cells were utilising TMAO when it is present (i.e. all cells express GFP) or whether only a sub-population reduce TMAO (i.e. only a proportion of the cells express GFP) (Figure 4.1).

4.2. Creation of $P_{torC\text{-}gfp}$, *ΔtorC* reporter strain (*E. coli* JRG6705) using λred recombineering

In order to investigate the metabolic mode of individual *E. coli* K-12 cells in the presence of excess TMAO, a $P_{torC\text{-}gfp}$, *ΔtorC* fusion strain (*E. coli* JRG6705) was created using λred recombineering (Datsenko and Wanner, 2000) (Section 2.8.11) (Table 2.1). As a result of transducing the reporter fusion into the strain, the chromosomal *torC* gene was replaced with genes coding for green fluorescent protein (GFP) and kanamycin resistance from the plasmid p-DOC-G (Lee *et al.*, 2009) (Table 2.2). Creation of the P_{torC}-gfp fusion, thus led to the knockout of the *torC* gene (Figure 4.2a). Once the genotype of the reporter strain had been

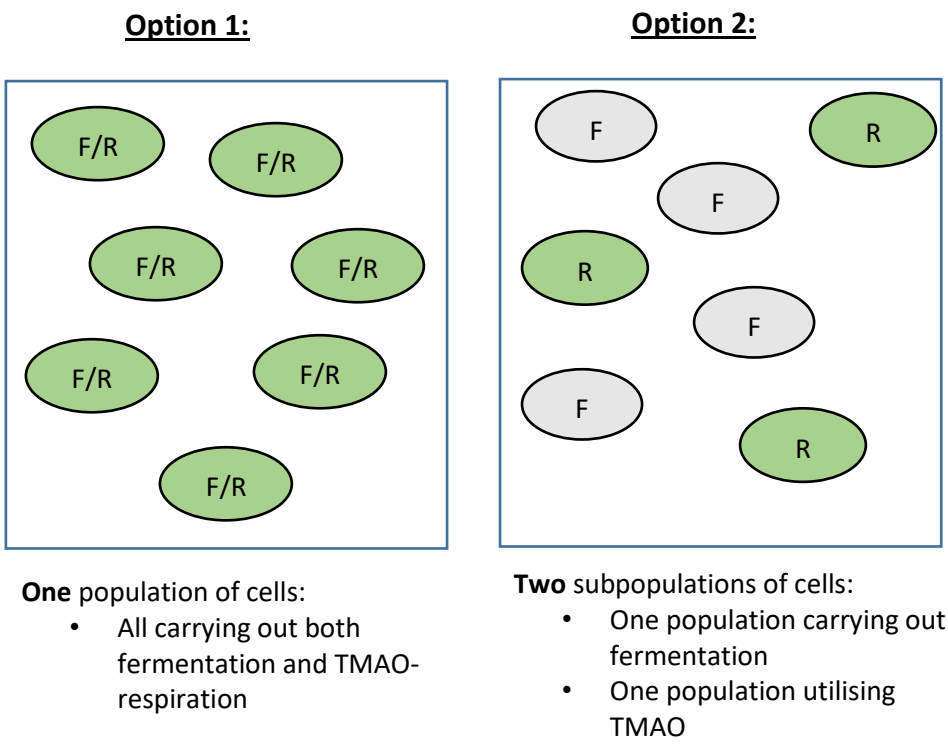


Figure 4.1. Schematic diagram showing the possible behaviours of individual cells in a mixed metabolic culture

Option 1, occurs when all cells within the population are carry out both fermentation and TMAO-respiration simultaneously; option 2, occurs when there are two sub-populations of cells each carrying out a different mode of metabolism with only one sub-population utilising TMAO. Ovals represent individual cells within the whole population; F, represents cells carrying out fermentation; R, represents cells utilising TMAO. Cells are coloured green when TMAO is being utilised as this will express *gfp* via the reporter fusion ($P_{tor\text{-}gfp}$).

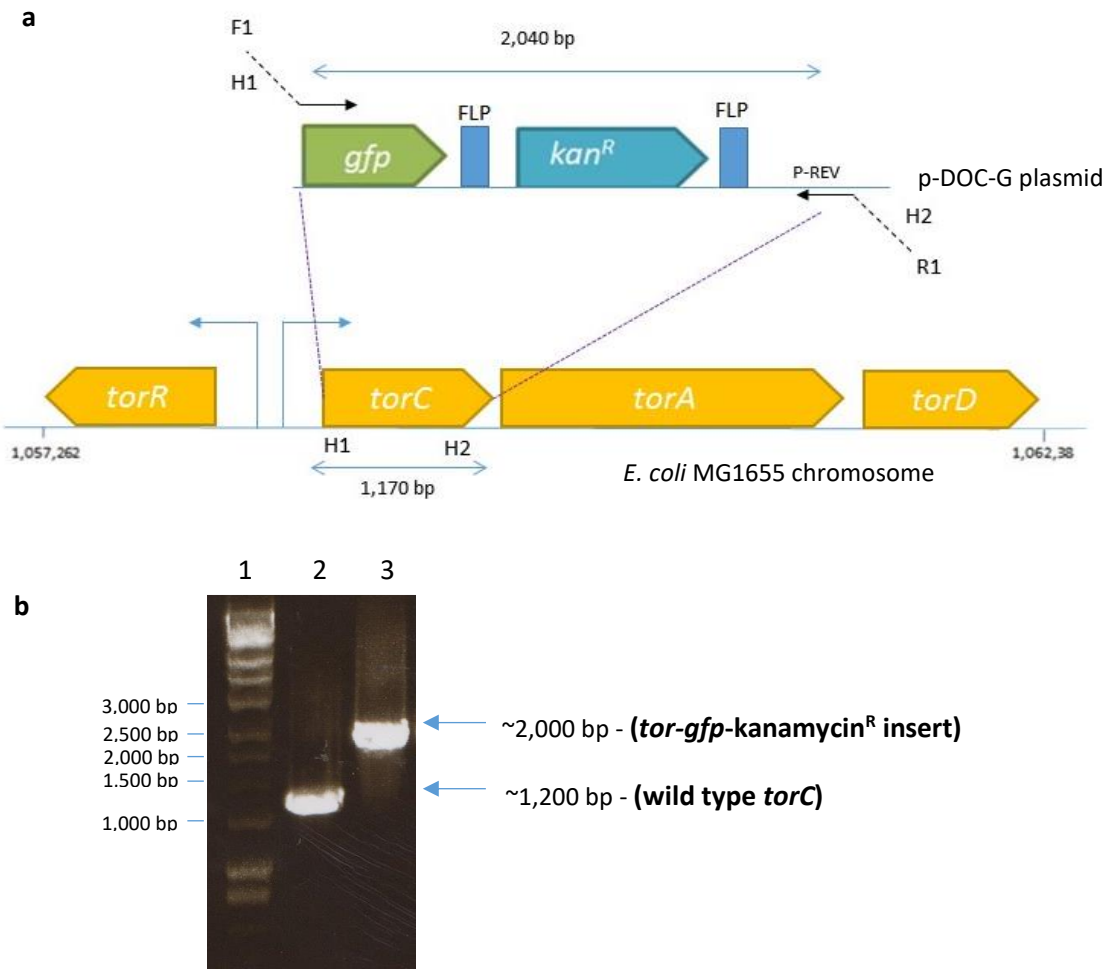


Figure 4.2a and b. Creation of the $P_{torC\text{-}gfp}$ fusion using homologous recombination

a. A schematic diagram representing the creation of the $P_{torC\text{-}gfp}$ fusion in the *E. coli* MG1655 genome (JRG6705) created using homologous recombination with the p-DOC-G plasmid (Section 2.8.11) (Table 2.2). The *gfp* and kanamycin resistance genes are encoded for on the p-DOC-G plasmid. Numbers represent either size of the indicated genes or position of the genes on the *E. coli* MG1655 chromosome. Oligonucleotide primer (F1 and R1) annealing sites are shown with dotted black lines highlighting the regions containing homology to the *torC* gene with these sites also being shown on the *E. coli* MG1655 chromosome (H1 and H2). FLP are Flp recombination sites. Dotted purple lines represent where homologous recombination events take place.

b. An agarose (1%) gel showing PCRs using TorC-fusion seq F and R oligonucleotide primers (Appendix 1) and either *E.coli* JRG6705 or *E.coli* MG1655 genomic DNA as the template. Lane 1, Hyperladder I (Bio Line); Lane 2, PCR using *E.coli* MG1655 genomic DNA; Lane 3, PCR using *E.coli* JRG6705 genomic DNA. Sizes of PCR products are indicated by the arrows and sizes of DNA markers are indicated.

verified by PCR (Figure 4.2b) and DNA sequencing, the phenotype was confirmed using NMR (Section 2.12). The reduction of TMAO by anaerobic cultures of both wild type and *E.coli* JRG6705 after ~ 18 h growth at 37°C was analysed. Wild type cells, which had a fully functioning TorCA respiratory chain were able to reduce the majority of the TMAO added to the culture, as shown by 13-fold decrease in the concentration of TMAO in the culture supernatant and production of TMA (TMAO: 0 h, 42.3 mM; 18 h, 3.3 mM; TMA: 0 h, 0 mM; 18 h, 32.0 mM) (Figure 4.3). However, *E. coli* JRG6705 cultures were unable to reduce TMAO to the same level as the wild type, with only 1.6-fold decrease in TMAO over the same time period (TMAO: 0 h, 42.9 mM; 18 h, 27.2 mM; TMA: 0 h, 0 mM; 18 h, 10.4 mM) (Figure 4.3). Therefore, replacing *torC* with *gfp* in the chromosome of *E. coli* JRG6705 resulted in an overall 8.2-fold decrease in TMAO reduction when compared to wild type. Creation of the fusion does not completely render the strain unable to reduce TMAO, as seen by the presence of TMA in the supernatants of *E. coli* JRG6705 cultures. Reduction of TMAO in this strain will still be carried out by the two other TMAO reduction systems present in *E. coli* (TorZY and DmsABC) (Bilous and Weiner, 1988; Gon *et al*., 2000). As these systems are expressed at low levels, they account for the smaller amount of TMAO reduction seen here (Figure 4.3). It is important to note that the decrease in TMAO reduction associated with *E. coli* JRG6705 cells will mean that less TMAO is converted to TMA, thus the signal for induction of $P_{tor\text{-}gfp}$ will persist.

4.3. Analysis of *torCAD* expression in response to TMAO over time

Continuous cultures of *E.coli* JRG6705 cells were grown in 1 L chemostat vessels under anaerobic conditions as previously described (Section 2.3.1). Cells were grown in order to use fluorescence microscopy and flow cytometry to observe the expression of *torCAD* in individual cells in response to TMAO.

4.3.1. The expression of *torCAD* in response to TMAO does not occur in every cell (all cells are not GFP-positive) and the mean level of expression changes (intensity of GFP signal) over time

Flow cytometry experiments were carried out on fixed cultures, removed from the chemostat at the indicated time points (Sections 2.3.1, 2.14.1 and 2.14.2). Cells were washed and fixed under aerobic conditions to allow for GFP maturation. Samples were analysed with an Attune Autosampler using the BL1 detector for GFP detection (Section 2.14.5). Flow cytometry experiments were gated using a sample of GFP-negative cells (no TMAO, anaerobic conditions) such that a cell would only qualify as being GFP-positive if the BL1

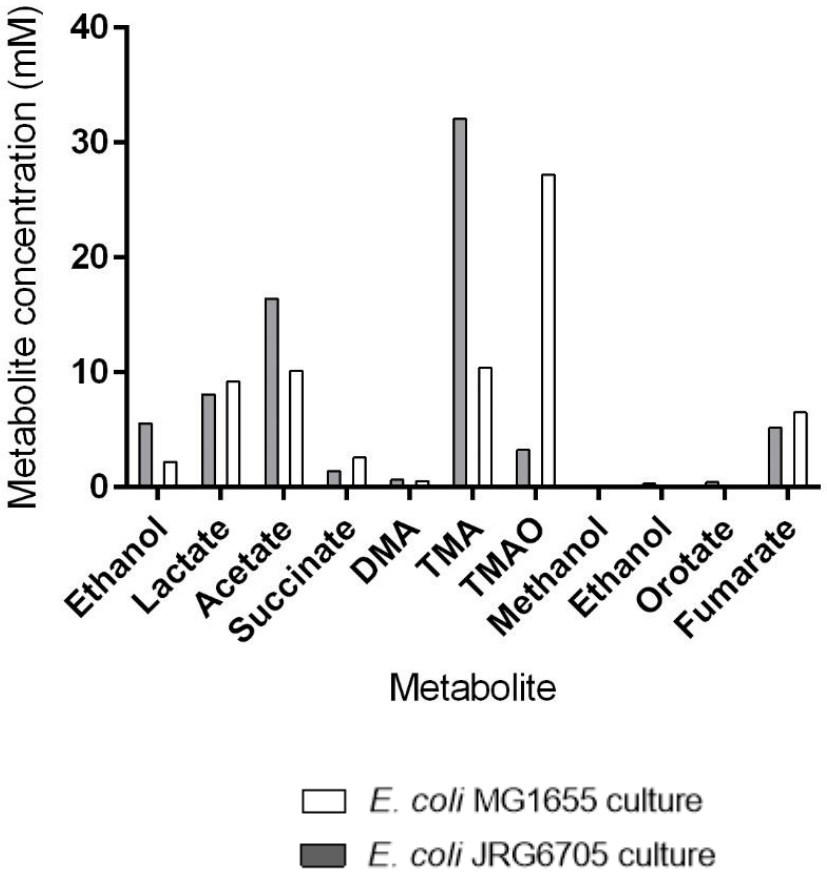


Figure 4.3. <u>Cultures of *E. coli* JRG6705 exhibit impaired TMAO reduction</u>

NMR was used to quantify the reduction of TMAO (~40 mM) after ~18 h of growth at 37°C under anaerobic conditions by cultures of both *E. coli* JRG6705 and wild type. The concentrations of the metabolites (mM) in the cell-free supernatants of both cultures are shown. Grey bars represent the metabolite profile of the *E. coli* MG1655 culture; white bars represent the *E. coli* JRG6705 culture. The metabolite concentrations are given for one experiment.

intensity for that cell was higher than that of the GFP-negative sample. By applying a gate in this way, it was seen that the number of individual cells expressing GFP increased the longer the culture was exposed TMAO. This is shown by an increase in the percentage of GFP-positive cells in the sample over time (Figures 4.4 and 4.5). However, after 60 min exposure to TMAO, it can be seen that only ~42% of the cells are expressing GFP, suggesting that approximately half of the population are not utilising TMAO. This is despite the signal (TMAO) persisting, due to decreased reduction of TMAO by the strain.

In order to further explore this, fluorescence microscopy was carried out on the same chemostat samples (Section 2.3.1). Samples were washed and fixed with 1% (v/v) paraformaldehyde to poly-lysine coated glass slides as described (Sections 2.14.1, 2.14.2 and 2.14.3). Individual cells from continuous cultures were visualised using both differential interference contrast (DIC) and fluorescence microscopy using a FITC filter set, to visualise GFP expression (Section 2.14.4). Samples withdrawn from the chemostat before the addition of TMAO did not have an observable GFP signal. It was not until 60 min after TMAO addition to the chemostat that GFP expression was observed (Figure 4.6). Expression of GFP appears to be uniform across the population, with every cell appearing to display similar intensities. However, expression of *gfp* appears to be is relatively low in these images, which leads to a "noisy" background, as there is little contrast between the GFP signal and the background fluorescence. Around ~200,000 molecules of GFP per cell are considered high enough for conventional one-colour detection even in samples with high background autofluorescence (Wendland and Bumann, 2002). The number of GFP molecules present in the samples studied here is likely to be low as a chromosomal fusion was used, leading to low contrast between background fluorescence and cellular expression of GFP.

4.4. Analysis of *torCAD* transcription in response to different concentrations of TMAO

Batch cultures of *E. coli* JRG6705 cells were grown in under anaerobic conditions in phosphate-buffered Evans medium at 37°C as previously described (Section 2.14.1). Cells were exposed to different concentrations of TMAO for 1 h before being analysed by fluorescence microscopy and flow cytometry to observe the expression of *torCAD* in individual cells in response to different concentrations of TMAO.

4.4.1. Expression of *torCAD* in response to low concentrations of TMAO is not uniform across the population

In order to explore whether cell-to-cell variability in *torCAD* expression is dependent on concentration of TMAO, *E.coli* JRG6705 cells were grown for 1 h in the presence of different

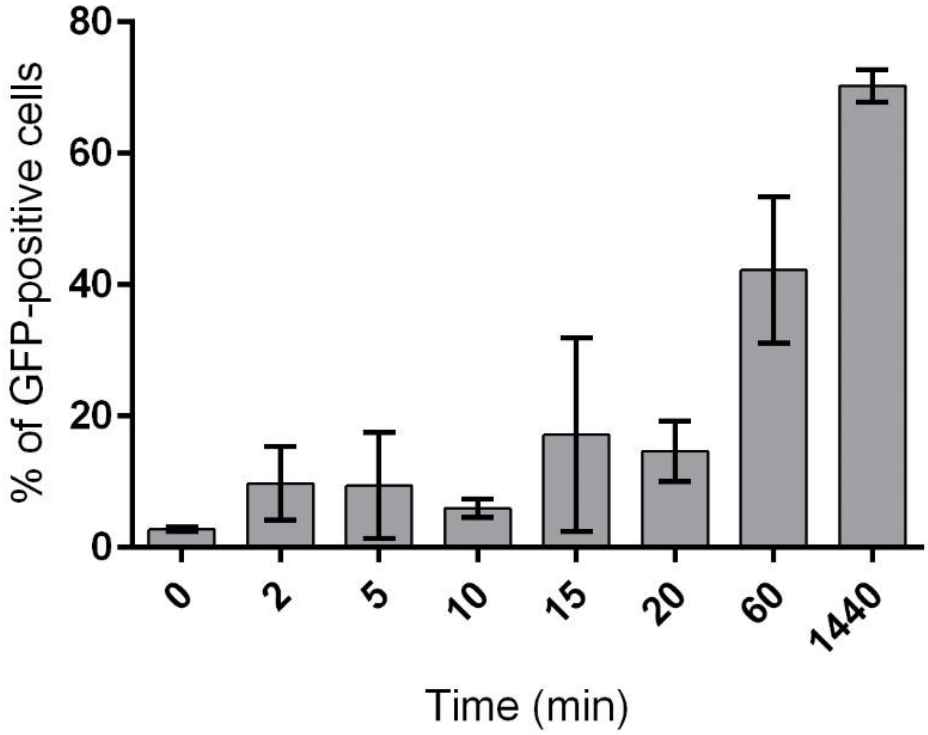


Figure 4.4. <u>The percentage of GFP-positive cells present in continuous cultures of *E. coli* JRG6705 in response to TMAO over time</u>

Using FlowJo analysis software, a GFP-negative sample (0 min) was used for gating, so that any cell detected by the BL1 filter as having a higher GFP signal than that of the GFP-negative sample was deemed to be expressing GFP (GFP-positive). The graph shows the % of GFP-positive cells present in each sample before (0 min) and after TMAO addition (2 – 1,440 min). The averages and standard deviations of three biological replicates are shown.

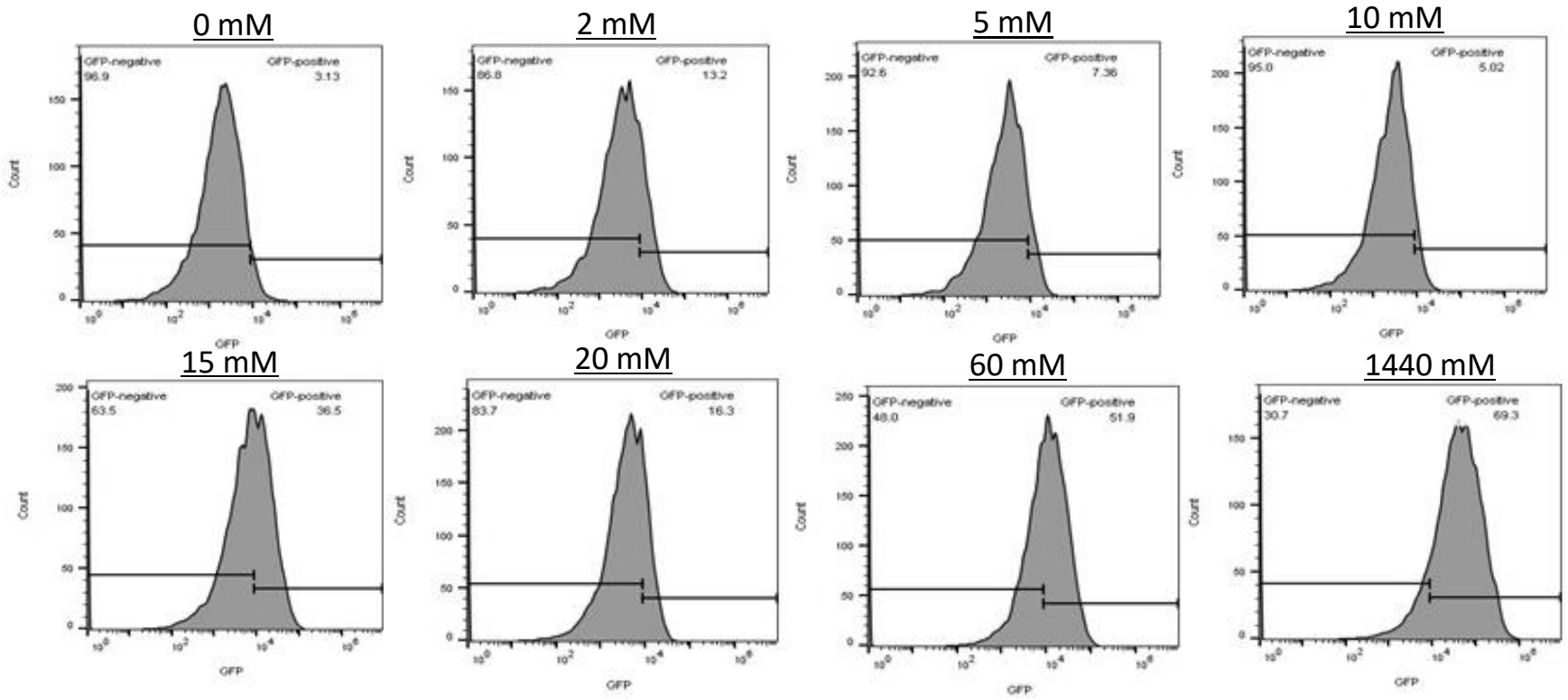


Figure 4.5. <u>Gated GFP-positive and GFP-negative sub-populations are present in continuous *E. coli* JRG6705 cultures exposed to TMAO</u>

FlowJo analysis was carried out on fixed cell samples (over time) of continuous cultures exposed to TMAO. A gate was placed on the 0 min sample (pre-TMAO), so that any cell with a higher GFP emission was classified as GFP-positive. Numbers shown are the percentage of cells that were deemed either GFP-positive or GFP-negative in the 10,000 cells sampled. Data shown are from one replicate that is representative of three biological repeats. Flow cytometry files and analysis for all three replicates are given on the provided 'Supplementary Data' CD.

Merged DIC and FITC images

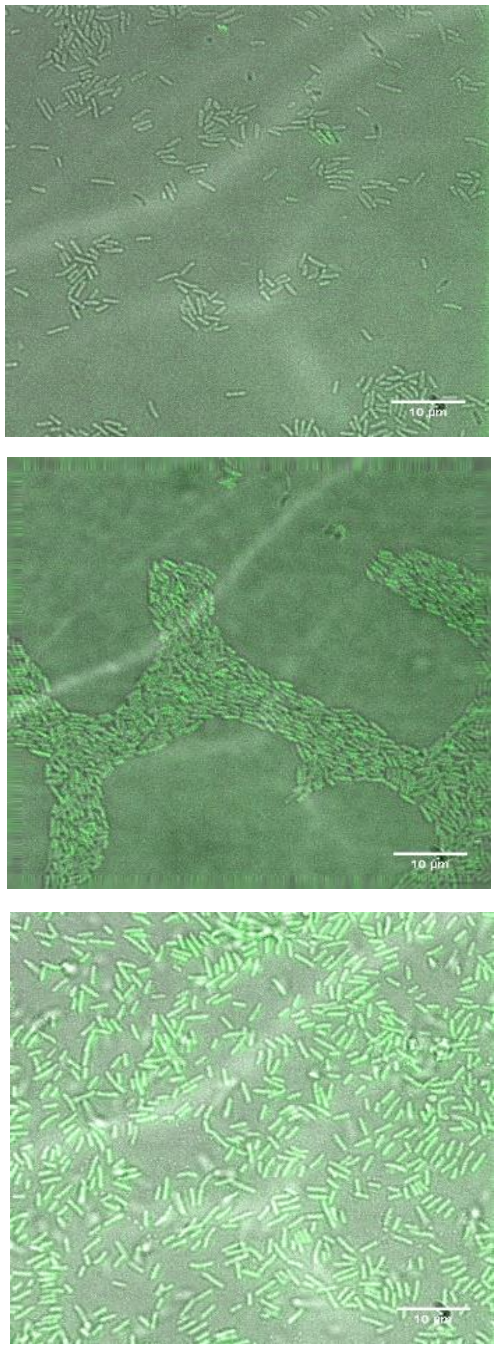


Figure 4.6. DIC and FITC images of samples from a continuous culture of *E. coli* JRG6705, showing expression of *torCAD* in response to TMAO over time

Anaerobic continuous cultures of *E. coli* JRG6705 ($P_{tor\text{-}gfp}$) were exposed to the alternative electron acceptor TMAO (~40 mM) under anaerobic conditions. Fixed cells from the culture were analysed using fluorescence microscopy to monitor expression of the *torCAD* operon in response to TMAO. DIC and FITC images are shown for one representative biological replicate. Scale bars in all images represent 10 µm. Image files for all three biological replicates are given on the provided 'Supplementary Data' CD.

concentrations of TMAO (0-40 mM) under anaerobic conditions before the levels of GFP expression in cultures were analysed by flow cytometry and fluorescence microscopy.

Flow cytometry experiments showed that when the GFP-positive gate was placed using the 0 mM TMAO (GFP-negative) sample, there was an increase in the number of GFP-positive cells present in cultures exposed to increasing concentrations of TMAO (Figure 4.7). When cultures were exposed to a lower concentrations of TMAO, i.e. 10 mM, only 7.7% of cells were deemed to be GFP-positive (92.3% being GFP-negative). Even after exposure to the highest concentration of TMAO (40 mM) for 1 h, two distinct sub-populations were present (GFP-positive and GFP-negative) in the sample. Only ~45% of *E. coli* JRG6705 cells were in the GFP-positive gate, thus less than half of the total population were expressing *torCAD* in response to a persistent concentration of TMAO (Figure 4.8).

When DIC and FITC microscopy was carried out on the same cultures, a number of individual cells were seen to not be expressing *torCAD* in response to the presence of 10 mM TMAO (Figure 4.9). This further confirms the presence of two sub-populations as shown by flow cytometry measurements. These subpopulations were also present when cells were grown in concentrations of TMAO lower than 40 mM (20-30 mM)(Figure 4.10). This is in contrast to cells grown in 40 mM TMAO for ~1 h, where the whole population appear to have a more uniform expression of *torCAD* (all cells are expressing GFP)(Figure 4.6).

4.5. Deletion of the *E. coli* K-12 *dam* gene results in an increase in the number of individual cells expressing *torCAD* expression in response to low concentrations of TMAO

As variation between the expression of *torCAD* in individual cells of a population exposed to lower concentrations of TMAO (2-10 mM) was observed by both flow cytometry and fluorescence microscopy, the mechanism responsible for this cell-to-cell variation was sought. Analysis of the intergenic region between *torC* and *torR* genes (divergent promoter region) revealed the presence of an *E. coli* DNA adenine methyltransferase (Dam) recognition site (Figure 4.11) overlapping the -10 element of P_{torC}.

In order to investigate the effect of methylation of the GATC site present in P_{tor} region of the *E. coli* MG1655 genome on the expression of *torCAD* operon, a commercially available *dam* mutant strain (*E. coli* ER2925, Table 2.1) was transduced with the $P_{tor\text{-}gfp}$ construct (from *E. coli* JRG6705) (Figure 4.12a and b) to create strain *E. coli* JRG6842 (Table 2.1). A *dam* mutant is unable to methylate DNA at GATC sites and thus, could promote transcription activation. If so, cultures of *E. coli* JRG6842 grown in the presence of TMAO under anaerobic conditions should uniformly express GFP (all cells green) even when low concentrations of TMAO are

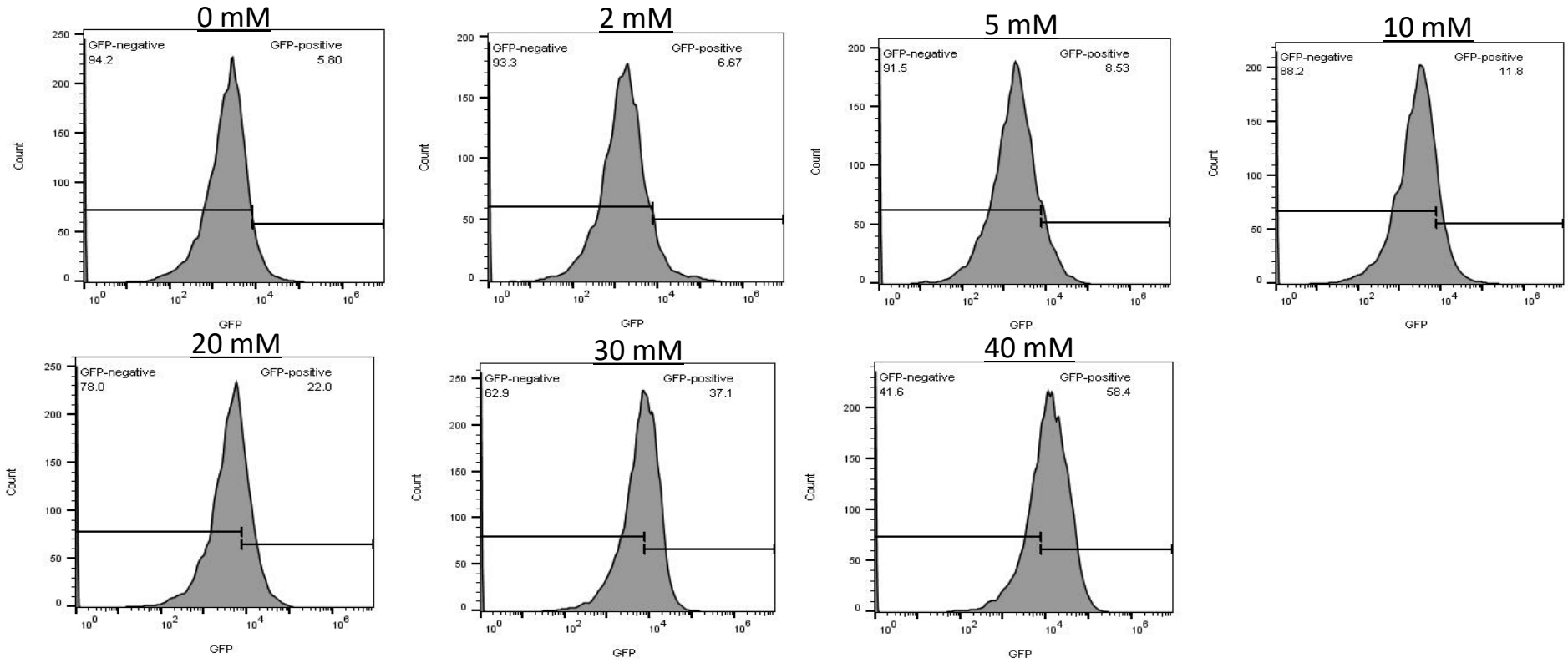


Figure 4.7. <u>The percentage of GFP-positive cells present in batch cultures of *E. coli* JRG6705 exposed to different concentrations of TMAO</u>

FlowJo analysis was carried out on fixed cells from cultures of *E. coli* JRG6705 exposed to different concentrations of TMAO (0-40 mM). A GFP-negative gate was placed using the 0 mM TMAO sample. Numbers shown are the percentage of cells that we deemed either GFP-positive or GFP-negative in 10,000 cells sampled. Data shown are from one replicate that is representative of 3 biological repeats. Flow cytometry files and analysis for all three replicates are given on the provided 'Supplementary Data' CD.

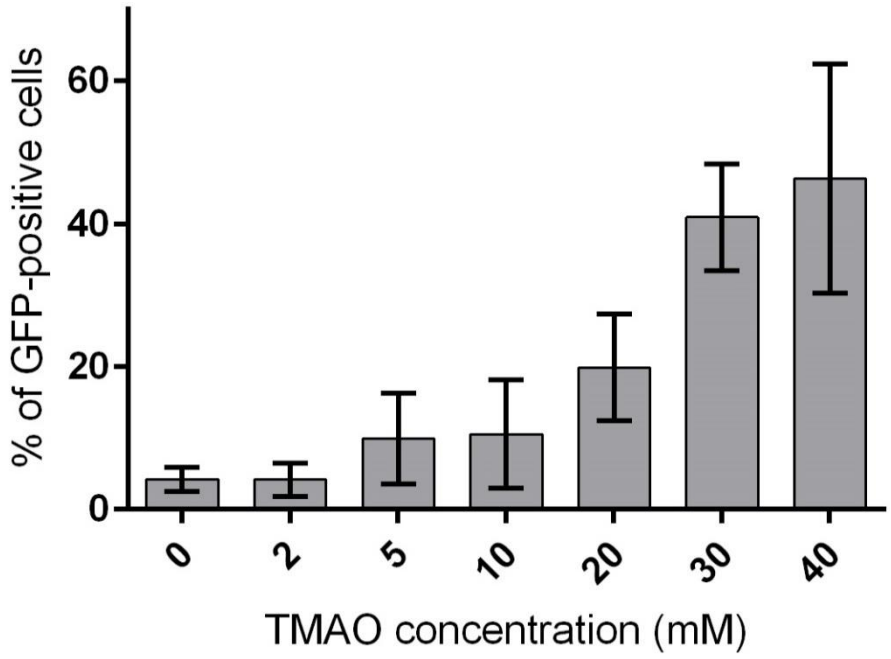


Figure 4.8. Gated GFP-positive and GFP-negative sub-populations present in batch cultures of *E. coli* JRG6705 exposed to different concentrations of TMAO

Using FlowJo software, a GFP-negative sample (0 mM TMAO) was used for gating so that any cell detected by the BL1 filter as having a higher GFP signal than that of the GFP-negative sample was deemed to be GFP-positive. The graph shows the percentage of GFP-positive cells present in each fixed sample of culture exposed to different concentrations of TMAO (0-40 mM) for 1 h. The average and standard deviation of three biological replicates are shown.

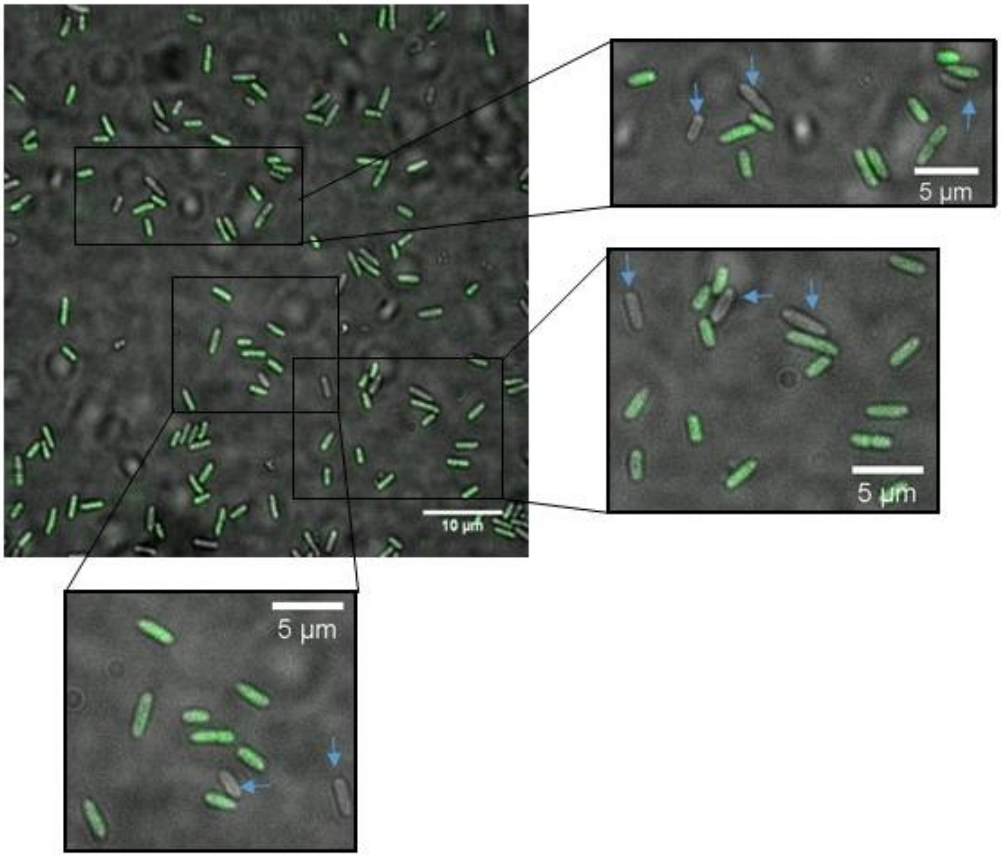


Figure 4.9. <u>Expression of *torCAD* by individual cells in response to TMAO (10 mM) is not uniform</u>

The DIC and corresponding FITC images of *E. coli* JRG6705 batch cultures grown in the presence of 10 mM TMAO under anaerobic conditions were merged. The black squares highlight the heterogeneous expression of *gfp* in individual cells in 3 fields of view. Scale bars are shown for each field of view.

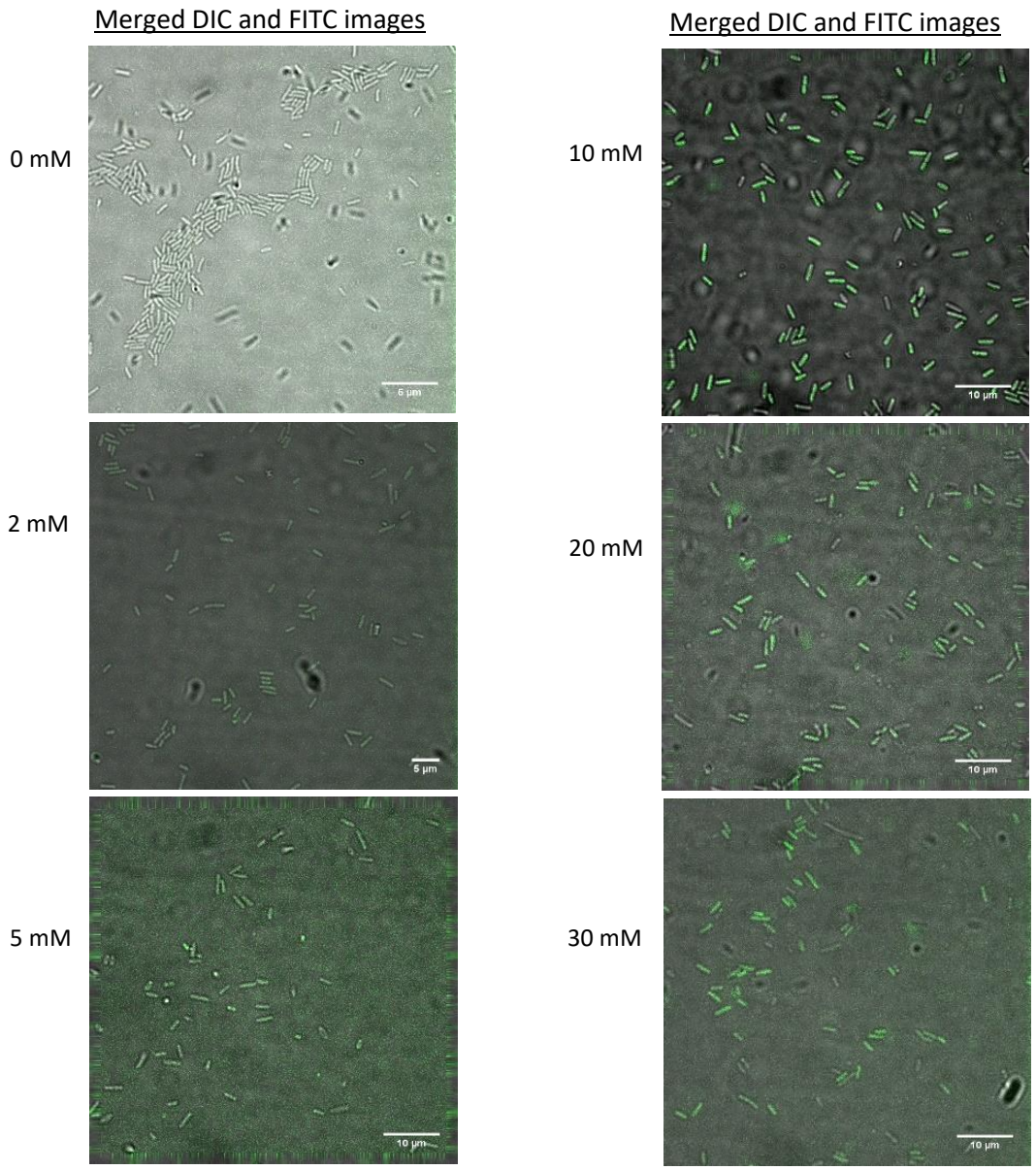


Figure 4.10. DIC and FITC images of batch cultures of *E. coli* JRG6705, showing the expression of *torCAD* ($P_{tor\text{-}gfp}$) in response to different concentrations of TMAO

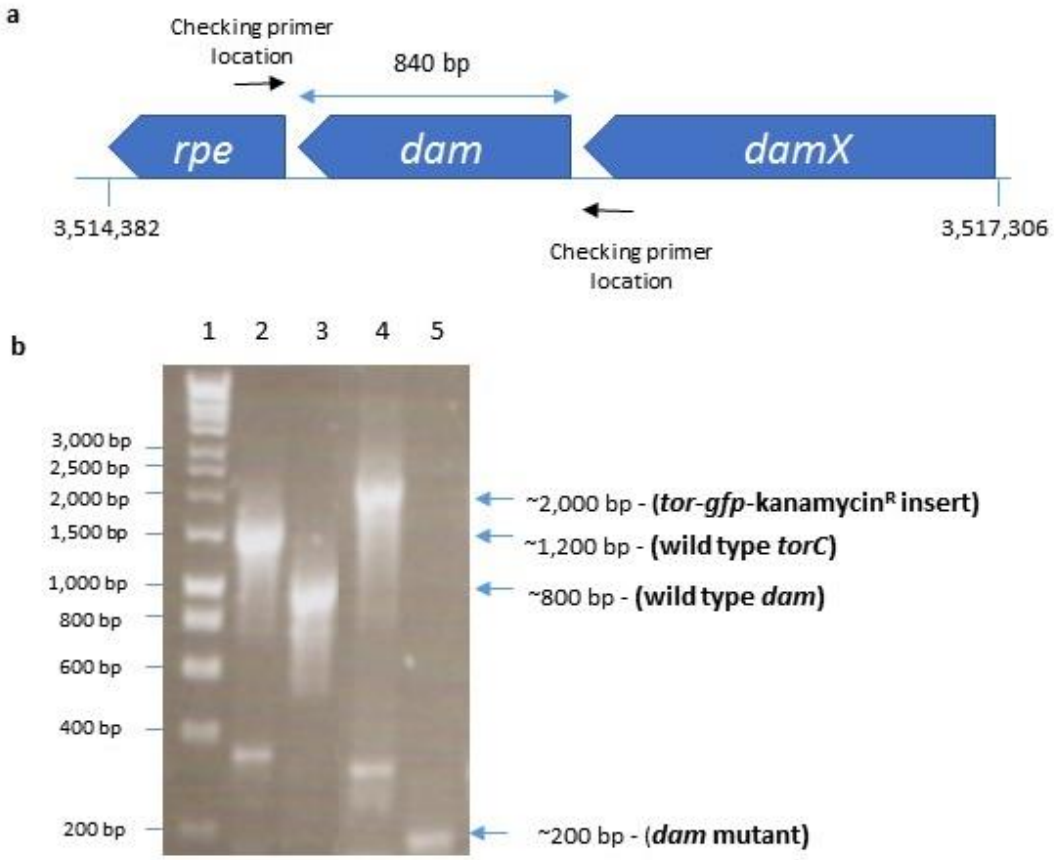
Anaerobic batch cultures of *E. coli* JRG6705 were exposed to different concentrations (0-30 mM) of the electron acceptor TMAO for 1 h before being fixed and analysed using fluorescence microscopy. DIC and FITC images are shown for one representative biological replicate. Scale bars in all images represent 10 µm. Image files for all three biological replicates are given on the provided 'Supplementary Data' CD.

TCAGAGGGTTTTACTCATTCTGTTCATATCTGTTCATATTCTGCCGTAAGCCGTTCATCCTGACCAGT

GCCGCTGTTCATATTTGCTCATTAAGATCGCTTCACTAAACCATAATTCTACAGGGGTTATTatg

4.11. Transcriptional features of the intergenic region between *torR* and *torC* on the *E. coli* MG1655 chromosome

The intergenic region between the divergent *torR* and *torC* genes on the *E. coli* MG1655 chromosome is shown with the transcriptional features highlighted. This region contains four direct repeats of a decametric consensus motif. These motifs are called *Tor* boxes (shown in red and underlined) and boxes 1, 2 and 4 have exactly the same decameric sequence (CTGTTCATAT). However, box 3 only matches 7 of the 10 bases of the motif (C**C**GTTCAT**CC**; bold bases represent mismatched bases of the motif) (Simon *et al.*, 1995). The transcriptional start site is shown (yellow). The -10 and -35 regions are underlined and the Dam recognition site is highlighted in green. The start codon of the *torC* gene is shown in green.



4.12a and b. <u>Confirmation of that the wild type dam gene is not present in the chromosome of *E. coli* JRG6843</u>

a. The position and size of the *E. coli* MG1655 wild type *dam* gene on the chromosome is shown, along with checking primers and where they anneal relative to the *dam* gene.

b. An agarose (1%) gel showing PCR reactions amplification using TorC-fusion seq F and R or dam check F and R primer pairs (Appendix 1) with either *E. coli* JRG6842 or *E. coli* MG1655 genomic DNA as the template. Lane 1, Hyperladder I (Bio Line); Lane 2, *E. coli* MG1655 genomic DNA as template and *tor* primer pair; Lane 3, *E. coli* MG1655 genomic DNA and *dam* primer pair; Lane 4, *E. coli* JRG6842 genomic DNA and *tor* primer pairs; Lane 5, *E. coli* JRG6842 genomic DNA and *dam* primer pair. Sizes of PCR products and fragments they represent are indicated by the arrows and sizes of DNA markers are indicated.

present, where expression of the reporter by wild type cultures is not uniform across the population.

Flow cytometry was carried out on *E. coli* JRG6842 cultures exposed to either no TMAO or different concentrations of TMAO (10 or 40 mM) for 1 h. When compared to previous wild type cultures grown under the same conditions (Section 4.4.1), *dam* mutant cultures have a higher percentage of individual cells expressing *torC* in response to lower concentrations of TMAO (10 mM) (Figure 4.13). The difference in the percentage of cells is highly significant (p <0.01). However, the number of individual cells expressing *torC* in response to a higher concentration of TMAO (40 mM) is comparable in wild type and *E. coli* JRG6842 cultures.

4.6. Discussion

The work reported in this chapter was carried out in order to determine whether the mixed metabolism observed in continuous cultures of wild type *E. coli* in the presence of sub-optimal concentrations of TMAO, are a result of the whole population of cells carrying out fermentation and TMAO respiration simultaneously or whether two distinct sub-populations carrying out only one of these modes of metabolism are present (Figure 4.1).

Here, work suggests that there are distinct two sub-populations of cells that either do or do not initiate the transcription of the *torCAD* operon in response to the presence of TMAO under anaerobic conditions. Expression of *torCAD* in response to persistent excess concentrations of TMAO (40 mM) is not uniform. Thus, the single-cell distributions of *torCAD* transcription in response to TMAO under anaerobic conditions are not uniform, particularly when cultures are exposed to low concentrations of TMAO. This is similar to previous reports for cultures of *E. coli* grown in the presence of TMAO under aerobic conditions (Roggiani and Goulian, 2015).

Initially, continuous cultures of *E. coli* JRG6705 were grown and maintained as previously described for the wild type strain. Fluorescence microscopy suggested that the expression of GFP occurred ~15 min and onwards after the addition of TMAO and was uniform (all cells expressing GFP) across the culture. However, when GFP expression was analysed using flow cytometry, only ~40% of cells were GFP-positive 60 min after the addition of TMAO. This suggests that only a proportion of the cells in the cultures from the original experiments (Section 3.2) were initiating *torCAD* transcription, even when excess TMAO was present in the environment (Table 3.1). Previously, it was shown that *E. coli* K-12 respires TMAO in both the presence and absence of oxygen (Ansaldi *et al.*, 2007).

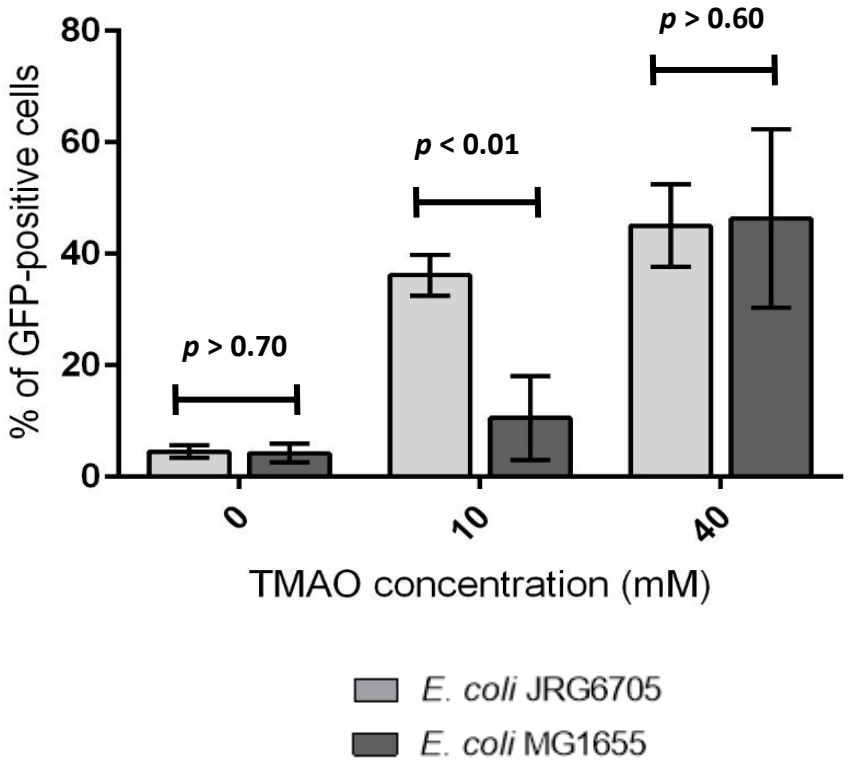


Figure 4.13. <u>The percentage of GFP-positive cells with a *dam* mutation is increased when exposed to lower concentrations of TMAO</u>

The percentage of GFP-positive cells is shown for both *dam* mutant (light grey) and wild type (dark grey) cultures when exposed to different concentrations of TMAO for 1 h is shown. The graph is the average of three biological replicates and their standard deviations. The *p* values are shown between wild type and *dam* mutants exposed to the same concentration of TMAO. The wild type bars are the same as those shown in Figure 4.8. Flow cytometry files and analysis for all three replicates are given on the provided 'Supplementary Data' CD.

In more recent work, it was shown that the average level of *torCAD* transcription in cultures containing TMAO grown in both anaerobic and aerobic conditions was similar (Roggiani and Goulian, 2015). However, *torCAD* transcription in aerobic cultures was found to be highly heterogeneous between individual cells, compared to a relatively homogeneous increase in fitness that could be due to the production of alkaline TMA by TMAO reduction counteracting acidification of the environment (Ansaldi *et al.*, 2007). Even if only a subset of cells in a population is performing the TMAO reduction, this mechanism could be effective (Roggiani and Goulian, 2015). Another hypothesis reflects a "bet-hedging strategy" whereby the cost of expressing the TMAO respiratory chain under aerobic conditions is balanced against the benefit of pre-expressed TorCAD should the environment rapidly become anaerobic. However, both these hypotheses apply to variability under aerobic conditions and here variability is seen under anaerobic conditions.

The Dam enzyme encoded by the *dam* gene is responsible for the methylation of nearly all GATC sites in the *E. coli* genome (Geier and Modrich, 1979; Ringquist and Smith, 1992; Wang and Church, 1992). The methylation state of bacterial promoter region can affect whether transcription of the corresponding gene occurs. This plays a role in the regulation of a number of genes in *E. coli* (Løbner-Olesen *et al.*, 2003; Oshima *et al.*, 2002; Riva *et al.*, 2004), including *dnaA* (Campbell and Kleckner, 1990). Expression of various genes in other bacterial species have been found to be reliant on DNA methylation by Dam, for example, a *Salmonella typhimurium dam* mutant was found to be totally avirulent (Heithoff *et al*., 1999). The presence of a GATC sequence in a promoter could affect expression of the gene through inhibition or promotion of repressor or activator binding (Marinus and Løbner-Olesen, 2014). DNA methylation can also influence chromosome replication by regulating the recruitment of the initiator of replication to the origin of replication (Boye and Løbner-Olesen, 1990). Also, DNA methylation is a post-replicative process, which means that newly synthesised strands of DNA are undermethylated and older strands are fully methylated (Marinus, 1976). Thus, DNA methylation can act as a regulator of DNA repair by informing the methyl-directed mismatch repair system about the age of the DNA (Glickman and Radman, 1980).

The Dam methylase transfers a methyl group from *S*-adenosylmethionine to the N6 position of the adenine nucleotide in the sequence GATC (Geier and Modrich, 1979; Marinus and Morris, 1973). As *E. coli dam* mutants are viable this suggests that no essential genes are solely dependent on Dam methylation (Marinus and Løbner-Olesen, 2014). Rather, Dam methylation can be considered to play a role in the fine tuning of basal levels of gene expression either by acting directly at GATCs in regulatory sequences or acting as a timing

switch or in an epigenetic fashion (Marinus and Løbner-Olesen, 2014). A small fraction of the approximately 20,000 GATC sites in the *E. coli* chromosome are totally or partially non-methylated in any given growth state and environmental condition (Casadesús and Low, 2006). The ability of Dam to methylate GATC sites is dependent upon competition between both Dam and specific DNA binding proteins. The binding of proteins primarily at gene regulatory regions result in DNA methylation patterns, which are altered by growth conditions that affect regulatory protein level(s) and/or DNA binding properties (Casadesús and Low, 2006). The *E. coli* ER2925 strain (Table 2.1) has several chromosomal mutations, including mutations in the two main DNA methylase enzymes (Dam and Dcm methylases). These mutations render the strain unable to methylate DNA. The $P_{tor\text{-}gfp}$ reporter construct was transduced into the chromosome of *E. coli* ER2925 and variability in individual cells' expression of *torCAD* in response to TMAO was analysed. Here, it was seen that the number of individual cells expressing *torCAD* in response to lower concentrations of TMAO (10 mM) was increased in a *dam* mutant cultures compared to wild type cultures (Figure 4.13). However, this increase compared to wild type was not observed when cultures were exposed to 40 mM TMAO. This suggests that methylation of the *torCAD* promoter may be involved in the regulation of the operon. It may be involved in the fine tuning of *torCAD* expression, which is more pronounced in cultures exposed to lower concentrations of TMAO. An increase in percentage is not seen in cultures grown in higher TMAO concentrations, which could mean that the maximum expression of *torCAD* is occurring in the wild type and thus an increase in expression in a *dam* mutant is not observable.

It is possible that a number of mechanisms, as well as DNA methylation are responsible for the cell-to-cell variation in *torCAD* expression described here. Variability in *torCAD* expression could result from changes in numbers of TorS and TorT molecules present in individual cells (Roggiani and Goulian, 2015). Ribosome profiling data were used to infer protein abundance and indicates that both TorT and TorS are expressed at low levels (~4 molecules/cell for both) (Li *et al.*, 2014). It is feasible that stochastic effects may arise from different cellular abundances of TorS and TorT across the population, produced by random partitioning during cell division and gene expression (Elowitz *et al.*, 2002; Huh and Paulsson, 2011; Ozbudak *et al.*, 2002). For example, if the number of TorS molecules happens to exceed the number of TorT molecules in an individual cell, then excess TorS that is not bound to TorT will be insensitive to TMAO (Roggiani and Goulian, 2015). This would lead to the dephosphorylation of TorR~P via reverse phosphotransfer (Ansaldi *et al.*, 2001), thus producing lower levels of TorR~P. It has previously been observed that overexpression of

torS represses the transcription of the *torCAD* operon (Ansaldi *et al.*, 2001). Cells that lack TorS molecules will also have low levels of TorR~P because of both spontaneous hydrolysis and dilution of pre-exsisting TorR~P from cell growth (Roggiani and Goulian, 2015). Thus, it can be expected that cells with no TorS (or a TorS/TorT ratio too low) or when TorS numbers exceed TorT (a TorS/TorT ratio too high) should have lower than average TorR~P and thus the system is sensitive to changes in the numbers of both proteins (Roggiani and Goulian, 2015). These stochastic changes could also mean that there is a maximum number of cells in a population that can carry out TMAO reduction. This could explain why an increase in the percentage of cells expressing *torCAD* in response to a high concentration of TMAO (40 mM) is not seen in a *dam* mutant, when it is in cultures exposed to lower concentrations of TMAO (10 mM). However, whilst both the stochastic changes and the role of *dam* in *torCAD* regulation described here are likely to contribute to the variability in *torCAD* expression, at present, other mechanisms that might produce cell-to-cell variability cannot be ruled out.

5. The mechanism of formaldehyde sensing by the *E. coli* transcriptional regulator FrmR

5.1. Introduction

The detrimental effects of toxic chemicals must be withstood by all organisms. Many organisms generate the toxic chemical formaldehyde endogenously (Chen *et al*., 2016; Yu *et al*., 2015). Formaldehyde is generated: (i) as an intermediate of methylotrophic metabolism; (ii) in the degradation of glycine, either by the glycolytic byproduct methylglyoxal or by Fenton chemistry; (iii) in the degradation of haem, during iron acquisition by some Gram-positive bacteria; (iv) by lipid peroxidation of sugars; (v) by the demethylation of histones; (vi) as a product of methylated DNA repair by AlkB; and (vii) by the action of *N*-methyltryptophan oxidase (SolA)(Chen *et al*., 2016; Koyama and Ohmori, 1996; Trewick *et al*., 2002). As a consequence, biological systems are often exposed to both endogenous and exogenous sources of formaldehyde.

The toxic effects of formaldehyde are mediated via the modification of essential cell components, including proteins and DNA, thus leading to cellular dysfunction. The formation of DNA-DNA and DNA-protein cross-links, as well as covalent DNA monoadducts are genotoxic effects mediated by formaldehyde (Chen *et al*., 2016; Lai *et al*., 2016; Yu *et al*., 2015). Formaldehyde is also able to covalently modify proteins, inhibiting their functions (Metz *et al*., 2004; Xie *et al*., 2016). The broad chemical reactivity of formaldehyde, which leads to life-threatening damage, has driven the evolution of mechanisms to detoxify formaldehyde, in order to counteract its detrimental effects (Chen *et al*., 2016). It is important to maintain specific response regulators in the cell that can sense the presence of formaldehyde, in order to regulate expression of these detoxification systems, to induce protection before significant damage occurs.

When *E. coli* K-12 adapts to the presence of the alternative electron acceptor TMAO, formaldehyde is produced, which results in the induction of the *frmRAB* operon, as shown by work previously described here (Chapter 3). An inability to respond (by induction of the *frmRAB* operon) to this endogenously produced formaldehyde resulted in growth inhibition, rather than growth promotion, when anaerobic cultures of *E. coli* were exposed to TMAO (Section 3.6). The *frmRAB* operon encodes a regulator protein, FrmR; a formaldehyde dehydrogenase, FrmA; and an *S*-formylglutathione hydrolase, FrmB (Herring and Blattner, 2004). In *E. coli* and many other organisms, fomaldehyde present in the cytosol reacts with the major reductant glutathione, yielding *S*-hydroxymethylglutathione, which is then oxidised by FrmA to *S*-formylglutathione, before being hydrolysed to non-toxic formate and glutathione by FrmB (Gonzalez *et al*., 2006; Harms *et al*., 1996). *S*-Nitrosoglutathione, a

product of nitrosative stress, is also a substrate for FrmA (Liu *et al*., 2001). Therefore, the proteins coded by the *frmRAB* operon permit the bacterium to sense and detoxify formaldehyde.

As previously stated, the first gene of the *E. coli frmRAB* operon codes for a regulatory protein, FrmR. *Escherichia coli* FrmR is a member of the RcnR/CsoR family of transcriptional repressors (Giedroc and Arunkumar, 2007; Higgins and Giedroc, 2014). Many characterised members of this family are metal-ion sensors; the properties of Cu(I) (CsoR, RicR), Ni(II) (InrS) and Ni(II)/Co(II) (RcnR, DmeR) sensors in diverse bacterial species have been reported, including those of the sulfite/sulfide sensor CstR (Foster *et al*., 2014; Grossoehme *et al*., 2011; Iwig *et al*., 2008; Iwig *et al*., 2006; Liu *et al*., 2007). Members of the family whose role is metal-sensing possess characteristic amino acid signatures (W-X-Y-Z fingerprint), which is associated with metal binding. CsoR binds Cu(I) via the X-Y-Z residues, whilst RcnR and InrS use all four positions to coordinate their cognate metal (Figure 5.1a). Whilst the Ni(II)/Co(II)-responsive and Cu(I)-responsive family members are relatively well-characterised, little is known about signal perception and DNA-binding mechanisms of the remaining proteins (FrmR and CstR), although they retain at least two residues of the W-X-Y-Z fingerprint. All family members share one absolutely conserved residue (Cys35) and one highly conserved residue (His60) within the W-X-Y-Z fingerprint (Iwig *et al*., 2008; Ma *et al*., 2009). Although amino acid residue variation within the fingerprint is implicated in signal specificity by coupling the metal-coordination preferences to the network connecting the metal- and DNA-binding residues of the protein (Higgins *et al*., 2012); the roles of specific residues in formaldehyde sensing are unclear.

Salmonella enterica serovar Typhimurium FrmR (*Se*FrmR) retains two of the metal-binding ligands (Cys35 and His60, with Glu instead of His64 and Pro instead of His2) that are present in the *S. enterica* Ni(II)-binding regulator RcnR (Osman *et al*., 2015). Recently, *in vitro* studies showed that *Se*FrmR binds Co(II), Cu(I) and Zn(II), but the binding affinities were weaker than those of the biological sensors of these metal ions in *Salmonella* (Osman *et al*., 2015). Thus, *Se*FrmR was outcompeted by CueR, RcnR, ZntR and ZurR for their cognate metals. Furthermore, *in vivo Se*FrmR was able to act as a formaldehyde-responsive transcription regulator, but was unable to act as a metal ion sensor. Even more recent work has shown that formaldehyde-sensing by *Se*FrmR is direct, specific and requires two residues of the W-X-Y-Z fingerprint (Pro2 and Cys35A; Figure 5.1b) (Osman *et al.*, 2016). The crystal structure of an E64H mutant was solved and revealed that the protein consists of a homotetrameric

a

Group	Signature				Ligand/Signal
	W	X	Y	Z	
RcnR	His	**Cys**	His	His	**Ni(II)/Co(II)**
InrS	His	**Cys**	His	Cys	**Ni(II)**
CsoR	X	**Cys**	His	Cys	**Cu(I)**
FrmR	Pro	**Cys**	His	Thr/Glu	**formaldehyde**
CstR	X	**Cys**	X	Cys	**sulfide**

b

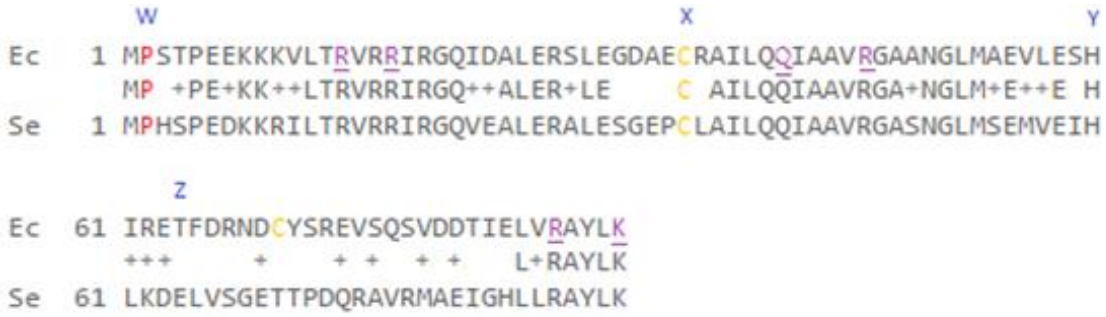


Figure 5.1a and b. The 'metal-binding' fingerprints of members of the CsoR/RcnR family and amino acid sequences of *E. coli* FrmR and *S. enterica* FrmR

a. Amino acid residues in the W-X-Y-Z fingerprint of CsoR/RcnR family proteins and the signals perceived by the indicated proteins.

b. Alignment of the *E. coli* (Ec) and *S. enterica* FrmR proteins (Se). Identical (single letter code) and similar (+) residues, Pro2 (red font), Cys35 and Cys70 (yellow font) are indicated. Residues of the W-X-Y-Z fingerprint (blue font) are indicated. Residues in purple font and underlined have been implicated in DNA-binding in other CsoR/RcnR family members.

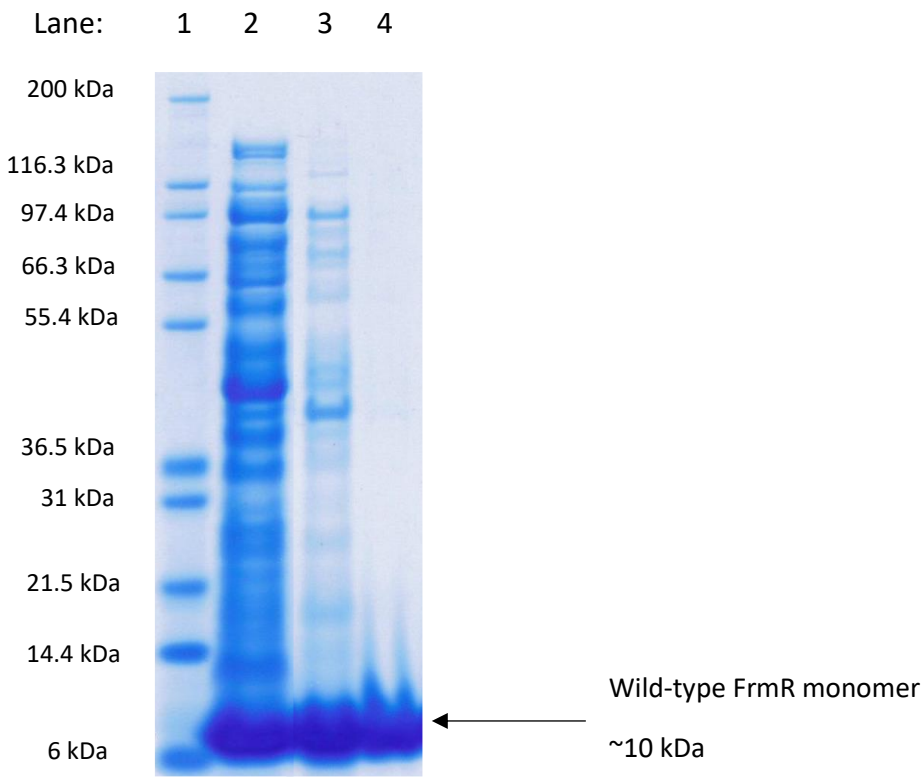
disc with a positively-charged surface region, which is predicted to participate in DNA binding (Osman *et al.*, 2016). Amino acid sequence alignment of the *E. coli* K-12 and *Se*FrmR proteins reveals strong conservation in the first 63 amino acids (67% identical, 89% similar) but weaker conservation thereafter (21% identical; 39% similar; Figure 5.1b). *Escherichia coli* FrmR retains Cys35 and His60 residues, but has a Thr residue at position 64, as well as an additional Cys residue (Cys70) that is absent from *Se*FrmR (Figure 5.1b). The biological imperative to mount an effective response to both exogenous and endogenous sources of formaldehyde, as evidenced by the observation that the *frmRAB* operon was essential for adaptation of *E. coli* to growth on TMAO and the differences between the *E. coli* and *Se*FrmR proteins, prompted an investigation of *E. coli* FrmR.

This is worthy of investigation because the mechanism(s) used by regulatory proteins to perceive and respond to formaldehyde are very poorly understood. Importantly, FrmA acts upon the GSH-formaldehyde adduct, raising the question of the identity of the molecular signal to which *Ec*FrmR responds and the mechanism of signal perception. Hence, the aim of the work reported here was to provide new insight into the mechanism used by biological systems to sense the presence of the highly reactive and toxic chemical, formaldehyde. *In vivo* and *in vitro* data show that FrmR senses formaldehyde directly with no metal-dependence via the formation of methylene bridges between Pro2 and Cys35 residues of adjacent FrmR subunits.

5.2. Properties of the isolated *E. coli* FrmR

The transcriptional regulator, FrmR and its mutants were purified from the overexpression strains *E. coli* JRG6782 (wild type), JRG6783 (C35A) and JRG6784 (P2A) by Dr Svetlana E. Sedelnikova (Department of Molecular Biology and Biotechnology, University of Sheffield, UK) (Figure 5.2a) (Section 2.10.3.1). The volume at which FrmR and its mutants were eluted from a calibrated size exclusion chromatography column was ~76 ml, which was indicative of a homotetramer (with a molecular weight of ~40 kDa) (Figure 5.2b), which is typical of other members of the CsoR/RcnR family (Iwig *et al.*, 2008). The concentration and authenticity of purified FrmR was then verified by total amino acid analysis (Section 2.10.9). Determining the concentration of purified FrmR using the A_{280}/A_{320} (Section 2.10.4.2) was found to underestimate the actual concentration of FrmR, as measured by total amino acid analysis, by 7.6%. Therefore, the concentration of FrmR was determined using A_{280}/A_{320} multiplied with a correction factor of 1.08. Liquid chromatography-mass spectrometry (LC-

a



b

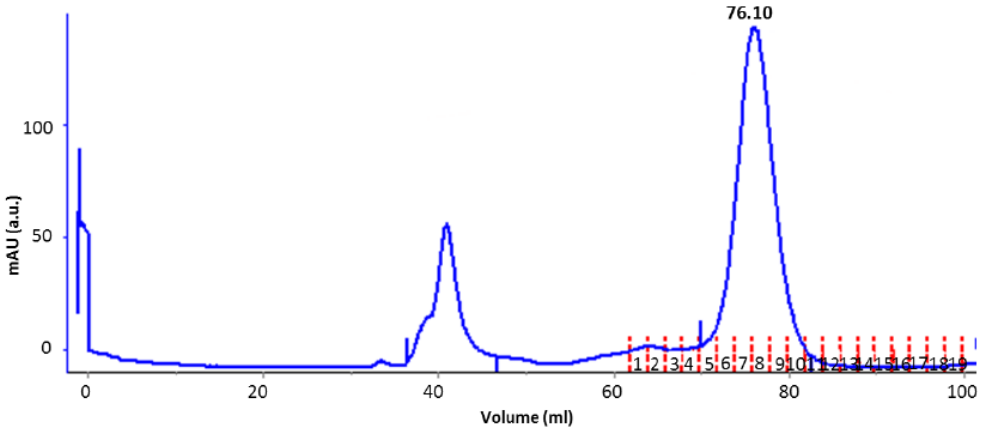


Figure 5.2a and b. <u>Purification of FrmR protein</u>

a. A representative analysis by SDS-PAGE of the purification of wild-type FrmR protein. Lane 1, Mark12^{TM} unstained standards (Thermo Scientific), sizes are shown; Lane 2, cell-free extract of induced culture; Lane 3, after HiTrap Heparin HP purification; Lane 4, after HiLoad Superdex 200 gel fitration (final purified FrmR preparation).

b. Absorbance was monitored at 280 nm for all fractions eluted from the gel filtration column (fractions are shown by the red dashed lines). An increase in absorbance (at 280 nm) occurred at a volume of 76.1 ml, which was indicative of a homotetramer of ~40 kDa.

MS) (Section 2.10.10) showed that the isolated protein lacked an *N*-terminal methionine (which was also confirmed by *N*-terminal amino acid sequencing carried out by Dr Jeff Iwig (Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, USA)) and had the expected monomeric molecular mass of 10186.5 Da. The isolated wild type FrmR protein was found to be metal-free as judged by ICP-MS analysis (Section 2.10.11).

In order to determine the number of free sulfhydryl groups in the purified protein, reaction with DTNB (Section 2.10.8) was carried. The number of reactive thiols per monomer of wild type FrmR was found to be 1.86 ± 0.21 (expected number of reactive thiols was 2.00 as two cysteine residues are present in the monomer). However, the number of reactive thiols per monomer decreased to 1.02 ± 0.08 upon aerobic storage of FrmR (>48 h). Therefore, this suggests that the protein can adopt an oxidized form with two intermolecular disulfide bonds per tetramer.

Recent research has shown that the *Se*FrmR protein is able to bind Zn(II) (Osman *et al*., 2015). Intrinsic fluorescence experiments were carried out in order to investigate the potential Zn(II)-binding properties of *E. coli* FrmR, (Section 2.10.9). Titration of FrmR with Zn(II) resulted in changes in intrinsic fluorescence at 304 nm (Figure 5.3). The specific binding isotherm obtained after subtraction of the linear non-specific binding phase, indicated that the reduced FrmR tetramer was able to bind ~4 atoms of Zn(II) (Figure 5.3), reaching saturation after the [Zn(II)]:[tetrameric FrmR] ratio exceeded four. This response is indicative of high affinity Zn(II) binding, due to its non-hyperbolic nature. Equivalent titrations with Mn(II) were carried out. In contrast to Zn(II) binding, titrations with Mn(II) resulted in non-specific binding response, which did not reach saturation even at a [Mn(II):[tetrameric FrmR] ratio of 19 (data not shown).

5.3. FrmR binds directly to the *frmRAB* promoter to repress transcription

Electrophoretic mobility shift assays (EMSAs) were attempted with purified FrmR protein and frmRAB promoter DNA. However, the procedure could not be optimised and never led to any reliable results, which was similar to what collaborators had experienced at Durham University. Thus, Bio-Layer Interferometry (BLItz) measurements were used to investigate the interactions between the FrmR protein and *frmRAB* promoter DNA (P_{frm}). The K_d for FrmR binding to immobilised P_{frm} DNA was ~220 nM, arising from an overall on-rate constant (k_f) of ~13000 M^{-1} s^{-1} and an overall off-rate constant (k_r) of 0.003 s^{-1} at 20°C (Figure 5.4a, Table 5.1). This interaction was specific because the K_d for FrmR binding to an immobilised

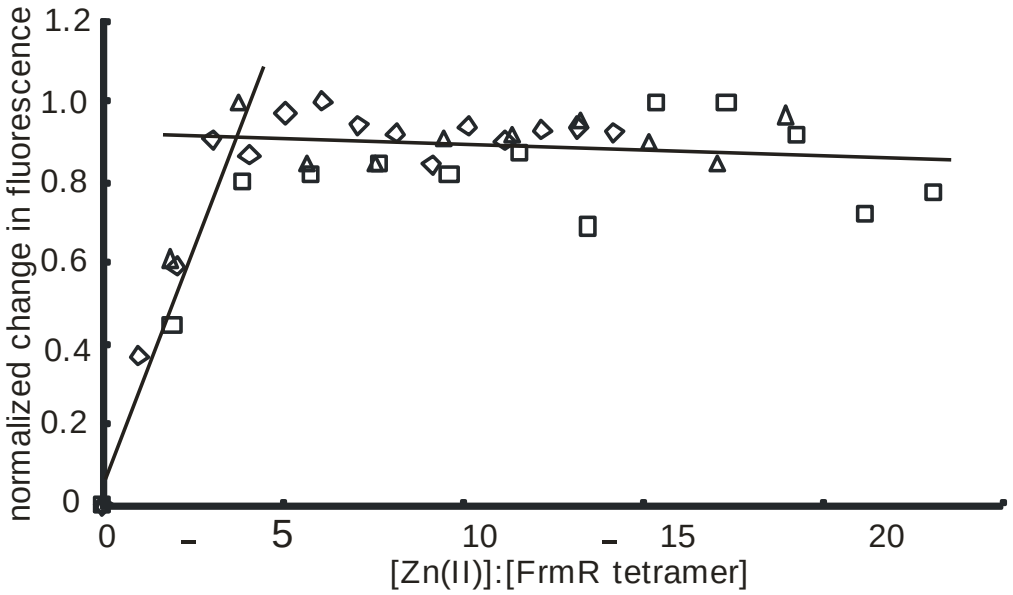


Figure 5.3. The *E. coli* FrmR tetramer binds Zn(II)

Changes in intrinsic fluorescence for three titrations (indicated by different symbols) were normalised and plotted against the Zn(II):FrmR tetramer ratio. The solid lines show linear regression fits for the data points during the phase when fluorescence was changing and for the titration end points.

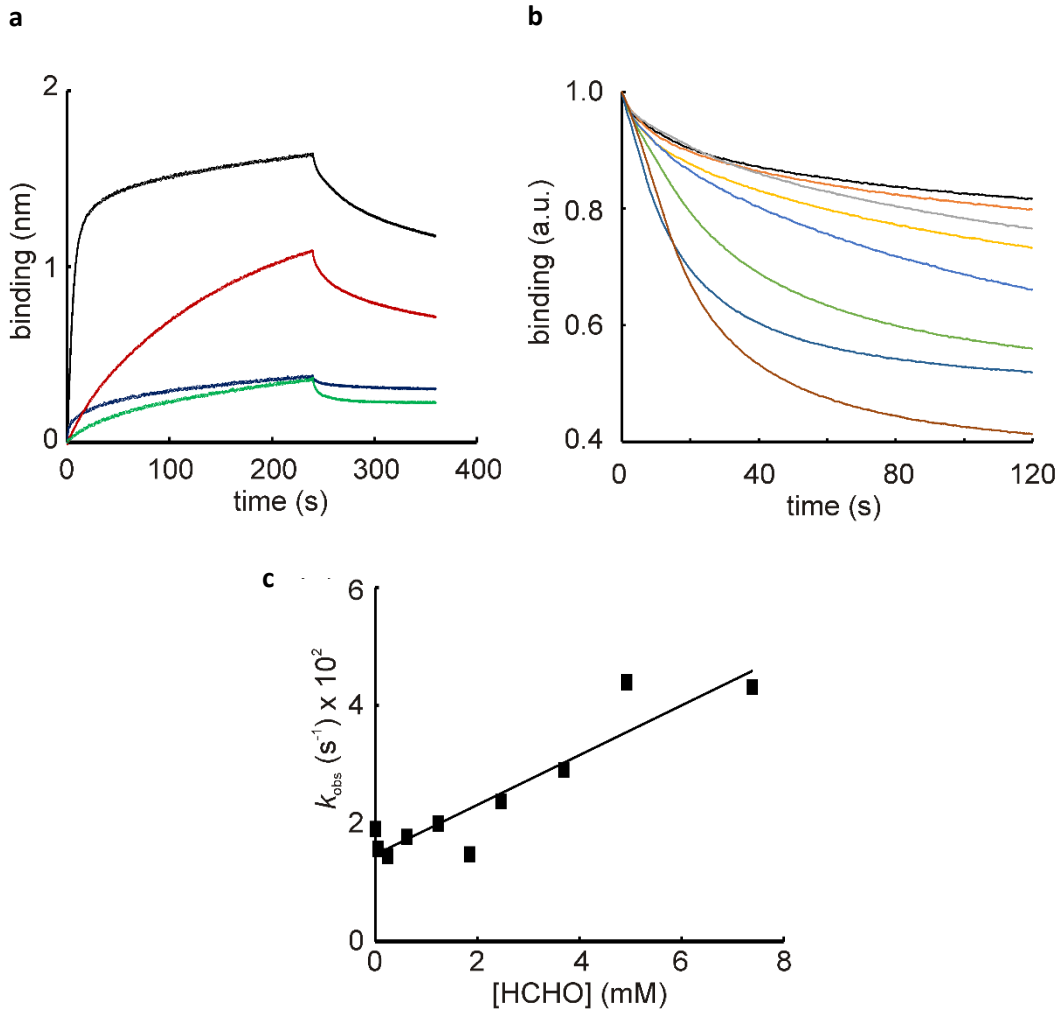


Figure 5.4a, b and c. **Formaldehyde enhances dissociation of the P_{frm}-FrmR complex**

a. Bio-Layer Interferometry (BLItz) assays. Traces for the interaction between biotin-labeled P_{frm} DNA immobilised on a streptavidin probe with FrmR (6.16 μM tetramer, black line, 0.88 μM tetramer; red line), FrmR (0.88 μM tetramer) pre-treated with 200-fold molar excess of formaldehyde (blue line), and FrmR binding at a non-target DNA (P_{ydhY}, 0.88 μM FrmR tetramer; green line) are shown. Analysis of the data was carried out using the BLItz Pro 1.1.0.31 software using global fitting and correcting both association and dissociation curves.

b. Pre-formed P_{frm}-FrmR complexes were exposed to different concentrations (0, black; 0.05 mM, orange; 0.25 mM, grey; 0.62 mM, yellow; 1.25 mM, blue; 3.69 mM, green; 4.92 mM, dark blue; 7.38 mM, brown) of formaldehyde and dissociation curves were recorded.

c. Single exponential fits to formaldehyde dissociation curves were used to obtain the observed rate constants (k_{obs}) which were plotted against formaldehyde concentration to obtain the apparent second order rate constant (k_{app}).

Table 5.1. Rate constants for FrmR and DNA interactions

Reaction	k_f (M^{-1} s^{-1})[a]	k_r (s^{-1})[a]	K_d (nM)	k_{app} (M^{-1} s^{-1})
FrmR + P_{frm} ↔ P_{frm}-FrmR	13000 ± 390	0.0028 ± 0.000086	220	-
FrmR-$(Zn)_4$ + P_{frm} ↔ P_{frm}-FrmR-$(Zn)_4$	5660 ± 165	0.003 ± 0.000078	520	-
FrmR + P_{ydhY} ↔ P_{ydhY}-FrmR	1950 ± 1000	0.007 ± 0.00017	3600	-

Reaction	k_f (M^{-1} s^{-1})[a]	k_r (s^{-1})[a]	K_d (nM)	k_{app} (M^{-1} s^{-1})
P_{frm}-FrmR + HCHO → FrmR-HCHO + P_{frm}	-	-	-	4.2
P_{frm}-FrmR-$(Zn)_4$ + HCHO → FrmR-$(Zn)_4$-HCHO + P_{frm}	-	-	-	0.7

[a] Value and standard error.

Using the BLItz Pro software, rate constants obtained for the reactions shown. In order to find the apparent second order rate constant (*k*app), the BLItz data from dissociation curves were fitted to a single exponential function to obtain the observed first order rate constants (*k*obs), which were then plotted against formaldehyde concentration.

unrelated *E. coli* promoter (*ydhY*) fragment was ~3600 nM (Table 5.1). Pre-treatment of FrmR with formaldehyde for 3 min abolished binding to P_{frm} (Table 5.1). Exposure of the pre-formed P_{frm}-FrmR complex to increasing concentrations of formaldehyde resulted in the dissociation of the complex (Figure 5.4b). The speed of the dissociation of the complex increased with the presence of increasing concentrations of formaldehyde. The dissociation curves were fitted to a single exponential function and dissociation of the P_{frm}-FrmR complex exhibited a linear dependence on formaldehyde concentration, with a rate constant of ~4 M^{-1} s^{-1} at 20°C (Figure 5.4c, Table 5.1). Furthermore, *in vitro* transcription reactions showed that the synthesis of the *frmRAB* transcript was inhibited in the presence of FrmR and that this inhibition was relieved when FrmR was treated with formaldehyde (Figure 5.5). *In vitro* transcription reactions with the *E. coli ndh* gene showed that transcription of this gene was not affected by the presence of either FrmR or formaldehyde. Therefore, the *in vivo* and *in vitro* data show that FrmR is a repressor of *frmRAB* expression, which specifically responds to formaldehyde leading to the dissociation of the FrmR-P_{frm} complex, allowing the expression of the formaldehyde detoxification system.

5.3.1. The role of Zn(II) in FrmR binding to P_{frm} DNA

The ability of the *E. coli* FrmR and SeFrmR to bind Zn(II) (Section 5.2 and (Osman *et al*., 2015)) raises the possibility that formaldehyde sensing is metal-dependent. Therefore, the effect of Zn(II) on the ability of *Ec*FrmR to bind to DNA (P_{frm}) was assessed by BLItz (Table 5.1). Pre-treatment of the *Ec*FrmR tetramer with four molar equivalents of Zn(II) increased the K_d for binding at P_{frm} ~2-fold due to a decrease in the rate of DNA-binding (Table 5.1). However, in response to formaldehyde exposure the dissociation of the pre-formed Zn(II)-loaded P_{frm}–FrmR complex was ~7-fold lower than that observed in the absence of Zn(II), suggesting that Zn(II) blocks amino acid residues required for formaldehyde sensing. Thus, the presence of Zn(II) would be antagonistic to the expression of the detoxification machinery. Therefore, formaldehyde-sensing is not metal- (Zn(II)-) dependent and, it is likely that Zn(II) could only inhibit *in vivo E. coli* FrmR activity under conditions when Zn(II) homeostasis is severely perturbed (Figure 5.6; Table 5.1)(Osman *et al*., 2015).

5.4. Identification of *E. coli* FrmR residues necessary for formaldehyde sensing

Site-directed mutagenesis of the amino acids of the W-X-Y-Z fingerprint has revealed the importance of these residues for the function of CsoR/RcnR family proteins (Figure 5.1a). Previous β-galactosidase activity measurements carried out by Dr Jeff Iwig (Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, USA)) showed that two mutants were unable to respond to formaldehyde, FrmR(P2A) and

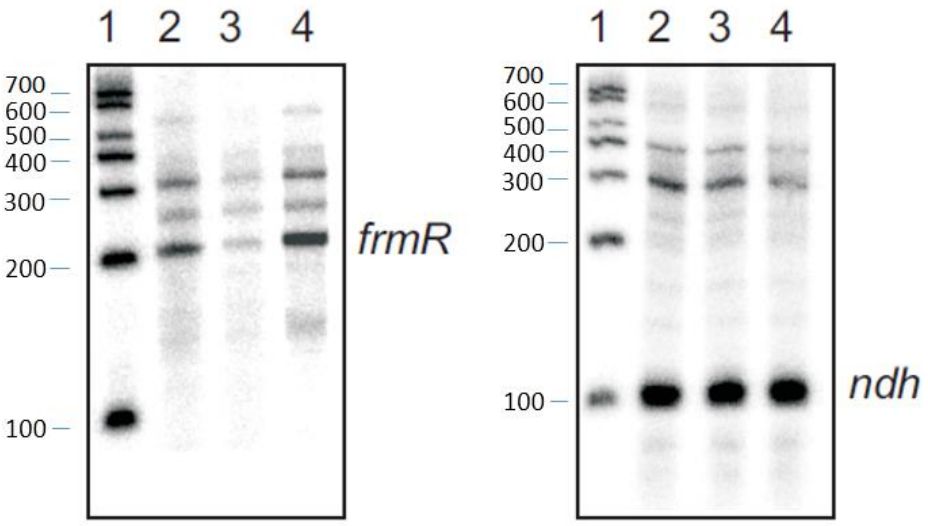


Figure 5.5. <u>Inhibition of *frmRAB* transcription by FrmR *in vitro* is relieved by formaldehyde</u>

In vitro transcription reactions for FrmR with P_{frm} and P_{ndh} DNA. Left panel, P_{frm}; right panel, P*ndh*. Lane 1, RNA size markers, sizes are shown; Lane 2, no FrmR; Lane 3, 1 nM FrmR tetramer; Lane 4, 1 nM FrmR tetramer pre-treated with 200-molar excess formaldehyde. The locations of the *frmR* and *ndh* are indicated. The contents of the lanes are the same for *frmR* and *ndh* gels. FrmR was reduced with 10 mM dithiothreitol and when required treated with 200-fold molar excess of formaldehyde for 5 min at room temperature, before quenching with 10 mM glycine. After this incubation, transcription was initiated by the addition of 2 µl solution containing UTP at 50 µM; ATP, CTP and GTP at 1 mM; and 2.5 µCi of [α-^{32}P]UTP (800 Ci mmol^{-1}, PerkinElmer Life Sciences), followed by incubation for 15 min at 37°C in the presence of the formaldehyde-treated or untreated FrmR protein.

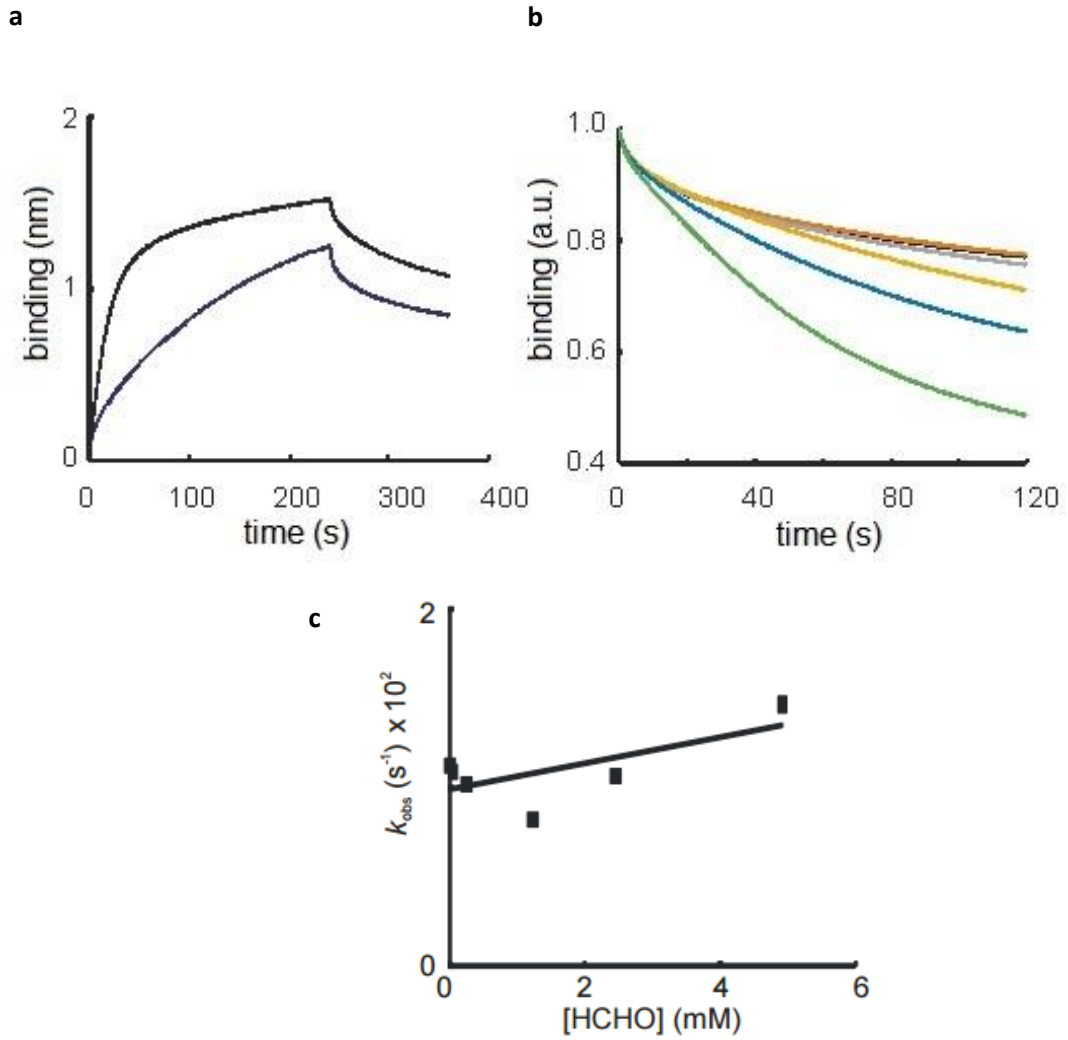


Figure 5.6a, b and c. <u>Zn(II)-loaded FrmR exhibits an impaired response to formaldehyde</u>

a. Bio-Layer Interferometry (BLItz) assays. Traces for interaction of biotin-labelled P_{frm} DNA immobilized on a streptavidin probe with FrmR (2.5 µM tetramer, black line) or FrmR(Zn(II))$_4$ (2.5 µM tetramer; dark blue line) are shown. Analysis of the data was carried out using the BLItz Pro 1.1.0.31 software using global fitting and correcting both association and dissociation curves.

b. Pre-formed P_{frm}-FrmR(Zn(II))4 complexes were exposed to different concentrations (0, black; 0.05 mM, orange; 0.25 mM, grey; 1.23 mM, yellow; 2.46 mM, blue; 4.92 mM, green) of formaldehyde and dissociation curves were recorded.

c. Single exponential fits to the data shown in (**b**) were used to obtain the observed rate constants which were plotted against the concentration of formaldehyde to obtain the apparent second order rate constant.

FrmR(C35A), suggesting that Pro2 and Cys35 (W and X positions in the CsoR/RcnR family fingerprint; Figure 5.1a) are essential for FrmR to perceive formaldehyde. Additionally, replacement of the only other cysteine residue (Cys70) in FrmR did not impair the response to formaldehyde. Based on this work, which suggested that Pro2 and Cys35 are required for FrmR to respond to formaldehyde, growth experiments were carried out with *E. coli* strains expressing FrmR(P2A) (JRG6825) and FrmR(C35A) (JRG6826) in place of wild type FrmR (Section 2.5). Cultures expressing these variants exhibited enhanced sensitivity to formaldehyde when compared to cultures expressing wild type FrmR (JRG6726) (Figure 5.7). This *in vivo* data is consistent with the low *frmRAB* expression observed in the reporter fusion experiments described above, which presumably arises from permanent repression *frmRAB* expression (Figure 5.7).

5.5. Reaction of the *E. coli* FrmR with formaldehyde

Liquid chromatography-mass spectrometry (LC-MS) was used to analyse the modification of FrmR by formaldehyde. In the absence of any formaldehyde, the major species corresponded to the FrmR monomer lacking the *N*-terminal methionine (10,186.60 Da; predicted mass 10,186.5 Da), with a disulfide-linked dimer (20,371.02 Da; predicted mass 20,373.2 Da) present in lower amounts (Table 5.2). FrmR was exposed to formaldehyde (8-fold molar excess) for 3 min and LC-MS carried out. This revealed a new species of molecular mass 20,396.87 Da, corresponding to an FrmR dimer plus an additional mass of 23.67 Da, along with a species corresponding to a tetramer with an additional mass of 22.30 Da (40768.73 Da) and the FrmR momomer (10,187.02 Da) (Table 5.2). When FrmR was incubated with Zn(II) (4 Zn(II) per FrmR tetramer), before or after formaldehyde treatment, dimeric species of molecular masses 20,396.99 Da and 20,396.95 Da were detected, respectively. The changes in mass described here are consistent with the formation of formaldehyde-mediated FrmR dimers linked by intermolecular methylene bridges (-CH_2-) (net mass gain of 2 x 12 Da per FrmR dimer). The formation of a modified tetramer (40,768.73 Da) upon exposure to fomaldehyde suggests that each subunit of the FrmR tetramer participates in only one methylene bridge (total of two in the tetramer) and that these cross-linked dimers are held together by one disulfide bond (Table 5.2).

As FrmR(P2A) and FrmR(C35A) mutants were unable to respond to formaldehyde *in vivo* (Figure 5.7), formaldehyde modification of the purified mutant proteins was examined. Both variants of the protein eluted in a volume indicative of a tetramer upon size exclusion chromatography, suggesting that the oligomeric state of the protein was not affected by either mutation. In either the absence or presence of formaldehyde, LC-MS showed that the

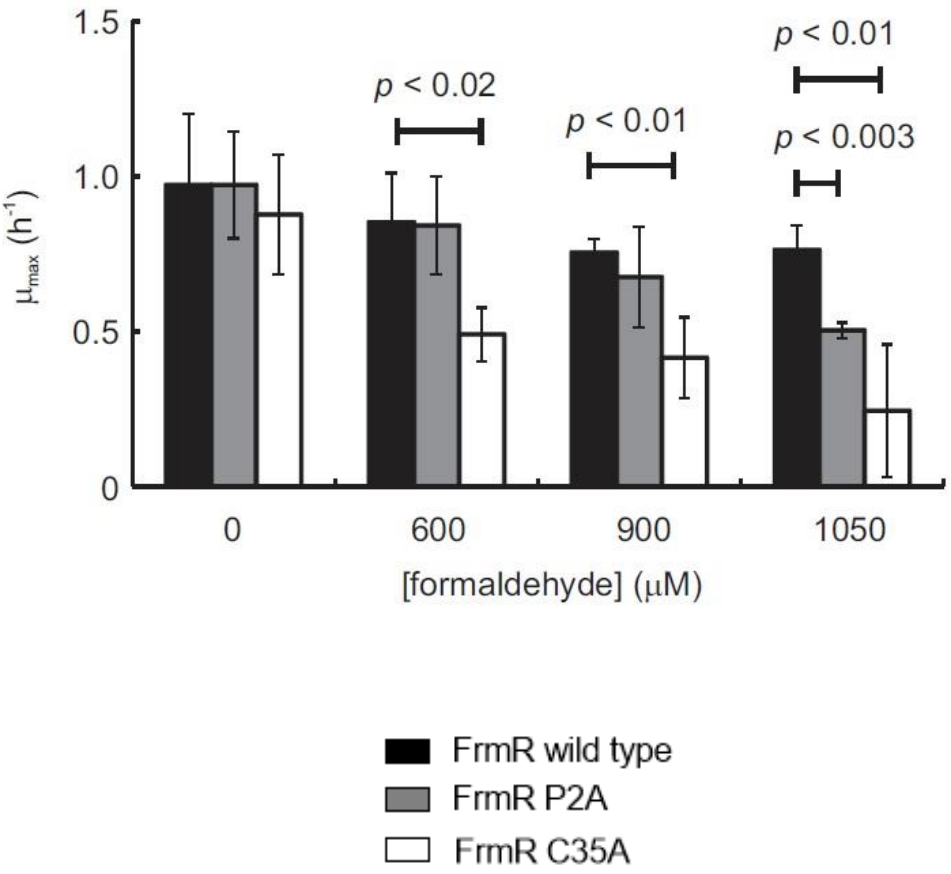


Figure 5.7. <u>Identification of formaldehyde-insensitive FrmR protein variants *in vivo*.</u>

Maximum growth rates (μ_{max}) of *E. coli* MG1655 *frmRAB* transformed with plasmids expressing the *frmRAB* operon from P_{frm} under the control of wild type FrmR (closed bars), FrmR(P2A) (grey bars) or FrmR(C35A) (open bars) cultured in the presence of the indicated initial concentrations of formaldehyde. The mean and standard deviations (n=3) and the *p* values for one-tailed t-tests are shown.

Table 5.2. Liquid chromatography mass spectrometry analyses of FrmR proteins treated with formaldehyde

Protein sample	Measured mass (Da)	Relative abundance (a. u.)	Mass difference (Da)[a]	Comment[b]
FrmR wild type				
FrmR	10,186.60	2.0×10^7	0.1	FrmR monomer
	20,371.02	4.4×10^6	-2.18	FrmR disulfide-linked dimer
HCHO-treated *FrmR*	10,187.02	1.8×10^6	0.52	FrmR monomer
	20,396.87	0.2×10^6	23.67	FrmR dimer with 2 methylene bridges
	40,768.76	2.0×10^5	22.30	Two FrmR dimers, each with 1 methylene bridge, forming a tetramer by 1 disulfide bond
HCHO-treated *FrmR* plus Zn(II)	10,187.01	1.8×10^6	0.51	FrmR monomer
	20,396.95	0.3×10^6	23.75	FrmR dimer with 2 methylene bridges
Zn(II)-treated *FrmR* plus HCHO	10,187.04	1.8×10^6	0.54	FrmR monomer
	20,396.99	0.4×10^6	23.79	FrmR dimer with 2 methylene bridges
FrmR P2A				
HCHO-treated *FrmR*(P2A)	20,319.69	4.0×10^5	-1.11	FrmR(P2A) disulfide-linked dimer
HCHO-treated *FrmR*(P2A)[c]	10,190.71	6.75×10^5	29.91	Hydroxymethylated FrmR(P2A) monomer
FrmR C35A				
HCHO-treated *FrmR*(C35A)	20,307.40	2.8×10^6	-1.40	FrmR(C35A) disulfide-linked dimer
HCHO-treated *FrmR*(C35A)[c]	20,307.21	1.7×10^5	-1.59	FrmR(C35A) disulfide-linked dimer

[a]Difference between predicted mass and measured mass

[b]Interpretation of mass difference; disulfide bond, -2.2 Da; methylene bridge, net mass gain 12 Da; hydroxymethylation, net mass gain 30 Da

[c]Exposure to formaldehyde without quenching

FrmR and its variants were analysed by liquid chromatography-mass spectrometry before and after treatment with formaldehyde.

disulfide-linked FrmR(P2A) dimers (20,319.69 Da; predicted mass 20,320.8 Da) was present (Table 5.2). When FrmR P2A was exposed to formaldehyde without quenching, a momomeric FrmR(P2A) species (10,190.17 Da) with an additional 29.91 Da, equivalent to an hydroxymethyl adduct was detected. However, for FrmR(C35A) a disulfide-linked dimer (20,307.4 Da; predicted mass 20,308.8 Da) was detected, suggesting this mutant protein was not modified in the presence of formaldehyde with or without quenching (Table 5.2). The LC-MS data suggests that the first step in formaldehyde perception by *E. coli* FrmR is reaction with Cys35 (Table 5.2) and is consistent with a mechanism in which FrmR senses formaldehyde by the formation of two methylene bridges between Pro2 and Cys35 residues of adjacent subunits, which is supported by *in vivo* data (Section 5.4). Also, FrmR is oxidized in air, forming dimers linked by intermolecular disulfide bonds.

5.6. Discussion

In vivo and *in vitro* experimental data presented here show that the *E. coli* FrmR transcriptional repressor specifically reacts with the toxic chemical formaldehyde, resulting in the formation of two inter-subunit methylene bridges between the *N*-terminal Pro2 and Cys35 residues of tetrameric FrmR and dissociation of the P_{frm}-FrmR complex.

The ability of the *E. coli* FrmR to bind Zn(II) was also investigated. Here, it is shown that four atoms of Zn(II) can bind to a FrmR tetramer (Section 5.2), which is consistent with other members of the CsoR/RcnR family members (Foster *et al*., 2012; Iwig *et al*., 2008; Ma *et al*., 2009). Recently, *Se*FrmR was shown to bind 4 atoms of Zn(II) (Osman *et al*., 2015). A $K_{Zn(II)}$ of the Zn(II) binding sites in *Se*FrmR was estimated to be ~1.7 x 10^{-10} M, which suggests that Zn(II) affinity of FrmR is much weaker than the cognate Zn(II) sensors in *S. enterica* (ZntR and Zur) (Osman *et al*., 2015). The presence of Zn(II) was shown to impair DNA-binding by FrmR and also slow down dissociation of the P_{frm}-FrmR complex in response to formaldehyde (Section 5.3.1). Metal sensors of the CsoR/RcnR family only bind DNA in the absence of their cognate metal(s), thus, these results contrast those (Foster *et al*., 2012; Iwig *et al*., 2006; Liu *et al*., 2007). Recently, it was revealed that the binding of Cu(I) to its cognate sensor CsoR from *Streptomyces lividans* results in the constriction of the tetramer, via a bulge in the α2 helices (Porto *et al*., 2015). Dissociation of the CsoR-DNA complex occurs in response to this reduced flexibility of the protein (Porto *et al*., 2015). Thus, it could be hypothesised that *E. coli* FrmR become more rigid when Zn(II) is bound, leading to impaired DNA binding. The conserved Cys35 and His60 residues (Figure 5.1b) are the metal-binding ligands in many members of the CsoR/RcnR family and thus, are likely to be involved in Zn(II) binding (Iwig *et*

al., 2008; Liu *et al*., 2007). As the Cys35 has been implicated in formaldehyde sensing here, the involvement of the Cys35 residue in Zn(II) binding could explain the reduced response to formaldehyde by Zn(II)-bound FrmR.

Alongside the experiments described here, the crystal structure of formaldehyde-exposed FrmR was solved to a resolution of 2.7 Å (PDB: 5LBM) by Dr Claudine Bisson (Department of Molecular Biology and Biotechnology, University of Sheffield, UK). FrmR oligomerises to form a disc-like tetramer, like other members of the CsoR/RcnR family of metal-sensors, which is constructed from two homodimers, each of which forms one face of the disc. Each of the subunits was found to consist of three helices (α1, residues 2-30; α2, residues 35-68; and α3, residues 73-91), which were linked by two short loops (L1, residues 31-34; and L2, residues 69-72) that are arranged as a flattened S-shape (Figure 5.8). The α3 helices slot together at the homodimer interface, such that each face of the tetramer is formed from a platform of five parallel helices with a hole at the centre (Figure 5.9a). This arrangement differs from that seen in CsoR and *Se*FrmR(E64H), where the equivalent α3 helix is domain swapped onto the opposite face of the tetramer (Chang *et al*., 2014; Osman *et al.*, 2016). This difference between *E. coli* FrmR and *Se*FrmR(E64H) could be a consequence of the differences in amino acid sequence in the region spanning the terminus of α2, loop 2 and the beginning of α3 (Figure 1b).

Two different dimers were observed in the formaldehyde-treated FrmR tetramer. An unmodified homodimer with a disordered *N*-terminal region (residues Pro2-Lys8) comprised one face of the tetramer (A/B face). There was no electron density corresponding to Zn(II) or other metal ions. However, located in similar positions in 3D space to those of the CsoR proteins from *Geobacillus thermodenitrificans* and *Thermus thermophilus* the W-X-Y-Z fingerprint residues in the unmodified homodimer are likely to constitute the FrmR Zn(II) binding site (Chang *et al*., 2014; Sakamoto *et al*., 2010). The other face of the tetramer (A´/B´ face) is sandwiched against the A/B face. In this face, electron density is present for residues Pro2-Lys8 in each of the subunits, which forms an extension to the *N*-terminal end of α1. The Pro2´-N atom is located within ~2.5 Å of the Cys35 atom of the corresponding subunit on the opposite face of the tetramer (Figure 5.9a), which supports the hypothesis that these residues are the site of crosslinking for the reaction with formaldehyde. In further support of this, extra electron density between these two atomic positions was observed, indicating the presence of the formaldehyde cross-link and a methylene bridge (-CH_2-) (Figure 5.9b). Thus, the structure of the crosslinked FrmR agrees with LC-MS data, which shows the formation of formaldehyde-mediated FrmR dimers linked by intermolecular methylene

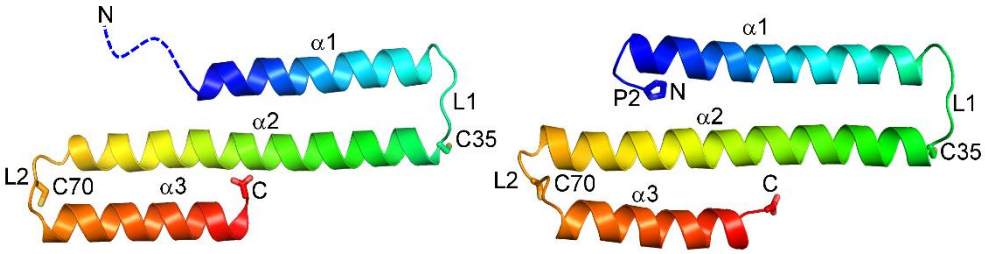


Figure 5.8. Structure of *E. coli* FrmR

Cartoon representations of uncross-linked (left) and cross-linked (right) FrmR monomers coloured blue (*N*-terminal) to red (*C*-terminal). Secondary structure elements (α-helices, α1 to α3; loops, L1 and L2) are labelled and the amino acid residues (single letter code, P2, C35 and C70) that are involved in cross-linking and disulfide bond formation are shown as sticks. The disordered *N*-terminal region in the uncross-linked subunit is represented by the blue dashed line.

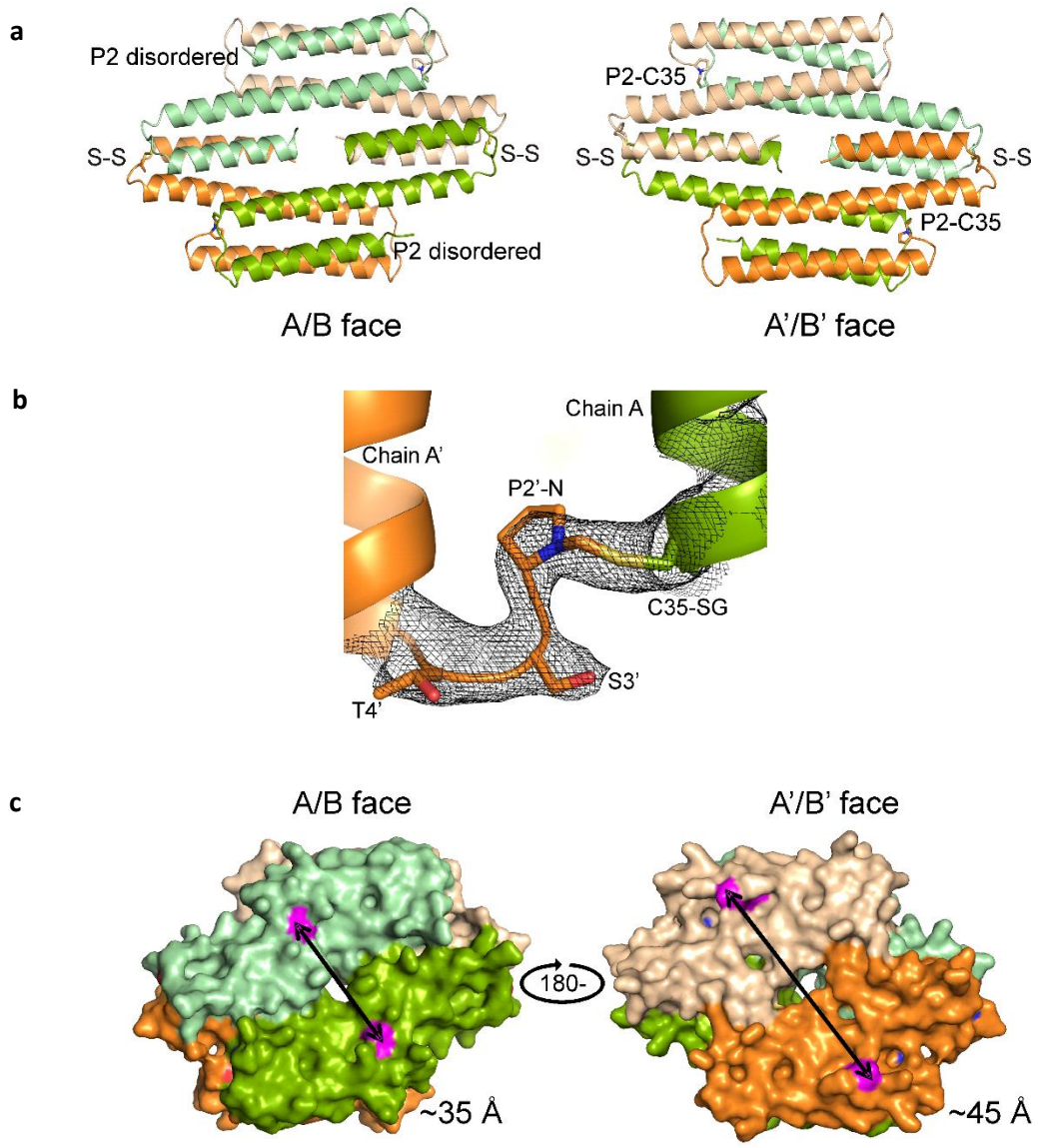


Figure 5.9a, b and c. <u>Comparison between the uncross-linked and cross-linked faces of the *E. coli* FrmR tetramer</u>

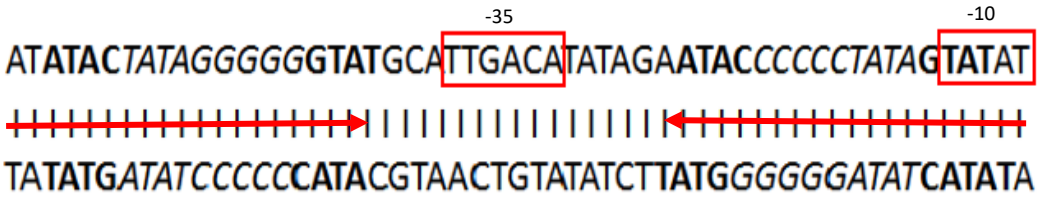
- **a.** The arrangement of the helices on each face of the tetramer, position of the methylene bridge and the disulfide bond are shown. The homodimer (A/B) on the uncross-linked face is drawn in shades of green and on the cross-linked face (A'/B') in shades of orange.
- **b.** A section of the 2Fo-Fc map between chains A and A´ obtained when the coordinates for Pro2 and the methylene bridge were omitted from the refinement (black mesh, contoured at 1σ). Residues are indicated by their single letter codes.
- **c.** Expansion of the surface upon cross-linking (black double headed arrow drawn between Arg14, highlighted in pink). The homodimer (A/B) on the uncross-linked face is drawn in shades of green and on the cross-linked face (A'/B') in shades of orange.

bridges (-CH_2-) (Table 5.2). Unlike the *E. coli* FrmR, the *N*-terminal region of *Se*FrmR(E64H) is visible in the absence of the methylene bridge (Osman *et al*., 2016). With the exception of the domain swapped α3, the structures of the uncross-linked (A/B) face of *E. coli* FrmR superposes well on the *Se*FrmR structure.

DNA-binding data from BLItz shows that *in vitro* FrmR binds to P_{frm} creating a P_{frm}-FrmR complex (Section 5.3). Dissociation of the P_{frm}-FrmR complex was seen in the presence of formaldehyde. The speed of this dissociation increased with increasing formaldehyde concentration. The dissociation of the P_{frm}-FrmR complex in the presence of formaldehyde enables the expression of the FrmAB enzymes to detoxify the toxic formaldehyde. It has previously been suggested that CsoR/RcnR family members recognise their target DNA via a combination of: (i) specific interactions with the flanking inverted repeats; (ii) shape selectivity, resulting from the central GC tract's propensity to adopt the A-DNA form; (iii) non-specific binding to distant DNA that might result in DNA-wrapping (Iwig and Chivers, 2009; Tan *et al*., 2014). The P_{frm} region contains a large inverted repeat centered at -29 relative to the predicted transcriptional start site. Tandem FrmR-binding sites consisting of 9 bp GC-rich tracts flanked by ATAC/GTAT inverted repeats are contained in this region (Figure 5.10a). The DNA-binding site of *E. coli* RcnR also contains a G-tract, which is flanked by AT-rich regions. It was previously shown that an RcnR binds to this region of the promoter sequence, using footprinting and isothermal titration calorimetry (Iwig and Chivers, 2009). It is highly likely that FrmR binds to the C- or G-tract within its promoter region, as other members of the CsoR/RcnR family also contain similar regions within their promoters (Foster *et al*., 2012; Rubio-Sanz *et al*., 2013).

Modeling of the P_{frm}-FrmR complexes with A- and B-form DNA (Figure 5.10b) was carried out (by Dr Claudine Bisson), as RcnR-binding sites in the *E. coli rcnR-A* impart some A-form DNA character on this region (Tan *et al*., 2014). Models with both A- and B-form DNA suggested that residues (Lys10, Arg14, Arg16 and Arg17) form two positively-charged protrusions (~35 Å apart) on the A/B face can pack into the major grooves of the DNA. The *Se*FrmR(E64H) positively-charged patches were also separated by ~35 Å. The ~45 Å that separates these two patches on the *E. coli* FrmR A′/B′ face precludes convincing interactions with either A- or B-form DNA (Figure 5.9c). In addition, the crosslink also increases the distance between the cluster of residues within the positively-charged protrusions by ~10 Å, thus interacting with the major groove is no longer possible, breaking the FrmR-P_{frm} complex (Figure 5.9c).

a



b

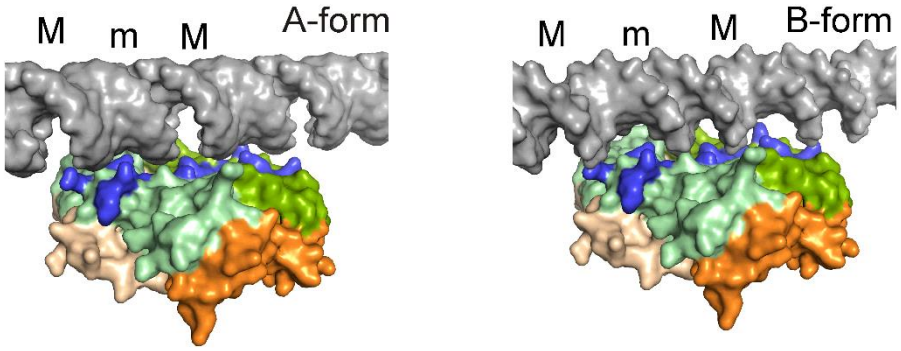


Figure 5.10a and b. <u>The FrmR binding sites in *E. coli* P_{frm} and modeling the P_{frm}–FrmR complex</u>

- **a.** The DNA sequence of the P_{frm} contains tandem FrmR binding sites consisting of ATAC/GTAT inverted repeats (bold) separated by G/C-rich tracts (italic) that form a larger inverted repeat (convergent red arrows). The -10 and -35 sites of P_{frm} are also show in red boxes.
- **b.** One of the FrmR binding sites of P_{frm} (dark grey) was modelled as A-form (left) and B-form (right) DNA as shown in the lower image. FrmR is shown as surface representation with subunits colored in shades of green (uncross-linked A/B face) and orange (cross-linked A´/B´ face), with the amino acid side-chains on the A/B face that are implicated in DNA-binding highlighted in blue. The major (M) and minor (m) grooves are indicated.

These differences suggest a mechanism for dissociation of the FrmR-P_{frm} in response to formaldehyde, as shown by BLItz data, and thus the de-repression of the *frmRAB* promoter.

Therefore, *In vivo* and *in vitro* experimental evidence show that the FrmR specifically reacts with the toxic chemical formaldehyde, resulting in the formation of inter-molecular methylene bridges between adjacent *N*-terminal Pro2 and Cys35 residues. The conservation of His60 and its proximity to Cys35´ in the adjacent subunit suggests that it could act to abstract a proton from the thiol group to facilitate reaction with formaldehyde forming an *S*-hydroxymethyl adduct, which is then subject to nucleophilic attack by Pro2 (secondary amine) to produce the methylene bridge (Table 5.2, Figure 5.11). FrmR is the first example of a CsoR/RcnR family protein where the asymmetry of the tetramer in the crystal structure reveals the conformational changes induced by signal perception that lead to de-repression of target promoters. As formaldehyde-induced methylene bridge formation expands the A´/B´ surface of FrmR resulting in a similar change in size to that observed for *G. thermodenitrificans* CsoR upon Cu(I) binding (Chang *et al*., 2014), suggesting that the signal transduction mechanism that controls binding to and release from DNA may be conserved throughout the CsoR/RcnR family.

Future work could focus on the role of the second Cys residue (Cys70), which is not conserved in *Se*FrmR. The Cys70 forms a disulfide bond linking loops 2 of the A/B´ and the B/A´ chains. The effect of the disulfide bond on sensing formaldehyde has not be assessed here. However, there is potential for FrmR to act as a redox sensor, through the formation of a disulfide bond, which is of interest because glutathione-depletion and oxidative stress is associated with exposure to formaldehyde in higher organisms (Levovich *et al*., 2008; Yano *et al*., 2009; Zhang *et al*., 2013). Further work to determine the possible role of glutathione (*S*-hydroxymethylglutathione is formed in *E. coli* exposed to formaldehyde) to better understand the role of oxidative stress in regulating FrmR activity *in vivo* could be undertaken, as well as determining the structure of the P_{frm}-FrmR complex.

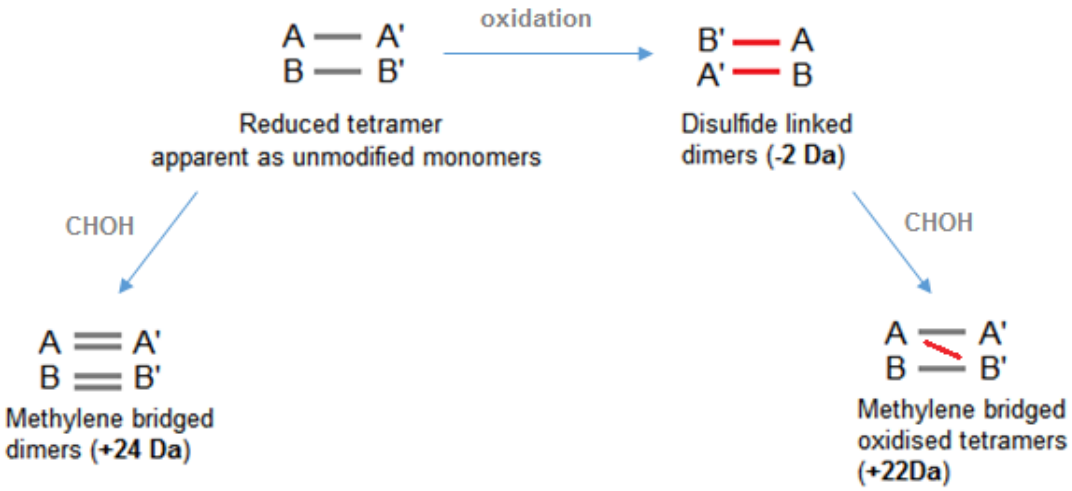


Figure 5.11. Schematic of the species of *E. coli* FrmR formed by oxidation and reaction with formaldehyde detected by LC-MS

Subunits of the FrmR tetramers are represented by upper case letters (A and B on one surface and A' and B' on the other). Disulfide bonds (red lines) and methylene bridges (grey lines) are shown linking FrmR subunits. The associated mass changes are shown in bold font and are those detected by LC-MS.

6. Regulation of TMAO respiration in *Escherichia coli* MG1655 in the presence of electron acceptors with higher redox potentials

6.1. Introduction

Escherichia coli can use several compounds as exogenous electron acceptors, including oxygen, nitrate, nitrite, DMSO, TMAO and fumarate (Gennis and Stewart, 1996; Unden and Bongaerts, 1997). Respiration in *E. coli* is a complex process, thus, to allow biosynthesis of the appropriate respiratory system, expression of the corresponding operons is controlled according to presence of environmental electron acceptors (Unden *et al*., 1994). The prevailing regulatory rule, whereby pathways with high ATP or growth yields are favoured, leads to the induction of the most energy-efficient respiratory system. Oxygen is the preferred electron acceptor and represses the terminal reductases of anaerobic respiration (Gennis and Stewart, 1996; Unden and Bongaerts, 1997). In anaerobic respiration, nitrate represses other terminal reductases, such as DMSO-, TMAO- and fumarate reductases (Unden and Bongaerts, 1997). Energy conservation is maximal with oxygen and lowest with fumarate. The main regulators involved in the anaerobic response are the two-component regulatory system ArcBA and the transcription factor FNR (Lin and Iuchi, 1991; Unden *et al*., 1994). FNR directly senses the absence of oxygen (Khoroshilova *et al*., 1997) and is the master activator of the nitrate-, nitrite-, DMSO- and fumarate reductases (Constantinidou *et al*., 2006). Whereas, ArcB monitors the redox state of the quinone pool (Georgellis *et al*., 2001). The TorCA respiratory system permits the reduction of TMAO, but neither regulated by FNR nor by ArcBA as the other anaerobic respiratory systems are (Pascal *et al*., 1984; Simon *et al*., 1994).

Unlike other alternative electron acceptors that are reduced by *E. coli,* previous work has demonstrated that both the expression of *torCAD* and TMAO respiration occurred in the presence of the preferred electron acceptor, oxygen (Ansaldi *et al*., 2007). Cells grown aerobically on rich medium (LB broth) in the presence of glycerol and TMAO had a higher growth yield than cells grown in glycerol without TMAO (Ansaldi *et al*., 2007). It was proposed that an advantage of TMAO reduction during aerobiosis is the production of alkaline TMA, which could enhance growth by increasing the pH of the culture. This is important in *E. coli*, as acetate is synthesised as a by-product of metabolism.

However, TMAO respiration has been shown to be repressed ~two-fold in the presence of nitrate, the preferred anaerobic electron acceptor (Takagi *et al*., 1981), with this repression not occurring in a *narL* mutant (Iuchi and Lin, 1987). A more recent microarray analysis suggested that *torCAD* transcript levels are decreased in the presence of nitrate

(Constantinidou *et al*., 2006). Nitrate-repression of *torCAD* would be consistent with the current dogma that the most energetically efficient mode is adopted because the redox potential of the NO_3^-/NO_2^- pair is +420 mV compared to +130 mV for the TMAO/TMA pair (Simon *et al*., 1994). However, the mechanism of nitrate-mediated regulation of *torCAD* is unknown. The *E. coli* two-component system NarXL is the key regulatory system involved in nitrate-sensing. NarX acts as a sensory kinase whilst NarL, a 23.9 kDa protein, is a response regulator capable of transcriptional repression or activation of target genes (Williams and Stewart, 1997).

In this chapter, TMAO respiration by *E. coli* K-12 in the presence of the preferred electron acceptor (oxygen) is investigated. The mechanism of nitrate-mediated regulation of *torCAD* and whether this regulation is direct, are also explored in relation to the prevailing regulatory rule.

6.2. TMAO is not reduced to TMA by aerobic batch cultures of *E. coli* MG1655 grown in defined minimal medium with glucose or glycerol as the sole carbon source

The energy yield of aerobic respiration is far greater than that of anaerobic respiration using any of the alternative electron acceptors (Section 6.1). The advantage for the cell to reduce TMAO in the presence of oxygen is not clear, even though TMAO reduction by wild type *E. coli* K-12 in aerobic conditions has previously been reported (Ansaldi *et al*., 2007). Similar growth experiments to those reported by Ansaldi *et al.* (2007) were carried out, but using glucose and glycerol as carbon sources in minimal medium (Section 2.3.2). Glucose can be respired aerobically and is fermentable in the absence of an electron acceptor (Unden and Bongaerts, 1997). Even under fully aerobic conditions *E. coli* excretes acetate when glucose is provided in excess, probably due to an imbalance in glucose uptake and the capacity of the Krebs cycle to futher metabolize acetylCoA (Farmer and Liao, 1997). However, metabolism of glycerol by *E. coli* requires the presence of external electron acceptors and thus is a non-fermentable carbon source (Lin, 1976; Quastel and Stephenson, 1925; Quastel *et al*., 1925). *Escherichia coli* MG1655 cultures were grown aerobically in buffered Evans minimal medium with either glucose (10 mM) or glycerol (20 mM) as the carbon source. The electron acceptor TMAO (40 mM) was added to cultures where indicated. Growth on both carbon sources was followed over 8 h and the pH of the medium was measured (Figure 6.1).

For glucose-grown cultures, the final cell density was slightly lower (~two-fold) in the absence of TMAO. In contrast, the presence of TMAO did not increase the growth yield when

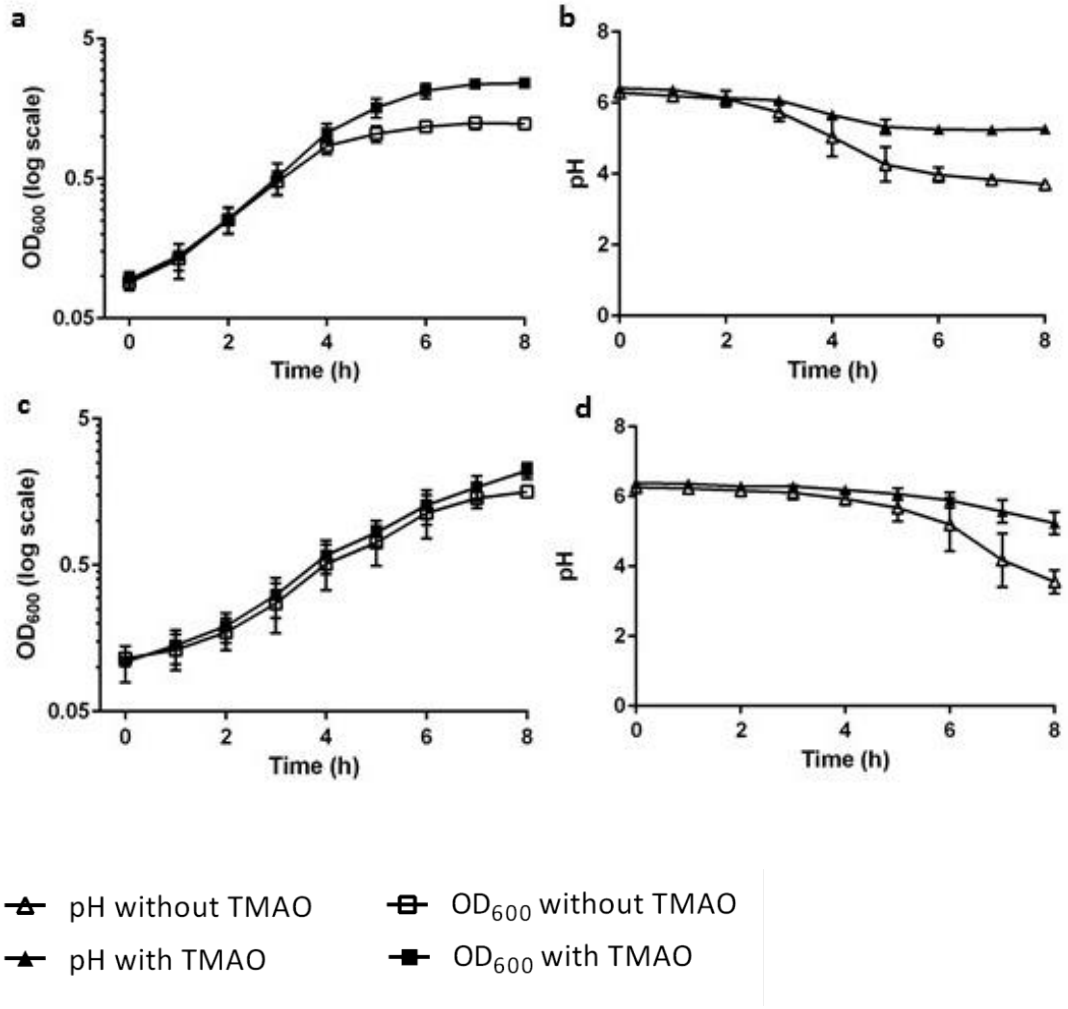


Figure 6.1. <u>Growth curves of *E. coli* MG1655 cultures and their pH under aerobic conditions in the presence of the alternative electron acceptor TMAO, with either glucose or glycerol as the electron donor</u>

Aerobic growth curves of *E. coli* MG1655 cultures in Evans minimal medium, with either 10 mM glucose (**a** and **b**) or 20 mM glycerol (**c** and **d**) as the electron donor, thus providing the cells with equivalent amounts of carbon. The cells were grown either with (closed symbols) or without 40 mM TMAO (open symbols), as indicated. Samples were removed every hour for 8 hours, with OD_{600} (squares (**a** and **c**)) and pH readings (triangles (**b** and **d**)) being recorded. Average values are shown for 3 biological replicates and errors are standard deviations from the mean.

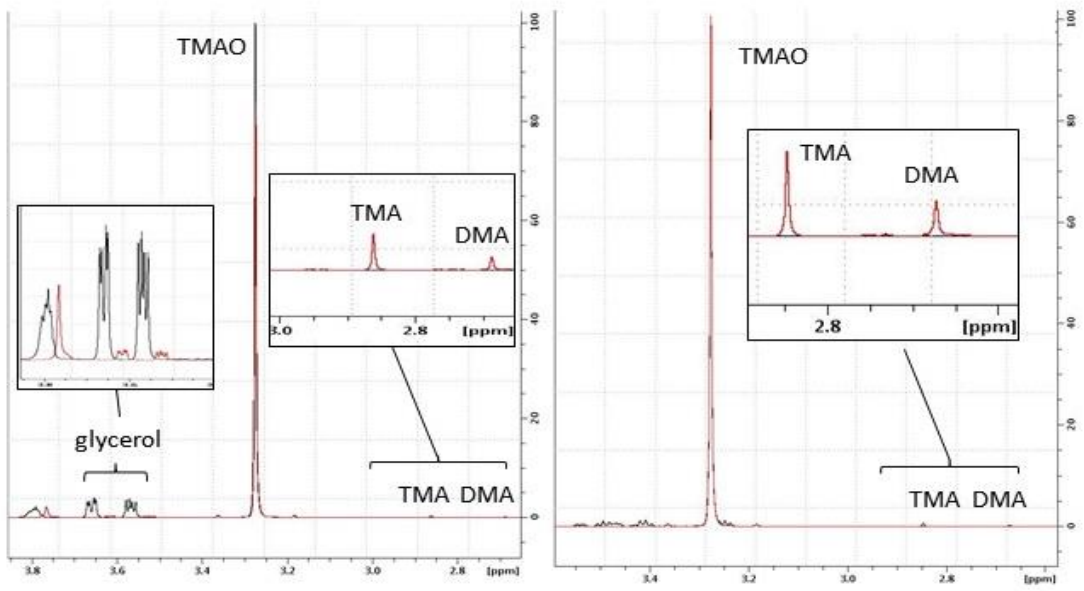
glycerol was supplied as the carbon source. However, pH of the cultures was maintained closer to neutrality for 8 h when TMAO was present in both glucose and glycerol cultures, when compared to aerobic growth on both carbon sources in the absence of TMAO (Figure 6.1). In contast to the observations of Ansaldi *et al*. (2007) with aerobic cultures grown in glycerol (67 mM) and TMAO (10 mM) supplemented LB broth, NMR measurements for both the glucose (10 mM) and glycerol (20 mM) cultures analysed here revealed that very little TMAO (40 mM) reduction occurred (Figure 6.2a). Peak-integration indicated that only 0.21 and 0.23 mM TMA was produced from TMAO after 8 h growth of aerobic growth on glucose and glycerol, respectively. The quantity of TMA produced in oxygenated cultures under these experimental conditions was far lower than had previously reported (3.6 mM over 6 h) (Ansaldi, *et al.* 2007) (Figure 6.2b). Thus, under these conditions, the maintenance of the pH of cultures grown in the presence of TMAO is not due to the production of alkaline TMA.

At time zero both glucose and glycerol cultures contained ~0.37 mM acetate, presumably originating from the inocula (1 % (v/v) of an overnight aerobic culture grown in buffered Evans media). In the absence of TMAO after 8 hours, the glucose-grown cultures contained ~3.5 ± 0.39 mM acetate, whereas the glycerol cultures produced less acetate ~0.75 ± 0.57 mM (for one of the replicate cultures acetate was estimated to be ~0.1 mM, whereas the other two cultures contained ~1.1 mM). This showed that some of the carbon source was incompletely oxidized and acetate could contribute to lowering the pH of the cultures (Figure 6.1). In the presence of TMAO, rather than producing acetate the glucose grown cultures essentially maintained the original concentrations (0.28 ± 0.18 mM), whereas the glycerol-grown cultures utilised acetate (~0.1 ± 0.12 mM). Although preliminary, this data suggests that under aerobic conditions in a defined minimal medium, the presence of TMAO promotes the utilisation of acetate in cultures containing glucose or glycerol. Nevertheless, it is clear that regulation of TMAO respiration by *E. coli* under aerobic conditions is likely to be more complex than previously thought and the growth promoting effects of aerobic TMAO reduction observed for the glucose-grown cultures reported here (Figure 6.1) are likely to be a result of maintaining the culture pH closer to neutral (in part by TMA) and by promoting the utilisation of acetate and/or the inhibition of acetate production.

6.3. Regulation of the TMAO reductase-encoding *torCAD* operon by the nitrate response regulator NarL

In order to investigate the mechanism of nitrate-mediated regulation of *tor*, the nitrate response regulator protein NarL was overproduced, purified and used in combination with

a



b

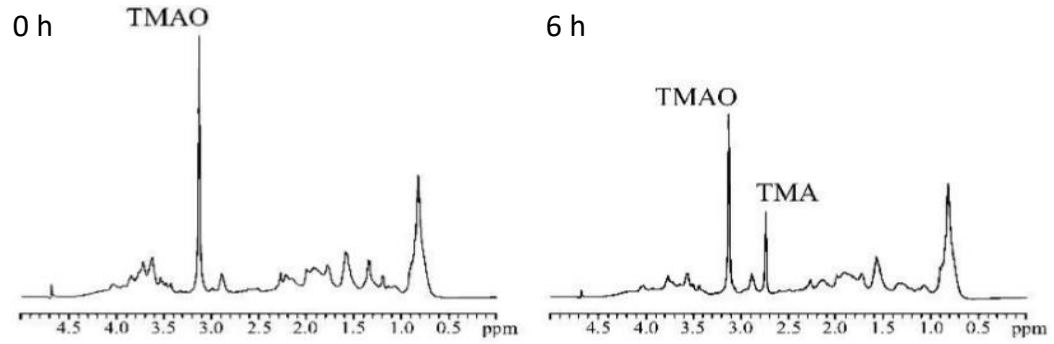


Figure 6.2. <u>Reduction of TMAO under aerobic conditions by *E. coli* MG1655 is lower than previously reported</u>

- **a.** Comparison of spectra of cultures grown on glycerol (20 mM) or glucose (10 mM) and TMAO (40 mM) at 0 hours (black) and 8 hours (red). Peak identities are are shown where known. Chemical shift is given in ppm, whilst the y-axes are identically scaled but show percentage peak positions relative to the 0 hour TMAO peak. Due to peaks shifting caused by pH changes during culture the 8 hour curve has been shifted to allow overlay. No carbon 13 decoupling was used therefore ^{13}C satellites are visible. Insets show scaled up versions of the glycerol and TMA/DMA regions of the spectra.
- **b.** Previously published NMR spectra showing TMAO reduction over 6 h under aerobic conditions with glycerol as the carbon source (Ansaldi *et al*., 2007; Figure 5).

torCAD promoter DNA (P_{tor}) in EMSA reactions, to look for an interaction between purified NarL protein and P_{tor} DNA.

6.3.1. Overproduction, purification and phosphorylation of *E. coli* NarL

NarL was overexpressed and purified using a MBP-NarL fusion protein construct. Once overproduced (Section 2.10.1.2), the protein was purified using affinity chromatography with a prepacked and equilibrated amylose resin column. Cell-free extracts of cultures expressing the protein fusion were added to the column and eluted, after a series of wash steps (Figure 6.3a), with 10 mM maltose (Figure 6.3b) (Sections 2.10.2.2 and 2.10.3.2). MBP-NarL protein was purified and exhibited the expected molecular weight of ~67 kDa (Figure 6.3b). Analysis of the final eluent using SDS-PAGE (Section 2.10.6), suggested that purified MBP-NarL was the most abundant species (~80% purity). The contaminating proteins may be proteolytically cleaved MBP-NarL or prematurely terminated MBP-NarL proteins. Purified NarL was phosphorylated for use in EMSAs, as phosphorylation is required in order to expose the *C*-terminal DNA-binding domain of the transcription factor, which is otherwise concealed by the *N*-terminal domain (Zhang *et al*., 2003). When the *N*-terminal domain is phosphorylated, a conformational change occurs via a hinge-like mechanism, which enables the *C*-terminal domain to bind to DNA as a dimer and affect transcription of its target genes (Zhang *et al*., 2003). NarL was phosphorylated by incubation of the purified protein with carbamoyl phosphate (100 mM) (Section 2.10.5).

6.3.2. Using EMSAs to study P_{tor} DNA and NarL interaction

The Thermo Scientific LightShift Chemiluminescence EMSA kit was used for EMSA experiments, which uses a nonisotopic method to detect DNA-protein interactions (Section 2.11.3). Biotin end-labelled DNA containing the binding site of interest was incubated with purified protein. This reaction was then subjected to gel electrophoresis on a native polyacrylamide gel and transferred to a nylon membrane. The biotin end-labelled DNA was then detected using Streptavidin-Horseradish Peroxidase Conjugate and the chemiluminescent substrate.

6.3.2.1. Nitrate-mediated regulation of *torCAD* does not directly involve NarL

In the *E. coli* MG1655 genome, the divergent *torC* and *torR* promoter regions are both present in the 133 bp intergenic region. Oligonucleotide primers were designed to amplify this region, with the forward primer annealing ~150 bp upstream of the ATG start site of the *torC* gene (Appendix 1). The reverse primer was biotin-labelled and designed to bind ~25 bp downstream of the *torC* start site (Figure 6.4) (Appendix 1). Sequence analysis of the *torCAD*

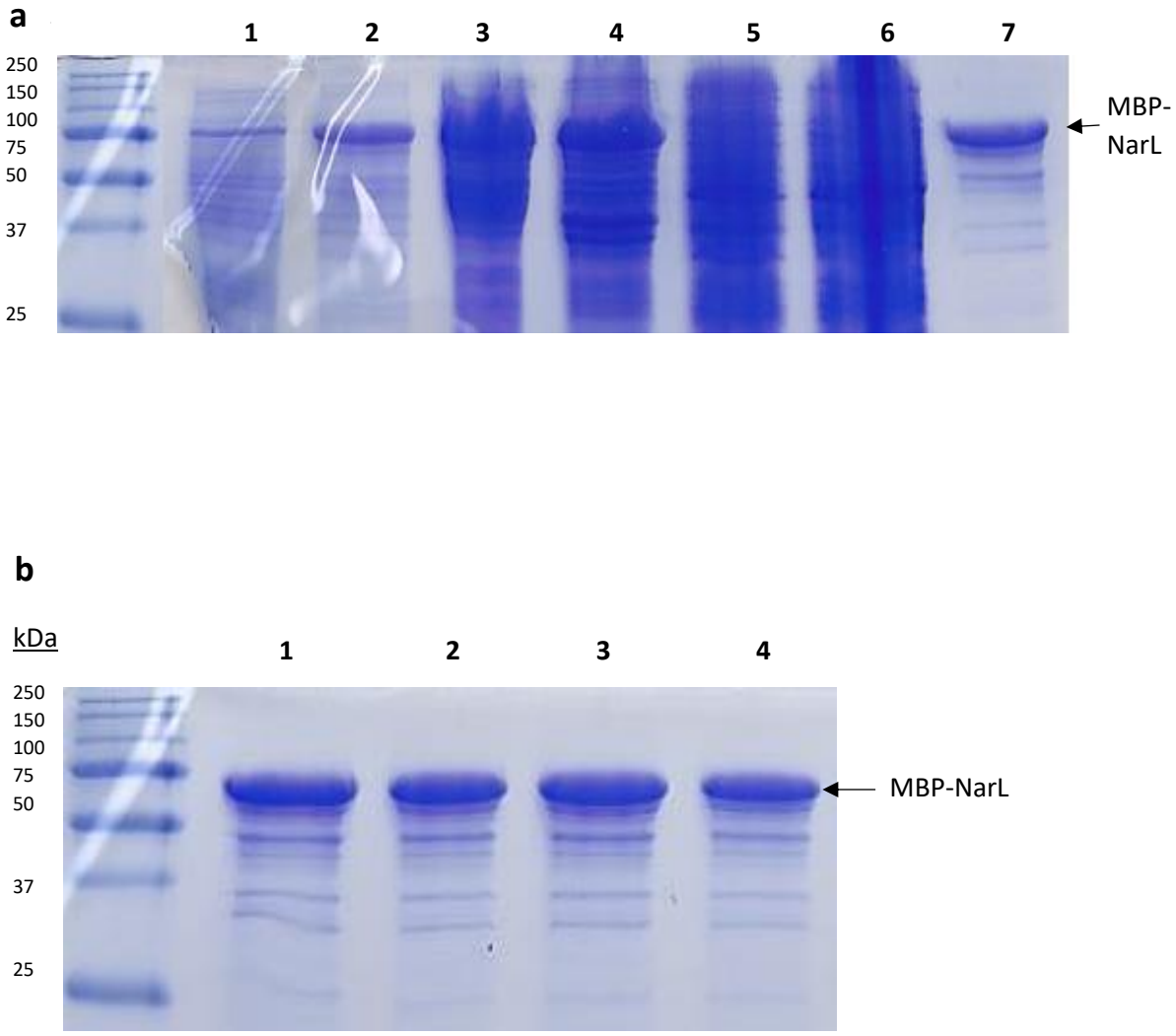


Figure 6.3. **SDS-Polyacrylamide gel electrophoresis analysis of MBP-NarL purification**

a. Polypeptide profiles of samples (10 µl): Lane 1, uninduced culture; Lane 2, induced culture; Lane 3, cell-free extract; Lane 4, pellet containing insoluble proteins; Lane 5, amylose resin loading flow through; Lane 6 – 7, column washing at start and end.

b. Lanes 1- 4, aliquots (10 µl) of fractions of MBP-NarL eluted from the amylose resin with 10 mM maltose. The positions of protein are indicated. Markers used for both gels are Bio-Rad Precision Plus Protein All Blue Protein Standards.

CCCTCCTGAGTGAAGTAGGATTGTAATCGCGCCTGGGTAACCGGCTCATCTTCAA^{-199}CAATAACAA
TGTGATGTGGCATCAGAGG^{-178}GTTTTACTCATT**CTGTTCATATCTGTTCATAT**TCTGCCGTAAG**CCG**
TTCATCCTGACCAGTGCCG**CTGTTCATAT**TTGCTCATTAAGATCGCTTCACTAAACCATAATTCTACA
GGGGTTATTatgCGGAAA^{+7}CTCTGGAACGCGCTACGCCG^{+25}ACCCAGTGCTCGTTGGTCGGTACTG
GCGCTGGTCGCAATTG

Figure 6.4. Sequence of the P_{tor} region in *E. coli* MG1655 with transcriptional features highlighted

The DNA sequence of the intergenic region of *torR* and *torC* of *E. coli* MG1655, which contains the *torC* and *torR* promoter regions and where transcription of both genes occurs bidirectionally, is shown. The TorR binding sites within the promoter are shown in bold text. The ATG start codon for *torC* is shown in green and the transcription start point is shown in red. The -10 and -35 regions are underlined. No definitive NarL consensus sequences were identified in the *torCAD* promoter region (consensus sequence: TACYYMT; Y, C or T residue; M, A or C residue (Darwin *et al.*, 1996)). Oligonucleotide primers (sequences shown in Appendix 1) were designed to bind to ~150 bp upstream of the ATG start site and ~25 bp downstream of the start site, the locations of primer binding are highlighted in yellow in the above sequence. The downstream oligonucleotide was biotin-labelled for detection.

promoter region revealed that a NarL consensus binding site sequence was not present (Figure 6.4). In order to determine whether NarL was able to directly regulate *torCAD* expression by binding to the intergenic region (P_{tor}) to prevent expression in the presence of nitrate/nitrite, EMSAs were carried out.

Firstly, 20 fmol P_{tor} DNA was incubated with increasing concentrations (0-20,000 nM) of MBP-NarL protein. When incubated with poly-deoxy-inosinic-deoxy-cytidylic acid (poly [d(I-dC)]) (1 µg), non-specific competitor DNA to mop up non-specific interactions, 'smearing' was observed in reactions containing higher concentrations of MBP-NarL (5,000 – 20,000 nM) (Figure 6.5). It was possible that the smear resulted from an increased probability of non-specific interactions occurring, due to the presence of excess protein in the assay. To act as a comparator, unphosphorylated MBP-NarL was then used in EMSA reactions. As the DNA-binding domain of the NarL transcription factor is not accessible when it is unphosphorylated, there should be no shift (Zhang *et al*., 2003). The EMSA reactions containing unphosphorylated MBP-NarL also resulted in 'smearing' at high concentrations of MBP-NarL (5,000 – 20,000 nM) (Figure 6.6). The lack of any clear shift in P_{tor} mobility with phosphorylated MBP-NarL compared to the unphosphorylated protein samples, suggested that any binding was due to non-specific interactions at high protein concentrations.

To ensure that NarL was being fully phosphorylated, EMSAs were repeated with reactions containing phosphorylated MBP-NarL protein (as before) and an excess of carbamoyl phosphate (1 M). This did not improve binding of MBP-NarL to P_{tor} (Figure 6.7).

The presence of the maltose binding protein (MBP) (43.4 kDa), which is fused to NarL for protein purification could prevent NarL binding to P_{tor} DNA by steric hindrance. In order to release NarL from the MBP-NarL, the protein was incubated with 0.2 ng enterokinase overnight at room temperature (Section 2.10.3.2). MBP-NarL treated with enterokinase was then phosphorylated for use in EMSA reactions. Once again, a mobility shift was only observed at the higher concentrations of protein (5,000 – 20,000 nM) (Figure 6.8).

Verification of the ability of the EMSA method to detect a positive shift with NarL was sought using *ydhY* promoter DNA (P_{ydhY}). The *ydhY* gene is part of an operon, which is known to be regulated by the NarXL system and encodes an oxidoreductase. NarL binds to the *ydhY* promoter and represses transcription of the gene, when nitrate is present in the environment (Partridge *et al*., 2008). The binding site for NarL in the *ydhY* promoter is composed of two heptamers, separated by two base-pairs (Darwin, *et al*., 1997). Oligonucleotide primers were designed to anneal ~200 bp upstream of the ATG start codon

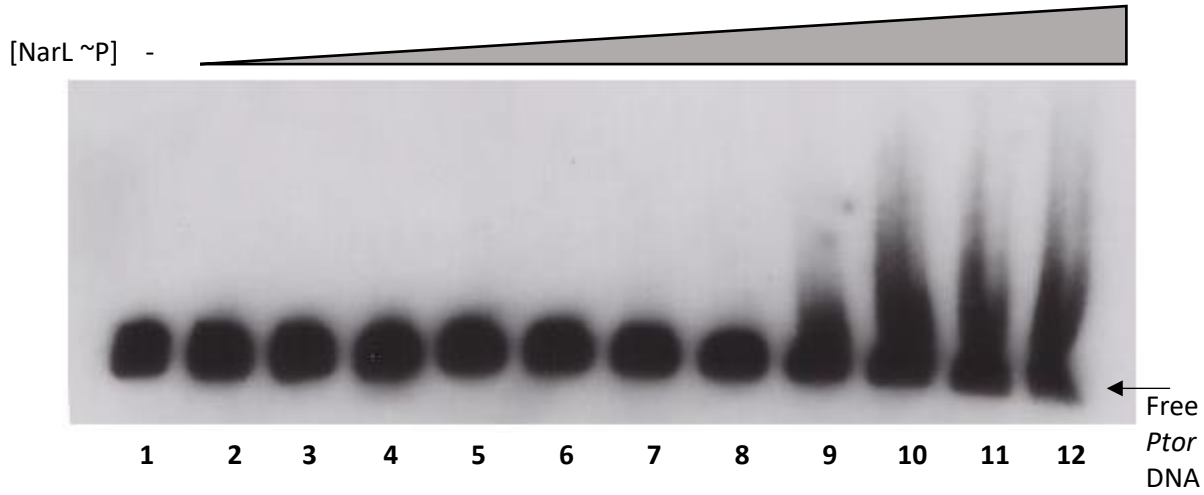


Figure 6.5. Phosphorylated NarL does not bind P_{tor} DNA

Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (1 µg) different concentrations of phosphorylated purified NarL protein (0 – 20,000 nM) as follows: Lane 1, no NarL; Lane 2, 10 nM NarL; Lane 3, 50 nM NarL; Lane 4, 100 nM NarL; Lane 5, 200 nM NarL; Lane 6, 500 nM NarL; Lane 7, 1,000 nM NarL; Lane 8, 2,000 nM NarL; Lane 9, 5,000 nM NarL; Lane 10, 10,000 nM NarL; Lane 11, 15,000 nM NarL; Lane 12, 20,000 nM NarL. The position of free biotin-labelled P_{tor} DNA is indicated. In the lanes (9 - 12) with reactions containing a high concentration of NarL (5,000 – 20,000 nM), 'smearing' was observed.

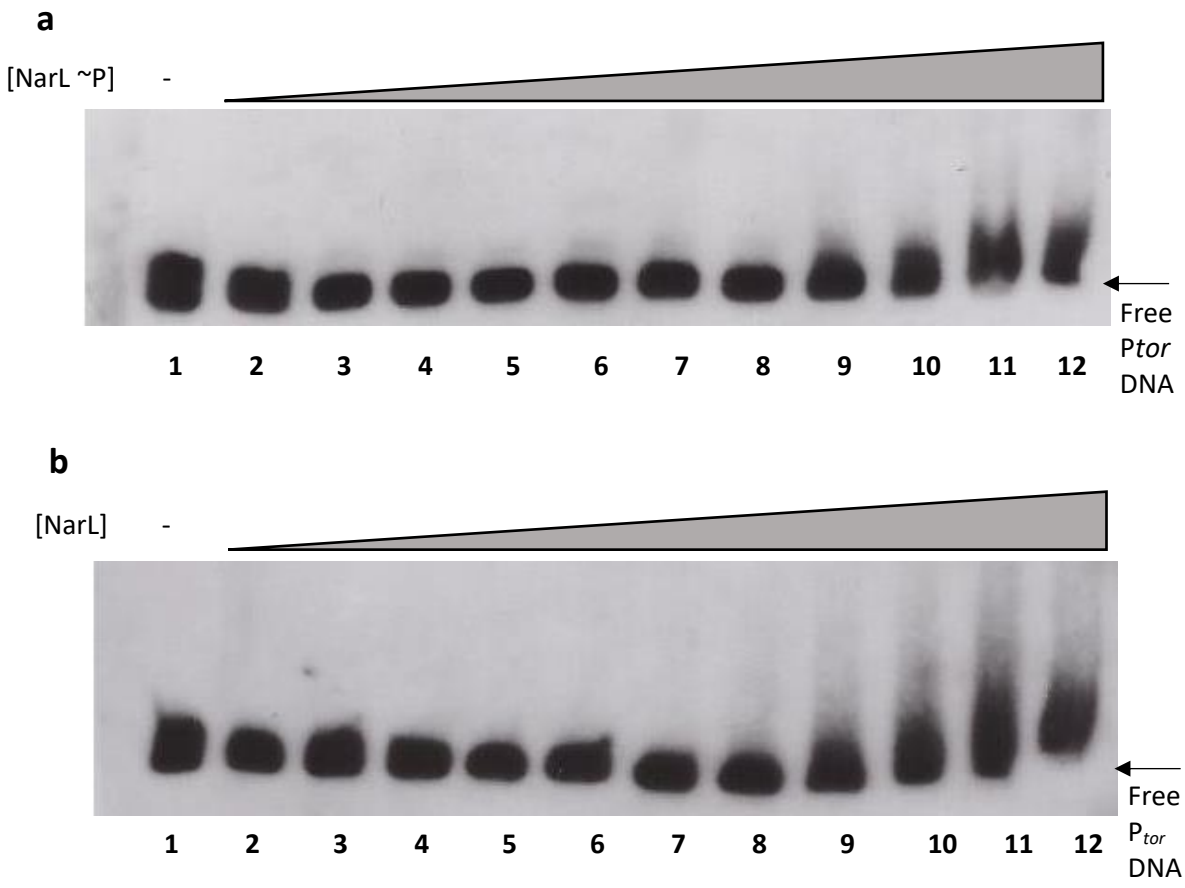


Figure 6.6. <u>Phosphorylation of NarL does not affect the ability of the protein to bind P_{tor}</u>

a. Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of carbamoyl phosphate-phosphorylated purified NarL protein (0 – 20,000 nM).

b. Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of unphosphorylated purified NarL protein (0 – 20,000 nM) as follows: Lane 1, no NarL; Lane 2, 10 nM NarL; Lane 3, 50 nM NarL; Lane 4, 100 nM NarL; Lane 5, 200 nM NarL; Lane 6, 500 nM NarL; Lane 7, 1,000 nM NarL; Lane 8, 2,000 nM NarL; Lane 9, 5,000 nM NarL; Lane 10, 10,000 nM NarL; Lane 11, 15,000 nM NarL; Lane 12, 20,000 nM NarL. The concentration of NarL in each lane is the same for both gels. The position of free biotin-labelled P_{tor} DNA is indicated.

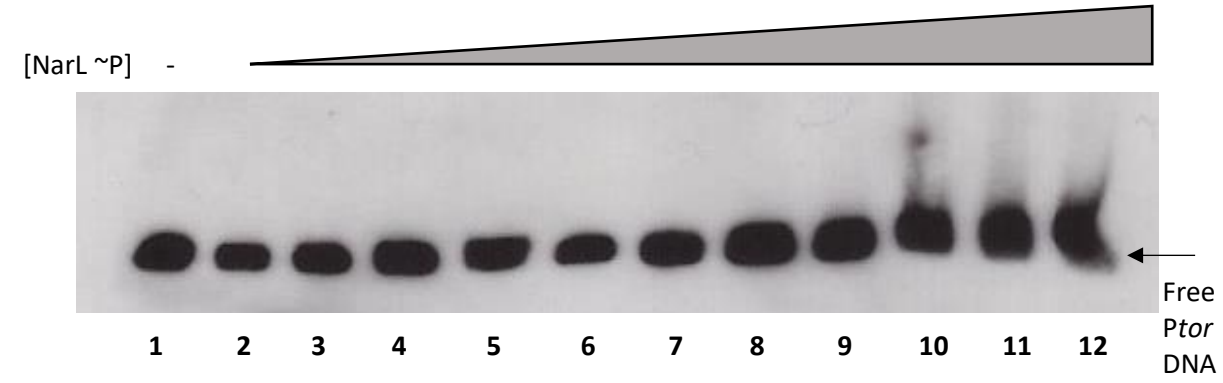


Figure 6.7. <u>The lack of NarL phosphorylation is not a factor preventing DNA binding</u>

Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of carbamoyl phosphate-phosphorylated purified NarL protein (0 – 20,000 nM). The reactions were also incubated with excess carbamoyl phosphate (1 M) to ensure NarL was fully phosphorylated before complexes were separated on an acrylamide (6%) gel. Lane 1, no NarL; Lane 2, 10 nM NarL; Lane 3, 50 nM NarL; Lane 4, 100 nM NarL; Lane 5, 200 nM NarL; Lane 6, 500 nM NarL; Lane 7, 1,000 nM NarL; Lane 8, 2,000 nM NarL; Lane 9, 5,000 nM NarL; Lane 10, 10,000 nM NarL; Lane 11, 15,000 nM NarL; Lane 12, 20,000 nM NarL. The position of free biotin-labelled P_{tor} DNA is indicated.

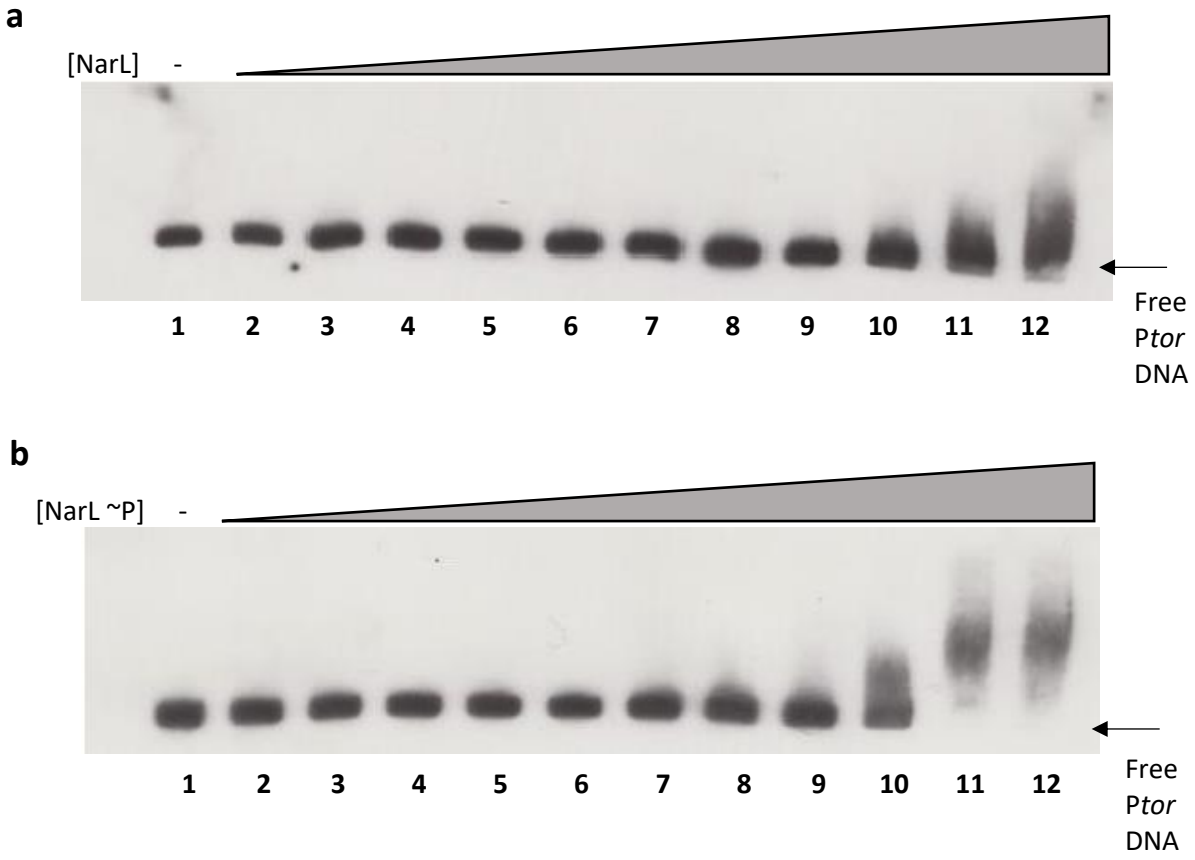


Figure 6.8. <u>Removing the MBP tag from purified NarL does not permit binding at P_{tor} at physiological NarL concentrations</u>

a. Prior to preparing EMSA reactions, purified MBP-NarL was treated with enterokinase to remove the MBP tag. Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of unphosphorylated purified enterokinase-treated NarL protein (0 – 20,000 nM).

b. Prior to preparing EMSA reactions, purified MBP-NarL was treated with enterokinase to remove the MBP tag. Biotin-labelled promoter fragment of *torCAD* (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of phosphorylated purified enterokinase-treated NarL protein (0 – 20,000 nM) as follows: Lane 1, no NarL; Lane 2, 10 nM NarL; Lane 3, 50 nM NarL; Lane 4, 100 nM NarL; Lane 5, 200 nM NarL; Lane 6, 500 nM NarL; Lane 7, 1,000 nM NarL; Lane 8, 2,000 nM NarL; Lane 9, 5,000 nM NarL; Lane 10, 10,000 nM NarL; Lane 11, 15,000 nM NarL; Lane 12, 20,000 nM NarL. The concentration of NarL in each lane is the same for both gels. The position of free biotin-labelled P_{tor} DNA is indicated.

of the *E. coli* MG1655 *ydhY* gene and the biotin-labelled reverse primer ~25 bp downstream of the start codon (Figure 6.9) (Appendix 1). An EMSA, with phosphorylated MBP-NarL, P_{tor} DNA and poly [d(I-dC)] (0.5 µg) were compared to a reaction containing P_{ydhY} DNA and MBP-NarL (200 nM) (Figure 6.10). An interaction between P_{ydhY} DNA (200 nM) and NarL was observed, demonstrating that DNA binding can be successfully detected using this method. The reaction containing P_{ydhY} DNA produced an obvious shift compared to reactions containing P_{tor} DNA. MBP-NarL is capable of binding *ydhY* DNA with ~100X greater affinity than to the P_{tor}. Therefore, it is likely that NarL is not directly involved in repressing the expression of the *torCAD* TMAO reductase system in the presence of nitrate, as it does not bind to P_{tor} DNA with high affinity.

6.4. Discussion

Escherichia coli, like other is a facultative anaerobes, is able to respire using several alternative electron acceptors, thus, regulation of the respiratory network is highly complex. The prevailing regulatory rule leads to the induction of the most energetically-efficient respiratory system, whereby pathways with high ATP or growth yields are favoured. Thus, the presence of oxygen prevents the induction of anaerobic respiratory systems. However, it was previously reported that TMAO reduction is carried out in the presence of the preferred electron acceptor, oxygen (Ansaldi *et al*., 2007).

It was shown here using growth studies that although aerobic cultures reached a higher cell density in the presence of TMAO, very little reduction of TMAO to TMA was detected in the cultures (Figures 6.1 and 6.2). It was previously suggested that the production of alkaline TMA can help mitigate some of the acidification from the production of acetate that occurs in aerobic conditions (Ansaldi *et al.*, 2007). Oxygen limitation could account for the differences in TMAO reduction reported here, compared to previously published work, where aerobic cultures reached an OD_{600} ~7 (Ansaldi *et al.,* 2007). Whereas, in the experiments described here, cultures reached a maximum OD_{600} ~2. This difference in OD and thus oxygen availibity, could have resulted in the increased TMAO reduction reported by Ansaldi *et al.* (2007).

If the expression of an anaerobic respiratory system during aerobiosis does not offer any further respiratory capacity, then it should provide the cell with some other form of selective advantage (Ansaldi *et al.,* 2007). The only known anaerobic respiratory system of *E. coli* to be expressed during aerobiosis is the nitrate reductase NRZ. The *narUZYWV* operon is weakly expressed during the exponential phase as it is under the control of the alternative

$^{-199}$GTTAAGGGCTCAGAATAATCAC^{-178}ATTATTTTTTCTCCTAACCATAAAGGATTAGTTTATCGGCG

TTAATCTTTTCTCATACACTATCGGGTTAGGGAAAATATCTGGTGTTGGGGACAGAGTCAGCAGCG

TATTAGTTGAATGCTTAATATCAGGCTGCTGAATTGTCAGCACTGTGTTTGCGAAGAGGATGCGAA

ATAA**atg**AA^{+3}CCCGGTTGATCGTCCACTATTAG^{+25}

Figure 6.9. Sequence of the *ydhY* promoter region in *E. coli* MG1655 with transcriptional features highlighted

The DNA sequence of the *yhdY* promoter region of *E. coli* MG1655 is shown. NarL binding to this gene has previously been shown (Partridge *et al*., 2008) and thus will act as a positive control in the EMSA experiments. The ATG start codon of *ydhY* is shown in green and the transcription start point is shown in red. The -10 and -35 regions are underlined. The NarL consensus binding sequence is shown in red and underlined where base pairs match exactly (consensus sequence: TACYYMT; Y, C or T residue; M, A or C residue (Darwin *et al*., 1996)). Oligonucleotide primers (sequences shown in Appendix 1) were the same as those used previously (Partridge *et al*., 2008) and the locations of primer binding are highlighted in yellow in the above sequence. The downstream oligonucleotide was biotin-labelled for detection.

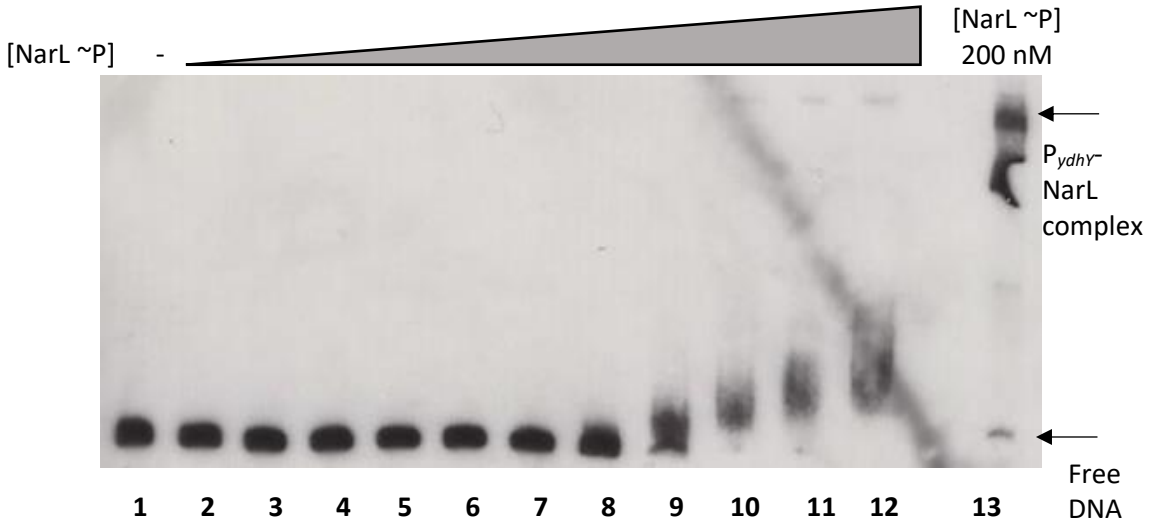


Figure 6.10. <u>NarL binds P_{ydhY} DNA at a physiologically relevant concentration, but does not bind to P_{tor} DNA</u>

Biotin-labelled promoter fragment of P_{tor} (20 fmol) was incubated for 20 min with poly [d(I-dC)] (0.5 µg) different concentrations of phosphorylated purified NarL protein (0 – 20,000 nM) as follows: Lane 1, no NarL; Lane 2, 10 nM NarL; Lane 3, 50 nM NarL; Lane 4, 100 nM NarL; Lane 5, 200 nM NarL; Lane 6, 500 nM NarL; Lane 7, 1,000 nM NarL; Lane 8, 2,000 nM NarL; Lane 9, 5,000 nM NarL; Lane 10, 10,000 nM NarL; Lane 11, 15,000 nM NarL; Lane 12, 20,000 nM NarL. The position of free biotin-labelled DNA is indicated. Lane 13 contains Biotin-labelled promoter fragment of *ydhY* (20 fmol), incubated for 20 min with poly [d(I-dC)] (0.5 µg) and phosphorylated purified NarL protein (200 nM). The position of bound *ydhY*-NarL on the membrane is indicated.

sigma factor RpoS (Chang *et al*., 1999; Clegg *et al*., 2006). It has been suggested that the *E. coli* NRZ could play an advantageous role during nutrient starvation (Clegg *et al*., 2006). Although reduction of TMAO was not observed in the experiments reported here, the pH of the cultures grown in the presence of TMAO under aerobic conditions was maintained closer to neutrality for longer, conferring a fitness advantage on cultures growing in the presence of both TMAO and the preferred electron acceptor, oxygen. The maintenance of culture pH could be due to the production of a small amount of alkaline TMA (in part) and also the utilisation of acetate and/or the inhibition of acetate production. However, how TMAO lowers acetate levels is still unclear.

TMAO reduction is apparently repressed in the presence of nitrate (the preferred anaerobic electron acceptor), with repression of *torCAD* not occurring in a *narL* mutant (Iuchi and Lin, 1987). Another study also showed that TMAO respiration was repressed in the presence of nitrate (Takagi *et al*., 1981). A more recent analysis of gene expression showed that *torCAD* transcript levels were decreased in the presence of nitrate (Constantinidou *et al*., 2006). However, earlier studies suggested that contradictory to prevailing rule, in which respiration in bacteria is hierarchically regulated by the most energy efficient electron acceptor, the presence of nitrate did not inhibit TMAO respiration (Pascal *et al*., 1984). As the *torCAD* promoter appears to lack putative NarL binding sites, there is much uncertainty surrounding the subject of NarL repression of the TorCA system (Simon *et al*., 1994) (Figure 6.4).

NarL is unlikely to directly regulate TMAO reductase expression, as suggested by the experiments detailed here. However, it may involved in the indirect regulation of TMAO reduction. One hypothesis could be that nitrate is detected by the NarXL two-component system, with activated NarL then inducing the expression of an transcription factor (unknown thus far), which plays a role in the expression of the TorCA system. Thus, by activating expression of a regulator protein an essential cofactor required for activity of the regulator, NarL may act as an indirect regulator of TMAO reduction by *E. coli* (Figure 6.11).

Another route of indirect regulation may be that when nitrate is present, NarL requires the presence of TorR in order to regulate transcription. TorR activates the expression of *torCAD* when TMAO is present. In the presence of TMAO, TorR binds to the *torCAD* promoter and triggers TorCA synthesis. If nitrate is available in the environment, NarL may bind to TorR, leading to a conformational change in TorR, which could then result in the repression of TMAO reduction by the TorCA system (Figure 6.12).TorR protein was not present in the

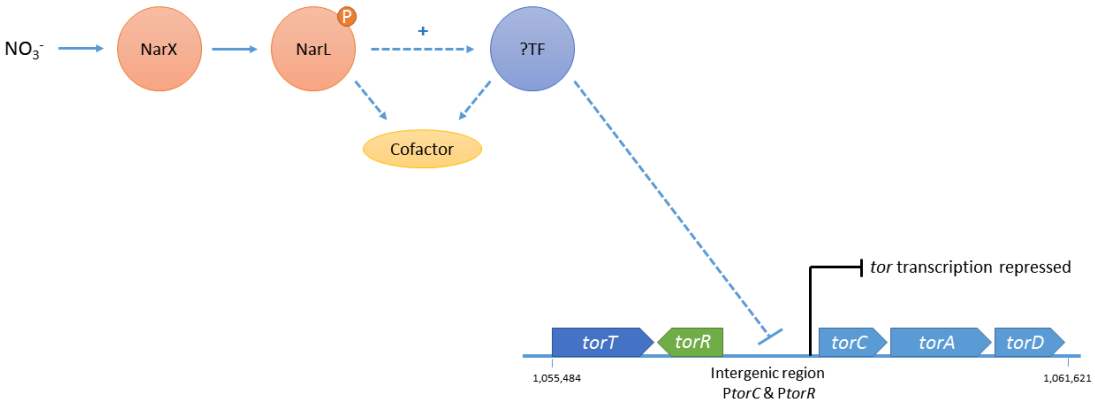


Figure 6.11. A model showing the indirect NarL-mediated regulation of *torCAD* by an unknown transcription factor

This shows a potential model of indirect NarL-mediated regulation of the *torCAD* operon. NarX detects nitrate in the environment and activates NarL. NarL is then either able to activate expression of a transcription factor of uknown identity ("?TF"), which can repress *torCAD* gene expression. Alternatively, NarL activates the expression of a cofactor which is required by the unknown transcription factor, which can then go on to repress the expression of the *torCAD* operon. Gene colour indicates operon membership.

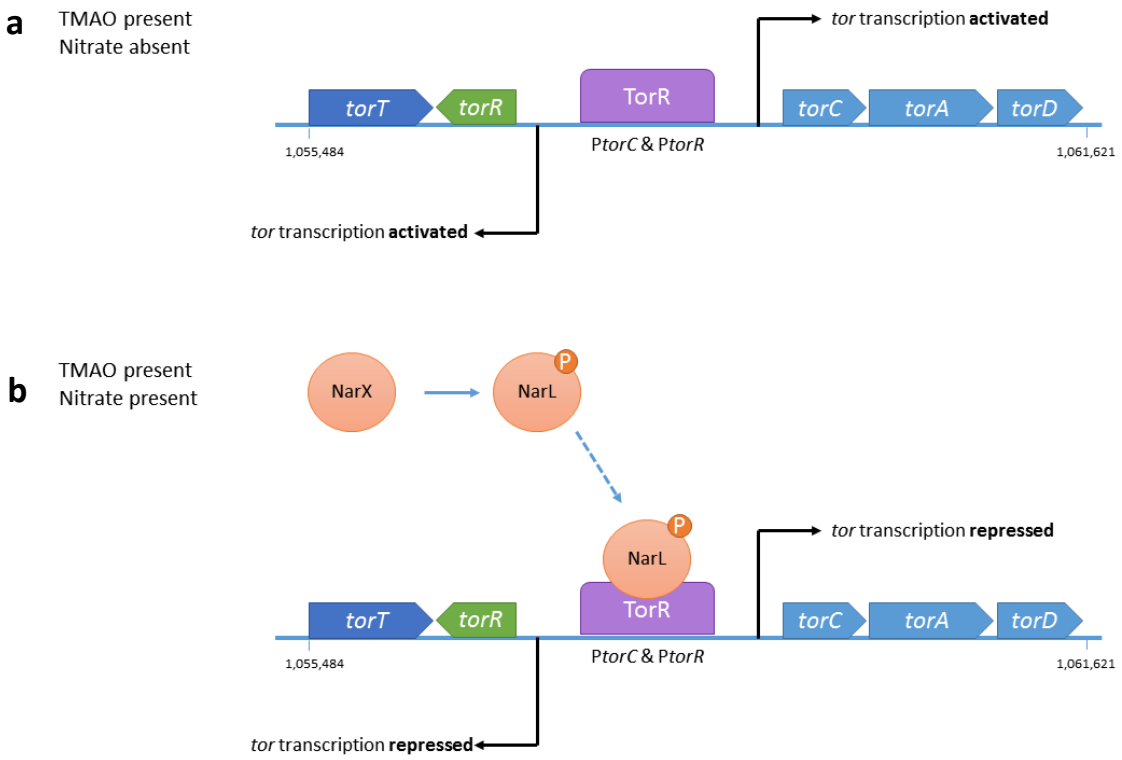


Figure 6.12. A model showing the indirect NarL-mediated regulation of *torCAD* which requires the presence of TorR for repression in the presence of nitrate

A summary of hypothesis that indirect regulation of the *torCAD* operon occurs, whereby TorR is required for NarL binding.

a. When TMAO is present and nitrate is absent, TorR binds to the *torCAD* promoter and activates the expression of the *torCAD* operon, resulting in TMAO reduction.

b. When both TMAO and nitrate are present in the environment, nitrate is detected in the environment by NarX, this eventually results in the phosphorylation of NarL. Phosphorylated NarL may be able to bind to TorR, whether bound to P_{tor} or not, causing a conformational change in TorR which may prevent TorR-mediated activation of the *torCAD* operon, meaning TMAO respiration is repressed. Gene colour indicates operon membership.

binding reactions described here, so binding through this mechanism would not have been detected. Identification of any unknown protein involved in the theoretical NarL-mediated indirect regulatory system could be the basis of further research. It is important to determine whether NarL is capable of interacting with TorR and if this process alters the conformation of TorR, which could indirectly affect *torCAD* expression. The role of the NarP protein on *torCAD* expression also needs to be studied. NarQ is a sensory kinase which detects nitrate in the environment and phosphorylates NarP to regulate NarP-regulated operons (Stewart, 1994; Wang *et al*., 1999). NarP is homologous to NarL in respect to its role as a nitrate-responsive transcriptional regulator. The NarP regulon overlaps the NarL regulon because both NarL and NarP can bind to a NarP site, which is a 7-2-7 inverted repeat sequence where the seven conserved bases form the NarL heptamers (Darwin *et al*., 1997). It is possible that NarP may form part of a regulatory system by binding to the *torCAD* promoter and repressing its transcription, therefore EMSAs with the protein should be carried out. However, there are no candidate NarP sites in the *torR-torCAD* intergenic region. This work shows that the simple hierarchical control based on substrate redox potential does not always apply.

7. Summary of findings and future directions

7.1. Summary of key findings

- New insights into the adaptive processes that occur when *E. coli* K-12 switches metabolic mode from fermentative growth to anaerobic growth with TMAO were gained.
- A key unexpected event was the detection of DMA in continuous culture supernatants and formaldehyde production by cell-free extracts in response to TMAO, which suggests that *E. coli* K-12 possesses TMAO demethylase activity (EC4.1.2.32).
- The concentration of formaldehyde produced by *E. coli* led to induction of the formaldehyde detoxification enzymes (FrmAB) and inhibited the growth of an isogenic *frmRAB* mutant when grown in the presence of excess TMAO.
- Genes involved in copper Cu(I) resistance were downregulated in response to TMAO, which could be accounted for by inhibition of Cu(II) reduction by TMAO reduction.
- The expression of the *torCAD* in response to the presence of TMAO under anaerobic conditions is not uniformly distributed across the population of cells.
- When an *E. coli* K-12 *dam* mutant was exposed to 10 mM TMAO, expression of *torCAD* expression across the population was more uniform, suggesting that methylation is involved in regulation of *torCAD* transcription.
- The *E. coli* FrmR protein acts as a repressor by specifically binding to the P_{frm} region, with the presence of formaldehyde causing dissociation of the P_{frm}-FrmR complex and relieving repression.
- FrmR specifically reacts with the formaldehyde via formation of two inter-molecular methylene bridges between Pro2 and Cys35 residues of adjacent subunits.
- Cultures expressing FrmR P2A and C35A mutations exhibited enhanced sensitivity to formaldehyde when compared to wild type FrmR.
- The presence of TMAO increases growth yield under aerobic conditions, as the pH of the culture is maintained closer to neutrality. However, production of alkaline TMA is not solely responsible for this pH balancing.
- NarL does not directly bind the P_{tor} region of *E. coli* K-12 and any nitrate-mediated regulation of TMAO reduction possibly occurs via an indirect route.

7.2. Future directions

In the future, a number of directions could be taken to build upon the insights of the work described here. It would be interesting to use other alternative electron acceptors (e.g. nitrate, nitrite, DMSO) in the experiments described here to characterise the different dynamic adaptive processes that *E. coli* K-12 utilises in order to respond to other electron acceptors.

Work could be carried out to identify the source of the TMAO demethylase activity in *E. coli* K-12. It would be interesting to see if there is a dedicated TMAO demethylase enzyme present in *E. coli* or whether another demethylase enzyme present in *E. coli* is able recognise and demethylate TMAO, in the presence of excess TMAO. If there is a dedicated enzyme and this was isolated, characterisation of the protein could be carried out.

Further investigation into the cell-to-cell variability of *torCAD* expression in response to the presence of TMAO could be carried out. It would be interesting to carry out time-lapse microscopy with live cells, to see the initiation of *torCAD* expression in individual cells upon the addition of TMAO. Fluorescence-activated cell sorting (FACS) could also be carried out to separate 'responders to TMAO' from 'non-responders' and see if non-responsders can eventually become responsive to TMAO over time.

Another route would involve further characterisation of the *E. coli* FrmR protein. Future work could focus on the role of the second Cys residue (Cys70), which is not conserved in *Se*FrmR. Solving the structure of FrmR revealed that Cys70 forms a disulfide bond linking loops 2 of the A/B´ and the B/A´ chains. The effect of this disulfide on the ability of all four subunits of the FrmR tetramer to undergo formaldehyde modification has not be assessed. Further work to determine the possible role of glutathione (*S*-hydroxymethylglutathione is formed in *E. coli* exposed to formaldehyde) to better understand the role of oxidative stress in regulating FrmR activity *in vivo* could be undertaken. Solving the structure of the P_{frm}-FrmR DNA-bound complex could also reveal insights into the DNA-binding mechanism of FrmR.

Future work could also look into the reasons why TMAO appears to sit outside of the hierarchical control of respiration in *E. coli.* It would be intesting to use a transcriptional reporter strain for *torC* to see when TorCA is expressed under aerobic conditions and rule out whether a lack of oxygen availability could have produced the TMAO reduction results previously published (Ansaldi *et al.,* 2007). Attempting to identify the means by which NarL indirectly regulates TMAO reduction could also be carried out. The role of the NarP protein

on *torCAD* expression also needs to be studied. NarQ is a sensory kinase which detects nitrate in the environment and phosphorylates NarP to regulate NarP-regulated operons (Stewart, 1994; Wang *et al*., 1999). NarP is homologous to NarL in respect to its role as a nitrate-responsive transcriptional regulator. It is possible that NarP may form part of a regulatory system by binding to the *torCAD* promoter and repressing its transcription, therefore EMSAs with the protein should be carried out. However, there are no candidate NarP sites in the *torR-torCAD* intergenic region.

8. References

Alexeeva, S. (2000). Molecular physiology of responses to oxygen in *Escherichia coli*. *PhD Thesis*. Amsterdam, the Netherlands: University of Amsterdam.

Alvarez, A.F., Rodriguez, C., and Georgellis, D. (2013). Ubiquinone and menaquinone electron carriers represent the yin and yang in the redox regulation of the ArcB sensor kinase. *Journal of Bacteriology*, **195**: 3054-3061.

Ansaldi, M., Jourlin-Castelli, C., Lepelletier, M., Théraulaz, L., and Méjean, V. (2001). Rapid dephosphorylation of the TorR response regulator by the TorS unorthodox sensor in *Escherichia coli*. *Journal of Bacteriology,* **183:** 2691-2695.

Ansaldi, M., Theraulaz, L., and Mejean, V. (2004). TorI, a response regulator inhibitor of phage origin in *Escherichia coli*. **Acta Crystallographica** Section D: Biological, **101**: 9423-9428.

Ansaldi, M., Théraulaz, L., Baraquet, C., Panis, G., and Méjean, V. (2007). Aerobic TMAO respiration in *Escherichia coli*. *Molecular Microbiolology*, **66**: 484-494.

Asif, H.M., Rolfe, M.D., Green, J., Lawrence, N.D., Rattray, M., and Sanguinetti, G. (2010). TFInfer: a tool for probabilistic inference of transcription factor activities. *Bioinformatics,* 26: 2635-2636.

Bäckhed, F. (20013). Meat-metabolizing bacteria in atherosclerosis. *Nature medicine*, 1: 533-534.

Bae, S., Ulrich, C.M., Neuhouser, M.L., Malysheva, O., Bailey, L.B., Xiao, L., Brown, E.C., Cushing-Haugen, K.L., Zheng, Y., Cheng, T.Y., et al. (2014). Plasma choline metabolites and colorectal cancer risk in the Women's Health Initiative Observational Study. *Cancer Research*, **74**: 7442-7452.

Bain, M.A., Faull, R., Fornasini, G., Milne, R.W., and Evans, A.M. (2006). Accumulation of trimethylamine and trimethylamine-*N*-oxide in end-stage renal disease patients undergoing haemodialysis. *Nephrology Dialysis Transplant*, **21:** 1300-1304.

Baraquet, C., Theraulaz, L., Guiral, M., Lafitte, D., Mejean, V., and Jourlin-Castelli, C. (2006). TorT, a member of a new periplasmic binding protein family, triggers induction of the Tor respiratory system upon trimethylamine *N*-oxide electron-acceptor binding in *Escherichia coli*. *Journal of Biological Chemistry,* **281**: 38189-38199.

Barrett, C.M., Ray, N., Thomas, J.D., Robinson, C., and Bolhuis, A. (2003). Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, **304**: 279-284.

Barrett, E.L., and Kwan, H.S. (1985). Bacterial reduction of trimethylamine oxide. *Annual Review of Microbiology*, **39:** 131-149.

Bartsch, H., and Montesano, R. (1984). Relevance of nitrosamines to human cancer. *Carcinogenesis,* **5:** 1381-1393.

Bauer, C.E., Elsen, S., and Bird, T.H. (1999). Mechanisms for redox control of gene expression. *Annual Review of Microbiology*, **53**: 495-523.

Beatty, S.A., and Gibbons, N.E. (1937). Measurement of spoilage in fish. *Journal of the Biological Board of Canada*, **3:** 77-91.

Bekker, M., Alexeeva, S., Laan, W., Sawers, G., Teixeira de Mattos, J., and Hellingwerf, K. (2010). The ArcBA two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool *Journal of Bacteriology,* **192**: 746-754.

Bell, A.I., Cole, J.A., and Busby, S.J. (1990). Molecular genetic analysis of an FNR-dependent anaerobically inducible *Escherichia coli* promoter. *Molecular Microbiology*, **4:** 1753-1763.

Bell, J.D., Lee, J.A., Lee, H.A., Sadler, P.J., Wilkie, D.R., and Woodham, R.H. (1991). Nuclear magnetic resonance studies of blood plasma and urine from subjects with chronic renal failure: identification of trimethylamine-*N*-oxide. *Biochimica* et *Biophysica Acta*, **1096**, 101-107.

Bennion, B.J., and Daggett, V. (2004). Counteraction of urea-induced protein denaturation by trimethylamine *N*-oxide: a chemical chaperone at atomic resolution. *Proceedings of the National Academy of Sciences USA,* **101**: 6433-6438.

Berks, B.C. (1996). A common export pathway for proteins binding complex redox cofactors? *Molecular Microbiology,* **22:** 393-404.

Bettenbrock, K., Bai, H., Ederer, M., Green, J., Hellingwerf, K.J., Holcombe, M., Kunz, S., Rolfe, M.D., Sanguinetti, G., Sawodny, O., et al. (2014). Towards a systems level understanding of the oxygen response of *Escherichia coli*. *Advances in Microbial Physiology*, **64:** 65-114.

Bilous, P.T., and Weiner, J.H. (1988). Molecular cloning and expression of the *Escherichia coli* dimethyl sulfoxide reductase operon. *Journal of Bacteriology,* **170:** 1511-1518.

Bock, A., and Sawers, G. (1996). *Escherichia coli* and Salmonella: cellular and molecular biology. In, F.C. Neidhardt, ed. (Washington, DC, USA: American Society for Microbiology), pp. 262-282.

Bordi, C., Theraulaz, L., Mejean, V., and Jourlin-Castelli, C. (2003). Anticipating an alkaline stress through the Tor phosphorelay system in *Escherichia coli*. *Molecular Microbiology*, **48:** 211-223.

Boye, E., and Løbner-Olesen, A. (1990). The role of dam methyltransferase in the control of DNA replication in *E. coli*. *Cell,* **62:** 981-989.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, **7**: 248-254.

Browning, D.F., Cole, J.A., and Busby, S.J. (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.

Campbell, J.L., and Kleckner, N. (1990). *E. coli oriC* and the *dnaA* gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell*, **62:** 967-979.

Casadesús, J., and Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiology and Molecular Biology Reviews*, **70:** 830-856.

Chang, F.M., Coyne, H.J., Cubillas, C., Vinuesa, P., Fang, X., Ma, Z., Ma, D., Helmann, J.D., García-de los Santos, A., Wang, Y.X., et al. (2014). Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR). *Journal of Biological Chemistry,* **289**: 19204-19217.

Chang, L., Wei, L.I., Audia, J.P., Morton, R.A., and Schellhorn, H.E. (1999). Expression of the *Escherichia coli* NRZ nitrate reductase is highly growth phase dependent and is controlled by RpoS, the alternative vegetative sigma factor. *Molecular Microbiology,* **34**: 756-766.

Chen, N.H., Djoko, K.Y., Veyrier, F.J., and McEwan, A.G. (2016). Formaldehyde Stress Responses in Bacterial Pathogens. Frontiers in Microbiology, **7**: 257.

Clegg, S.J., Jia, W., and Cole, J.A. (2006). Role of the *Escherichia coli* nitrate transport protein, NarU, in survival during severe nutrient starvation and slow growth. *Microbiology,* **152**: 2091-2100.

Human Microbiome Project Consortium. (2012). A framework for human microbiome research. *Nature,* **486:** 215-221.

Constantinidou, C., Hobman, J.L., Griffiths, L., Patel, M.D., Penn, C.W., Cole, J.A., and Overton, T.W. (2006). A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. *Journal of Biological Chemistry,* **281:** 4802-4815.

Cotter, P.A., and Gunsalus, R.P. (1989). Oxygen, nitrate, and molybdenum regulation of dmsABC gene expression in *Escherichia coli*. *Journal of Bacteriology,* **171:** 3817-3823.

Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J., and Palsson, B.O. (2004). Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, **429:** 92-96.

Crack, J., Green, J., and Thomson, A.J. (2004). Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR). *Journal of Biological Chemistry,* **279:** 9278-9286.

Crack, J.C., Green, J., Cheesman, M.R., Le Brun, N.E., and Thomson, A.J. (2007). Superoxide-mediated amplification of the oxygen-induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR. *Proceedings of the National Academy of Sciences USA,* **104:** 2092-2097.

Czjzek, M., Dos Santos, J.P., Pommier, J., Giordano, G., Mejean, V., and Haser, R. (1998). Crystal structure of oxidized trimethylamine *N*-oxide reductase from Shewanella massilia at 2.5 A resolution. Journal *of* Molecular Biology, **284**: 435-447.

D'Angelo, R., Esposito, T., Calabrò, M., Rinaldi, C., Robledo, R., Varriale, B., and Sidoti, A. (2013). FMO3 allelic variants in Sicilian and Sardinian populations: trimethylaminuria and absence of fish-like body odor. *Gene,* **515:** 410-415.

D'Argenio, V., and Salvatore, F. (2015). The role of the gut microbiome in the healthy adult status. *Clinica Chimica Acta,* **451**: 97-102.

D'Mello, R., Hill, S., and Poole, R.K. (1995). The oxygen affinity of cytochrome bo' in *Escherichia coli* determined by the deoxygenation of oxyleghemoglobin and oxymyoglobin: Km values for oxygen are in the submicromolar range. *Journal of Bacteriology,* **177:** 867-870.

D'mello, R., Hill, S., and Poole, R.K. (1996). The cytochrome *bd* quinol oxidase in *Escherichia coli* has an extremely high oxygen affinity and two oxygen-binding haems: implications for regulation of activity *in vivo* by oxygen inhibition. *Microbiology,* **142:** 755-763.

Darwin, A., Hussain, H., Griffiths, L., Grove, J., Sambongi, Y., Busby, S., and Cole, J. (1993). Regulation and sequence of the structural gene for cytochrome c552 from *Escherichia coli*: not a hexahaem but a 50 kDa tetrahaem nitrite reductase. *Molecular Microbiology,* **9:** 1255-1265.

Darwin, A.J., Li, J., and Stewart, V. (1996). Analysis of nitrate regulatory protein NarL-binding sites in the fdnG and narG operon control regions of *Escherichia coli* K-12. *Molecular Microbiology,* **20:** 621-632.

Darwin, A.J., Tyson, K.L., Busby, S.J., and Stewart, V. (1997). Differential regulation by the homologous response regulators NarL and NarP of *Escherichia coli* K-12 depends on DNA binding site arrangement. *Molecular Microbiology,* **25:** 583-595.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences USA,* **97:** 6640-6645.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* **505:** 559-563.

de Graef, M.R., Alexeeva, S., Snoep, J.L., and Teixeria de Mattos, M.J. (1999) The steady-state internal redox state (NADH/NAD) reflects the external redox state and is correlated with catabolic adaptation in *Escherichia coli*. *Journal of Bacteriology,* **181**:2351–2357.

Doucet, C., Dutheil, D., Petit, I., Zhang, K., Eugene, M., Touchard, G., Wahl, A., Seguin, F., Milinkevitch, S., Hauet, T., et al. (2004). Influence of colloid, preservation medium and trimetazidine on renal medulla injury. *Biochimica et Biophysica Acta,* **1673:** 105-114.

Drapal, N., and Sawers, G. (1995). Purification of ArcA and analysis of its specific interaction with the pfl promoter-regulatory region. *Molecular Microbiology,* **16:** 597-607.

Dumas, M.E., Maibaum, E.C., Teague, C., Ueshima, H., Zhou, B., Lindon, J.C., Nicholson, J.K., Stamler, J., Elliott, P., Chan, Q., et al. (2006). Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP Study. *Analytical Chemistry*, **78:** 2199-2208.

Eiglmeier, K., Honoré, N., Iuchi, S., Lin, E.C., and Cole, S.T. (1989). Molecular genetic analysis of FNR-dependent promoters. *Molecular Microbiology,* **3:** 869-878.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. *Science,* **297:** 1183-1186.

Farmer, W.R., and Liao, J.C. (1997). Reduction of aerobic acetate production by *Escherichia coli*. *Applied Environmental Microbiology,* **63:** 3205-3210.

Foster, A.W., Patterson, C.J., Pernil, R., Hess, C.R., and Robinson, N.J. (2012). Cytosolic Ni(II) sensor in cyanobacterium: nickel detection follows nickel affinity across four families of metal sensors. *Journal of Biological Chemistry,* **287:** 12142-12151.

Foster, A.W., Pernil, R., Patterson, C.J., and Robinson, N.J. (2014). Metal specificity of cyanobacterial nickel-responsive repressor InrS: cells maintain zinc and copper below the detection threshold for InrS. *Molecular Microbiology,* **92:** 797-812.

Fu, X.Y., Xue, C.H., Miao, B.C., Liang, J.N., Li, Z.J., and Cui, F.X. (2006). Purification and characterization of trimethylamine-*N*-oxide demethylase from jumbo squid (*Dosidicus gigas*). *Journal of Agricultural and Food Chemistry,* **54:** 968-972.

Fukami, K., Yamagishi, S., Sakai, K., Kaida, Y., Yokoro, M., Ueda, S., Wada, Y., Takeuchi, M., Shimizu, M., Yamazaki, H., et al. (2015). Oral L-carnitine supplementation increases trimethylamine-*N*-oxide but reduces markers of vascular injury in hemodialysis patients. *Journal of Cardiovascular Pharmacology,* **65:** 289-295.

Futai, M., Nakanishi-Matsui, M., Okamoto, H., Sekiya, M., and Nakamoto, R.K. (2012) Rotational catalysis in proton pumping ATPases: from *E. coli* F-ATPase to mammalian V-ATPase. *Biochimica et Biophysica Acta,* **1817**: 1711–1721.

Gao, X., Liu, X., Xu, J., Xue, C., Xue, Y., and Wang, Y. (2014). Dietary trimethylamine *N*-oxide exacerbates impaired glucose tolerance in mice fed a high fat diet. *The Journal of* Bioscience *and Bioengineering,* **118:** 476-481.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Barioch, A. (2005). Protein identification and Analysis Tools on the ExPASy Server. *In The Proteomics Protocols Handbook*, Walker, J.M., ed. (New York, USA: Humana Press): 571-607.

Geier, G.E., and Modrich, P. (1979). Recognition sequence of the dam methylase of *Escherichia coli* K12 and mode of cleavage of Dpn I endonuclease. *Journal of Biological Chemistry,* **254:** 1408-1413.

Genest, O., Mejean, V., and Iobbi-Nivol, C. (2009). Multiple roles of TorD-like chaperones in the biogenesis of molybdoenzymes. *FEMS Microbiology Letters*, **297:** 1-9.

Genest, O., Neumann, M., Seduk, F., Stöcklein, W., Méjean, V., Leimkühler, S., and Iobbi-Nivol, C. (2008). Dedicated metallochaperone connects apoenzyme and molybdenum cofactor biosynthesis components. *Journal of Biological Chemistry,* **283:** 21433-21440.

Gennis, R.B., and Stewart, V. (1996). *Escherichia coli* and Salmonella: cellular and molecular biology. In, F.C. Neidhardt, ed. (Washington, DC, USA: American Society for Microbiology): 217-261.

Georgellis, D., Kwon, O., and Lin, E.C. (1999). Amplification of signaling activity of the arc two-component system of *Escherichia coli* by anaerobic metabolites. An *in vitro* study with different protein modules. *Journal of Biological Chemistry,* **274:** 35950-35954.

Georgellis, D., Kwon, O., and Lin, E.C. (2001). Quinones as the redox signal for the arc two-component system of bacteria. *Science,* **292:** 2314-2316.

Gerasimova, A.V., Rodionov, D.A., Mironov, A.A., and Gel'fand, M.S. (2001). Computer analysis of regulatory signals in bacterial genomes. Fnr binding segments. *Journal of Molecular Biology,* **35:** 1001-1009.

Giedroc, D.P., and Arunkumar, A.I. (2007). Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Transactions*, **7:** 3107-3120.

Glickman, B.W., and Radman, M. (1980). *Escherichia coli* mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proceedings of the National Academy of Sciences USA*, **77:** 1063-1067.

Goldstein, L., and Funkhouser, D. (1972). Biosynthesis of trimethylamine oxide in the nurse shark, Ginglymostoma cirratum. *Molecular and Integrative* Physiology *of* Comparative Biochemistry *and* Physiology, **42:** 51-57.

Gon, S., Giudici-Orticoni, M.T., Mejean, V., and Iobbi-Nivol, C. (2001a). Electron transfer and binding of the c-type cytochrome TorC to the trimethylamine *N*-oxide reductase in *Escherichia coli*. *Journal of Biological Chemistry,* **276:** 11545-11551.

Gon, S., Jourlin-Castelli, C., Theraulaz, L., and Mejean, V. (2001b). An unsuspected autoregulatory pathway involving apocytochrome TorC and sensor TorS in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA,* **98:** 11615-11620.

Gon, S., Patte, J.C., Mejean, V., and Iobbi-Nivol, C. (2000). The *torYZ (yecK bisZ)* operon encodes a third respiratory trimethylamine *N*-oxide reductase in *Escherichia coli*. *Journal of Bacteriology,* **182:** 5779-5786.

Gonzalez, C.F., Proudfoot, M., Brown, G., Korniyenko, Y., Mori, H., Savchenko, A.V., and Yakunin, A.F. (2006). Molecular basis of formaldehyde detoxification. Characterization of two *S*-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *Journal of Biological Chemistry,* **281:** 14514-14522.

Gonzalez, R., Tao, H., Purvis, J.E., York, S.W., Shanmugam, K.T., and Ingram, L.O. (2003). Gene array-based identification of changes that contribute to ethanol tolerance in ethanologenic *Escherichia coli*: comparison of KO11 (parent) to LY01 (resistant mutant). *Biotechnology Progress,* **19:** 612-623.

Govantes, F., Albrecht, J.A., and Gunsalus, R.P. (2000). Oxygen regulation of the *Escherichia coli* cytochrome d oxidase (*cydAB*) operon: roles of multiple promoters and the Fnr-1 and Fnr-2 binding sites. *Molecular Microbiology,* **37:** 1456-1469.

Gralla, J. & Collado-Vides, J. (1996). Organisation and function of transcription regulatory elements*. In Regulation of Gene Expression in Esherichia coli,* Lin, E.C.C. and Lynch, A.S., eds. (Austin, Texas, USA: Chapman and Hall): 1232–1245.

Green, J., Irvine, A.S., Meng, W., and Guest, J.R. (1996). FNR-DNA interactions at natural and semi-synthetic promoters. *Molecular Microbiology,* **19:** 125-137.

Green, J., Scott, C., and Guest, J.R. (2001). Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Advances in Microbial Physiology*, **44:** 1-34.

Gross, C. A., M. Lonetto, and R. Losick. (1992). Bacterial sigma factorsTranscriptional regulation. *In Experiments in molecular genetics* (Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press): 129–176.

Grossoehme, N., Kehl-Fie, T.E., Ma, Z., Adams, K.W., Cowart, D.M., Scott, R.A., Skaar, E.P., and Giedroc, D.P. (2011). Control of copper resistance and inorganic sulfur metabolism by paralogous regulators in Staphylococcus aureus. *Journal of Biological Chemsitry,* **286:** 13522-13531.

Guest, J.R., Green, J., Irvine, A.S. and Spiro, S. (1996). The FNR modulon and FNR-regulated gene expression. *In Regulation of Gene Expression in Esherichia coli,* Lin, E.C.C. and Lynch, A.S., eds. (Austin, Texas, USA: Chapman and Hall): 317-321.

Gunsalus, R.P. (1992). Control of electron flow in *Escherichia coli*: coordinated transcription of respiratory pathway genes. *Journal of Bacteriology,* **174:** 7069-7074.

Gutheil, W.G., Holmquist, B., and Vallee, B.L. (1992). Purification, characterization, and partial sequence of the glutathione-dependent formaldehyde dehydrogenase from *Escherichia coli*: a class III alcohol dehydrogenase. *Biochemistry,* **31:** 475-481.

Harms, N., Ras, J., Reijnders, W.N., van Spanning, R.J., and Stouthamer, A.H. (1996). S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *Journal of Bacteriology,* **178:** 6296-6299.

Hatzixanthis, K., Clarke, T.A., Oubrie, A., Richardson, D.J., Turner, R.J., and Sargent, F. (2005). Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. *Proceedings of the National Academy of Sciences USA*, **102:** 8460-8465.

Hauet, T., Baumert, H., Gibelin, H., Godart, C., Carretier, M., and Eugene, M. (2000). Citrate, acetate and renal medullary osmolyte excretion in urine as predictor of renal changes after cold ischaemia and transplantation. *Clinical Chemistry* and *Laboratory Medicine,* **38:** 1093-1098.

Heithoff, D.M., Sinsheimer, R.L., Low, D.A., and Mahan, M.J. (1999). An essential role for DNA adenine methylation in bacterial virulence. *Science,* **284:** 967-970.

Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content of individual bacteria. *Applied Environmental Microbiology,* **50:** 1251-1257.

Hempfling, W.P., and Mainzer, S.E. (1975) Effects of varying the carbon source limiting growth on yield and maintenance characteristics of *Escherichia coli* in continuous culture. *Journal of Bacteriology,* **123:** 1076–1087.

Hernandez, D., Addou, S., Lee, D., Orengo, C., Shephard, E.A., and Phillips, I.R. (2003). Trimethylaminuria and a human FMO3 mutation database. *Human Mutation*, **22:** 209-213.

Herring, C.D., and Blattner, F.R. (2004). Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *Journal of Bacteriology,* **186:** 6714-6720.

Higgins, K.A., Chivers, P.T., and Maroney, M.J. (2012). Role of the *N*-terminus in determining metal-specific responses in the *E. coli* Ni- and Co-responsive metalloregulator, RcnR. *Journal of American Chemical Society,* **134:** 7081-7093.

Higgins, K.A., and Giedroc, D. (2014). Insights into Protein Allostery in the CsoR/RcnR Family of Transcriptional Repressors. *Chemistry Letters*, **43:** 20-25.

Hille, R., Retey, J., Bartlewski-Hof, U., and Reichenbecher, W. (1998). Mechanistic aspects of molybdenum-containing enzymes. *FEMS Microbiology Reviews*, **22:** 489-501.

Hoskisson, P.A., and Hobbs, G. (2005). Continuous culture--making a comeback? *Microbiology,* **151:** 3153-3159.

Huh, D., and Paulsson, J. (2011). Non-genetic heterogeneity from stochastic partitioning at cell division. *Nature Genetics*, **43:** 95-100.

Hussain, H., Grove, J., Griffiths, L., Busby, S., and Cole, J. (1994). A seven-gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria. *Molecular Microbiology,* **12:** 153-163.

Ilbert, M., Mejean, V., Giudici-Orticoni, M.T., Samama, J.P., and Iobbi-Nivol, C. (2003). Involvement of a mate chaperone (TorD) in the maturation pathway of molybdoenzyme TorA. *Journal of Biological Chemistry,* **278:** 28787-28792.

Ingledew, W.J., and Poole, R.K. (1984). The respiratory chains of *Escherichia coli*. *Microbiology Reviews*, **48:** 222-271.

Iobbi-Nivol, C., Pommier, J., Simala-Grant, J., Mejean, V., and Giordano, G. (1996). High substrate specificity and induction characteristics of trimethylamine-*N*-oxide reductase of *Escherichia coli*. *Biochimica et Biophysica Acta*, **1294:** 77-82.

Iuchi, S., Cameron, D.C., and Lin, E.C. (1989). A second global regulator gene (arcB) mediating repression of enzymes in aerobic pathways of *Escherichia coli*. *Journal of Bacteriology,* **171:** 868-873.

Iuchi, S., and Lin, E.C. (1987). The *narL* gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine *N*-oxide reductase operons in *Escherichia coli*. Proc Natl Acad Sci U S A 84, 3901-3905.

Iuchi, S., and Lin, E.C. (1988). arcA (dye), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. *Proceedings of the National Academy of Sciences USA,* **85:** 1888-1892.

Iuchi, S., and Lin, E.C. (1993). Adaptation of *Escherichia coli* to redox environments by gene expression. *Molecular Microbiology,* **9:** 9-15.

Iuchi, S., and Lin, E.C.C. (1995). Signal Transduction in the Arc System for Control of Operons Encoding Aerobic Respiratory Enzymes. *In 'Two-component signal transduction'*. (Washington, USA: American Society for Microbiology).

Iwig, J.S., and Chivers, P.T. (2009). DNA recognition and wrapping by *Escherichia coli* RcnR. *Journal of Molecular Biology,* **393:** 514-526.

Iwig, J.S., Leitch, S., Herbst, R.W., Maroney, M.J., and Chivers, P.T. (2008). Ni(II) and Co(II) sensing by *Escherichia coli* RcnR. *Journal of the American Chemical Society,* **130:** 7592-7606.

Iwig, J.S., Rowe, J.L., and Chivers, P.T. (2006). Nickel homeostasis in *Escherichia coli* - the *rcnR-rcnA* efflux pathway and its linkage to NikR function. *Molecular Microbiology,* **62:** 252-262.

Jeon, Y., Lee, Y.S., Han, J.S., Kim, J.B., and Hwang, D.S. (2001). Multimerization of phosphorylated and non-phosphorylated ArcA is necessary for the response regulator function of the Arc two-component signal transduction system. *Journal of Biological Chemistry,* **276:** 40873-40879.

Jordan, P.A., Thomson, A.J., Ralph, E.T., Guest, J.R., and Green, J. (1997). FNR is a direct oxygen sensor having a biphasic response curve. *FEBS Letters*, **416:** 349-352.

Jourlin, C., Ansaldi, M., and Mejean, V. (1997). Transphosphorylation of the TorR response regulator requires the three phosphorylation sites of the TorS unorthodox sensor in *Escherichia coli*. *Journal of Molecular Biology,* **267:** 770-777.

Jourlin, C., Bengrine, A., Chippaux, M., and Mejean, V. (1996a). An unorthodox sensor protein (TorS) mediates the induction of the tor structural genes in response to trimethylamine *N*-oxide in *Escherichia coli*. *Molecular Microbiology,* **20:** 1297-1306.

Jourlin, C., Simon, G., Pommier, J., Chippaux, M., and Mejean, V. (1996b). The periplasmic TorT protein is required for trimethylamine *N*-oxide reductase gene induction in *Escherichia coli*. *Journal of Bacteriology,* **178:** 1219-1223.

Kang, Y., Weber, K.D., Qiu, Y., Kiley, P.J., and Blattner, F.R. (2005). Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function *Journal of Bacteriology,* **187:** 1135-1160.

Keseler, I.M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., Fulcher, C., Huerta, A.M., Kothari, A., Krummenacker, M., et al. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Research*, **41:** D605-612.

Khoroshilova, N., Popescu, C., Münck, E., Beinert, H., and Kiley, P.J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proceedings of the National Academy of Sciences USA,* **94:** 6087-6092.

Kiley, P.J., and Beinert, H. (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. FEMS Microbiology Reviews, **22:** 341-352.

Klig, L.S., Carey J. and Yanofsky C. (1988). *trp* repressor interactions with the *trp aroH* and *trpR* operators. Comparison of repressor binding *in vitro* and repression *in vivo*. *Journal of Molecular Biology*, **202**:769–777.

Koeth, R.A., Wang, Z., Levison, B.S., Buffa, J.A., Org, E., Sheehy, B.T., Britt, E.B., Fu, X., Wu, Y., Li, L., et al. (2013). Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature Medicine,* **19:** 576-585.

Koyama, Y., and Ohmori, H. (1996). Nucleotide sequence of the *Escherichia coli solA* gene encoding a sarcosine oxidase-like protein and characterization of its product. *Gene,* **181:** 179-183.

Kuczler, F.J., Nahrwold, D.L., and Rose, R.C. (1977). Choline influx across the brush border of guinea pig jejunum. *Biochimica et Biophysica Acta*, **465:** 131-137.

Lai, Y., Yu, R., Hartwell, H.J., Moeller, B.C., Bodnar, W.M., and Swenberg, J.A. (2016). Measurement of Endogenous versus Exogenous Formaldehyde-Induced DNA-Protein Crosslinks in Animal Tissues by Stable Isotope Labeling and Ultrasensitive Mass Spectrometry. *Cancer Research,* **76:** 2652-2661.

Lazazzera, B.A., Beinert, H., Khoroshilova, N., Kennedy, M.C., and Kiley, P.J. (1996). DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. .*Journal of Biological Chemistry,* **271:** 2762-2768.

Lee, A.I., Delgado, A., and Gunsalus, R.P. (1999). Signal-dependent phosphorylation of the membrane-bound NarX two-component sensor-transmitter protein of *Escherichia coli*: nitrate elicits a superior anion ligand response compared to nitrite. *Journal of Bacteriology,* **181:** 5309-5316.

Lee, D.J., Bingle, L.E., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J., and Hobman, J.L. (2009). Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiology*, **9:** 252.

Lenz, E.M., Bright, J., Wilson, I.D., Hughes, A., Morrisson, J., Lindberg, H., and Lockton, A. (2004). Metabonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *Journal of Pharmaceutical and Biomedical Analysis*, **36:** 841-849.

Lever, M., George, P.M., Slow, S., Bellamy, D., Young, J.M., Ho, M., McEntyre, C.J., Elmslie, J.L., Atkinson, W., Molyneux, S.L., et al. (2014). Betaine and Trimethylamine-*N*-Oxide as Predictors of Cardiovascular Outcomes Show Different Patterns in Diabetes Mellitus: An Observational Study. *PLoS One,* **9:** e114969.

Levovich, I., Nudelman, A., Berkovitch, G., Swift, L.P., Cutts, S.M., Phillips, D.R., and Rephaeli, A. (2008). Formaldehyde-releasing prodrugs specifically affect cancer cells by

depletion of intracellular glutathione and augmentation of reactive oxygen species. *Cancer Chemotherapy Pharmacology,* **62:** 471-482.

Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell,* **157:** 624-635.

Li, X., Chen, Y., Liu, J., Yang, G., Zhao, J., Liao, G., Shi, M., Yuan, Y., He, S., Lu, Y., et al. (2012). Serum metabolic variables associated with impaired glucose tolerance induced by high-fat-high-cholesterol diet in *Macaca mulatta*. *Experimental Biology and Medicine*, **237:** 1310-1321.

Lidbury, I., Murrell, J.C., and Chen, Y. (2014). Trimethylamine *N*-oxide metabolism by abundant marine heterotrophic bacteria. *Proceedings of the National Academy of Sciences USA,* **111:** 2710-2715.

Liffourrena, A.S., Salvano, M.A., and Lucchesi, G.I. (2010). *Pseudomonas putida* A ATCC 12633 oxidizes trimethylamine aerobically via two different pathways. *Archives in Microbiology,* 192: 471-476.

Lijinsky, W., Keefer, L., Conrad, E., and Van de Bogart, R. (1972). Nitrosation of tertiary amines and some biologic implications. *Journal of the National Cancer Institute,* **49:** 1239-1249.

Lin, E.C. (1976). Glycerol dissimilation and its regulation in bacteria. *Annual Reviews in Microbiology*, **30:** 535-578.

Lin, E.C., and Iuchi, S. (1991). Regulation of gene expression in fermentative and respiratory systems in *Escherichia coli* and related bacteria. *Annual Reviews in Genetics*, **25:** 361-387.

Liu, L., Hausladen, A., Zeng, M., Que, L., Heitman, J., and Stamler, J.S. (2001). A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans. *Nature,* **410:** 490-494.

Liu, T., Ramesh, A., Ma, Z., Ward, S.K., Zhang, L., George, G.N., Talaat, A.M., Sacchettini, J.C., and Giedroc, D.P. (2007). CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nature Chemical Biology,* **3:** 60-68.

Liu, X., and De Wulf, P. (2004). Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *Journal of Biological Chemistry,* **279:** 12588-12597.

Lundstrom, R.C., and Racicot, L.D. (1983). Gas chromatographic determination of dimethylamine and trimethylamine in seafoods. *Journal of the Association of Official Analytical Chemists*, **66:** 1158-1163.

Lunn, J.C., Kuhnle, G., Mai, V., Frankenfeld, C., Shuker, D.E., Glen, R.C., Goodman, J.M., Pollock, J.R., and Bingham, S.A. (2007). The effect of haem in red and processed meat on the endogenous formation of *N*-nitroso compounds in the upper gastrointestinal tract. *Carcinogenesis,* **28:** 685-690.

Lupachyk, S., Watcho, P., Stavniichuk, R., Shevalye, H., and Obrosova, I.G. (2013). Endoplasmic reticulum stress plays a key role in the pathogenesis of diabetic peripheral neuropathy. *Diabetes,* **62:** 944-952.

Lynch, A.S., and Lin, E.C. (1996). Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli*: characterization of DNA binding at target promoters. *Journal of Bacteriology,* **178:** 6238-6249.

Løbner-Olesen, A., Marinus, M.G., and Hansen, F.G. (2003). Role of SeqA and Dam in *Escherichia coli* gene expression: a global/microarray analysis. *Proceedings of the National Academy of Sciences USA*, **100:** 4672-4677.

Ma, J., Pazos, I.M., and Gai, F. (2014). Microscopic insights into the protein-stabilizing effect of trimethylamine *N*-oxide (TMAO). *Proceedings of the National Academy of Sciences USA*, **111:** 8476-8481.

Ma, Z., Jacobsen, F.E., and Giedroc, D.P. (2009). Coordination chemistry of bacterial metal transport and sensing. *Chemical Reviews*, **109:** 4644-4681.

Malpica, R., Franco, B., Rodriguez, C., Kwon, O., and Georgellis, D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proceedings of the National Academy of Sciences USA*, **101:** 13318-13323.

Marinus, M.G. (1976). Adenine methylation of Okazaki fragments in *Escherichia coli*. *Journal of Bacteriology,* **128:** 853-854.

Marinus, M.G., and Løbner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus*, **6**.

Marinus, M.G., and Morris, N.R. (1973). Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. *Journal of Bacteriology,* **114:** 1143-1150.

Martıínez-Antonio, A., Salgado, H., Gama-Castro, S., Gutie´rrez-Rı´os, R. M., Jime´nezJacinto, V., and Collado-Vides, J. (2003). Environmental conditions and transcriptional regulation in *Escherichia coli*: A physiological integrative approach. *Biotechnology and Bioengineering*, **84:** 743–749.

McCrindle, S.L., Kappler, U., and McEwan, A.G. (2005). Microbial dimethylsulfoxide and trimethylamine-*N*-oxide respiration. *Advances in Microbial Physiology,* **50:** 147-198.

Mejean, V., Iobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M., and Pascal, M.C. (1994). TMAO anaerobic respiration in *Escherichia coli*: involvement of the tor operon. *Molecular Microbiology,* **11:** 1169-1179.

Mettert, E.L., and Kiley, P.J. (2005). ClpXP-dependent proteolysis of FNR upon loss of its O2-sensing [4Fe-4S] cluster. *Journal of Molecular Biology*, **354:** 220-232.

Metz, B., Kersten, G.F., Hoogerhout, P., Brugghe, H.F., Timmermans, H.A., de Jong, A., Meiring, H., ten Hove, J., Hennink, W.E., Crommelin, D.J., et al. (2004). Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *Journal of Biological Chemistry*, **279:** 6235-6243.

Miller, C.A., Corbin, K.D., da Costa, K.A., Zhang, S., Zhao, X., Galanko, J.A., Blevins, T., Bennett, B.J., O'Connor, A., and Zeisel, S.H. (2014). Effect of egg ingestion on trimethylamine-*N*-oxide production in humans: a randomized, controlled, dose-response study. *American Journal of Clinical Nutrition*, **100**: 778-786.

Miller, J.H. (1972). Experiments in molecular genetics. (Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press).

Mondal, J., Stirnemann, G., and Berne, B.J. (2013). When does trimethylamine *N*-oxide fold a polymer chain and urea unfold it? *Journal of Physical Chemistry B*, **117:** 8723-8732.

Moore, J.O., and Hendrickson, W.A. (2012). An asymmetry-to-symmetry switch in signal transmission by the histidine kinase receptor for TMAO. *Structure,* **20:** 729-741.

Moore, L.J., and Kiley, P.J. (2001). Characterization of the dimerization domain in the FNR transcription factor. *Journal of Biological Chemistry*, **276:** 45744-45750.

Myers, K.S., Yan, H., Ong, I.M., Chung, D., Liang, K., Tran, F., Keleş, S., Landick, R., and Kiley, P.J. (2013). Genome-scale analysis of *escherichia coli* FNR reveals complex features of transcription factor binding. *PLoS Genetics,* **9:** e1003565.

Méjean, V., Iobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M., and Pascal, M.C. (1994). TMAO anaerobic respiration in *Escherichia coli*: involvement of the *tor* operon. *Molecular Microbiology*, **11:** 1169-1179.

Nash, T. (1953). The colorimetric estimation of formaldehyde by means of the Hantzsch reaction. *Biochemical Journal*, **55:** 416-421.

Neidhardt, F.C., Ingraham, J.L., and Schaechter, M. (1990). Physiology of the Bacterial Cell. (Sunderland, Massachusetts, USA: Sinauer Associates): 4.

Norat, T., Bingham, S., Ferrari, P., Slimani, N., Jenab, M., Mazuir, M., Overvad, K., Olsen, A., Tjønneland, A., Clavel, F., et al. (2005). Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. *Journal of the National Cancer Institute*, **97:** 906-916.

Oshima, T., Wada, C., Kawagoe, Y., Ara, T., Maeda, M., Masuda, Y., Hiraga, S., and Mori, H. (2002). Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*. *Molecular Microbiology,* **45:** 673-695.

Osman, D., Piergentili, C., Chen, J., Chakrabarti, B., Foster, A.W., Lurie-Luke, E., Huggins, T.G., and Robinson, N.J. (2015). Generating a Metal-responsive Transcriptional Regulator to Test What Confers Metal Sensing in Cells. *Journal of Biological Chemistry,* **290:** 19806-19822.

Osman, D., Piergentili, C., Chen, J., Sayer, L. N., Uson, I., Huggins, T. G., Robinson, N. J. & Pohl, E. (2016). The effectors and sensory sites of formaldehyde-responsive regulator FrmR and metal-sensing variant. *Journal of Biological Chemistry,* **doi: 10.1074/jbc.M116.745174**.

Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nature Genetics*, **31:** 69-73.

Park, D.M, Sohail Akhtar, M.D., Ansari, A.Z., Landick, R. and Kiley, P.J. (2013). The bacterial response regulator ArcA uses a diverse binding site architecture to regulate carbon oxidation globally. *PLoS Genetics.* **9:** 1-18.

Partridge, J.D., Sanguinetti, G., Dibden, D.P., Roberts, R.E., Poole, R.K., and Green, J. (2007). Transition of *Escherichia coli* from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components. *Journal of Biological Chemistry*, **282:** 11230-11237.

Partridge, J.D., Scott, C., Tang, Y., Poole, R.K., and Green, J. (2006). *Escherichia coli* transcriptome dynamics during the transition from anaerobic to aerobic conditions. *Journal of Biological Chemistry*, **281:** 27806-27815.

Pascal, M.C., Burini, J.F., and Chippaux, M. (1984). Regulation of the trimethylamine *N*-oxide (TMAO) reductase in *Escherichia coli*: analysis of tor::Mud1 operon fusion. *Molecular Genome and Genetics*, **195:** 351-355.

Pascal, M.C., Lepelletier, M., Giordano, G., and Chippaux, M. (1991). A regulatory mutant of the trimethylamine *N*-oxide reductase of *Escherichia coli* K12. *FEMS Microbiology Letters*, **62:** 297-300.

Patschkowski, T., Bates, D.M., and Kiley, P.J. (2000). Mechanism for sensing and responding to oxygen deprivation. *In Bacterial Stress Response*. (Washington, USA: ASM Press).

Pommier, J., Mejean, V., Giordano, G., and Iobbi-Nivol, C. (1998). TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine *N*-oxide reductase enzyme (TorA) in *Escherichia coli*. *Journal of Biological Chemistry*, **273:** 16615-16620.

Popescu, C.V., Bates, D.M., Beinert, H., Münck, E., and Kiley, P.J. (1998). Mössbauer spectroscopy as a tool for the study of activation/inactivation of the transcription regulator FNR in whole cells of *Escherichia coli*. *Proceedings of the National Academy of Sciences USA,* **95:** 13431-13435.

Porto, T.V., Hough, M.A., and Worrall, J.A. (2015). Structural insights into conformational switching in the copper metalloregulator CsoR from *Streptomyces lividans*. *Acta Crystallographica* Section D: Biological Crystallography, 71, 1872-1878.

Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiology Reviews*, **57:** 50-108.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature,* **464:** 59-65.

Quastel, J.H., and Stephenson, M. (1925). Further Observations on the Anaerobic Growth of Bacteria. *Biochemical Journal*, **19:** 660-666.

Quastel, J.H., Stephenson, M., and Whetham, M.D. (1925). Some Reactions of Resting Bacteria in Relation to Anaerobic Growth. *Biochemical Journal*, **19:** 304-317.

Rabin, R.S., and Stewart, V. (1993). Dual response regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in *Escherichia coli* K-12. *Journal of Bacteriology*, **175:** 3259-3268.

Rasmussen, L.G., Winning, H., Savorani, F., Toft, H., Larsen, T.M., Dragsted, L.O., Astrup, A., and Engelsen, S.B. (2012). Assessment of the effect of high or low protein diet on the human urine metabolome as measured by NMR. *Nutrients,* **4:** 112-131.

Rigault, C., Le Borgne, F., and Demarquoy, J. (2006). Genomic structure, alternative maturation and tissue expression of the human BBOX1 gene. *Biochimica et Biophysica Acta*, **1761:** 1469-1481.

Rigault, C., Mazué, F., Bernard, A., Demarquoy, J., and Le Borgne, F. (2008). Changes in l-carnitine content of fish and meat during domestic cooking. *Meat Science*, **78:** 331-335.

Ringquist, S., and Smith, C.L. (1992). The *Escherichia coli* chromosome contains specific, unmethylated *dam* and *dcm* sites. *Proceedings of the National Academy of Sciences USA*, **89:** 4539-4543.

Riva, A., Delorme, M.O., Chevalier, T., Guilhot, N., Hénaut, C., and Hénaut, A. (2004). Characterization of the GATC regulatory network in *E. coli*. *BMC Genomics*, **5**: 48.

Rodriguez, C., Kwon, O., and Georgellis, D. (2004). Effect of D-lactate on the physiological activity of the ArcB sensor kinase in *Escherichia coli*. *Journal of Bacteriology,* **186:** 2085-2090.

Roggiani, M., and Goulian, M. (2015). Oxygen-Dependent Cell-to-Cell Variability in the Output of the *Escherichia coli* Tor Phosphorelay. *Journal of Bacteriology,* **197:** 1976-1987.

Rolfe, M.D., Ocone, A., Stapleton, M.R., Hall, S., Trotter, E.W., Poole, R.K., Sanguinetti, G., Green, J. (2012). Systems analysis of transcription factor activities in environments with stable and dynamic oxygen concentrations. *Open Biology,* **2:** 120091.

Rolfe, M.D., Ter Beek, A., Graham, A.I., Trotter, E.W., Asif, H.M., Sanguinetti, G., de Mattos, J.T., Poole, R.K., and Green, J. (2011). Transcript profiling and inference of *Escherichia coli* K-12 ArcA activity across the range of physiologically relevant oxygen concentrations. *Journal of Biological Chemistry*, **286:** 10147-10154.

Rubio-Sanz, L., Prieto, R.I., Imperial, J., Palacios, J.M., and Brito, B. (2013). Functional and expression analysis of the metal-inducible *dmeRF* system from *Rhizobium leguminosarum bv. viciae*. *Applied Environmental Microbiology*, **79:** 6414-6422.

Sakamoto, K., Agari, Y., Agari, K., Kuramitsu, S., and Shinkai, A. (2010). Structural and functional characterization of the transcriptional repressor CsoR from *Thermus thermophilus* HB8. *Microbiology,* **156:** 1993-2005.

Salmon, K., Hung, S.P., Mekjian, K., Baldi, P., Hatfield, G.W., and Gunsalus, R.P. (2003). Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. *Journal of Biological Chemistry*, **278:** 29837-29855.

Salmon, K.A., Hung, S.P., Steffen, N.R., Krupp, R., Baldi, P., Hatfield, G.W., and Gunsalus, R.P. (2005). Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. *Journal of Biological Chemistry*, **280:** 15084-15096.

Sambasivarao, D., Scraba, D.G., Trieber, C., and Weiner, J.H. (1990). Organization of dimethyl sulfoxide reductase in the plasma membrane of *Escherichia coli*. *Journal of Biological Chemistry*, **172:** 5938-5948.

Sanford, P.A., and Smyth, D.H. (1971). Intestinal transfer of choline in rat and hamster. *Journal of Physiology,* **215:** 769-788.

Sanguinetti, G., Lawrence, N.D., and Rattray, M. (2006). Probabilistic inference of transcription factor concentrations and gene-specific regulatory activities. *Bioinformatics,* **22:** 2775-2781.

Sarma, R., and Paul, S. (2013). Exploring the molecular mechanism of trimethylamine-*N*-oxide's ability to counteract the protein denaturing effects of urea. *Journal of Physical Chemistry*, **117:** 5691-5704.

Sawers, G. (1999). The aerobic/anaerobic interface. *Current Opinion in Microbiology,* **2:** 181-187.

Scott, C., and Green, J. (2002). Miscoordination of the iron-sulfur clusters of the anaerobic transcription factor, FNR, allows simple repression but not activation. *Journal of Biological Chemistry*, **277:** 1749-1754.

Sheard, N.F., and Zeisel, S.H. (1986). An *in vitro* study of choline uptake by intestine from neonatal and adult rats. *Paediatric Research*, **20:** 768-772.

Shimokawa, O., and Ishimoto, M. (1979). Purification and some properties of inducible tertiary amine *N*-oxide reductase from *Escherichia coli*. *Journal of Biochemistry*, **86:** 1709-1717.

Simon, G., Jourlin, C., Ansaldi, M., Pascal, M.C., Chippaux, M., and Méjean, V. (1995). Binding of the TorR regulator to cis-acting direct repeats activates tor operon expression. *Molecular Microbiology*, **17:** 971-980.

Simon, G., Mejean, V., Jourlin, C., Chippaux, M., and Pascal, M.C. (1994). The *torR* gene of *Escherichia coli* encodes a response regulator protein involved in the expression of the trimethylamine *N*-oxide reductase genes. *Journal of Bacteriology*, **176:** 5601-5606.

Slonczewski, J., and Foster, J.W. (2009). Microbiology: an evolving science (New York, N.Y.: W.W. Norton and Co.).

Sotelo, C.G., Piñeiro, C., and Pérez-Martín, R.I. (1995). Denaturation of fish proteins during frozen storage: role of formaldehyde. *Zeitschrift für Le0bensmittel-Untersuchung und – Forschung*, **200:** 14-23.

Spiro, S. (1994). The FNR family of transcriptional regulators. *Antonie Van Leeuwenhoek*, **66:** 23-36.

Spiro, S., and Guest, J.R. (1991). Adaptive responses to oxygen limitation in *Escherichia coli*. *Trends in Biochemical Science*, **16:** 310-314.

Straney, S. B., and Crothers, D.M. (1987). Lac repressor is a transient gene activating protein. *Cell,* **51**:699–707.

Stewart, V. (1993). Nitrate regulation of anaerobic respiratory gene expression in *Escherichia coli*. *Molecular Microbiology,* **9:** 425-434.

Stewart, V. (1994). Dual interacting two-component regulatory systems mediate nitrate- and nitrite-regulated gene expression in *Escherichia coli*. *Research in Microbiology*, **145:** 450-454.

Suwa, A. (1909). Reduction of Trimethylamine oxide. *Archiv im* Gesamten *Physiologie,* **128:** 421-426.

Svensson, B.G., Akesson, B., Nilsson, A., and Paulsson, K. (1994). Urinary excretion of methylamines in men with varying intake of fish from the Baltic Sea. *Journal of Toxicology and Environmental Health*, **41:** 411-420.

Takagi, M., Ishimoto, M. (1983). *Escherichia coli* mutants defective in trimethylamine *N*-oxide reductase. *FEMS Microbiology Letters,* **17:** 247-250.

Takagi, M., Tsuchiya, T., and Ishimoto, M. (1981). Proton translocation coupled to trimethylamine *N*-oxide reduction in anaerobically grown *Escherichia coli*. *Journal of Bacteriology*, **148:** 762-768.

Takeuchi, K., Hatanaka, A., Kimura, M., Seki, N., Kimura, I., Yamada, S., and Yamashita, S. (2003). Aspolin, a novel extremely aspartic acid-rich protein in fish muscle, promotes iron-mediated demethylation of trimethylamine-*N*-oxide. *Journal of Biological Chemistry,* **278:** 47416-47422.

Tan, B.G., Vijgenboom, E., and Worrall, J.A. (2014). Conformational and thermodynamic hallmarks of DNA operator site specificity in the copper sensitive operon repressor from *Streptomyces lividans*. *Nucleic Acids Research*, **42:** 1326-1340.

Tang, H., Rothery, R.A., Voss, J.E., and Weiner, J.H. (2011). Correct assembly of iron-sulfur cluster FS0 into *Escherichia coli* dimethyl sulfoxide reductase (DmsABC) is a prerequisite for molybdenum cofactor insertion. *Journal of Biological Chemistry*, **286:** 15147-15154.

Tang, W.H., Wang, Z., Levison, B.S., Koeth, R.A., Britt, E.B., Fu, X., Wu, Y., and Hazen, S.L. (2013). Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *New England Journal of Medicine*, **368:** 1575-1584.

Thauer, R.K., Jungermann, K., and Decker, K. (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriological Reviews*, **41:** 100-180.

Thelander L. (1973) Physicochemical characterization of ribonucleoside diphosphate reductase from *Escherichia coli*. *Journal of Biological Chemistry*, **248:** 4591–4601

Thomas, J.D., Daniel, R.A., Errington, J., and Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Molecular Microbiology*, **39:** 47-53.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature,* **419:** 174-178.

Trotter, E.W., Rolfe, M.D., Hounslow, A.M., Craven, C.J., Williamson, M.P., Sanguinetti, G., Poole, R.K., and Green, J. (2011). Reprogramming of *Escherichia coli* K-12 metabolism during the initial phase of transition from an anaerobic to a micro-aerobic environment. *PLoS One,* **6:** e25501.

Trøseid, M., Ueland, T., Hov, J.R., Svardal, A., Gregersen, I., Dahl, C.P., Aakhus, S., Gude, E., Bjørndal, B., Halvorsen, B., et al. (2015). Microbiota-dependent metabolite trimethylamine-*N*-oxide is associated with disease severity and survival of patients with chronic heart failure. *Journal of International Medicine*, **277:** 717-726.

Tyson, K.L., Bell, A.I., Cole, J.A., and Busby, S.J. (1993). Definition of nitrite and nitrate response elements at the anaerobically inducible *Escherichia coli nirB* promoter: interactions between FNR and NarL. *Molecular Microbiology*, **7:** 151-157.

Tyson, K.L., Cole, J.A., and Busby, S.J. (1994). Nitrite and nitrate regulation at the promoters of two *Escherichia coli* operons encoding nitrite reductase: identification of common target heptamers for both NarP- and NarL-dependent regulation. *Molecular Microbiology,* **13:** 1045-1055.

Ufnal, M., Jazwiec, R., Dadlez, M., Drapala, A., Sikora, M., and Skrzypecki, J. (2014). Trimethylamine-*N*-oxide: a carnitine-derived metabolite that prolongs the hypertensive effect of angiotensin II in rats. *Canadian Journal of Cardiology*, **30:** 1700-1705.

Ufnal, M., Zadlo, A., and Ostaszewski, R. (2015). TMAO: A small molecule of great expectations. *Nutrition,* **31:** 1317-1323.

Unden, G., Becker, S., Bongaerts, J., Holighaus, G., Schirawski, J., and Six, S. (1995). O_2-sensing and O_2-dependent gene regulation in facultatively anaerobic bacteria. *Archives in Microbiology*, **164:** 81-90.

Unden, G., Becker, S., Bongaerts, J., Schirawski, J., and Six, S. (1994). Oxygen regulated gene expression in facultatively anaerobic bacteria. *Antonie Van Leeuwenhoek,* **66:** 3-22.

Unden, G., and Bongaerts, J. (1997). Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochimica et Biophysica Acta,* **1320:** 217-234.

Ussher, J.R., Lopaschuk, G.D., and Arduini, A. (2013). Gut microbiota metabolism of L-carnitine and cardiovascular risk. *Atherosclerosis,* **231:** 456-461.

Valentin-Hansen, P., L. Søgaard-Andersen, and H. Pedersen. (1996). A flexible partnership: the CytR anti-activator and the cAMP-CRP activator protein, comrades in transcription control. *Molecular Microbiology*, **20**: 461–466.

Volbeda, A., Darnault, C., Renoux, O., Nicolet, Y., and Fontecilla-Camps, J.C. (2015). The crystal structure of the global anaerobic transcriptional regulator FNR explains its extremely fine-tuned monomer-dimer equilibrium. *Science Advances*, **1:** e1501086.

Volentini, S.I., Farías, R.N., Rodríguez-Montelongo, L., and Rapisarda, V.A. (2011). Cu(II)-reduction by *Escherichia coli* cells is dependent on respiratory chain components. *Biometals,* **24:** 827-835.

Wang, H., Tseng, C.P., and Gunsalus, R.P. (1999). The *napF* and *narG* nitrate reductase operons in *Escherichia coli* are differentially expressed in response to submicromolar concentrations of nitrate but not nitrite. *Journal of Bacteriology*, **181:** 5303-5308.

Wang, M.X., and Church, G.M. (1992). A whole genome approach to *in vivo* DNA-protein interactions in *E. coli*. *Nature,* **360:** 606-610.

Wang, Z., Klipfell, E., Bennett, B.J., Koeth, R., Levison, B.S., Dugar, B., Feldstein, A.E., Britt, E.B., Fu, X., Chung, Y.M., et al. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature,* **472:** 57-63.

Wang, Z., Tang, W.H., Buffa, J.A., Fu, X., Britt, E.B., Koeth, R.A., Levison, B.S., Fan, Y., Wu, Y., and Hazen, S.L. (2014). Prognostic value of choline and betaine depends on intestinal microbiota-generated metabolite trimethylamine-*N*-oxide. *European Heart Journal*, **35:** 904-910.

Weiner, J.H., MacIsaac, D.P., Bishop, R.E., and Bilous, P.T. (1988). Purification and properties of *Escherichia coli* dimethyl sulfoxide reductase, an iron-sulfur molybdoenzyme with broad substrate specificity*. Journal of Bacteriology*, **170:** 1505-1510.

Wendland, M., and Bumann, D. (2002). Optimization of GFP levels for analyzing Salmonella gene expression during an infection. *FEBS Letters*, **521:** 105-108.

West, A.A., Shih, Y., Wang, W., Oda, K., Jaceldo-Siegl, K., Sabaté, J., Haddad, E., Rajaram, S., Caudill, M.A., and Burns-Whitmore, B. (2014). Egg n-3 fatty acid composition modulates biomarkers of choline metabolism in free-living lacto-ovo-vegetarian women of reproductive age. *Journal of Academic Nutrition and Diet*, **114:** 1594-1600.

Williams, S.B., and Stewart, V. (1997). Discrimination between structurally related ligands nitrate and nitrite controls autokinase activity of the NarX transmembrane signal transducer of *Escherichia coli* K-12. *Molecular Microbiology*, **26:** 911-925.

Winstone, T.M., Tran, V.A., and Turner, R.J. (2013). The hydrophobic region of the DmsA twin-arginine leader peptide determines specificity with chaperone DmsD. *Biochemistry*, **52:** 7532-7541.

Wissenbach, U., Kroger, A., and Unden, G. (1990). The specific functions of menaquinone and demethylmenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamine *N*-oxide and nitrate by *Escherichia coli*. *Archives in Microbiology*, **154:** 60-66.

Wissenbach, U., Ternes, D., and Unden, G. (1992). An *Escherichia coli* mutant containing only demethylmenaquinone, but no menaquinone: effects on fumarate, dimethylsulfoxide, trimethylamine *N*-oxide and nitrate respiration. *Archives in Microbiology*, **158:** 68-73.

Woltjer, R.L., McMahan, W., Milatovic, D., Kjerulf, J.D., Shie, F.S., Rung, L.G., Montine, K.S., and Montine, T.J. (2007). Effects of chemical chaperones on oxidative stress and detergent-insoluble species formation following conditional expression of amyloid precursor protein carboxy-terminal fragment. *Neurobiology of Disease*, **25:** 427-437.

Wu, H., Tyson, K.L., Cole, J.A., and Busby, S.J. (1998). Regulation of transcription initiation at the *Escherichia coli nir* operon promoter: a new mechanism to account for co-dependence on two transcription factors. *Molecular Microbiology*, **27:** 493-505.

Xie, M.Z., Shoulkamy, M.I., Salem, A.M., Oba, S., Goda, M., Nakano, T., and Ide, H. (2016). Aldehydes with high and low toxicities inactivate cells by damaging distinct cellular targets. *Mutation Research*, **786:** 41-51.

Yahr, T.L., and Wickner, W.T. (2001). Functional reconstitution of bacterial Tat translocation *in vitro*. *EMBO Journal*, **20:** 2472-2479.

Yamamoto, I., Okubo, N., and Ishimoto, M. (1986). Further characterization of trimethylamine *N*-oxide reductase from *Escherichia coli*, a molybdoprotein. *Journal of Biochemistry*, **99:** 1773-1779.

Yancey, P.H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *Journal of Experimental Biology*, **208:** 2819-2830.

Yancey, P.H., Fyfe-Johnson, A.L., Kelly, R.H., Walker, V.P., and Auñón, M.T. (2001). Trimethylamine oxide counteracts effects of hydrostatic pressure on proteins of deep-sea teleosts. *Journal of Experimental Zoology,* **289:** 172-176.

Yano, T., Takigami, E., Yurimoto, H., and Sakai, Y. (2009). Yap1-regulated glutathione redox system curtails accumulation of formaldehyde and reactive oxygen species in methanol metabolism of Pichia pastoris. *Eukaryotic Cell*, **8:** 540-549.

Yu, R., Lai, Y., Hartwell, H.J., Moeller, B.C., Doyle-Eisele, M., Kracko, D., Bodnar, W.M., Starr, T.B., and Swenberg, J.A. (2015). Formation, Accumulation, and Hydrolysis of Endogenous and Exogenous Formaldehyde-Induced DNA Damage. *Toxicology Science*, **146:** 170-182.

Zeisel, S.H. (1981). Dietary choline: biochemistry, physiology, and pharmacology. *Annual Reviews in Nutrition*, **1:** 95-121.

Zeisel, S.H., DaCosta, K.A., and Fox, J.G. (1985). Endogenous formation of dimethylamine. *Biochemical Journal*, **232:** 403-408.

Zeisel, S.H., daCosta, K.A., Youssef, M., and Hensey, S. (1989). Conversion of dietary choline to trimethylamine and dimethylamine in rats: dose-response relationship. *Journal of Nutrition*, **119:** 800-804.

Zeisel, S.H., Wishnok, J.S., and Blusztajn, J.K. (1983). Formation of methylamines from ingested choline and lecithin. *Journal of Pharmacology and Experimental Therapy*, **225:** 320-324.

Zhang, L., Nelson, K.J., Rajagopalan, K.V., and George, G.N. (2008). Structure of the molybdenum site of *Escherichia coli* trimethylamine *N*-oxide reductase. *Inorganic Chemistry*, **47:** 1074-1078.

Zhang, Y., Liu, X., McHale, C., Li, R., Zhang, L., Wu, Y., Ye, X., Yang, X., and Ding, S. (2013). Bone marrow injury induced via oxidative stress in mice by inhalation exposure to formaldehyde. *PLoS One,* **8:** e74974.

Zhu, Y., Jameson, E., Parslow, R.A., Lidbury, I., Fu, T., Dafforn, T.R., Schäfer, H., and Chen, Y. (2014). Identification and characterization of trimethylamine *N*-oxide (TMAO) demethylase and TMAO permease in *Methylocella silvestris* BL2. *Environmental Microbiology*, **16:** 3318-3330.

9. Appendix

Appendix 1. Oligonucleotide primers used in this study

Primer	Sequence (5' → 3')	T_m (°C)	Function
Frm knockout F	CCC CCT ATA GTA TAT TGC ATG CAG ATG ATG AGG TGC GAA GTG TAG GCT GGA GCT GCT TC	79.1	Forward primer for knockout of *frmRAB* operon in *E. coli* MG1655. Region in black has homology to the sequence immediately upstream of the start of the *frmR* protein coding region, red region is homologous to Kn-resistance cassette start (pKD4).
Frm knockout R	TAG GCC GGA TAA GGC GTT CAC GCC GCA TCC GGC AGT CGT GCT ATT ACA TAT GAA TAT CCT CCT TAG	80.6	Reverse primer for knockout of *frmRAB* operon in *E. coli* MG1655. Region in black has homology to the sequence immediately downstream of the end of the *frmB* protein coding region, red region is homologous to Kn-resistance cassette start (pKD4).
Frm knockout seq F	GAT TCC TTC TGC CGC CCG CTA TC	60.6	Forward primer that is homologous to the MG1655 chromosomal region upstream of *frmR*, to check, by sequencing, that the operon had been replaced with a Kn-resistance cassette from pKD4.
Frm knockout seq R	GTT TAT GGG CGA AGT GTA GAC CGG	59.1	Reverse primer that is homologous to the MG1655 chromosomal region downstream of *frmB*, to check, by sequencing, that the operon had been replaced with a Kn-resistance cassette from pKD4.
Frm In-Fusion F	TCG TCT TCA AGA ATT CGT ATT CAA CCG CTG ACT G	66.3	Forward primer for use with the In-Fusion Cloning Kit (Clontech) to clone the *frmRAB* operon with its native promoter into the EcoR1

			restriction site of the vector pBR322. The primers were designed to be homologous to the insert (black) and the overhangs of the cut vector (pBR322).
Frm In-Fusion R	TCA AAC ATG AGA ATT TTA TCA ACG CAT ATT CAG TTA	62.8	Reverse primer for use with the In-Fusion Cloning Kit (Clontech) to clone the *frmRAB* operon with its native promoter into the EcoR1 restriction site of the vector pBR322. The primers were designed to be homologous to the insert (black) and the overhangs of the cut vector (pBR322).
Frm In-Fusion seq F	CTA AGA AAC CAT TAT TAT CAT		Forward sequencing primer, which is homologous to a region of pBR322 upstream of the *frmRAB* insert.
Frm In-Fusion seq R	TGA GCG CAT TGT TAG ATT TC		Reverse sequencing primer, which is homologous to a region of pBR322 downstream of the *frmRAB* insert.
TorC-fusion F	GAT CGC TTC ACT AAA CCA TAA TTC TAC AGG GTT ATT ATG AAG GGC GAG GAG CTG TTC AC	84.4	Forward primer for replacing *torC* gene in MG1655 chromosomal DNA with *gfp* and Kn-resistance cassette from p-DOC-G, using λred recombineering. Region in black has homology to the sequence immediately upstream and including the start of the *torC* protein coding region, red region is homologous to the start of the *gfp* gene on p-DOC-G.
TorC-fusion R	TGC CAG AAA ACG CCG ACG TGA TGC CTG AAA GAG ATC GTT AAA TAT CCT CCT TAG TTC C	87.1	Reverse primer for replacing *torC* gene in MG1655 chromosomal DNA with *gfp* and Kn-resistance cassette from p-DOC-G, using λred recombineering. Region in black

			has homology to the sequence at the end of the *torC* protein coding region, red region is homologous to the P-REV sequence p-DOC-G, after the Kn-resistance cassette.
TorC-fusion seq F	GTA ATC GCG CCT GGG TAA	63.6	Forward sequencing primer, which is homologous to a region of MG1655 chromosome upstream of the *torC* gene.
TorC-fusion seq R	CCG GCA ATC ATT TTC GAC	64.1	Reverse sequencing primer, which is homologous to a region of MG1655 chromosome downstream of the *torC* gene.
Tor biotin R EMSA	BIO-CGG CGT AGC GCG TTC CAG AG	60.0	Biotin labelled (at 5' end) reverse primer for amplifying *tor* promoter region of MG1655 chromosome for EMSAs. The primer is homologous to the region between 7 – 26 bp after the start codon of *torC*.
Tor F 150bp EMSA	CAA TAA CAA TGT GAT GTG GCA TCA GAG G	58.5	Forward primer for amplifying *tor* promoter region of MG1655 chromosome for EMSAs. The primer is homologous to the region 150 bp upstream of the *torC* start codon.
YdhY biotin R EMSA	BIO-CTA ATA GTG GAC GAT CAA CCG GG	57.1	Biotin labelled (at 5' end) reverse primer for amplifying *ydhY* promoter region of MG1655 chromosome for EMSAs. The primer is homologous to the region between 3 - 25 bp after the start codon of *ydhY*.
YdhY F 200bp EMSA	GTT AAG GGC TCA GAA TAA TCA C	51.1	Forward primer for amplifying *ydhY* promoter region of MG1655 chromosome for EMSAs. The primer is homologous to the

			region 200 bp upstream of the *ydhY* start codon.
FrmR biotin R BLItz	Bio-TTT CTT CTC TTC CGG AGT ACT G	58.0	Biotin labelled (at 5' end) reverse primer for amplifying *frmR* promoter region of MG1655 chromosome for BLItz experiments. The primer is homologous to the region between 3 - 24 bp after the start codon of *frmR* gene.
FrmR F BLItz	GGC CTT CCC TGC CGA TTA G	61.0	Forward primer for amplifying *frmR* promoter region of MG1655 chromosome for BLItz experiments. The primer is homologous to the region 169 bp upstream of the *frmR* gene start codon.
pGS2516 P2A SDM F	ATG ATG AGG TGC GAA ATG GCC AGT ACT CCG GAA GAG AAG	73.7	Forward primer used for SDM to create P2A mutant of *frmR* on pGS2516. Sequence of codon change is shown in red.
pGS2516 P2A SDM R	CTT CTC TTC CGG AGT ACT GGC CAT TTC GCA CCT CAT CAT	73.7	Reverse primer used for SDM to create P2A mutant of *frmR* on pGS2516. Sequence of codon change is shown in red.
pGS2516 C35A SDM F	CTG TTG GAG TAT GGC ACG GGC TTC GGC ATC ACC CTC CAG	75.0	Forward primer used for SDM to create C35A mutant of *frmR* on pGS2516. Sequence of codon change is shown in red.
pGS2516 C35A SDM R	CTG GAG GGT GAT GCC GAA GCC CGT GCC ATA CTC CAA CAG	75.0	Forward primer used for SDM to create C35A mutant of *frmR* on pGS2516. Sequence of codon change is shown in red.
pGS2516 SDM seq F	CTA AGA AAC CAT TAT TAT CAT	48.1	Forward primer used to sequence pGS2516 to check mutations were created in the expected codon. This primer sits upstream of *frmR* on the plasmid.

pGS2516 SDM seq R	TGA GCG CAT TGT TAG ATT TC	53.2	Reverse primer used to sequence pGS2516 to check mutations were created in the expected codon. This primer sits downstream of *frmR* on the plasmid.
PfrmR IVT F	GAA ATT CTG ATT CCT TCT GCC GC	60.6	Forward primer for PCR to amplify P_{frmR} region of MG1655. Primer is homologous to region 197 bp upstream of *frmR* start codon.
PfrmR IVT R	CGT TTC CCG GAT ATG GCT TTC AAG	62.7	Reverse primer for PCR to amplify P_{frmR} region of MG1655. Primer is homologous to region 192 bp downstream of *frmR* start codon. Transcription of this PCR product would result in a product of ~200 bp.
Pndh IVT F	ATG CCT GAT GCG CTT CTT ATC A	58.4	Forward primer for PCR to amplify *Pndh* region of MG1655. Primer is homologous to region 300 bp upstream of *ndh* start codon.
Pndh IVT R	CGT AGT CAA CGT GAC CCC C	61.0	Reverse primer for PCR to amplify *Pndh* region of MG1655. Primer is homologous to region 7 bp downstream of *ndh* start codon. Transcription of this PCR product would result in a product of ~100 bp.
dam check F	CAG ATG TCC AGG CCA AAA AC	57.3	Forward primer used in PCR to verify that the wild type dam gene is no longer present in the JRG6842 strain.
dam check R	GCA AGG ATT TCA GCA CCA TT	55.3	Reverse primer used in PCR to verify that the wild type dam gene is no longer present in the JRG6842 strain.

Oligonucleotide primers used in this study for RT-PCR

Primer	Sequence (5' → 3')	Tm (°C)	Function
Frm RT-PCR F	CTG GAA CGG TCG CTG GAG	60.5	RT-PCR; forward primer to analyse the expression of *frm* mRNA in response to the presence of environmental TMAO.
Frm RT-PCR R	ATC AGC CCA TTA GCC GCG	58.2	RT-PCR; reverse primer to analyse the expression of *frm* mRNA in response to the presence of environmental TMAO
gyrA RT-PCR F	ACC TTG CGA GAG AAA TTA CAC	63.1	RT-PCR; forward primer analyse the expression of the housekeeping *gyrA* mRNA in response to the presence of environmental TMAO
gyrA RT-PCR F	AAT GAC CGA CAT CGC ATA ATC	63.5	RT-PCR; reverse primer analyse the expression of the housekeeping *gyrA* mRNA in response to the presence of environmental TMAO

T_m = melting temperature