CARDIFF UNIVERSITY

PRIFYSGOL CAERDYDD

The Response of *Escherichia coli* to the Biocide Polyhexamethylene Biguanide

Submitted for the degree of Doctor of Philosophy at University of Wales by

Michael James Allen

2004

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed.....(candidate)

Date.....12/5/04

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated.

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed.....(candidate)

Date.....12/5/04

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed.....(candidate)

Date.....12/5/04

Abstract

Polyhexamethylene biguanide (PHMB) is a cationic surfactant commonly used as a biocide against a wide range of microorganisms in a variety of medical and industrial contexts. PHMB was bacteriostatic at low concentrations and bactericidal at high concentrations and hitherto its action was thought to be caused by disruption of the outer membrane and inner membrane respectively.

The effects of different concentrations of PHMB on the growth of *E. coli* in batch culture confirmed bactericidal action at high concentrations and bacteriostatic effects at low concentrations. Transcriptional (using microarrays and macroarrays) and protein (using 2-D PAGE) profiles were generated for *E. coli* cells exposed to bacteriostatic concentrations of PHMB. Genes found to be altered in their expression were associated not only with outer membrane, periplasmic space and inner membrane but also with cytoplasmic function.

Gene knock-out strains in which genes whose expression was sensitive to PHMB-exposure had been inactivated, were tested for PHMB sensitivity/resistance. Similarly, the strains bearing over-expression plasmids of the same genes in either the wild-type or the corresponding knock-out background were also assessed for the PHMB sensitivity/resistance phenotype. Collectively, the results showed the induction of members of the DNA and cytoplasmic protein damage responses, suggesting that effects of PHMB were not restricted to the cell envelope but also occurring in the cytoplasm. PHMB was shown to bind to nucleic acids *in vitro* in a highly cooperative manner and cause their precipitation. Hence, it was concluded that the dosage dependent effect observed in PHMB action could be attributed, not only to the disruption of inner membrane, but also to the interaction with cytoplasmic nucleic acids.

In addition, as this work was performed a unique opportunity arose to explore the variation in results obtained from differing statistical analysis systems and between microarray and macroarray experiments.

Acknowledgements

Firstly, I would like to thank my supervisors, Dr. Graham White and Dr. Andrew Morby, for their help, support and friendship. I would also like to acknowledge the many friends in the department, in particular in Lab 315, for making my lab work a thoroughly enjoyable experience, providing useful advice (sometimes) and for never letting me go for breakfast on my own.

I would also like to acknowledge the people at the EXGEN Project in Birmingham University, especially Dr. Jon Hobman, for their help with the microarray work. I would also like to thank Dr. Simon Andrews for allowing me access to the Mori clone collection and Avecia for supplying PHMB and its fluorescent homologue.

Special thanks go to Sara for keeping me sane whilst writing up this thesis. Lastly, I would like to thank my Mum, Dad and Sister to whom this thesis is dedicated. I will be eternally grateful for your encouragement and support.

Abbreviations

2-D PAGE	2-Dimensional polyacrylamide gel electrophoresis
Amp	Ampicillin
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
bp	Base pairs
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
Cb	Carbenicillin
CDS	Coding sequence
CHAPS	3-(3-Cholamidopropyl)diethyl-ammonio-1 propanesulphonate
Chl	Chloramphenicol
CSP	Cold shock protein
D_x	Optical attenuance at wave-length x nm
(k)Da	(kilo) Daltons
(c/g/ss/ds)DNA	(complementary/genomic/single-strand/double-strand) Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide 5'-triphosphate
DMSO	Dimethylsulphoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra-acetic acid
FC	Fold-change
FDR	False discovery rate
FSN	False significant number
g	Gravitational force
GFP	Green fluorescent protein
GPC	Gel permeation chromatography
HEPES	4-(2-Hydroxyethyl)-1-piperazine
HSP	Heat shock protein
IEF	Isoelectric focusing
IM	Inner membrane
IPG	Immobilised pH gradient
IPTG	Isopropyl-D-thiogalactopyranoside
IR	Induction ratio
kb	Kilo base pairs
Km	Kanamycin
KO	Knock-out
LB	Luria-Bertani
LPS	Lipopolysaccharide
MALDI-TOF	Matrix assisted laser desorption ionisation, time-of-flight
MS	Mass spectrometry
MIAME	Minimum Information About Microarray Experiments

MIC	Minimal inhibitory concentration
Mw	Molecular weight
MOPS	3-(N-morpholino) propanesulphonic acid
n/a	Not available
NADP(H)	Nicotinamide adenine dinucleotide phosphate (reduced form)
nt	Nucleotide
OASL	O-acetyl serine (thiol) lyase
OM	Outer membrane
(ds)ORF	(down-stream) Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PHMB	Polyhexamethylene biguanide
(m/r/t)RNA	(messenger/ribosomal/transfer) Ribonucleic acid
RNase	Ribonuclease
RNA-P	RNA polymerase
SAM	Significance analysis of microarrays
SD	Standard deviation
SDS	Sodium dodecylsulphate
SSC	Saline sodium citrate
SSPE	Saline sodium phosphate EDTA
TBE	Tris Borate EDTA
TE	Tris EDTA
TSS	Transfer and storage solution
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultra-violet
UFA	Unsaturated fatty acid
v/v	Volume / volume
w/v	Weight / volume

Guide to Figures Page

Guide to Tables Page

Contents

Chapter 1: Introduction

1.1 Antibacterial agents

Microorganisms have evolved to prosper in a range of environmental conditions, having been isolated from habitats as diverse as the Antarctic and hot-vents (at least 113°C), at pH values approaching zero and at salinities found in saturated brines (Storz and Hengge-Aronis, 2000). It is therefore unsurprising that microorganisms can thrive in biologically permissive man-made environments such as breweries, swimming pools, hospitals and kitchens. The use of antimicrobial and antibiotic compounds is steadily increasing in response to both the real and the perceived threat of microbial contamination and infection. Antibacterial agents can be found in a wide variety of places ranging from surface cleaners, cosmetics, toothpaste, to socks and chopping boards and are used to inhibit bacterial growth (bacteriostatic activity) or destroy bacteria completely (bactericidal activity). There are two main classes of antibacterial agent: antibiotics and biocides.

1.1.1 Antibiotics

Antibiotics are naturally occurring substances, usually secondary metabolites, (or semi-synthetic derivatives) that display selective toxicity for certain microorganisms (Mann and Crabbe, 1998). The bases of selective toxicity are the biochemical and physiological differences that exist between organisms. A good example is the cell wall of bacteria. A peptidoglycan layer accounts for approximately 5% and 40% of the cell wall of Gram-negative and Gram-positive bacteria, respectively. No equivalent structure is present in eukaryotic cells. Thus, it is a target for the major ß-lactam series of antibiotics that inhibit peptidoglycan biosynthesis by irreversibly binding to the active site of transpeptidase, the enzyme responsible for completing the cross linking of glycine to D-alanine. Although antibiotics have been produced naturally by some organisms for millions of years, it is only in the last 60 years or so that they have been exploited by humans. However, due to heavy overuse bacteria are fast becoming antibiotic-resistant.

1.1.2 Biocides

Biocides are chemicals that display a broad spectrum of antimicrobial activity and are commonly used as disinfectants, antiseptics and preservatives (Yeates, 2002). Because they act at multiple sites, biocides often lead to a general loss of cellular function. A simple alteration in one cellular component is unlikely to give rise to biocide resistance given the multiple mechanism of toxicity.

The use of biocides for food preservation is well known and has been practised for thousands of years. Salt, sugar, smoke, ethanol and vinegar have all been used extensively to prolong the lifespan of perishable goods. Ancient Romans were even known to keep beverages (e.g. drinking water and milk) in silver vessels in order to keep them fresh (Knight and Cook, 2002). Today, the biocide market is extremely profitable (Hauthal, 1992). According to the Global Biocides Report, carried out by the Biocide Information Services (BIS), an estimated $3,380 million was spent worldwide on biocides in 2001 (BIS Global Biocides Report, 2002). The growing use of biocides as an aid to improved hygiene has led to a 12% growth in this market during the last 3 years (Knight and Cook, 2002).

Biocides range from the structurally simple (e.g. hydrogen peroxide, metals and salt) to more complex organics such as the biguanides and diamidines. Some of the various compounds used as biocides and some of their applications are listed in Table 1.1. Biocides of different types are commonly used in combinations to increase activity in a synergistic way. For example, biocides containing simple inorganic chemicals such as sodium hypochlorite or hydrochloric acid have their biocidal activity increased by the addition of membrane-disrupting surfactants (**surface active agents**).

Table 1.1 Examples and uses of biocides. (Table adapted from O'Malley, 2000)

Class	Example	Applications
Alcohols	Ethanol	Preservatives, hard surface disinfection
Aldehydes	Glutaraldehyde	Disinfection of surgical equipment
Anilides	Trichlorocarbanilide 'Triclocarban'	Antisepsis
Biguanides	Polyhexamethylene biguanide	Disinfection of swimming pools
Bis-phenols	2,4,4'-Trichloro-2 hydroxydiphenol 'Triclosan'	Surgical scrubs
Diamidines	Dibromopropamidine	Topical wound treatment
Halogen releasing	Sodium hypochlorite	Hard surface disinfection
Halophenols	Chloroxylenol 'Dettol'	Antiseptics
Quaternary ammonium compounds	Benzalkonium chloride	Antiseptics, preservatives
Peroxygens	Hydrogen peroxide	Disinfection, sterilisation
Phenolics	Cresol	Antiseptics
Silver compounds	Silver nitrate	Wound treatment

1.2 Surfactants as Biocides

Cationic, anionic and nonionic surfactants are all important industrial compounds. Anionics have the biggest share due to their importance in cleaning and hygiene products, followed by the nonionics that are used mainly as laundry detergents and agricultural wetting agents (Hauthal, 1992). Cationics and amphoterics make up the remainder. Anionics and nonionics are commonly used in disinfection products but it is the cationics that are particularly useful for their biocidal properties.

1.2.1 Cationic Biocides

Cationic biocides are used as antiseptics, disinfectants and preservatives in a range of preparations including cosmetics, paints and pharmaceuticals. It has already been noted that quaternary ammonium compounds and biguanides possess both surfactant and biocidal properties (Table 1.1). The biguanide based surfactant polyhexamethylene biguanide is the focus of this study. Biguanides have been used as biocides for over 50 years (Curd and Rose, 1946). Initially molecules containing a single biguanide moiety were used as anti-malarial drugs. The molecule with the greatest effect was an anti-malarial known as proguanil (Figure 1.1). Biguanide derivatives, such as phenoform, buformin and metformin, have also been used to prevent and treat insulin resistance in people with Type II diabetes (Mehnert, 2001). Metformin became the most frequently prescribed oral anti-diabetic drug in the USA within 12 weeks of receiving US marketing approval (Bloomgarden, 1996).

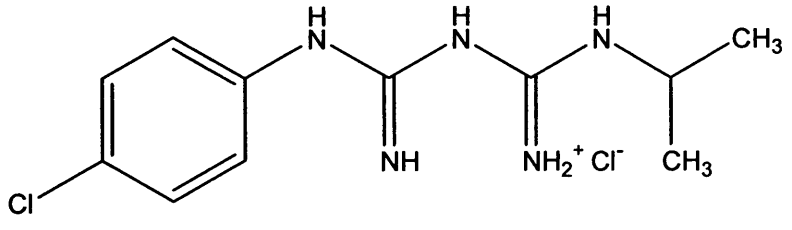


Figure 1.1 Structure of Proguanil

It was a decade after their discovery as a treatment for malaria that the antibacterial properties of biguanides were first exploited (Rose and Swain, 1956). Rose and Swain synthesised 1,6-di(N^5-*p*-chlorophenol-N^1-biguanido)hexane which

was a potent antibacterial agent and contained not one, but two biguanide groups. The molecule, commonly known under the brand name chlorhexidine (Figure 1.2), is now a very widely used antibacterial agent found in mouthwashes, toothpastes and antiseptics.

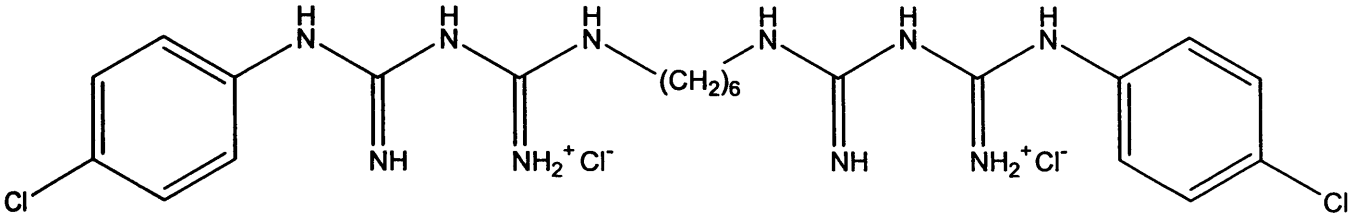


Figure 1.2 Structure of Chlorhexidine

1.2.2 Polyhexamethylene Biguanide

The increased antibacterial activity of the bisbiguanide chlorhexidine over monomeric biguanides stimulated the development of polymeric biguanides containing repeating biguanide groups linked by hexamethylene chains known as **poly**hexa**methylene** **biguanide** (PHMB) (Davies *et al.*, 1968). PHMB is a mixture of polymers having the general formula shown in Figure 1.3 with n varying from 2 to 35, with a mean of 5.5 (Gilbert *et al.*, 1990a; Gilbert *et al.*, 1990b). Because of the method of synthesis, the end groups in any given molecule can be a combination of an amine, cyanoguanide or guanide (Figure 1.4).

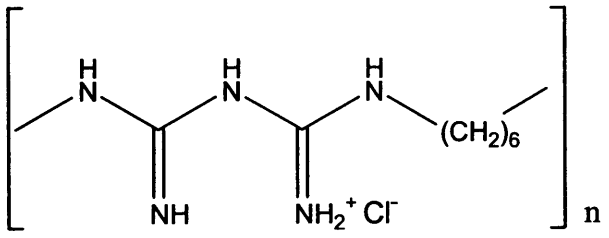


Figure 1.3 Structure of polyhexamethylene biguanide (PHMB)

PHMB is marketed under the trade name Vantocil, has a broad spectrum of activity against both Gram-positive and Gram-negative bacteria and very importantly has very low mammalian toxicity (Broxton *et al.*, 1983; Jones, 1995). The concentrations of Vantocil PHMB needed to prevent the growth of a range of problem microorganisms (i.e. the minimum inhibitory concentration) are shown in Table 1.2. PHMB is very versatile and is used in a diverse range of end-use products including liquid disinfectants, soaps, tabletised products and impregnated on surfaces (Yeates, 2002). PHMB has recently been recognised as an approved biocide under the latest European Biocidal Products Directive (EBPD) and therefore meets the requirements of human safety and technical performance (European Biocidal Products Directive, May 1998, 98/8/EC; Rasmussen and Kappes, 2000).

A wide variety of specialised PHMB usage has arisen including contact lens solutions to combat *Acanthamoeba castellanii* (a free living protozoan which can cause severe eye infections), in cooling water systems to combat *Legionella* (the cause of Legionnaires' disease), against *Canidida albicans* (the cause of superficial and systemic candidosis) and even as an egg shell sanitising agent to combat *Salmonella typhimurium* and *Salmonella heidelberg* infection in chickens (Khunkitti *et al.*, 1998a; Kusnetsov *et al.*, 1997; Jones *et al.*,1995; Cox *et al.*, 1999).

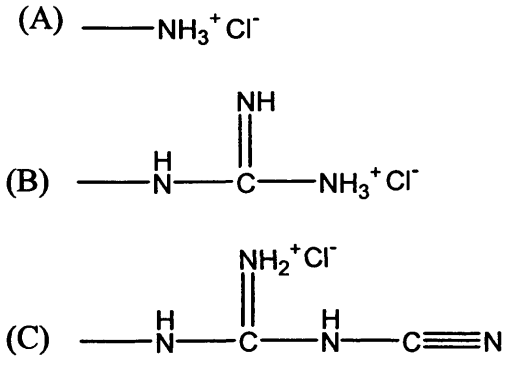


Figure 1.4 PHMB end groupings: A; amine, B; guanide C; cyanoguanide

Table 1.2 Minimum inhibitory concentrations (MIC) of Vantocil against various microorganisms. Vantocil contains 20% w/v of PHMB. Data taken from Avecia web page on 2/06/03 (www.avecia.com/biocides/applications/disinfection/efficacy.htm).

Microorganism	MIC of Vantocil (ppm, v/v)
Bacteria	
Bacillus subtilis	5
Bacillus cereus	20
Clostridium difficile	0.5
Enterobacter cloacae	20
Multiple Antibiotic Resistant *Enterococcus faecium*	30
Escherichia coli 0157:H7	5
Legionella pneumophila	200
Listeria monocytogenes	45
Mycobacterium tuberculosis	25
Proteus vulgaris	200
Pseudomonas aeruginosa	100
Pseudomonas putida	25
Salmonella choleraesius	55
Salmonella typhimurium	8
Staphylococcus aureus	1
Epidemic Multiple Antibiotic Resistant *S. aureus* (EMRSA)	30
Streptococcus faecalis	25
Streptococcus lactis	25
Vibrio cholerae Non 0:1	5
Yersinia enterocolitica	300
Fungi	
Aspergillus niger	750
Trycophyton mentagrophytes	25
Yeasts	
Endomycopsis albicans	300
Saccharomyces cerevisiae	100
Rhodotorula rubra	25
Amoeba	
Acanthamoeba polyphaga	5
Viruses	
Vaccinia	300
Herpes simplex type 1	50
Rotavirus	200

1.3 Mechanism of Action of PHMB

PHMB causes bacteriostatic or bactericidal effects in *E. coli* (and other bacteria) depending upon the dosage level (Davies *et al.*, 1968; Broxton *et al.*, 1983). It has been shown to increase the permeability of the outer membrane of *E. coli* at low concentrations and to increase the permeability of both the outer and cytoplasmic membrane at higher concentrations (Yasuda *et al.*, 2003). The disruption of the cytoplasmic membrane is thought to be the crucial step in the bactericidal action of PHMB, leading to the diffusion and irreversible loss of essential cellular components (Broxton *et al.*, 1984c).

1.3.1 Interaction of PHMB with the Outer Membrane

The outer membrane of *E. coli* is a lipid bilayer consisting of lipopolysaccharides, phospholipids and proteins. The number of hydrocarbon chains in phospholipids is approximately equal to that of those contained in lipopolysaccharides (LPS) (Nikaido and Vaara, 1987). The distribution of phospholipids and LPS in the bilayer is asymmetric, with most (if not all) LPS being found in the outer leaflet and most of the phospholipid being located in the inner leaflet. Lateral interactions between the negatively charged proximal heads of LPS molecules and divalent cations (such as Mg^{2+} and Ca^{2+}) are very strong and create a rigid and impermeable LPS monolayer. It is the outer membrane that forms the primary defence barrier against PHMB. Up to 80% of PHMB may be excluded or bound by the outer membrane of *E. coli* at the minimum inhibitory concentration (Gilbert *et al.*, 1990a). The contribution of outer membrane proteins to this exclusion is considered negligible, with the majority of the interaction occurring between PHMB and lipopolysaccharide.

The binding of PHMB to high affinity binding sites is thought to cause the displacement of metal cations and thus the destabilisation and disruption of the outer membrane. The release and dissociation of Ca^{2+} from the membrane and the subsequent release of LPS from the outer membrane in *E. coli* exposed to bacteriostatic levels of PHMB provides evidence for this model (Yasuda *et al.*, 2003). This facilitates the entry of PHMB into the periplasmic space and allows its subsequent interaction with the negatively charged cytoplasmic membrane. In this

manner, PHMB itself is promoting its own uptake into the periplasm. Bactericidal activity increases with increasing polymerisation of PHMB. However, this effect is greater on spheroplasts (cells without an outer membrane) than on whole cells (providing further evidence that the outer membrane does act as a very effective exclusion barrier) (Hugo and Longworth, 1964; Davies *et al.*, 1968). However, the situation is complicated by the profound synergy that is found between high and low molecular weight PHMB (Gilbert *et al.*, 1990b). This has led to the hypothesis that the lower molecular weight PHMB binds to lipopolysaccharide causing the disruption of the outer membrane, allowing heavier molecular weight PHMB to gain access to the cytoplasmic membrane to cause greater disruption. The reduced capacity of PHMB to deal with increasing inoculum levels (i.e. a lower dose per cell) is likely to be caused by the increased adsorption of PHMB by the increased amount of lipopolysaccharide present on cell surfaces.

1.3.2 Interaction of PHMB with the Cytoplasmic Membrane

The cytoplasmic membrane of *E. coli* exhibits a greater protein content than the outer membrane, but contains no LPS (Cronan *et al.*, 1987). Damage to the outer membrane appears to be targeted to lipopolysaccharide and is reversible (Gilbert *et al.*, 1990a; Yasuda *et al.*, 2003). Damage to the cytoplasmic membrane is immediate and completed shortly after contact with PHMB (Broxton *et al.*, 1983). However, a rapid recovery (< 1 min) is possible upon removal of PHMB (at bacteriostatic concentrations) indicating that PHMB is likely to be causing a physical disorganisation of the cytoplasmic membrane (Broxton *et al.*, 1984c). Higher concentrations are likely to cause a physical reorganisation leading to total membrane disruption.

It is thought that PHMB interacts with acidic phospholipids in the membrane. PHMB has been shown to bind preferentially with acidic phospholipid head groups (Ikeda *et al.*, 1984), which could lead to a localised phase separation between the two phospholipids types (neutral and acidic) (Ikeda *et* al., 1985; Cronan and Rock, 1987). The effects of this localised phase separation are further enhanced because PHMB is polymerised. The localised phase separation could lead to the phospholipids abandoning a bilayer phase structure and assuming the energetically favourable

hexagonal phase, which would lead to the loss of the permeability barrier (Anonymous, Avecia website). The hexagonal phase tends to be concentrated around points of maximum charge density, with integral membrane proteins providing appropriate sites. The initial interaction at sub-lethal levels may be localised around such protein sites causing them to lose function due to the change in their boundary phospholipid. This could explain the loss of function of enzymes such as ATPase at sub-lethal concentrations of PHMB (Broxton *et al.*, 1983).

Moreover, in addition to affecting cell permeability, PHMB also causes the precipitation of the cellular cytoplasmic contents such as sugars and proteins when present at high enough concentrations (Khunkitti *et al*, 1998a). The rate of membrane disruption increases with increasing PHMB concentrations up to a point until the rate of leakage of cytoplasmic contents then falls back to pre-treatment levels. Cytoplasmic components of these cells are found to be precipitated (Davies *et al.*, 1968).

1.4 Bacterial Stress Responses

Given the high toxicity of PHMB, it is a reasonable hypothesis that bacterial cells experience stress when exposed to PHMB, and respond to that stress. Thus, as will emerge shortly (Section 1.11), the aim of this thesis is to study the response of *E. coli* to the presence of PHMB. In order to do this, it is necessary to understand the mechanisms that *E. coli* uses to respond to stress. Since it is clear that PHMB acts primarily at the outer membrane and cytoplasmic membrane (Section 1.3), the following sections describe the signalling pathways that *E. coli* uses to detect and respond to damage in the outer membrane, the periplasmic space and the cytoplasmic membrane. The general stress response is also discussed, since it can be triggered by a wide range of stimuli. The response of *E. coli* to stress caused by the presence of acid, alkali and other reactive species are discussed because PHMB may cause damage similar to these agents. The heat- and cold-shock response are also discussed because these stresses can induce changes in the composition of the inner and outer membrane (apparently the main target site for PHMB action).

Essentially all microorganisms other than intracellular parasites and symbionts face ever-changing environmental conditions (Bremer and Kramer, 2000). Environmental stress can be caused by many factors including nutrient limitation (leading to stationary phase), high or low temperatures, chemical stress and physical stress. The ability to sense and respond to potentially lethal changes in the environment is a trait crucial to the survival of any organism (Foster, 2000).

In the majority of cases, the bacterial response leads to the transcriptional activation of genes whose products cope with the environmental insult (Ramos *et al.*, 2001). Gene regulators respond to specific signals (environmental or cellular) by stimulating or inhibiting gene expression (at the transcriptional or translational level) or by modifying their protein products. There are two types of bacterial stress responses: general stress responses and specific stress responses.

General stress responses, usually controlled through a single or a few master regulators, serve to prevent damage rather than repairing it (Hengge-Aronis, 2002a). They typically provide protection from a wide variety of environmental stresses

('cross-protection') and ensure the cell's short-term survival. However, a specific stress response allows growth (or at least survival) when a cell is confronted with a specific environmental insult by dealing directly with the insult. Specific proteins are induced that act to eliminate the stress agent or repair cellular damage produced by the stress. Although cells can persist or continue to grow under the stressful conditions, they are not generally more resistant to other stresses. Specific stress responses typically involve highly integrated networks of physiological and genetic adaptation mechanisms.

1.5 The General Stress Response in *E. coli*

The general stress response of *E. coli* is controlled by the master regulator σ^S and can be triggered by a range of environmental stresses including non-optimal temperatures, starvation, high osmolarity and acidic pH (Storz and Hengge-Aronis, 2000; Hengge-Aronis, 2002b). σ^S, encoded by *rpoS*, is a sigma subunit of RNA polymerase and is a close relative of the housekeeping, *rpoD* encoded, σ^{70}. Rapidly growing *E. coli* cells contain very little σ^S. However, in response to acute stresses, σ^S levels increase dramatically. So far, around seventy genes have been found to be under the control of σ^S (Hengge-Aronis, 2000a; Hengge-Aronis, 2002b). *E. coli rpoS* mutants lack a typical general stress response and are rapidly killed by oxidative stress, exposure to high temperatures (over 50°C), high osmolarity and low pH (Hengge-Aronis, 2002a).

All known acute stresses that affect σ^S levels act at the post transcriptional level of σ^S control (Hengge-Aronis, 2000b). Stresses such as low temperature stimulate σ^S synthesis, stresses such as high temperature interfere with σ^S degradation and some stresses, such as high osmolarity, can affect both processes. Figure 1.5 shows how various stresses affect σ^S levels. It is thought that under non-stress conditions translation of *rpoS* mRNA is inhibited by the formation of extensive secondary structures. A structural rearrangement induced by certain stresses may lead to translational activation, although the mechanism for this translational control is not yet clear.

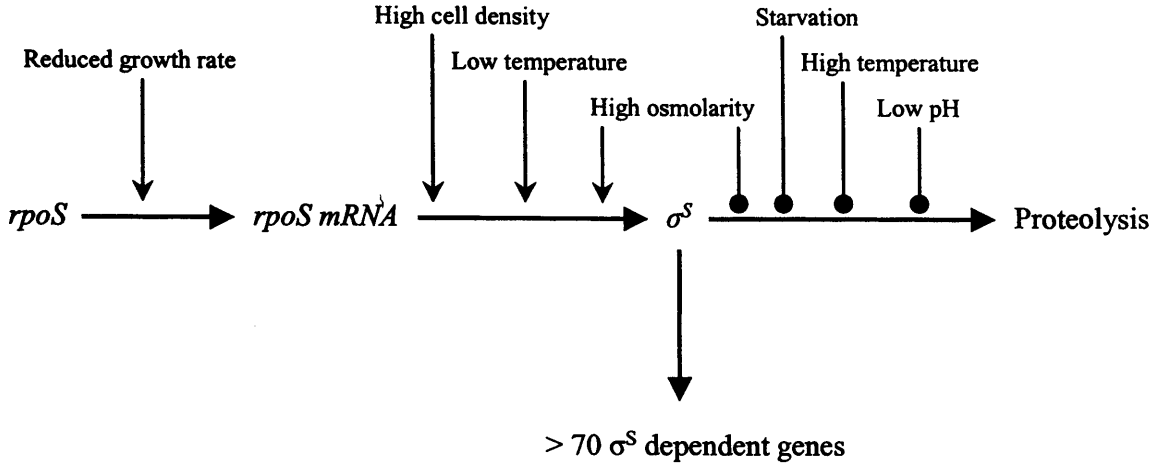


Figure 1.5 Control of σ^S levels. Arrow heads indicate stimulation, rounded heads indicate inhibition. Diagram adapted from Hengge-Aronis, 2000b.

The protease responsible for the degradation of σ^S is the multi-subunit ATP dependent ClpXP (Schweder *et al.*, 1996). ClpXP degrades σ^S completely and no stable degradation products have been observed (Zhou *et al.*, 2001). The response regulator RssB is also essential for σ^S degradation. A phosphorylated RssB binds to the *cis*-acting region of σ^S and delivers it to the ClpXP protease where σ^S becomes unfolded and degraded, RssB is released from the complex and is subsequently recycled. It is not know whether RssB is released from this complex in a phosphorylated state. Stressful conditions can lead to the de-phosphorylation of RssB, resulting in a reduced affinity for σ^S and its subsequent stabilisation. Whether the stress signal works by active de-phosphorylation or inhibition of re-phosphorylation (or even by the down regulation of cellular RssB levels) is yet to be determined since no sensor kinase or phosphatase for RssB have been identified. A diagrammatic view of this pathway is shown in Figure 1.6.

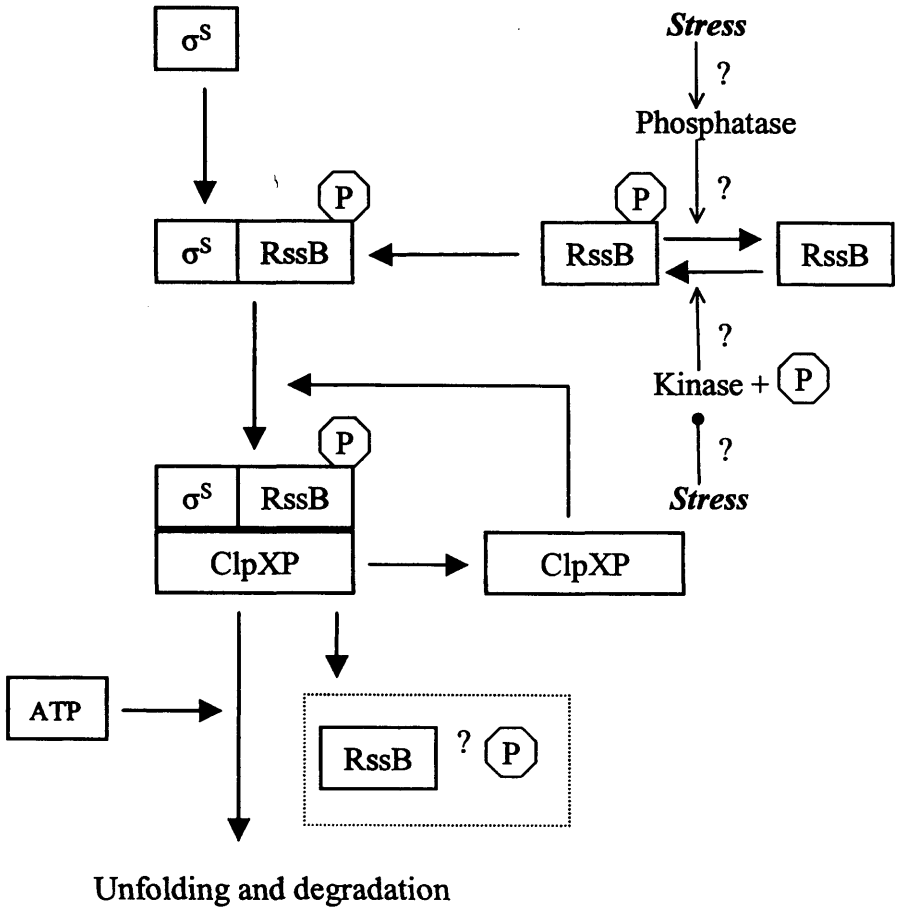


Figure 1.6 Pathway of σ^S recognition and degradation by the response regulator RssB and ClpXP protease. Adapted from Hengge-Aronis, 2002b.

The genes found to be under the control of σ^S confer stress resistance, change cellular morphology, redirect metabolism and affect virulence. Although σ^S concentration is elevated in response to many types of environmental insult, the biochemical and physiological consequences of its induction are not always the same since most σ^S-dependent genes are also affected by other co-regulators.

σ^S-dependent genes involved in oxidative stress include *xthA* (an exonuclease III involved in DNA repair), *dps* (DNA protection), *katG* and *katE* (catalases), *gar* (glutathione reductase) and *sodC* (a periplasmic superoxide dismutase) (Hengge-Aronis, 2000b). Many genes involved with acid resistance (see Section 1.9.1) are also found to be σ^S controlled such as *hdeAB* and the *gad* genes. However, the regulation of these genes is not fully understood because their

expression is tightly linked with non-σ^S-dependent mechanisms. Other genes of interest that are σ^S-dependent include the *osm* genes. While the function of *osmC* and *osmE* remain unknown, *osmB* encodes an outer membrane lipoprotein that has been implicated in cell aggregation (Jung *et al.*, 1989).

The *otsBS* operon is also under the control of σ^S and encodes the enzymes involved in the production of trehalose (Hengge-Aronis *et al.*, 1991). Trehalose is a non reducing disaccharide in which two glucose molecules are linked together in an α,α-1,1-glycosidic linkage (Elbein *et al.*, 2003) and has been shown to accumulate in response to heat shock, cold shock, oxidative stress, osmotic stress and during starvation (Kandror *et al.*, 2002). Trehalose biosynthesis is catalysed by the two enzymes trehalose-6-phosphate synthase and trehalose-6-phosphatase (encoded by *otsA* and *otsB* respectively) (Kaasen *et al.*, 1994). Both proteins have been shown to be both cold shock and heat shock proteins and are crucial for survival at both high and low temperatures (Hengge-Aronis, 1991; Kandror *et al.*, 2002). This appears to be unusual for heat-shock and cold-shock proteins. Normally heat-shock proteins are specifically repressed under cold conditions and cold-shock proteins are hardly detectable at 37°C and completely absent at higher temperatures (Yura *et al.*, 2000; Phadtare *et al.*, 2000). The versatility of trehalose function raises the question of its mechanism of action. Trehalose has been shown to function as a free radical scavenger (preventing oxidative damage), as a chemical chaperone (reducing denaturation and aggregation of proteins) and may also play a role in stabilising cell membranes (Welch and Brown, 1996).

1.6 Envelope Stress

The envelope of Gram-negative bacteria consists of two lipid bilayers (the inner and outer membranes) and a periplasmic space. The periplasmic space is much more viscous and densely laden with proteins than the cytoplasm and has a diffusion coefficient close to a 100-fold lower (Hengge-Aronis, 2000b). The envelope has many functions including solute transport, protein translocation, lipid biosynthesis and oxidative phosphorylation. However, since the envelope is in direct contact with the immediate environment it also serves as the primary protective barrier. Whereas conditions in the cytoplasm are predominantly reducing, the periplasm is predominantly oxidising. The high diversity in envelope associated functions leads to the presence of distinct sets of proteins found in the inner membrane, periplasmic space and the outer membrane. Envelope proteins have highly varied structures and can be either soluble, peripheral or integral membrane proteins or covalently modified lipoproteins.

The envelope is continually exposed to an ever-changing environment (e.g. changes in temperature, pH, osmolarity, chemicals) and as a result, the associated proteins can become damaged (Raivio and Silhavy, 2001). Damaged and misfolded proteins in the envelope will not function correctly and may place considerable stress upon the bacterial cell. In order to ensure that the resident envelope proteins are all correctly folded, bacteria enlist the help of "foldases", proteases and chaperones. Many types of "foldase" have been identified and have been shown to be essential for the correct folding of envelope proteins e.g. disulphide bond oxidoreductases catalyse the formation and isomerisation of disulphide bonds, while peptidyl-prolyl isomerases catalyse the *cis-trans* isomerisation around X-Pro peptide bonds (Raivio and Silhavy, 2000). Chaperones such as LolA and p20 ensure that the correct lipoproteins are incorporated into the outer membrane. Proteases, such as DegP (a periplasmic serine endoprotease), Tsp (a periplasmic protease) OmpT and OmpP (serine endoproteases), are used to destroy abnormally folded envelope proteins that cannot be refolded

Three envelope-specific stress responses have been identified in *E. coli* (Raivio and Silhavy, 1999). One controlled by σ^E and the other two by a two component regulatory pair; these are now briefly described in turn.

1.6.1 The Cpx Regulon

The Cpx envelope response is activated by a variety of envelope perturbations including alterations in pH, membrane composition and the presence of misfolded proteins (Raffa and Raivio, 2002). It is also closely linked to the biogenesis of P pili (Hung *et al.*, 2001). P pili are important virulence factors that allow bacteria to bind specifically to target cells (Bullitt and Makowski, 1995). The assembly of pili is complex and occurs post-secretionally i.e. after the subunits have been translocated across the cytoplasmic membrane (Hultgren *et al.*, 1991). The Cpx regulon is controlled by CpxA, CpxR and CpxP. CpxA is membrane bound and functions as an autokinase, a CpxR kinase and as a phosphatase of phosphorylated CpxR (Raivio and Silhavy, 1997). CpxR is a transcription factor which has been proposed to regulate around 14 genes to date (Raivio and Silhavy, 2001). CpxP is a small periplasmic inhibitor of CpxA. CpxP is thought to inhibit CpxA by interacting with its periplasmic sensing domain (Raivio *et al.*, 2000). The presence of misfolded proteins is thought to sequester CpxP, leading to the autophosphorylation of CpxA and the subsequent phosphorylation of CpxR (Ravio and Silhavy, 2001). Activated CpxR regulates *cpxP*, *cpxRA*, *degP* (protease), *dsbA*, *ppiA* and *ppiD* (pilus), *yihE*, *motABcheAW* (flagellar), *tsr* (chemoreceptor) and *rpoH* (σ^H).

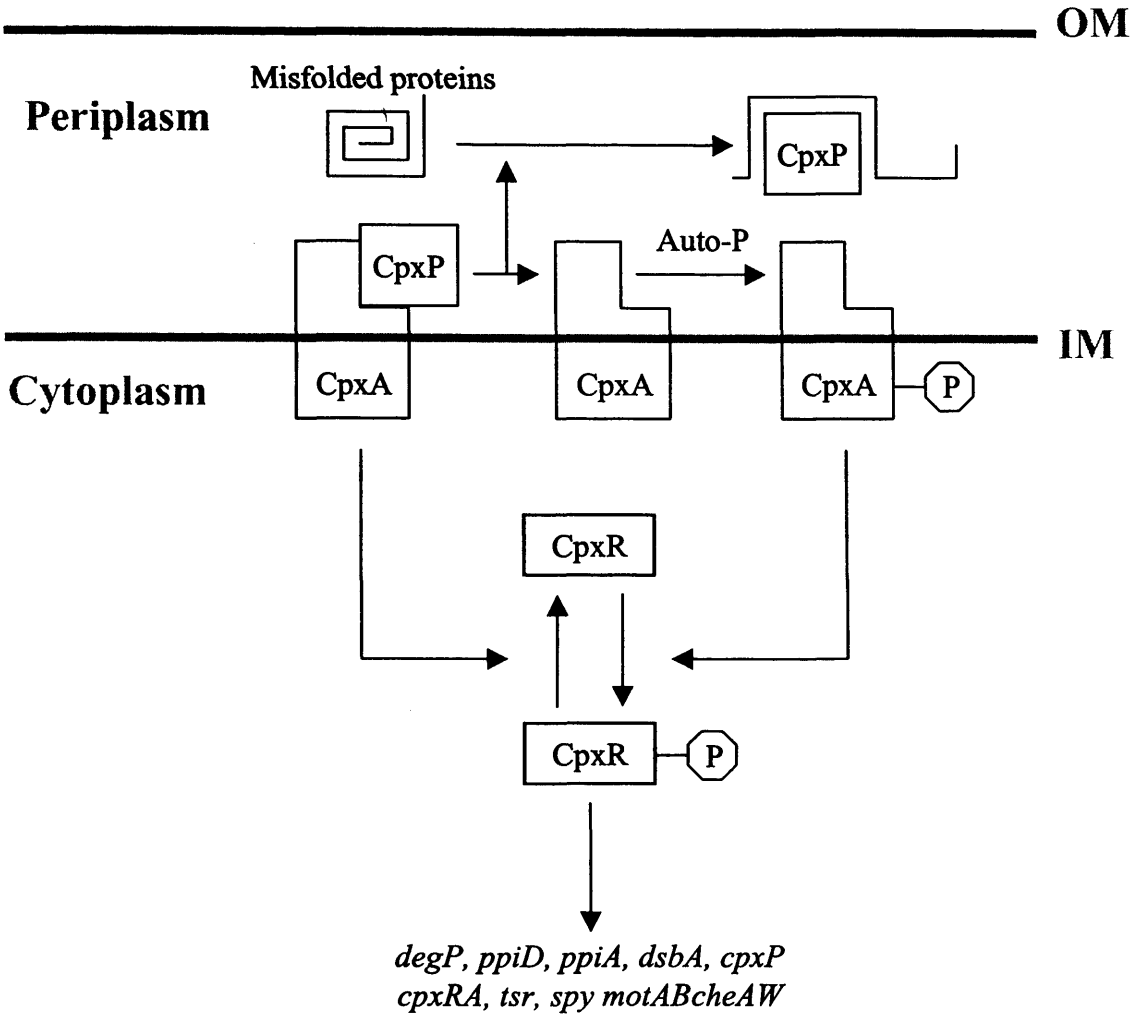


Figure 1.7 Current model of the Cpx envelope stress response. Auto-P denotes autophosphorylation. Diagram adapted from Raivio and Silhavy, 2001.

1.6.2 The Bae Response

Recently a third envelope stress signal transduction pathway has been identified (Raffa and Raivio, 2002). The BaeSR two component system mediates adaptation to stress and appears to overlap with the Cpx response (they both induce *spy* expression). Spy is a small, basic periplasmic-protein of unknown function. Spy has been found to be undetectable (immunologically) in intact cells, but is produced abundantly in spheroblasts (Hagenmaier *et al.*, 1997). However, it appears that the Bae pathway

does not respond to the same type of signal as the Cpx response and appears to recognise a distinct type of misfolded envelope protein. Only one gene (*spy*) has been identified, so far, as being Bae regulated but it is likely that others exist since *baeR* mutants are more sensitive than *spy* mutants to envelope perturbations (Raffa and Raivio, 2002).

1.6.3 σ^E Regulation

A heat stress response linked to the bacterial envelope is controlled by the *rpoE* encoded σ^E. The σ^E pathway can be activated by stresses such as heat and ethanol, which lead to the denaturation of proteins in the periplasm and outer membrane of the bacterial envelope (Rouviere *et al.*, 1995). The σ^E pathway does not respond to misfolded proteins in the cytoplasm and appears to be specifically activated by misfolded outer membrane and periplasmic proteins (Raivio and Silhavy, 2001).

The activity of σ^E is controlled by RseA and RseB (regulator of sigma **E**) (Yura and Nakahigashi, 1999). The precise nature of the σ^E activation signal is yet to be determined but it is thought that σ^E is inactivated by sequestration by the cytoplasmic N terminal domain of the transmembrane anti-sigma factor RseA (Collinet *et al.*, 2000). Misfolded proteins accumulated in the periplasm lead to degradation of RseA by DegS allowing σ^E to interact with RNA polymerase and to activate transcription from σ^E-dependent promoters. The periplasmic RseB interacts with the C terminal of RseA and modulates the affinity of RseA for σ^E.

There are at least 43 members of the σ^E regulon including *fkpA* (encoding a peptidyl prolyl *cis/trans* isomerase) and *degP* (encoding a periplasmic protease which is also a member of the Cpx regulon), which are activated to assist in re-folding or degradation of misfolded envelope proteins (Dartigalongue *et al.*, 2001). Recently a further 15 genes were identified as being potential members of the σ^E regulon (Rezuchova *et al.*, 2003). The induction of *rpoE* and the *rseABC* operon provides a mechanism of autoregulation that allows the pathway to be switched off rapidly upon alleviation of the envelope stress. The transcription of *rpoH* is also under the control of σ^E and provides an important link to the cytoplasmic heat stress response.

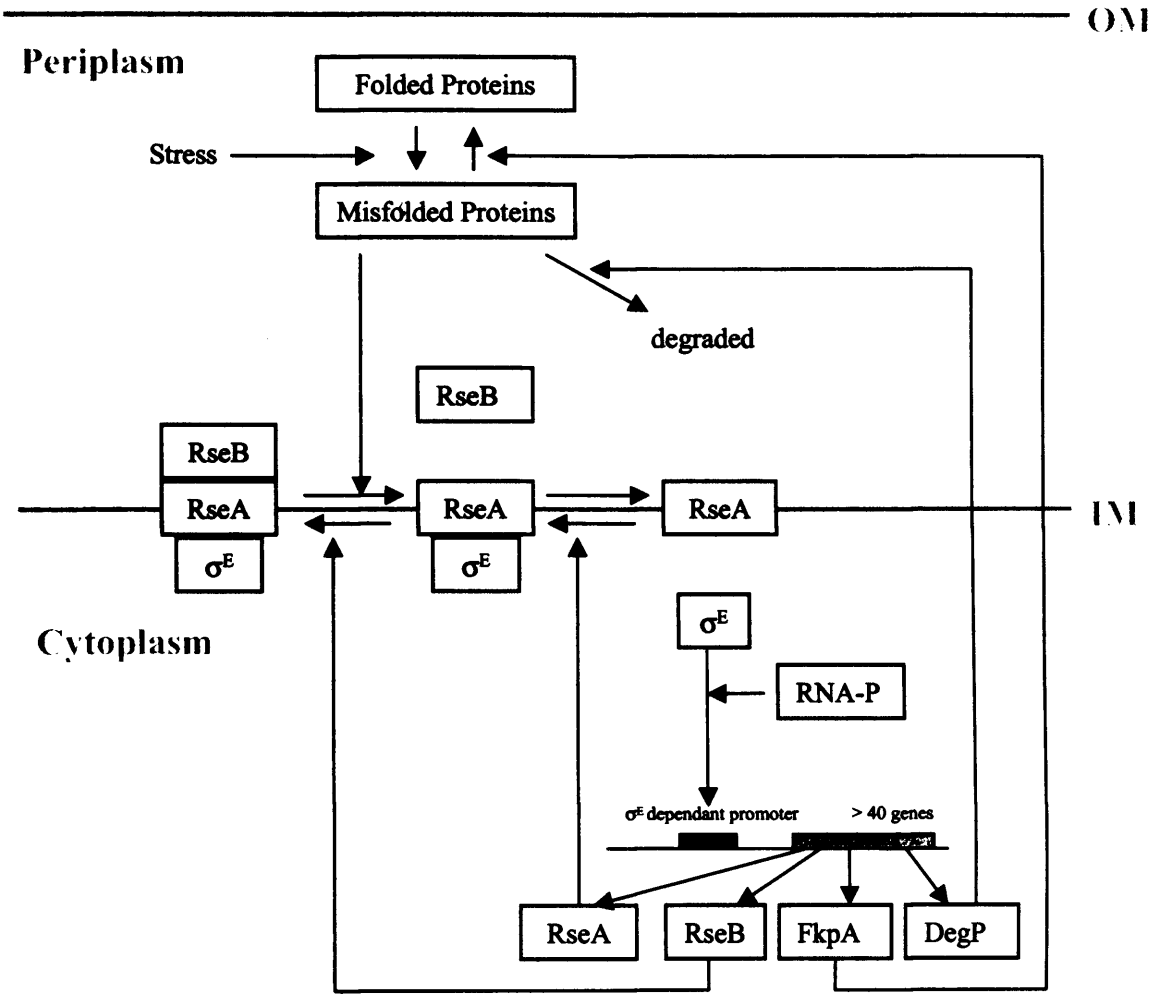


Figure 1.8 Model for the regulation of σ^E activity by RseA and RseB. Diagram adapted from Collinet *et al.*, 2000.

1.7 The Heat Shock Response

The so-called heat shock regulon is composed mainly of chaperones and proteases that coordinate the repair or degradation of misfolded proteins (Yura *et al.*, 1993). The heat shock response is governed by two master regulators: σ^{32} (encoded by *rpoH*) responds to damaged cytoplasmic proteins and σ^{E} (encoded by *rpoE*) responds to damaged envelope proteins (Yura and Nakahigashi, 1999). Due to the common sensing mechanism (the presence of damaged proteins), several forms of stress have been shown to induce various heat shock proteins (HSPs) including pH, osmolarity, UV and the presence of toxic substances (such as ethanol, heavy metals and antibiotics), oxidative stress and starvation (Neidhardt and Savageau, 1987). Indeed, most HSPs are synthesised under non-stress conditions at reduced rates and play fundamental roles in normal cell physiology in addition to their activity under stress conditions (Yura *et al.*, 1993).

1.7.1 σ^{32} Regulation

Upon temperature upshift, the cellular level of σ^{32} increases rapidly due to increased synthesis and stabilisation (Yura *et al.*, 2000). Under normal conditions σ^{32} is unstable (half-life of less than 1min) and is rapidly degraded by FtsH protease (Bertani *et al.*, 2001). A relatively modest heat shock (e.g. 30°C to 42°C) activates *rpoH* mRNA translation and stabilises σ^{32} whereas a more severe heat shock (e.g. > 50°C) can also activate *rpoH* transcription. The *rpoH* mRNA secondary structure plays a crucial role in the regulation of RpoH translation by acting as a thermosensor by blocking translation initiation (Morita *et al.*, 1999). The thermal melting of *rpoH* mRNA secondary structure at the initiation region allows ribosome entry and thus translation to occur. The molecular chaperones DnaK and DnaJ are the major factors regulating σ^{32} activity. DnaK and DnaJ form a stable complex with σ^{32} and prevent it binding to RNA polymerase. Following heat shock, the accumulation of abnormally folded proteins in the cytoplasm sequesters DnaK and DnaJ (which are themselves HSPs) thus preventing them from interacting with σ^{32} (Figure 1.9). Free cellular σ^{32} can then bind RNA polymerase and activate transcription from heat shock promoters. σ^{32} complexed with RNA polymerase is protected from degradation by FtsH.

A general function of a heat shock protein (HSP) is to monitor and respond to the state of protein folding. More than 30 genes encoding HSPs have been identified as being under the control of σ^{32} including σ^{70}, the isomerase PpiD, the proteases ClpP, ClpX and Lon, and the chaperones DnaK, DnaJ, HtpG, ClpB, GroEL and GroES (Yura *et al.*, 2000). These chaperones, proteases and foldases function to coordinate the repair or degradation of misfolded proteins caused by heat stress. Three HSP make up a chaperone team known as the DnaK-DnaJ-GrpE system (Wild *et al.*, 1996). Current models have DnaK-ATP binding to misfolded peptides, with this interaction being stabilised by DnaJ. DnaJ promotes the hydrolysis of the DnaK-bound-ATP to ADP. GrpE dissociates the bound ADP allowing the rebinding of another ATP and peptide release (Han and Christen, 2003).

A second team is known as the GroEL-GroES chaperone team (Walter, 2002). GroEL is found in cells in a large ribosome-sized structure of 14 GroEL subunits arranged as a double ringed structure (Grallert and Buchner, 2001). This large structure interacts with the seven membered ring GroES. Both systems are involved in the same general processes (i.e. the protein folding pathway) but there appears to be a distinction between their modes of action. The DnaK system prefers polypeptides in an extended conformation (including very early folding intermediates), whereas the GroEL system prefers the molten globule state (i.e. later intermediates) (Houry, 2001). The two systems may act sequentially in a protein folding pathway (Ben-Zvi and Goloubinoff, 2001).

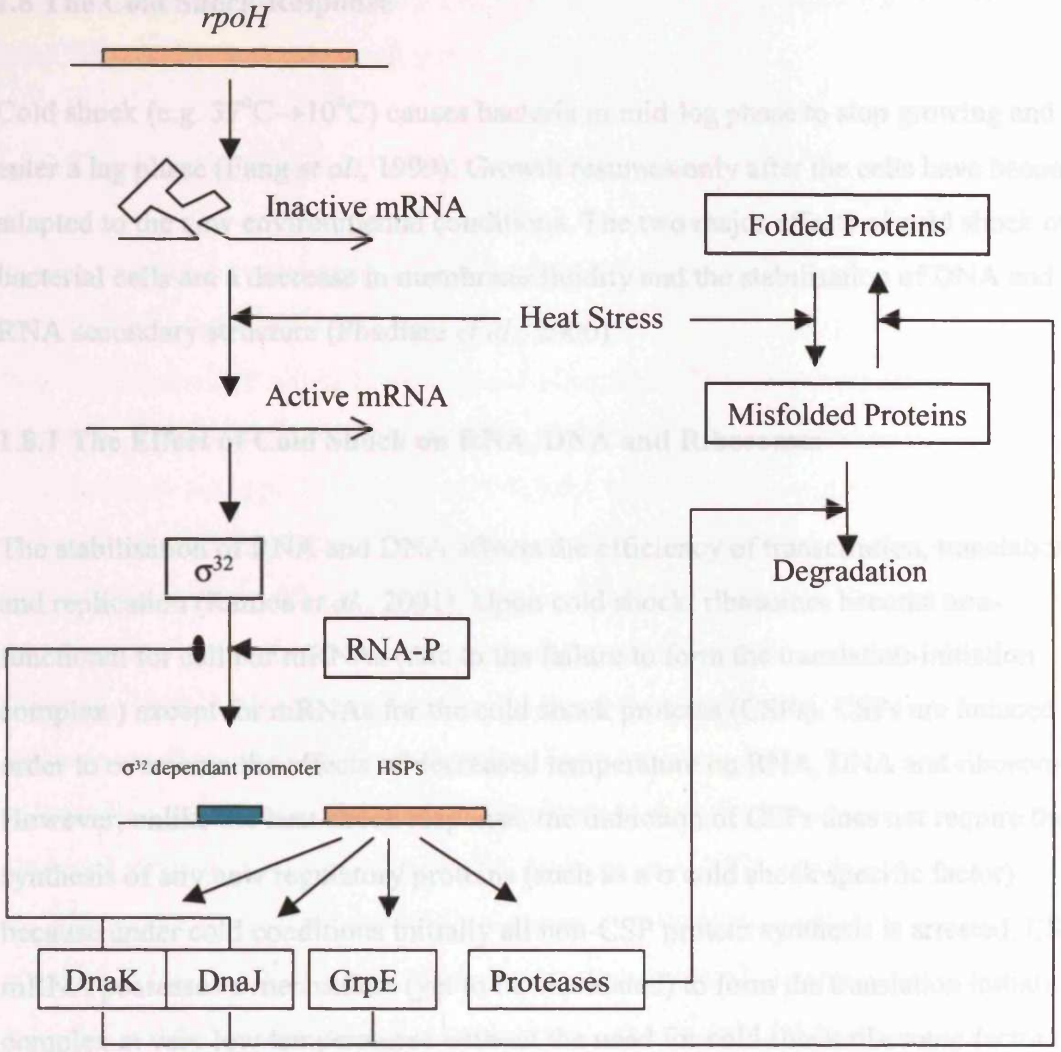


Figure 1.9 Regulation of σ^{32} activity. RNA-P denotes RNA polymerase. Diagram adapted from Yura and Nakahigashi 1999.

1.8 The Cold Shock Response

Cold shock (e.g. 37°C→10°C) causes bacteria in mid-log phase to stop growing and enter a lag phase (Fang *et al.*, 1999). Growth resumes only after the cells have become adapted to the new environmental conditions. The two major effects of cold shock on bacterial cells are a decrease in membrane fluidity and the stabilisation of DNA and RNA secondary structure (Phadtare *et al.*, 2000).

1.8.1 The Effect of Cold Shock on RNA, DNA and Ribosomes

The stabilisation of RNA and DNA affects the efficiency of transcription, translation and replication (Ramos *et al.*, 2001). Upon cold shock, ribosomes become non-functional for cellular mRNAs (due to the failure to form the translation-initiation complex) except for mRNAs for the cold shock proteins (CSPs). CSPs are induced in order to overcome the effects of decreased temperature on RNA, DNA and ribosomes. However, unlike the heat shock response, the induction of CSPs does not require the synthesis of any new regulatory proteins (such as a σ cold shock specific factor) because under cold conditions initially all non-CSP protein synthesis is arrested. CSP mRNA possesses a mechanism (yet to be elucidated) to form the translation initiation complex at very low temperatures without the need for cold shock ribosome factors.

There are around 15 CSPs in *E. coli* of which there are two classes (Thieringer *et al.*, 1998). Class I CSPs (e.g. CspA, CspB, CspG, CsdA, CspI and RbfA) are expressed at very low levels at 37°C and are dramatically induced after cold shock (Phadtare *et al.*, 2000). Class II (e.g. RecA, IF-2, H-NS) are present at 37°C and are only moderately induced after cold shock (less than ten fold). CspA, CspB, CspG, CsdA and CspI are thought to function as RNA chaperones by compensating for the higher stability in RNA secondary structure at lower temperatures. Also among the CSPs are many ribosome-associated proteins that are needed for the efficient formation of the translation initiation complex (involving normal cellular mRNA) at low temperatures e.g. CsdA (RNA unwinding), RbfA (ribosome binding factor), IF-2 (initiation factor) and NusA (transcription termination and antitermination). During the lag phase, unadapted ribosomes become converted to cold-adapted ribosomes by

acquiring these cold shock factors. Cold adapted ribosomes are then able to translate non CSP mRNAs and as a result growth can resume.

1.8.2 The effect of cold shock on membrane fluidity

In order to overcome the decrease in membrane fluidity caused by low temperatures bacteria have evolved a highly conserved response. The proportion of unsaturated fatty acyl chains (UFAs) in the membrane phospholipids is increased during a process known as 'homeoviscous adaption' (Sinensky *et al.*, 1974). Membrane lipids with saturated fatty acids pack together very tightly, whereas lipids with unsaturated fatty acids do not pack together very well because the *cis* double bonds cause bends in the chains that interfere with packing (Becker *et al.*, 1996). However, the resulting bends decrease the van der Waals interactions and therefore lower the transition temperature (Voet and Voet, 1995). Consequently, the larger the number of double bonds, the lower the melting point of the acyl chains. For example, a 2-fold increase in unsaturation (mainly by increased *cis*-vaccenate at the expense of palmitate) is observed in the glycophospholipids of *E. coli* grown at 17°C versus 37°C (Morein *et al.*, 1996). The enzyme β-ketoacyl-acyl carrier protein synthase II (encoded by the gene *fabA*) converts palmitoleic acid into *cis*-vaccenic acid. Although this enzyme is activated at low temperatures, it is not induced immediately following cold shock (Garwin *et al.*, 1980).

1.9 Chemical stress

The repair or degradation of damaged proteins is crucial for cell survival, but overcoming the causative environmental stress is equally crucial for further propagation and growth. In the case of physical stresses, such as hot and cold temperatures, little can be done directly by bacteria to change the environmental conditions. However, bacteria have greater success dealing with stress caused by changes in the chemical environment. Bacteria have evolved to deal directly with some environmental insults, as well as the damage caused by them. The following sections describe the systems *E. coli* uses to deal with the presence of high pH, low pH, metal ions and reactive oxygen species. These systems tend to be induced by one or more of the sensing pathways described earlier (e.g. σ^H, σ^S, σ^E)

1.9.1 pH Stress

The ability to survive extreme changes in the environmental pH is essential for the successful colonisation of the mammalian host by *E. coli* (Tramonti *et al.*, 2002). *E. coli* is faced with an extreme acid shock in the stomach, followed by a pancreatic induced base shock in the upper intestine. External pH can also be significantly altered as a by-product of bacterial metabolism. However, despite major changes (ranging from pH 4.4 to 9.2) in the external environmental pH, *E. coli* can still maintain an internal pH of around 7.6 during growth phase (Stancik *et al.*, 2002).

1.9.1.1 Acid stress

E. coli has three distinct systems that are used for acid resistance: a glucose-catabolite repressed system and two amino acid decarboxylase-dependent systems. The mechanism behind the glucose repressed acid resistance is unknown, but the amino acid decarboxylase systems are thought to consume protons that leak into the cell, through the decarboxylation or arginine and glutamate. The glutamate decarboxylase (Gad) system is the best studied of these three because of its major role in the acid resistance of enteric pathogens such as *E. coli, Shigella flexneri* and *Listeria monocytogenes* (Tramonti *et al.*, 2002). The Gad system is made up of *gadA*, the *gadBC* operon and the recently identified *gadX* and *gadW* (Ma *et al.*, 2002). Two

isoforms of glutamate decarboxylase are encoded by *gadA* and *gadB*, while *gadC* encodes a dedicated antiporter (Masuda and Church, 2003). Control of *gadA* and *gadBC* expression is complex and poorly understood. GadX activates the transcription of *gadA* and *gadBC*, while GadW is thought to act as a repressor of both *gadX* transcription and GadX function (Tramonti *et al.*, 2002; Tucker *et al.*, 2003). However, control of the system is not this simple, since GadW can also activate the transcription of *gadA* and *gadBC* when GadX is absent (but only at pH 8). The pH control of the system is thought to be linked to the pH control of RpoS (σ^S) production. σ^S activates the transcription of *gadX*. Further control of the system occurs through the actions of H-NS which is thought to repress the transcription of *gadX* as well as *gadA* and *gadBC*.

Other genes associated with acid resistance are *hdeA* and *hdeB* (**H**-NS **de**pendent **e**xpression) (Gajiwala and Burley, 2000). HDEA is thought to act as a chaperone preventing the aggregation of periplasmic proteins denatured by acid conditions (Gajiwala and Burley, 1999). The function of HDEB is unknown, but it is a structural homologue of HDEA and may form heterodimers with it in the periplasm.

1.9.2 Heavy Metal Stress

Most heavy metals are transition elements with incompletely filled *d* orbitals (Nies, 1999). These *d* orbitals allow the formation of complex compounds that may or may not be redox active. This allows heavy metal ions to play an important role as essential cofactors in metabolic pathways and they are therefore crucial for microbial growth. Since heavy metals are usually found in trace amounts in the environment microorganisms usually face a struggle in order to acquire them. Prokaryotes have developed two systems in order to scavenge heavy metal ions from their surroundings. The first type of system is non-specific, fast, driven by the chemiosmotic gradient across the cytoplasmic membrane and constitutively expressed. An example of this type of system is CorA, which accumulates nickel, cobalt, zinc and manganese. The second type of system is slower, has higher specificity and is often driven by ATP hydrolysis (sometimes in addition to the chemiosmotic gradient) (Nies and Silver, 1995). These specific systems are only used in times of need, starvation or in special metabolic situations.

However, some heavy metals are also common environmental pollutants and are toxic even at very low concentrations. Heavy metals become toxic because of the fast acting, constitutively expressed non-specific uptake systems. These uptake systems provide an open gate through which metal ions can freely diffuse to toxic concentrations (Nies and Silver, 1995).

1.9.2.1 The effect of Metal Ions on Cell Biochemistry

A supra-optimal cellular concentration of heavy metal ions can lead to enzyme inhibition, protein denaturation, biopolymer hydrolysis and uncontrollable redox cycling. Heavy metal ions (especially those with higher atomic numbers such as mercury, cadmium and silver) have a tendency to bind SH groups, including those of cysteine residues in proteins, causing an inhibition of activity in sensitive enzymes. Heavy metal cations can also bind to glutathione to form bisglutathionato complexes that tend to react with molecular oxygen to form oxidised bisglutathione, the metal cation and H_2O_2 (Kachur *et al.*, 1998). H_2O_2 itself is toxic to cells and causes oxidative stress (see Section 1.9.3), but also the bisglutathione has to be reduced again by an NADPH-dependent reaction. The metal cation can immediately bind two more glutathione molecules leading to uncontrollable redox cycling, placing considerable oxidative stress upon the cell.

Heavy metal oxyanions can also interfere with the metabolism of structurally related non-metals (e.g. chromate with sulphate) and the reduction of these heavy metal anions may lead to the production of free radicals (e.g. chromate). It is clear that heavy metals can place considerable stress upon bacteria when found at higher than optimal concentrations. However, since heavy metal ions cannot be degraded or modified like toxic organic compounds, there remain three possible mechanisms that bacteria can use to cope with metal ion stress: the metal ion can be excluded from the cell by an active pumping mechanism/permeability barrier, heavy metal cations can be sequestered by complex compounds or the ion may be reduced to a less toxic oxidation state.

Bacteria usually use a combination of two or all three of these basic mechanisms in order to cope with heavy metal stress. However, there are some

problems associated with these stress responses. Not all metals can be reduced by a cell. Metal compounds that have been reduced must be able to diffuse out of the cell (otherwise they might re-oxidise), but reduction products can be insoluble or even more toxic. The cell must therefore provide an effective efflux system to export reduced products if it detoxifies a compound by this method. The energy cost of complexation is huge. Complexation is only efficient at low concentrations of heavy metal cations when all the heavy metal in the immediate environment can be complexed by the bacterial population (this is not usually the case). However, the cycle of uptake and then efflux is completely futile and equally wasteful of resources.

1.9.3 Oxidative Stress

Oxidative stress can be caused by the presence of reactive oxygen species such as superoxide anion (e.g. $O_2^{\cdot-}$), hydrogen peroxide (H_2O_2) or hydroxyl radicals ($HO\cdot$) (Storz and Zheng, 2000). These species are created by bacteria themselves in the respiratory chain and are also produced upon exposure to radiation, metals and redox active drugs. They are highly reactive and lead to damage of protein, DNA and membranes. *E. coli* has many antioxidant defences including the oxyR and soxRS operons (Smirnova *et al.*, 2000).

1.9.3.1 The *oxyR* Regulon of *E. coli*

The *oxyR* regulon is induced by increased levels of H_2O_2 which leads to the expression of anti-oxidant genes. OxyR is activated by H_2O_2 by direct oxidation of two thiol groups allowing an intramolecular disulphide bond to form. The oxidised form can then activate transcription from responsive promoters. OxyR activated genes protect against oxidative stress caused by H_2O_2 by a variety of methods. Hydroperoxidase I and alkyl hydroperoxide reductase (encoded by *katG* and *ahpCF* respectively) protect by directly eliminating H_2O_2. Glutathione reductase (*gorA*), glutaredoxin 1 (*grxA*) and thioredoxin 2 (*trxC*) maintain the cellular thiol-disulphide balance. Dps protects against DNA damage and mutation and also sequesters iron (Dps is a ferritin homologue) (Grant *et al.*, 1998). The induction of *fur* (a ferric iron uptake repressor) prevents damage caused by $HO\cdot$ generated by H_2O_2 reacting with intracellular iron (the Fenton reaction) (Zhang *et al.*, 1999).

The induction of *oxyS* provides a link to other stress responses. OxyS is a small unique RNA that protects against mutagenesis through an unknown mechanism. OxyS RNA also affects the expression of *rpoS* and *fhlA* (Zheng *et al.*, 1998). RpoS and FhlA are transcriptional regulators and this link may serve to integrate the adaptive response with other cellular stress responses. OxyR also protects against stress caused by hypochlorous acid, organic solvents and reactive nitrogen species. OxyR can also cause the repression of a sub set of genes including *oxyR* itself, *fhuF* (a ferric iron reductase) and flu (an outer membrane protein) (Storz and Zheng, 2000).

1.9.3.2 The SoxRS Regulon of *E. coli*

The *soxRS* regulon is induced in response to elevated $O_2^{\cdot -}$ levels via a two stage process (Wu and Weiss, 1992). SoxR becomes converted into an active form (by oxidation) and enhances *soxS* transcription. SoxS then induces the expression of the genes of the SoxRS regulon (Tsaneva and Weiss, 1990). The gene products induced to combat oxidative stress directly include superoxide dismutase (*sodA*), DNA repair enzyme endonuclease IV (*nfo*) and glucose-6-phosphate dehydrogenase (*zwf*) (Nunoshiba *et al.*, 1992). Fur is also induced by SoxRS, which helps to stop the formation of $HO\cdot$, and provides an overlap with the OxyR regulon (Zheng *et al.*, 1999). The activation of an outer membrane protein (*tolC*) and an efflux pump (*acrAB*) may serve to export $O_2^{\cdot -}$ generating compounds.

1.10 Interactions Between Different Stress Responses

Bacteria have evolved different mechanisms for coping with different stresses. However, since many very different stresses can have similar effects on a bacterial cell (e.g. heat and chemicals can cause proteins to be denatured in the cytoplasm) specific and general stress responses should not be thought of as completely independent pathways. Heat, extreme pH and chemicals can all cause proteins to become denatured in the membrane. Metal ion stress can cause the inhibition of enzymes but it can also cause severe oxidative stress (like that caused by high H_2O_2 concentration or UV radiation) through the production of radicals. Different stresses

can therefore cause the induction of one or more stress responses to varying degrees and it is not surprising that many of the regulatory components of specific stress responses overlap with each other.

1.11 Aims

The aim of this thesis is to profile the response of *E. coli* to PHMB exposure. This was done at the transcriptional level using gene arrays and at the protein level using 2-dimensional gel electrophoresis. Genes and proteins of interest were studied further to determine their role in the response.

This investigation was designed to determine:

- **WHAT** is altered in response to PHMB (i.e. the genes and proteins)
- **WHERE** these alterations act (i.e. the cellular location)
- **WHEN** these alterations occur (i.e. the timing of the response)

And provide in an insight into:

- **HOW** these alterations occur (i.e. the regulation)
- **WHY** these alterations occur (i.e. the biochemical and physiological reasons)

The identification of systems that are altered in response to PHMB may also provide further clues to the mechanism of action of PHMB.

Chapter 2: General Materials and Methods

2.1 Materials

2.1.1 Bacterial Strains, Plasmids and Oligonucleotides

The bacterial strains and plasmids used in this study are shown in Table 2.1 and Table 2.2 respectively. All DNA oligonucleotides were purchased from Sigma-Genosys (Dorset, UK) or Gibco-BRL (Paisley, UK). Each was supplied at 50 nmol scale of synthesis with no additional modifications. On receipt, each oligonucleotide was resuspended in TE buffer to a final concentration of 1 $\mu g/\mu l$ and stored at -20°C. These 10 × stock solutions were diluted ten-fold in TE buffer before use in PCR reactions.

Table 2.1 *Escherichia coli* K12 derivatives used in this study

Strain	Genotype	Source
W3110	F^-, λ^-, IN(rrnD-rrnE), rph-1	CGSC# 4474
W3110 Δ*cpxP*	F^-, λ^-, IN(rrnD-rrnE), rph-1, Δ*cpxP*::Km	This study
W3110 Δ*b1228*	F^-, λ^-, IN(rrnD-rrnE), rph-1, Δ*b1228*::Km	This study
MG1655	F^-, λ^-, rph-1	Blattner *et al.*, 1997
MG1655 Δ*aceA*	F^-, λ^-, rph-1, Δ*aceA*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*aspA*	F^-, λ^-, rph-1, Δ*aspA*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*cbl*	F^-, λ^-, rph-1, Δ*cbl*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*cpxP*	F^-, λ^-, rph-1, Δ*cpxP*::Km	This study
MG1655 Δ*cysB*	F^-, λ^-, rph-1, Δ*cysB*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*cysK*	F^-, λ^-, rph-1, Δ*cysK*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*evgS*	F^-, λ^-, rph-1, Δ*evgS*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*flgJ*	F^-, λ^-, rph-1, Δ*flgJ*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*fliC*	F^-, λ^-, rph-1, Δ*fliC*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*gadW*	F^-, λ^-, rph-1, Δ*gadW*::Km	Ma *et al.*, 2002
MG1655 Δ*gadX*	F^-, λ^-, rph-1, Δ*gadX*::Km	Ma *et al.*, 2002
MG1655 Δ*gatC*	F^-, λ^-, rph-1, Δ*gatC*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*glpD*	F^-, λ^-, rph-1, Δ*glpD*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*hdeA*	F^-, λ^-, rph-1, Δ*hdeA*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*hns*	F^-, λ^-, rph-1, Δ*hns*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*htrL*	F^-, λ^-, rph-1, Δ*htrL*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*intB*	F^-, λ^-, rph-1, Δ*intB*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*lldP*	F^-, λ^-, rph-1, Δ*lldP*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*malK*	F^-, λ^-, rph-1, Δ*malK*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*manX*	F^-, λ^-, rph-1, Δ*manX*::Km	U.W. *E.coli* Genome Project

Table 2.1 Continued

Strain	Genotype	Source
MG1655 Δ*recX*	F⁻, λ⁻, rph-1, Δ*recX*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*rfbX*	F⁻, λ⁻, rph-1, Δ*rfbX*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*rhsE*	F⁻, λ⁻, rph-1, Δ*rhsE*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*tdcR*	F⁻, λ⁻, rph-1, Δ*tdcR*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*xseA*	F⁻, λ⁻, rph-1, Δ*xseA*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*ydhA*	F⁻, λ⁻, rph-1, Δ*ydhA*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*yebG*	F⁻, λ⁻, rph-1, Δ*yebG*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*yicJ*	F⁻, λ⁻, rph-1, Δ*yicJ*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*mglB*	F⁻, λ⁻, rph-1, Δ*mglB*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*osmB*	F⁻, λ⁻, rph-1, Δ*osmB*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*pflB*	F⁻, λ⁻, rph-1, Δ*pflB*::Km	U.W. *E.coli* Genome Project
MG1655 Δ*recA*	F⁻, λ⁻, rph-1, Δ*recA*::Km	U.W. *E.coli* Genome Project

Table 2.2 Plasmids used in this study. pCA24N and pMJA based plasmids are IPTG inducible containing His-tagged ORFs.

Plasmid	Relevant Characteristics	Source
pKD46	Recombination plasmid	Datsenko and Wanner, 2000
pKD4	Amplification of Km resistance	Datsenko and Wanner, 2000
pCR-Blunt	Blunt ended cloning vector	Invitrogen (Paisley, UK)
pCA24N-*aceA*	GFP fused, *aceA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*aspA*	GFP fused, *aspA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*b0499*	GFP fused, *b0499* over expression plasmid	Mori *et al.*, 2000
pCA24N-*b1458*	GFP fused, *b1458* over expression plasmid	Mori *et al.*, 2000
pCA24N-*b1459*	GFP fused, *b1459* over expression plasmid	Mori *et al.*, 2000
pCA24N-*b2854*	GFP fused, *b2854* over expression plasmid	Mori *et al.*, 2000
pCA24N-*cbl*	GFP fused, *cbl* over expression plasmid	Mori *et al.*, 2000
pCA24N-*cpxP*	GFP fused, *cpxP* over expression plasmid	Mori *et al.*, 2000
pCA24N-*cysB*	GFP fused, *cysB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*cysK*	GFP fused, *cysK* over expression plasmid	Mori *et al.*, 2000
pCA24N-*dnaK*	GFP fused, *dnaK* over expression plasmid	Mori *et al.*, 2000
pCA24N-*evgS*	GFP fused, *evgS* over expression plasmid	Mori *et al.*, 2000
pCA24N-*fis*	GFP fused, *fis* over expression plasmid	Mori *et al.*, 2000

Table 2.2 Continued

Plasmid	Relevant Characteristics	Source
pCA24N-*flgJ*	GFP fused, *flgJ* over expression plasmid	Mori *et al.*, 2000
pCA24N-*fliC*	GFP fused, *fliC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*gadW*	GFP fused, *gadW* over expression plasmid	Mori *et al.*, 2000
pCA24N-*gadX*	GFP fused, *gadX* over expression plasmid	Mori *et al.*, 2000
pCA24N-*gatC*	GFP fused, *gatC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*glpD*	GFP fused, *glpD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*hdeA*	GFP fused, *hdeA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*hns*	GFP fused, *hns* over expression plasmid	Mori *et al.*, 2000
pCA24N-*htrL*	GFP fused, *htrL* over expression plasmid	Mori *et al.*, 2000
pCA24N-*intB*	GFP fused, *intB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*lldP*	GFP fused, *lldP* over expression plasmid	Mori *et al.*, 2000
pCA24N-*malK*	GFP fused, *malK* over expression plasmid	Mori *et al.*, 2000
pCA24N-*manX*	GFP fused, *manX* over expression plasmid	Mori *et al.*, 2000
pCA24N-*mglB*	GFP fused, *mglB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*osmB*	GFP fused, *osmB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*pflB*	GFP fused, *pflB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rbsD*	GFP fused, *rbsD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*recA*	GFP fused, *recA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*recX*	GFP fused, *recX* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rfaL*	GFP fused, *rfaL* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rfbX*	GFP fused, *rfbX* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rfc*	GFP fused, *rfc* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rhsA*	GFP fused, *rhsA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rhsB*	GFP fused, *rhsB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rhsC*	GFP fused, *rhsC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rhsD*	GFP fused, *rhsD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*rhsE*	GFP fused, *rhsE* over expression plasmid	Mori *et al.*, 2000
pCA24N-*tdcR*	GFP fused, *tdcR* over expression plasmid	Mori *et al.*, 2000
pCA24N-*tnaA*	GFP fused, *tnaA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*tnaL*	GFP fused, *tnaL* over expression plasmid	Mori *et al.*, 2000
pCA24N-*uspA*	GFP fused, *uspA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*vacJ*	GFP fused, *vacJ* over expression plasmid	Mori *et al.*, 2000
pCA24N-*xseA*	GFP fused, *xseA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yadC*	GFP fused, *yadC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yahA*	GFP fused, *yahA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybbC*	GFP fused, *ybbC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybbD*	GFP fused, *ybbD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybfB*	GFP fused, *ybfB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybfC*	GFP fused, *ybfC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybfD*	GFP fused, *ybfD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ybfO*	GFP fused, *ybfO* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ydcC*	GFP fused, *ydcC* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ydcD*	GFP fused, *ydcD* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ydcE*	GFP fused, *ydcE* over expression plasmid	Mori *et al.*, 2000

Table 2.2 Continued

Plasmid	Relevant Characteristics	Source
pCA24N-*ydhA*	GFP fused, *ydhA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yebG*	GFP fused, *yebG* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yefI*	GFP fused, *yefI* over expression plasmid	Mori *et al.*, 2000
pCA24N-*ygeV*	GFP fused, *ygeV* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yhaB*	GFP fused, *yhaB* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yhhH*	GFP fused, *yhhH* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yhhI*	GFP fused, *yhhI* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yibA*	GFP fused, *yibA* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yibG*	GFP fused, *yibG* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yibJ*	GFP fused, *yibJ* over expression plasmid	Mori *et al.*, 2000
pCA24N-*yicJ*	GFP fused, *yicJ* over expression plasmid	Mori *et al.*, 2000
pMJA	Overexpression plasmid (ORFless)	This Study
pMJA-*aceA*	*aceA* over expression plasmid	This Study
pMJA-*aspA*	*aspA* over expression plasmid	This Study
pMJA-*b0499*	*b0499* over expression plasmid	This Study
pMJA-*b1458*	*b1458* over expression plasmid	This Study
pMJA-*b1459*	*b1459* over expression plasmid	This Study
pMJA-*b2854*	*b2854* over expression plasmid	This Study
pMJA-*cbl*	*cbl* over expression plasmid	This Study
pMJA-*cpxP*	*cpxP* over expression plasmid	This Study
pMJA-*cysB*	*cysB* over expression plasmid	This Study
pMJA-*cysK*	*cysK* over expression plasmid	This Study
pMJA-*dnaK*	*dnaK* over expression plasmid	This Study
pMJA-*evgS*	*evgS* over expression plasmid	This Study
pMJA-*fis*	*fis* over expression plasmid	This Study
pMJA-*flgJ*	*flgJ* over expression plasmid	This Study
pMJA-*fliC*	*fliC* over expression plasmid	This Study
pMJA-*gadW*	*gadW* over expression plasmid	This Study
pMJA-*gadX*	*gadX* over expression plasmid	This Study
pMJA-*gatC*	*gatC* over expression plasmid	This Study
pMJA-*glpD*	*glpD* over expression plasmid	This Study
pMJA-*hdeA*	*hdeA* over expression plasmid	This Study
pMJA-*hns*	*hns* over expression plasmid	This Study
pMJA-*htrL*	*htrL* over expression plasmid	This Study
pMJA-*intB*	*intB* over expression plasmid	This Study
pMJA-*lldP*	*lldP* over expression plasmid	This Study
pMJA-*malK*	*malK* over expression plasmid	This Study
pMJA-*manX*	*manX* over expression plasmid	This Study
pMJA-*mglB*	*mglB* over expression plasmid	This Study
pMJA-*osmB*	*osmB* over expression plasmid	This Study
pMJA-*pflB*	*pflB* over expression plasmid	This Study
pMJA-*rbsD*	*rbsD* over expression plasmid	This Study
pMJA-*recA*	*recA* over expression plasmid	This Study
pMJA-*recX*	*recX* over expression plasmid	This Study

Table 2.2 Continued

Plasmid	Relevant Characteristics	Source
pMJA-*rfaL*	*rfaL* over expression plasmid	This Study
pMJA-*rfbX*	*rfbX* over expression plasmid	This Study
pMJA-*rfc*	*rfc* over expression plasmid	This Study
pMJA-*rhsA*	*rhsA* over expression plasmid	This Study
pMJA-*rhsB*	*rhsB* over expression plasmid	This Study
pMJA-*rhsC*	*rhsC* over expression plasmid	This Study
pMJA-*rhsD*	*rhsD* over expression plasmid	This Study
pMJA-*rhsE*	*rhsE* over expression plasmid	This Study
pMJA-*tdcR*	*tdcR* over expression plasmid	This Study
pMJA-*tnaA*	*tnaA* over expression plasmid	This Study
pMJA-*tnaL*	*tnaL* over expression plasmid	This Study
pMJA-*uspA*	*uspA* over expression plasmid	This Study
pMJA-*vacJ*	*vacJ* over expression plasmid	This Study
pMJA-*xseA*	*xseA* over expression plasmid	This Study
pMJA-*yadC*	*yadC* over expression plasmid	This Study
pMJA-*yahA*	*yahA* over expression plasmid	This Study
pMJA-*ybbC*	*ybbC* over expression plasmid	This Study
pMJA-*ybbD*	*ybbD* over expression plasmid	This Study
pMJA-*ybfB*	*ybfB* over expression plasmid	This Study
pMJA-*ybfC*	*ybfC* over expression plasmid	This Study
pMJA-*ybfD*	*ybfD* over expression plasmid	This Study
pMJA-*ybfO*	*ybfO* over expression plasmid	This Study
pMJA-*ydcC*	*ydcC* over expression plasmid	This Study
pMJA-*ydcD*	*ydcD* over expression plasmid	This Study
pMJA-*ydcE*	*ydcE* over expression plasmid	This Study
pMJA-*ydhA*	*ydhA* over expression plasmid	This Study
pMJA-*yebG*	*yebG* over expression plasmid	This Study
pMJA-*yefI*	*yefI* over expression plasmid	This Study
pMJA-*ygeV*	*ygeV* over expression plasmid	This Study
pMJA-*yhaB*	*yhaB* over expression plasmid	This Study
pMJA-*yhhH*	*yhhH* over expression plasmid	This Study
pMJA-*yhhI*	*yhhI* over expression plasmid	This Study
pMJA-*yibA*	*yibA* over expression plasmid	This Study
pMJA-*yibG*	*yibG* over expression plasmid	This Study
pMJA-*yibJ*	*yibJ* over expression plasmid	This Study
pMJA-*yicJ*	*yicJ* over expression plasmid	This Study

Table 2.3 Oligonucleotides and their sequences used in this study. 'KO' in the name refers to oligonucleotides used in the production of gene knock-outs in *E. coli*. 'Check' in the name refers to oligonucleotides used in the verification of knock-out strains.

Oligonucleotide Name	Sequence
ygeQ forward KO	TATAAGGAGA TCAAGGTAAA GACCTAAAGC ATCATTTTGT GTAGGCTGGA GCTGC
ygeQ reverse KO	ATGCAATCCT TTTGTAACAA CCCAGGTAAA ACACCTTATG AATATCCTCC TTAGTTCC
yaiN forward KO	CAGCTATTTA AGATAGGCAC GAACCAGTTC AATAGTTGTG TAGGCTGGAG CTGC
yaiN reverse KO	AAATTTTATT CTCCAGTGTT ATATACTATA GGGGGGTATG AATATCCTCC TTAGTTCC
b1228 forward KO	GGATGCGCCT TCGCTTATCC GACCTACAGG GGAGGATATT GTTAGGCTGG AGCTGC
b1228 reverse KO	CACCATCACT TTCAAAAGTC CCTGAACTCT CAAGCGAATA TGAATATCCT CCTTAGTTCC
ycgW forward KO	CGATCATCGA AAACATGTAA TCTCTCCATG TGTTAAATAT TGTGTAGGCT GGAGCTGC
ycgW reverse KO	TAATTAATTT ATTTTTTTGA GGGGGGGGTA ATATACTCAT ATGAATATCC TCCTTAGTTC C
cpxP forward KO	CATGACTTTA CGTTGTTTTA CACCCCCTGA CGCATGTTTG TGTAGGCTGG AGCTGC
cpxP reverse KO	CTGACGCTGA TGTTCGGTTA AACTTATGCC GTCGAACATA TGAATATCCT CCTTAGTTCC
b3914 reverse KO	CCCCCACAGC ATGTGGGGGA AGACAGGGAT GGTGTCTATG AATATCCTCC TTAGTTCC
ygeQ forward check	GCAACCGACT TTAATCGGTC
ygeQ reverse check	TCTTCGCCGT AATACTTCCC
yaiN forward check	TTCATCTCTC GCTCTTCCTC
yaiN reverse check	GTTTTCCGAC CACATTCACC
b1228 forward check	GTTCACATAG ACCCTGCTTC G
b1228 reverse check	GCAGCGCTGA GTAATCCTTC
ycgW forward check	TGATGGAAGG CGCTAAGCTG
ycgW reverse check	AAAAACAACG GCCGTGCCAC
cpxP forward check	CTCCGAGGCA GAAATTACGT C
cpxP reverse check	GGGCCTGTTG CATAAGATCT C
b3914 reverse check	CCGCTATCAA CTGACGCTAG
K_1	CAGTCATAGC CGAATAGCCT
K_2	CGGTGCCCTG AATGAACTGC
K_T	CGGCCACAGT CGATGAATCC

2.1.2 Chemicals, Reagents and Laboratory Consumables

General laboratory chemicals were obtained from Fisher-Scientific (Leicester, UK) or Sigma-Aldrich (Poole, UK) and were of analytical grade or higher. Plastic ware and general laboratory consumables were obtained from Greiner (Stonehouse, UK) and Alpha Laboratories (Eastleigh, UK). All other reagents and consumable suppliers are listed in the text where appropriate. Type II ultrapure 18 MOhm water was obtained from a Milli Q 50 water purification unit (Millipore, Watford, UK).

2.1.2.1 Polyhexamethylene Biguanide (PHMB)

PHMB was kindly provided by Avecia (Manchester, UK), as a mixture of homologues with *n* (Figure 1.3) ranging from 2 to 15 and with a mean value of 5.5. Fluorescent-PHMB, also from Avecia, was of the same composition except that 1% of all biguanide units was replaced by the fluor, 6-amino-benzo[de]isoquinoline-1,3-dione, attached to flanking hexamethylene chains via the ring N(2) and amino substituent at position 6 (See Appendix D, Figure D.1). The synthetic route ensured that each homologue (n = 2 to 15) was labelled and this was confirmed by coincidence of UV and fluorescent elution profiles of effluents from GPC column chromatography and fluorescence was not quenched by added PHMB. The parent fluor, 1,8-naphthalimide, was obtained from Lancaster Chemicals (Morecambe, UK). The excitation and emission wavelengths were 340 nm and 395 nm respectively.

2.1.3 Enzymes

Restriction endonucleases, DNA ligase, Vent and Taq polymerase supplied with their appropriate buffers were obtained from New England Biolabs (Hitchin, UK). Bovine serum albumin (BSA), where required, was supplied with the enzyme.

2.1.4 Growth Media

Table 2.4 Composition of growth media

Medium	Composition (per litre)	Supplier
Luria-Bertani broth (Pre-mixed form)	10 g Tryptone 5 g Yeast extract 10 g NaCl	GibcoBRL (Paisley, UK)
Luria-Bertani broth Agar (Capsule form)	10 g Tryptone 5 g Yeast Extract 10 g NaCl 15 g Agar-B	Bio101 (London, UK)
SOC	20 g Tryptone 5 g Yeast extract 0.5 g NaCl 10 ml 1 M $MgCl_2$ 10 ml 1 M $MgSO_4$ 2 ml 20% (w/v) Glucose	GibcoBRL (Paisley, UK)

Media were prepared in accordance with the manufacturer's instructions.

2.1.5 Antibiotic Selection

Antibiotics were obtained from Melford Laboratories (Ipswich, UK) and were used in the following concentrations:

Kanamycin (Km)	50 μg ml^{-1}
Carbenicillin (Cb)	250 μg ml^{-1}
Ampicillin (Amp)	50 μg ml^{-1}
Chloramphenicol (Chl)	37 μg ml^{-1}

Kanamycin, carbenicillin and ampicillin were dissolved in water at 1000 × working concentration in water, sterilised by filtration (0.2 μm filter) and stored at –20°C. Chloramphenicol was dissolved in 70 % aqueous ethanol, sterilised by filtration (0.2 μm filter pore-size) and stored at –20°C.

2.1.6 Commonly used solutions

Table 2.5 Composition of commonly used solutions

Solution	Components	Supplier
10 × TBE	0.89 M Tris Borate 20 mM EDTA (disodium salt)	Fisher-Scientific (Leicester, UK)
DNA loading solution	0.05% (w/v) Bromophenol blue 50% (w/v) Sucrose 10 mM EDTA (disodium salt)	BDH (Poole, UK) Fisher-Scientific (Leicester, UK)
TfbI pH 5.8	2.9 g Potassium acetate 1.2 g $RbCl_2$ 2.3 g $CaCl_2$ 9.9 g $MnCl_2$ 150 ml Glycerol Acetic acid (to adjust pH)	All components from Sigma-Aldrich (Poole, UK) except glycerol (Fisher-Scientific, Leicester, UK)
TfbII pH 6.5	2.093 g MOPS 11.027 g $CaCl_2$ 1.2 g RbCl 150 ml Glycerol KOH (to adjust pH)	All components from Sigma-Aldrich (Poole, UK) except glycerol (Fisher-Scientific, Leicester, UK)
TE Buffer pH 8	10 mM-Tris/HCl 1 mM EDTA	Fisher-Scientific (Leicester, UK)
SSPE	0.18 M NaCl 10 mM NaH_2PO_4 1 mM EDTA, pH 8.0	Fisher-Scientific (Leicester, UK)
Denhardt's Reagent	0.02% (w/v) Ficoll 0.02% (w/v) Polyvinylpyrrolidone 0.02% (w/v) Bovine serum albumin	All components from Sigma-Aldrich (Poole, UK)

2.2 General Methods

2.2.1 Maintenance of Bacterial strains

All bacterial strains were cultured at 37°C unless otherwise stated. Liquid cultures were incubated with constant shaking at 200 rpm. For long term storage, liquid cultures were mixed 1:1 (v/v) with sterile glycerol, mixed thoroughly and stored at -80°C.

2.2.2 Treatment of Glassware

All flasks used during bacterial growth experiments involving PHMB were washed in concentrated nitric acid, rinsed twice in distilled water, air dried, rinsed with 2% dimethyldichlorosilane in 1,1,1-trichloroethane (BDH, Poole, UK), dried, baked at 130°C and then rinsed three times in distilled water.

2.2.3 Transformations

Transformations of *E. coli* cells with plasmid DNA was carried out using the rubidium chloride method (adapted from Hanahan, 1983) or the TSS method (described by Chung *et al.*, 1989) for producing chemically competent cells.

2.2.3.1 Preparation of Chemically Competent Cells: Rubidium Chloride Method

The strain to be transformed was grown overnight in 5 ml LB medium at 37°C with shaking, subcultured (5 ml) into 200 ml LB media (prewarmed to 37°C) and incubated at 37°C with shaking until optical attenuance (D_{600}) was 0.4. The culture was chilled on ice for 5 mins before harvesting by centrifugation at 6000 rpm for 5 mins at 4°C (Beckman rotor). The pellet was resuspended in 80 ml of ice-cold TfbI buffer. The resuspended cells were left on ice for 1 h before being harvested by centrifugation (6000 rpm, 5 mins, 4°C) and resuspended in 8 ml of ice cold TfbII. The cells were left on ice for 4 h, aliquoted and snap-frozen in liquid nitrogen. The 100 µl aliquots were stored at –80°C until needed.

2.2.3.2 Preparation of Chemically Competent Cells: TSS Method

E.coli cells were grown in LB broth to D_{600} 0.4 before harvesting by centrifugation (14,000 g, bench top centrifuge, 1 min). The pellet was resuspended in one-tenth original volume of ice-cold TSS solution (LB broth containing 10% (w/v) PEG, 5% (v/v) DMSO and 50 mM $MgCl_2$, pH 6.5), stored on ice for 10 minutes before being either snap-frozen in liquid nitrogen and stored at -80°C or transformed immediately using the method described below (Section 2.2.3.4).

2.2.3.3 Preparation of Electro Competent cells

E.coli W3110 containing pKD46 were grown in 100ml of LB broth (containing Cb) at 30°C to D_{600} 0.6. The culture was harvested by centrifugation (10,000 g, 4°C, 15 minutes) and the pellet washed twice in 100 ml ice-cold sterile MilliQ and then once in 3 ml 10% glycerol. The pellet was resuspended in 0.3 ml 10% glycerol, aliquoted and snap-frozen in liquid nitrogen. The 25 μl aliquots were stored at -80°C until needed.

2.2.3.4 Transformation of Bacterial Cells with Plasmid DNA

Chemically competent cells (produced by either of the methods described in Sections 2.2.3.1 or 2.2.3.2) were thawed on ice for 15 minutes and up to 50 ng plasmid DNA was added per 100 μl cell suspension. After incubation on ice for 30 minutes, the cells were heated at 42°C for 90 s and returned to the ice for a further 5 minutes. Expression of selectable markers was achieved by the addition of 400 μl LB media (containing the appropriate antibiotic) and incubated for 1-2 h at 37°C with shaking. The cells were harvested (1000 g, 4 mins) and resuspended in approximately half the original volume of LB media. The cell suspension was plated out onto LB agar plates (containing the appropriate antibiotic) and incubated at 37°C overnight.

2.2.3.5 Transformation of Bacterial Cells with Linear DNA

Electro competent cells (see Section 2.2.3.3) were thawed on ice for 30 minutes and transferred to an ice-cold 2mm electroporation cuvette (ThermoHybaid, London, UK).

Up to 100 ng linear DNA was added per 25 μl cells and the cells immediately electroporated (Micropulser, BioRad, Hemel Hempstead, UK) using the recommended settings (program EC2, 1.8 kV, 6.1 mS). SOC media was added (1ml) and the cells incubated for 1 h at 37°C with shaking. Cells were plated out onto LB agar plates (containing the appropriate antibiotic) and incubated at 37°C overnight.

2.2.4 Preparation of Nucleic Acid

2.2.4.1 Small Scale Plasmid Preparation

The plasmid bearing strain was grown to stationary phase overnight (37°C, 200 rpm) in 5 ml LB media containing the appropriate antibiotics. DNA extraction was performed using Qiagen Plasmid Mini Kits as described by the manufacturer (Qiagen, Crawley, UK).

2.2.4.2 Large Scale Plasmid Preparation

The plasmid bearing strain was grown to stationary phase overnight (37°C, 200 rpm) in 100 ml LB media containing the appropriate antibiotics. DNA extraction was performed using Qiagen Plasmid Mini Kits as described by the manufacturer (Qiagen, Crawley, UK).

2.2.4.3 Purification of Genomic DNA

Genomic DNA was prepared from *E. coli* strains according to the US Dept Commerce molecular biology protocol (http://micro.nwfsc.noaa.gov/protocols/Q-genomic.html). Strains were cultured overnight in 5 ml LB broth containing antibiotics where applicable. The cells were harvested by centrifugation (1,500 g, 10 mins, 4°C) and resuspended in 3.5 ml of buffer B1 (50 mM-Tris/HCl, 50 mM EDTA, 0.5% (v/v) Tween 20, 0.5% (v/v) Triton X-100, 200 μg/ml RnaseA). Lysozyme (80 μl of 100 mg/ml stock solution) and Proteinase K (100 μl of 20 mg/ml stock solution) were added and the solution incubated at 37°C for 30 minutes. Solution B2 (1.2 ml of 3 M guanidine hydrochloride, 20% (v/v) Tween 20) was added, mixed gently by inversion and incubated at 50°C for a further 30 minutes. A Qiagen 100 column (Qiagen,

Crawley, UK) was pre-equilibrated with 4 ml of buffer QBT (750 mM NaCl, 50 mM MOPS, 15% (v/v) ethanol, 0.15% (v/v) Triton X-100, pH 7.0). The lysate was diluted with an equal volume of QBT and allowed to pass through the column under gravity. The column was washed twice with 7.5 ml aliquots of buffer QC (1.0 M NaCl, 50 mM MOPS, 15% (v/v) ethanol, pH 7.0) and the genomic DNA eluted with 5 ml of buffer QF (1.25 mM NaCl, 50 mM-Tris/HCl, 15% (v/v) aqueous ethanol, pH 8.5). The eluted fraction was then precipitated as described below and resuspended in TE buffer to a final concentration of ~1 μg/μl and stored at -20°C.

2.2.4.4 Precipitation of DNA

One-tenth volume of sodium acetate (3M, pH 5.2) and two volumes propan-2-ol were added to the DNA solution, mixed thoroughly and the sample incubated at -20°C for 30 minutes. The DNA was pelleted by centrifugation (20,000 g, 10 minutes, room temperature), washed in 70% ice-cold ethanol and air dried for 10 minutes. The DNA was resuspended in an appropriate volume of MilliQ or TE buffer.

2.2.4.5 Isolation of Total RNA

Samples of cell culture (1 ml) were pipetted into 1.5 ml Eppendorf tubes and the cells were harvested by centrifugation (20,000g in a bench top centrifuge). Total RNA was isolated using an RNeasy miniprep kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. RNA was eluted in 30 μl RNAse free water (Sigma, Poole, UK). The RNA concentration was determined by absorbance at 260 nm using a Genequant spectrophotometer (Amersham, Buckinghamshire, UK).

2.2.5 DNA Manipulations

2.2.5.1 DNA Restriction Digests

DNA restriction was performed by incubation for 1-2 h with the appropriate restriction enzyme, buffer, and when required, BSA according to the manufacturer's instructions.

2.2.5.2 DNA Ligations

Ligations were performed using 100 units of T4 DNA ligase (NEB, Hitchin, UK) in the manufacturer's supplied buffer. Ligations were performed at 16°C for 12 h.

2.2.5.3 Polymerase Chain reaction

PCR reaction mixtures were prepared as follows:

Primers	1 μl of 100 ng/μl stock
DNA template	0.5 – 5 μl (depending on source: plasmid, gDNA, cell)
dNTPs	1 μl of 10mM dATP, dCTP, dGTP, dTTP mixture
10 × Buffer	2.5 μl
Mg^{2+}	0 – 4 mM ($MgCl_2$ for Taq, $MgSO_4$ for Vent)
Taq or Vent Polymerase	0.5 units
Water	to 25 μl

Vent polymerase was used for any DNA amplification required for cloning work, Taq polymerase was used only for PCR screening. Where cells were used as a source of DNA template, a colony was resuspended in 10 μl high-purity water (MilliQ) and incubated at 95°C for 15 minutes. A volume of 4 μl of this lysed cell suspension was added directly to the PCR reaction mixture as template. All PCR reactions were performed in either Techne Genius/Progene thermal cyclers or a MWG-Biotech Primus thermal cycler on the following program:

Initial incubation	96°C	5 minutes
30 cycles of:		
Denaturation	96°C	1 minute
Annealing	50-65°C	1 minute
Extension	72°C	1 minute per kb amplified sequence
Then:		
Final Incubation	72°C	10 minutes

2.2.5.4 Gel Purification of DNA

DNA bands were excised from the agarose gel using a scalpel. DNA was then extracted and purified using a Concert Rapid Gel Extraction System (Gibco BRL, Paisley, UK) according to the manufacturer's instructions. DNA was eluted in 30 µl of nuclease free water (Sigma, Poole, UK).

2.2.5.5 Purification of DNA

DNA amplified to single band purity by PCR and DNA from a restriction endonuclease digestion were purified of dNTPs, oligonucleotides and enzymes using Qiagen PCR Purification Kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. DNA was eluted in 30 μl TE buffer.

2.2.6 Agarose Gel Electrophoresis

2.2.6.1 Agarose electrophoresis of DNA

Gels containing 1-2% (w/v) agarose were prepared by adding solid agarose (Bioline) to 1 × TBE and heating to boiling point in a microwave oven. After cooling to approximately 45°C, ethidium bromide (Pharmacia Biotech, Buckinghamshire, UK) was added to a final concentration of 0.5 µg/ml and the gel poured into a casting tray containing well-forming combs and left to solidify at room temperature. DNA samples were loaded into wells in volumes up to 12 µl of which 20% (v/v) was DNA loading buffer. Electrophoresis was performed at a constant voltage of 200 V, 400 mA for 40 mins in 1 × TBE containing 0.5 µg/ml ethidium bromide. DNA bands were visualised by placing the gel on a UV transilluminator. Band sizes were estimated using 1kb ladder (Gibco BRL, Paisley, UK).

2.2.6.2 Agarose Electrophoresis of RNA

Agarose gel electrophoresis of RNA samples was performed as above except that prior to use all equipment was sterilised by soaking in 10% (v/v) hydrogen peroxide and rinsed in double autoclaved MilliQ water. The agarose gel and TBE buffer were autoclaved twice prior to use.

2.2.7 Lab-on-a-Chip Analysis of RNA

RNA to be used in microarray and macroarray analysis was analysed for purity and degradation using the Lab-on-a-chip system (Agilent, West Lothian, UK). This was carried out as a service by Miss Megan John at University of Wales College of Medicine, Cardiff, UK.

2.3 Genomic Methods

2.3.1 Transcriptional Profiling of *Escherichia coli* using Macroarrays

2.3.1.1 Chemicals and Reagents

RNase free DNase was obtained from Qiagen (Crawley, UK). M-MLV reverse transcriptase and M-MLV reverse transcriptase buffer were obtained from Promega (Southampton, UK). MicroSpin G-25 spin columns were obtained from Amersham Pharmacia (Buckinghamshire, UK). Panorama *E. coli* gene arrays and 3' ORF-specific oligonucleotides were obtained from Sigma-Genosys Biotechnologies (Poole, UK). [α^{33}P]dCTP (74-111 TBq/mmol) was from NEN Life Science Products (Beaconsfield, UK).

2.3.1.2 cDNA Synthesis and Labelling Conditions

Hybridisation probes were generated by standard cDNA synthesis. The protocol supplied by the manufacturer of the DNA arrays was suitable for achieving >60% incorporation of the ^{33}P-labelled nucleotide. Since it is not possible to purify bacterial mRNA from total RNA (i.e. by purification of polyadenylated mRNA as in eukaryotes), the labelling protocol takes into account the presence of rRNA and tRNA which constitute approximately 85% of the total RNA. The C-terminal primer set (4,290 ORF-specific C-terminal primers [Sigma-GenoSys, Poole,UK]) was used to generate the hybridisation probe in a standard first-strand cDNA synthesis. Briefly 1 μg of RNA was annealed to cDNA labelling primers (4 μl of stock solution), in a volume of 15 μl, heated to 90°C for 2 mins and then cooled to 42°C at a rate of 2°C min^{-1}.

To the annealed mRNA/primer-mixture; M-MLV reverse transcriptase (Promega, Southampton, UK), reverse transcriptase buffer, dATP, dGTP, dTTP and 0.74 MBq of [α-^{33}P]dCTP were added (total volume 30 μl) and this cDNA synthesis reaction mixture incubated at 42°C for 2 h (Table 2.6).

Table 2.6 Composition of RT-PCR reaction mixture.

Component	Stock Reagent	Volume per Reaction
RNA and *E. coli* Primer mix	-	15 µl
1x Reverse Transcriptase Buffer	5 ×	6 µl
333 µM dATP	10mM	1 µl
333 µM dGTP	10mM	1 µl
333 µM dTTP	10mM	1 µl
740 kBq [α-^{33}P] dCTP (74-111 TBq/mmol)	370 kBq/µl	2 µl
50 U M-MLV Reverse Transcriptase	25 U/µl	2 µl
Water	-	2 µl
	Final volume	30 µl

Unincorporated nucleotides were removed by gel filtration through a MicroSpin G-25 Sephadex column (Amersham-Pharmacia, Buckinghamshire, UK) according to the manufacturer's instructions.

2.3.1.3 DNA Macroarray Hybridisation

Each Panorama array contains 4,290 PCR-amplified ORFs from the *E. coli* K12 (MG1655) genome (Blattner *et al.*, 1997). All 4,290 ORFs are printed in duplicate at equal mass per spot onto positively charged, 12 cm × 24 cm, nylon membranes. The hybridisation and washing steps were carried out according to the manufacturer's instructions. Briefly, the arrays were prehybridised in hybridisation solution (5 × SSPE, 2% (w/v) SDS, 1 × Denhardt's Reagent, 100 µg/ml denatured herring sperm DNA) at 65°C for 1 h in a 30 cm by 3.5 cm roller bottle in a hybridisation oven. The entire cDNA probe, generated as described in Section 2.3.1.2, was added to 3 ml of hybridisation solution and the blot hybridised with this solution for 18 h at 65°C. The arrays were washed with washing solution (0.5 × SSPE, 0.2% (w/v) SDS) three times for 5 mins at room temperature and then three times for 20 minutes at 65°C. The arrays were partially air dried, wrapped in Saran Wrap and exposed to a 20 cm × 25 cm phosphoimager screen (Bio-Rad, Hemel Hempstead, UK) for 48 h. The image was captured on a Personal Imager FX (Bio-Rad, Hemel Hempstead, UK) using the PC based Quantity One software. Arrays were stripped for re-use by washing at 100°C with stripping solution as specified by the manufacturer.

2.3.2 Transcriptional Profiling of *Escherichia coli* using Microarrays

2.3.2.1 cDNA Synthesis and Labelling Conditions

Cy5 and Cy3 labelled hybridisation probes were generated using a CyScribe Post-labelling Kit (Amersham, Buckinghamshire, UK). This protocol has been developed as a two-step procedure. The first step involves the incorporation of amino allyl-dUTP (AA-dUTP) during cDNA synthesis. The second step involves chemically labelling the amino allyl modified cDNA using CyDye NHS-esters.

Briefly, 10 μg of RNA was annealed to random hexamers (1 μl of stock solution), in a volume of 12 μl, heated to 70°C for 5 mins and then cooled at room temperature for 10 minutes. To the annealed mRNA/primer-mixture; CyScript reverse transcriptase (1μl), reverse transcriptase buffer (4 μl), dNTPs (1 μl), 0.1 M DTT (2 μl), AA-dUTP (1 μl) were added (total volume 20 μl) and this reaction incubated at 42°C for 18 h. Degradation of mRNA was achieved by alkaline treatment. To each reaction, 2 μl 2.5 M NaOH was added and the mixture incubated at 37°C for 1 h, before the addition of 10 μl 2M HEPES (free acid form). Amino allyl modified cDNA was purified using a CyScribe GFX Purification kit (Amersham, Buckinghamshire, UK) according to the manufacturer's instructions.

Following purification, amino allyl cDNA (resuspended in 0.1 M sodium bicarbonate, pH 9.0) was added directly into one aliquot of CyDye NHS ester (Cy3 for control samples, Cy5 for test samples) and the mixture incubated in the dark, at room temperature for 4 h. Unreacted CyDye NHS-ester molecules were inactivated by adding 15 μl 4 M hydroxylamine and incubating for 15 minutes.
CyDye labelled cDNA was purified using a CyScribe GFX Purification Kit according to the manufacturer's instructions.

The absorbance of the purified labelled cDNA (60 μl) in 1 cm pathlength UV cuvettes (WPA, Cambridge, UK) was determined at 550 nm and 650 nm for control Cy3 and test Cy5 samples respectively (Ultrospec 3100*Pro*, Amersham, Buckinghamshire, UK). The amounts of Cy3 and Cy5 incorporated into cDNA were calculated using the equation:

$$\text{pmoles Cy3 or Cy5} = (A/E) \times (1/W) \times (Z) \times 10^{-6}$$

where: A = absorbance Cy3 at 550 nm or Cy5 at 650 nm, E = extinction coefficient for Cy3 or Cy5 (150 000 l mol^{-1} cm^{-1} at 550 nm Cy3 and 250 000 l mol^{-1} cm^{-1} at 650 nm for Cy5), W = path length in cm and Z = volume in microlitres.

The volumes required to provide 40 pmoles of either Cy3 or Cy5 were calculated and the corresponding Cy3 and Cy5 labelled cDNAs were combined in one tube in a 30 μl volume ready for hybridisation.

2.3.2.2 DNA Microarray hybridisation

The *E. coli* Microarray (EXGEN Project, Birmingham University) is described extensively in Section 7.1. Slides were pre-hybridised in 200 ml of prehybridisation solution (25% formamide, 5 × SSC, 10 mg/l BSA (Fraction V), 0.1% SDS) for 2 h at 42°C. Following pre-hybridisation, slides were washed briefly in water twice and then ethanol. Excess ethanol was shaken off and the slides dried by centrifugation. The hybridisation probe (Cy3/Cy5 cDNA probe, 25% formamide, 5 × SSC, 0.1% SDS, 8 μg Poly-A, 1 × Denhardts, total volume ~60 μl) was incubated at 95°C for 3 minutes and loaded onto the array area of a slide placed in a hybridisation chamber. A plastic coverslip (Hybrislip, Sigma, Poole, UK) was carefully placed over the array area, 10 μl water added to each well in the hybridisation chamber and the hybridisation chamber fully assembled. The chambers were placed flat in an airtight container lined with moist paper towels, the container covered with foil and incubated overnight at 42°C.

The hybridisation chambers were disassembled and the slides dipped into 200 ml prewarmed Wash Buffer I (2 × SSC, 0.1% SDS) in a black plastic trough until the coverslips fell off. After retrieval of the coverslips, the slides were washed for 2 minutes at 42°C with vigorous shaking. The slides were transferred into a black plastic trough containing 200 ml Wash Buffer II (0.2 × SSC) and washed at room temperature with vigorous shaking, before two final vigorous washes for 2 minutes in

Wash Buffer III (0.05 × SSC). Slides were dried by centrifugation and immediately scanned using a GenePi 400A Microarray Scanner (Axon Instruments).

2.3.3 Reduction of Experimental Variation in Macroarrays and Microarrays

To minimise variation in the cDNA reaction the same batch of primers was used and the triplicate cDNA reactions carried out simultaneously. The triplicate probes generated were hybridised to triplicate sets of arrays of the same batch number to minimise variation between individual arrays. In order to combat the problem of cross hybridisation, each ORF deemed to have its expression significantly altered was checked for potential cross-hybrids using a paralogue database (which can be downloaded at www.genome.edu/pub/expression/paratab.txt).

2.3.4 Identification of Significantly Expressed ORFs

Spot intensities on a given array were normalised by calculation of the intensity of each as a fraction of the total intensity of all spots taken together. In a standard analysis, the normalised intensities for each ORF in test and control arrays were compared and the induction ratio and fold increase/decrease calculated using Excel (Microsoft). Genes were considered to exhibit significantly changed expression if the log [induction ratio] was greater than 2 × standard deviation from the mean of the log [induction ratio] for all spots, in three separate experiments.
In a S.A.M. (Significance Analysis of Microarrays) analysis (see Section 3.3.2.2), normalised spot intensities were analysed using SAM Software Version 1.21, downloadable from http://www-stat.stanford.edu/~tibs/SAM (Tusher *et al.*, 2001).

2.4 Proteomic Methods

2.4.1 Isolation of *E. coli* Periplasmic proteins

E. coli W3110 was grown in 100 ml LB cultures, treated with PHMB (7.5 mg l^{-1}) when the D_{600} reached ~0.3 and 50 ml cells were harvested by centrifugation (4000 g, 20 mins, 4°C) when the D_{600} reached ~0.7. Cell pellets were resuspended in 20 ml of 30mM-Tris/HCl, 20% sucrose, pH 8.0. EDTA was added to a final concentration of 1 mM and then samples were incubated for 10 mins at room temperature with shaking. Cells were harvested by centrifugation (8000 g, 20 mins, 4°C), the supernatant was discarded and the pellet was resuspended in 1ml of ice-cold 5 mM $MgSO_4$ and shaken for 10 mins in an ice-bath. Cell debris was harvested by centrifugation (8000 g, 20 mins, 4°C), and the supernatant (referred to as the cold osmotic shock fluid) was transferred to a clean Eppendorf tube and stored at –80°C.

2.4.2 Determination of Protein Concentration

The protein content of samples was measured by the Bio-Rad protein assay (Bradford assay) based on the observation that the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm when binding to protein occurs. BSA was used as a standard protein. The dye reagent supplied by the manufacturer (Bio-Rad, Hemel Hempstead, UK) was five-fold diluted prior to use. For the calibration curve, 4 µl of 0, 0.2, 0.4, 0.8 and 1 mg/ml BSA standards were each mixed with 200 µl of dye reagent in a microtitre plate. After 5 min, the optical density (595 nm) was measured using a Thermomax Maxline plate reader. Suitably diluted protein samples (4 µl) were mixed with 200 µl of dye reagent and the absorbance determined as above.

2.4.3 2-Dimensional Gel Electrophoresis

2.4.3.1 Preparation of Protein Sample for 2-Dimensional Gel Electrophoresis

Cold osmotic shock fluid (100 μl) was cleaned of salts and surfactants using a PlusOne 2-D Clean Up Kit (Amersham-Pharmacia, Buckinghamshire, UK) according to the manufacturer's instructions. Protein pellets were resuspended in 100 µl rehydration solution (8 M urea, 2% (w/v) CHAPS, 2% (v/v) IPG Buffer, 40 mM DTT, 0.002 % (w/v) bromophenol blue).

2.4.3.2 First Dimension Isoelectric Focusing

Immobilised pH gradient isoelectric focusing (IPG-IEF) for 2-D electrophoresis was performed using the integrated IPGphor Isoelectric Focusing System (Amersham-Pharmacia, Buckinghamshire, UK). Protein samples (200 μg) were prepared for electrophoresis in 125 μl rehydration solution (containing 8 M urea, 2% (w/v) CHAPS, 0.01% (w/v) bromophenol blue, 0.28 % (w/v) DTT and 2% (v/v) IPG buffer; the last 2 components were added fresh prior to use). Each sample was placed in the IPG strip holder and the IPG strip (Immobiline DryStrip, pH 3-10 NL, 7 cm) positioned gel-side down such that the gel was coated in sample and air bubbles were absent. IPG cover fluid (0.8 ml) was pipetted over the strip and the strip holder replaced. The strip was allowed to rehydrate for 12 h at 20°C. Isoelectric focusing was performed on an IPGphor (BioRad, Hemel Hempstead, UK) at 50 mA per strip at constant temp of 20°C for a total of 1-15 kV hours. Focused IPG strips were stored at –70°C.

2.4.3.3 Second Dimension SDS-PAGE Electrophoresis

Prior to second dimension electrophoresis, IPG strips were equilibrated in SDS equilibration buffer 1, then buffer 2, for 15 min each. Buffer 1 contained 50 mM-Tris/HCl pH 8.8, 6 M urea, 30% (w/v) SDS, 0.01% (w/v) bromophenol blue and 0.1% (w/v) dithiothreitol. Buffer 2 was the same as buffer 1 except dithiothreitol was

replaced by 0.25% (w/v) iodoacetamide. After equilibration the IPG strips were blotted with moistened filter paper to remove excess equilibration solution.

SDS-PAGE was then performed using the Multiphor II unit. Electrophoresis temperature was maintained at 15°C via the Multitemp III unit. Anodic and cathodic buffer strips were positioned on a precast Excel Gel gradient SDS gel (8-18%) on the Multiphor II unit according to the manufacturer's instructions. IPG strips were placed, gel side down, on the SDS gel such that the IPG strip and cathodic buffer strip were parallel and 2-3 mm apart. Sample application pieces were placed on the SDS gel underneath the plastic tab formed by the overhanging gel support film at each end of the IPG strip. Protein markers, when used, were pipetted (up to 10 µl) onto extra application pieces placed on the gel surface just beyond the IPG strip.

SDS PAGE was carried out in 2 steps. The first, low current, step allowed migration of sample from IPG strip to SDS gel and was performed at 20 mA for 30 min. After the bromophenol blue dye front had moved away from the IPG strip by 2 mm, the IPG strip and application pieces were removed. The second, higher current phase, separated proteins on the SDS gel at 50 mA for 70 min.

2.4.3.4 Visualisation of Proteins

Gels were rinsed for 1 hour in fixing solution (40% ethanol, 10% acetic acid) before being stained using Coomassie R-350 dye (Amersham, Buckinghamshire, UK) for 10 minutes. Following staining, gels were rinsed briefly in MilliQ and destained in destaining solution (25% ethanol, 8% acetic acid) until the background of the gels became clear. For short term storage, gels were stored in 10% v/v methanol and for longer term storage, gels were stored in water/glycerol/methanol (8:1:1 by volume) at 4°C. Gel images were captured using a flat bed scanner.

2.4.3.5 Protein Identification

Protein spots of interest were excised from gels using a scalpel blade. Protein identification by in-gel tryptic digestion followed by mass fragmentation and mass

fingerprinting were carried out as a service by Dr Len Packman at the Cambridge Protein and Nucleic Acid Chemistry Facility, University of Cambridge.

2.5 Minimum Inhibitory Concentration Assay

E. coli strains were assessed for tolerance of PHMB by growth in the presence of increasing concentration of PHMB in minimum inhibitory concentration (MIC) assays. *E. coli* strains were grown in LB broth (5ml) for 18 h (containing 0.1mM IPTG and chl where appropriate). These cultures were used to inoculate fresh LB broth (containing 0.1mM IPTG and Chl where appropriate) to a D_{600} of precisely 0.100 (Genequant Pro, Amersham, Buckinghamshire, UK). This culture was dispensed (145 μl) into 96-well micro-titre plates containing PHMB (5 μl) at varying concentrations. Plates were incubated at 37°C, 200 rpm for 48 h and growth determined by monitoring D_{600}.

2.6 PHMB-Nucleic Acid Interaction Assays

For PHMB-nucleic acid interaction studies tRNA, double stranded DNA and single stranded DNA were used. A 100-bp sequence of DNA was randomly selected from the *Escherichia coli* K12 MG1655 genome. The sequence was part of the *yfiA* gene, from bases 3001 to 3100 of *Escherichia coli* MG1655, section 236 of the 400 in the complete genome (Blattner *et al.*, 1997). Single stranded (ss) and double stranded (ds) versions were purchased from Sigma (Poole, UK). A Hind III digest of lambda DNA marketed as a molecular weight marker (125 -23130 bp, (Promega, Southampton, UK) was used as a mixed-molecular weight sample of dsDNA. Yeast tRNA, from Sigma-Aldrich, was a mixture of approximately 250 different tRNAs ranging in size from 71 to 106 bases, mainly 72-74 bases long (70%), with the mode at 72 bases.

2.6.1 PHMB-Nucleotide Precipitation Assay

Aliquots of (100-150 µl) of stock aqueous solution of a nucleic acid (either dsDNA, ssDNA, Hind III-digested lambda DNA or yeast tRNA, containing the same concentration of nucleic acid expressed as nucleotides, typically 0.3 mM) were mixed with appropriate volumes (0-50 µl) of stock PHMB solutions to produce 150 µl final volumes containing 0.25 mM nucleotide and incremental concentrations of PHMB up to 0.3 mM biguanide units. After mixing for 10 s at 20°C , tubes were centrifuged at 20,000g for 25 min to sediment precipitated material. Supernatants were diluted with 0.35 ml of water and UV absorbances at 236 and 260 nm determined in 1 cm quartz cuvettes using a Hewlett Packard diode array spectrophotometer.

2.6.2 PHMB-Nucleotide Interaction Assay

Aliquots (3 µl) of either DNA or tRNA stock solution (1.67 mM nucleotides) were added to 2ml of fluorescent-PHMB (25 µM in biguanide units) in 3-ml quartz fluorimetry cuvettes (Hellman) with constant stirring at 30°C and after each addition, fluorescence measurements were made in a Cary Eclipse fluorimeter (Varian, Walton-on-Thames, UK). Excitation was for 0.5 s at 455 nm through a 20 nm slit in the vertical plane. Emissions were measured at 535 nm in the vertical and horizontal planes. For the two controls , the procedure was repeated except that either (a) aliquots of the stock nucleic acid solutions were replaced by the same aliquots of water, or (b) PHMB solution was replaced with a solution of the parent fluor, 1,8-naphthalimide, at 0.25 µM (equivalent to the total fluor concentration in the fluorescent-PHMB solution).

Chapter 3: Transcriptional analysis of the *E. coli* response to PHMB stress

3.1 Introduction

Gene array technology has previously been used successfully to study the global response of *E. coli* to many environmental stresses including UV exposure, hydrogen peroxide exposure, antibiotics, heavy metals, heat shock, cold shock, starvation, pH and osmotic stress (Courcelle *et al.*, 2001; Zheng *et al.*, 2001; Phadtree *et al.*, 2002; Brocklehurst and Morby,1999; Tao *et al.*, 1999; Tucker *et al.*, 2002; Weber and Jung, 2002). This chapter describes the transcriptional profiling of *E. coli* in order to identify genes differentially expressed in response to PHMB stress. To identify these genes, the gene expression profiles of exponentially growing *E. coli* W3110 during entry into, and recovery from, a PHMB induced period of stress were analysed.

3.2 Methods

Transcriptional profiling of *E. coli* using macroarrays was performed as described in Section 2.3.1. Briefly, RNA was isolated (Section 2.2.4.5), cDNA generated (Section 2.3.1.2) and hybridised to Panorama *E. coli* gene Arrays (Sigma-Genosys, Poole, UK) (Section 2.3.1.3). Spot intensity was determined (Section 2.3.1.3), the data was normalised and significantly altered ORFs were identified (Section 2.3.4).

3.3 Results

3.3.1 Effect of PHMB on the growth of *E. coli*

Bacterial cultures (100ml) were grown to early exponential phase (D_{600} ~0.25) in LB medium. PHMB was added at varying concentrations and the growth was monitored by optical attenuance. Figure 3.1 shows that, at all concentrations tested, PHMB altered the growth characteristics of *E. coli*. Concentrations ≥10 mg l^{-1} were bactericidal. The addition of 7.5 mg l^{-1} PHMB at D_{600} ~0.3 caused a temporary cessation in growth lasting approximately 4-5 h after which growth resumed. The effect of the addition of 7.5 mg l^{-1} PHMB to rapidly growing early exponential *E. coli* cultures was reproducible (See Figure 3.2) and this system was used as the basis for further experiments.

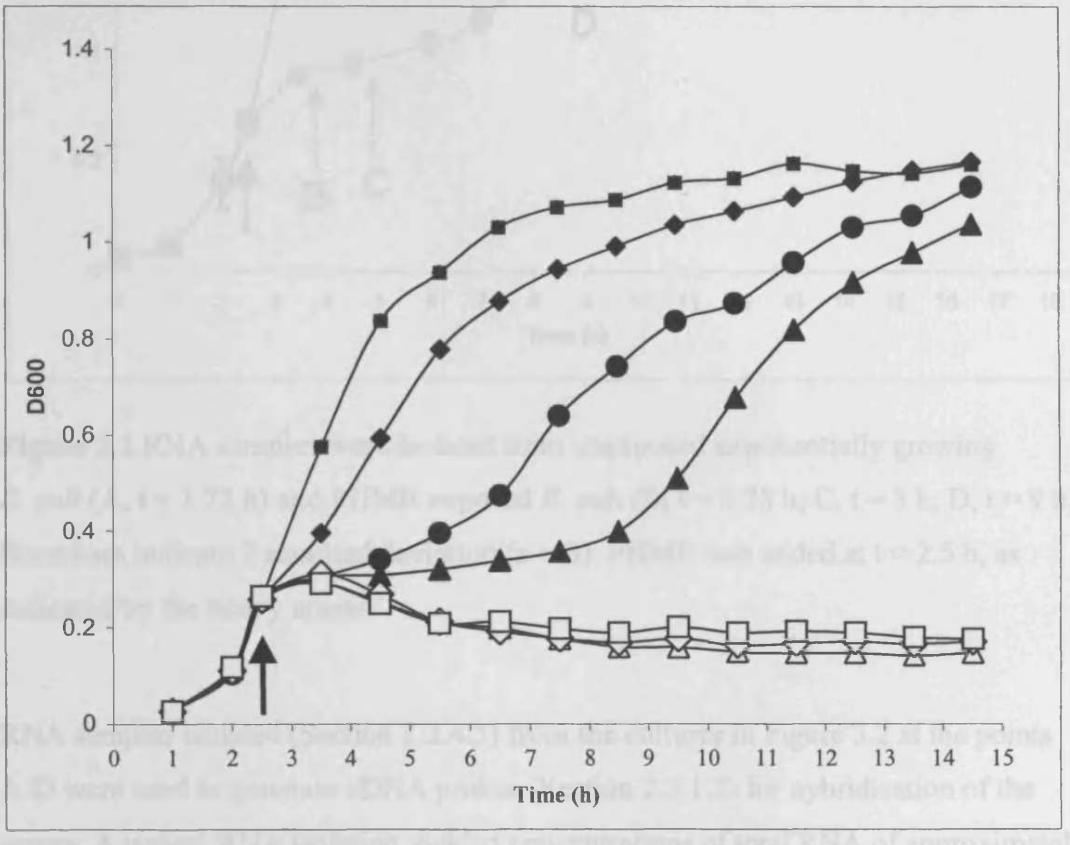


Figure 3.1 Effect of addition of PHMB on growth of *E. coli*. PHMB was added at early exponential phase (2.5 h) as indicated by the arrow at: filled squares, 0 mg/l; filled diamonds, 2.5 mg/l; filled circles, 5 mg/l; filled triangles, 7.5 mg/l; open triangles, 10 mg/l; open diamonds, 12.5 mg/l; open squares, 15 mg/l.

3.3.2 DNA Macroarray-based global transcription profiling of *E. coli* response to PHMB treatment

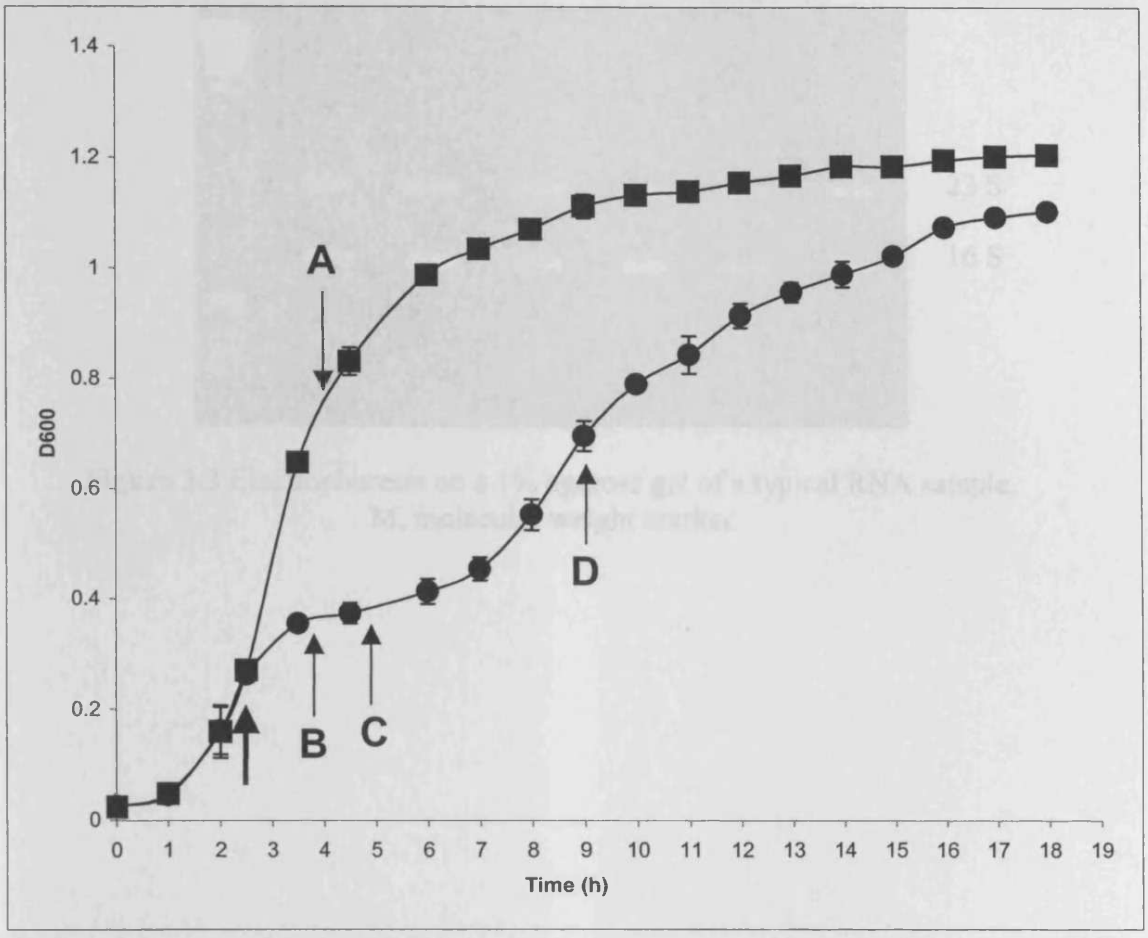


Figure 3.2 RNA samples were isolated from unexposed exponentially growing *E. coli* (A, t = 3.75 h) and PHMB exposed *E. coli* (B, t = 3.75 h; C, t = 5 h; D, t = 9 h). Error bars indicate 2 standard deviation (n = 3). PHMB was added at t = 2.5 h, as indicated by the heavy arrow.

RNA samples isolated (Section 2.2.4.5) from the cultures in Figure 3.2 at the points A-D were used to generate cDNA probes (Section 2.3.1.2) for hybridisation of the arrays. A typical RNA isolation yielded concentrations of total RNA of approximately 1 ng μl^{-1}. Samples separated on a 1% agarose gel showed the 23S and 16S ribosomal RNA bands clearly (Figure 3.3). The absence of high molecular-weight material shows that the RNA has not been contaminated with genomic DNA. A typical pair of arrays is shown in Figure 3.4

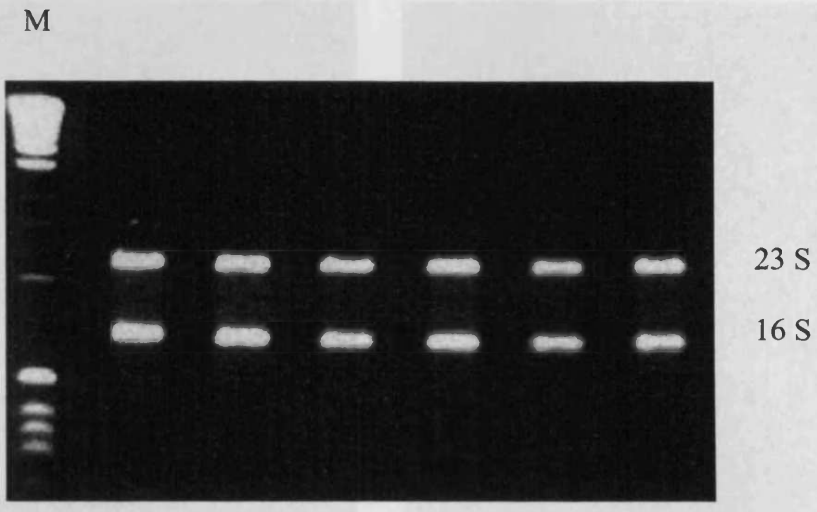


Figure 3.3 Electrophoresis on a 1% agarose gel of a typical RNA sample, M, molecular weight marker.

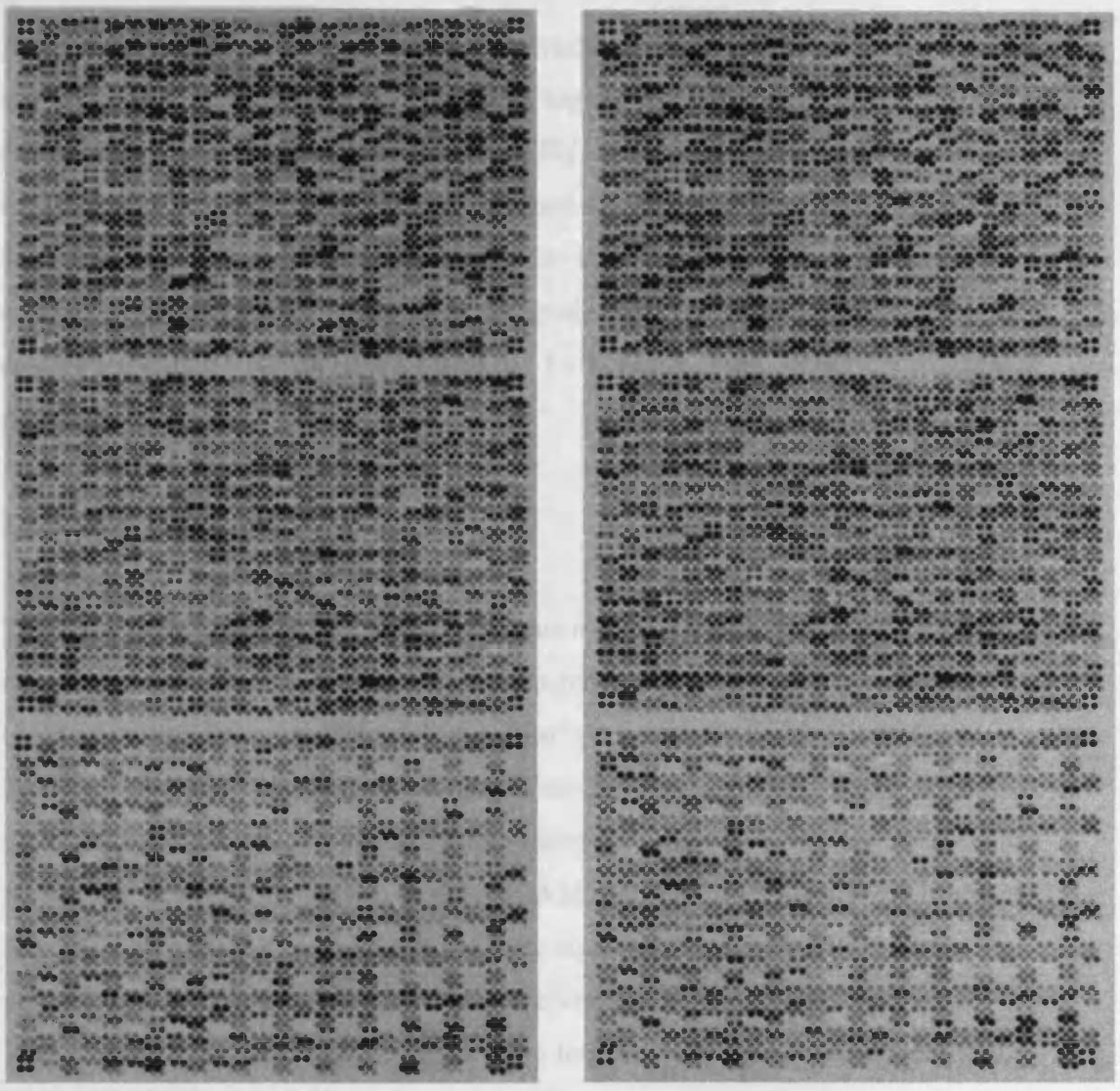


Figure 3.4 DNA arrays of the entire set of *E. coli* genes hybridised with probes generated from RNA extracted from cells growing at points A (left panel) and D (right panel) in Figure 3.2.

The intensity of each spot signal was quantified as described in Section 2.3.1.3. Two separate statistical analyses were performed on the normalised data: a standard analysis and a SAM analysis (see Section 2.3.4). Figure 3.5 shows in diagrammatic form, the five transcriptional profile comparisons that were performed in this experiment and are described in Section 3.3.5.

3.3.3 Standard Analysis

The standard analysis was performed as described in Section 2.3.4. Briefly, a gene was deemed to be significantly altered if its $\log_{10}$[IR] value was more than 2 standard deviations away from the mean of all $\log_{10}$[IR] values for the 4,290 ORFs in all three experiments. The full results from the standard analysis can be seen in Tables 3.1 to 3.10. In order to provide an estimate of the level of induction or repression, the average fold change values are shown alongside each gene deemed to significantly changed in expression (indicated with a '+' for induced genes and a '–' for repressed genes).

3.3.4 SAM Analysis

The array specific software SAM (Significance Analysis of Microarrays) was used for a second statistical analysis. SAM identifies genes with statistically significant changes in expression by assimilating a set of gene specific t tests (Tusher *et al.*, 2001). Each gene is assigned a score based on change in gene expression relative to the standard deviation of repeated measurements. For genes with scores greater than an adjustable threshold (known as delta), SAM then estimates the number of genes identified by chance (referred to as the false significant number, FSN) and the percentage of genes identified by chance (referred to as the false discovery rate, FDR). By altering the threshold smaller and larger sets of genes can be identified, with the FDR being calculated for each set. For the purposes of this study, delta values were chosen which created the largest set of genes possible which had a FSN of less than 1 i.e. in each set of genes identified as being significantly altered there is likely to be, at most, one false positive. The full results from this analysis can be seen in Appendix A.

3.3.5 Summary of Transcriptional Profile Comparisons

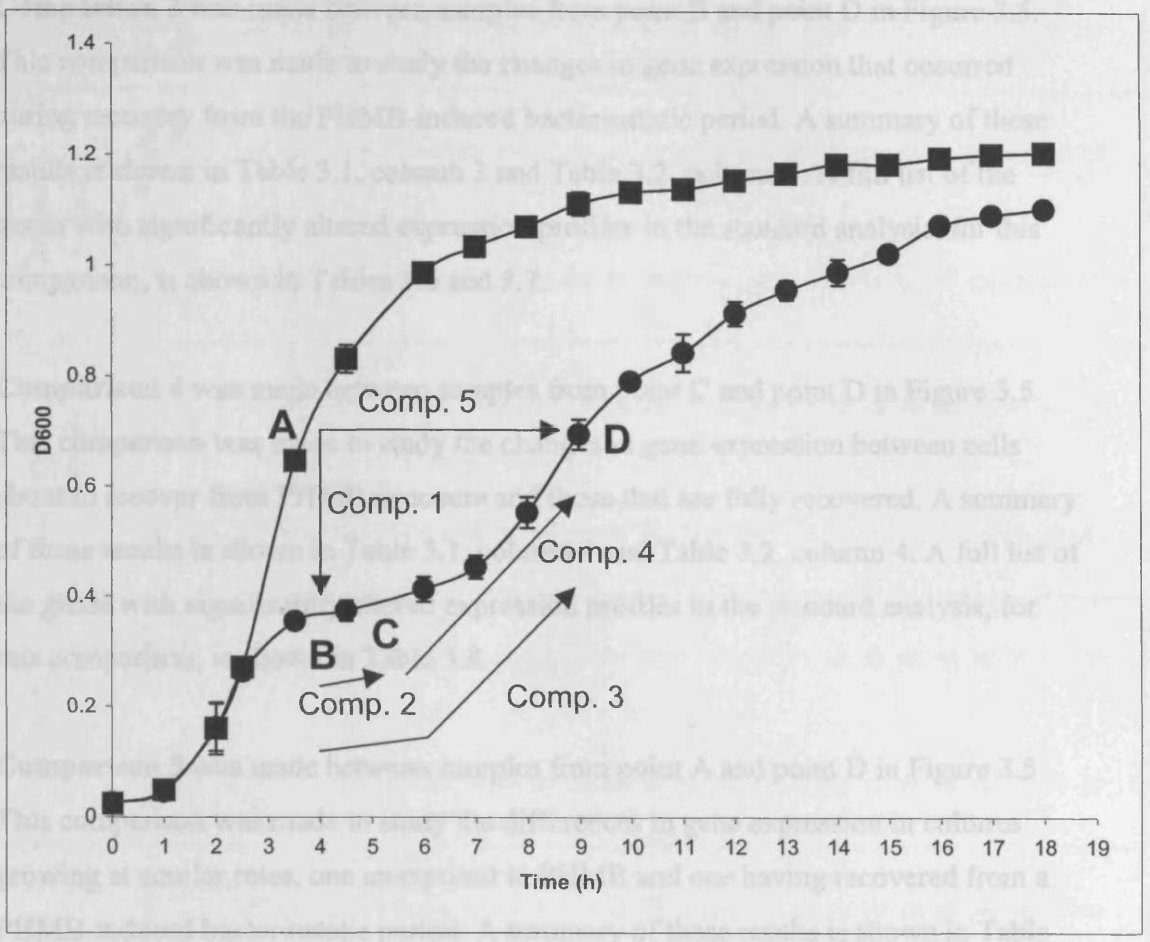


Figure 3.5 Diagrammatic view of the five expression-profile comparisons that were performed.

Comparison 1 was made between samples from point A and point B in Figure 3.5. This comparison was made to study the initial response to PHMB exposure. A summary of the results of this comparison is shown in Table 3.1, column 1 and Table 3.2, column 1. A full list of the genes with significantly altered expression profiles in the standard analysis, for this comparison, is shown in Tables 3.3 and 3.4.

Comparison 2 was made between samples from point B and point C in Figure 3.5. This comparison was made to study the changes in gene expression during the PHMB-induced bacteriostatic period. A summary of these results is shown in Table 3.1, column 2 and in Table 3.2, column 2. The single gene with significantly altered expression profile in the standard analysis, for this comparison, is shown in Table 3.5.

Comparison 3 was made between samples from point B and point D in Figure 3.5. This comparison was made to study the changes in gene expression that occurred during recovery from the PHMB-induced bacteriostatic period. A summary of these results is shown in Table 3.1, column 3 and Table 3.2, column 3. A full list of the genes with significantly altered expression profiles in the standard analysis, for this comparison, is shown in Tables 3.6 and 3.7.

Comparison 4 was made between samples from point C and point D in Figure 3.5. This comparison was made to study the changes in gene expression between cells about to recover from PHMB exposure and those that are fully recovered. A summary of these results is shown in Table 3.1, column 4 and Table 3.2, column 4. A full list of the genes with significantly altered expression profiles in the standard analysis, for this comparison, is shown in Table 3.8.

Comparison 5 was made between samples from point A and point D in Figure 3.5. This comparison was made to study the differences in gene expression in cultures growing at similar rates, one unexposed to PHMB and one having recovered from a PHMB-induced bacteriostatic period. A summary of these results is shown in Table 3.1, column 5 and Table 3.2, column 5. A full list of the genes with significantly altered expression profiles in the standard analysis, for this comparison is shown in Tables 3.9 and 3.10.

The full lists of the genes identified using SAM analysis for Comparisons 1,2,3,4 and 5 and the corresponding delta values and FDR values can be seen in Appendix A in Tables A.2, A.4, A.6, A.8 and A.10 respectively.

Table 3.1 Summary of the data from the standard analysis of the five expression profile comparisons shown in Figure 3.5. Genes are classified according to the initial classification of all 4,290 ORFS made by Blattner *et al.*, 1997

Functional Group		Number of genes with significantly altered transcript levels									
		Column1 A vs B		Column 2 B vs C		Column 3 B vs D		Column 4 C vs D		Column 5 A vs D	
	Total	Higher in A	Higher in B	Higher in B	Higher in C	Higher in B	Higher in D	Higher in C	Higher in D	Higher in A	Higher in D
Whole genome (*E. coli* K12)	4,290	**12**	**71**	**1**	**0**	**55**	**9**	**11**	**3**	**25**	**12**
Amino acid biosynthesis	131	**2**	**2**	0	0	**1**	**3**	0	**1**	**2**	**1**
Biosynthesis of cofactors, prosthetic groups and carriers	103	0	0	0	0	0	0	0	0	0	0
Carbon compound metabolism	130	0	0	0	0	0	0	0	0	0	0
Cell processes	188	**3**	**3**	0	0	**2**	**1**	0	**1**	**1**	0
Cell structure	224	0	**5**	0	0	**4**	0	0	0	**5**	**1**
Central intermediary metabolism	188	**1**	0	0	0	0	**1**	0	0	**1**	0
DNA replication, repair, restriction/ modification	115	0	**3**	0	0	**1**	0	**1**	0	0	**2**
Energy metabolism	242	**1**	**1**	0	0	**1**	**1**	0	0	**2**	0
Fatty acid and phospholipid metabolism	48	0	0	0	0	0	0	0	0	0	0
Hypothetical, unclassified, unknown	1,636	**2**	**40**	**1**	0	**34**	**1**	**8**	**1**	**5**	**8**
Membrane	13	0	**2**	0	0	0	0	0	0	0	0
Nucleotide biosynthesis and metabolism	58	0	0	0	0	0	0	0	0	0	0
Phage, transposon, or plasmid	87	0	**2**	0	0	1	0	**1**	0	0	0
Putative enzymes	251	0	0	0	0	0	0	0	0	**1**	0
Putative chaperones	9	0	**1**	0	0	0	0	0	0	0	0
Putative regulatory proteins	133	0	**4**	0	0	**5**	0	**1**	0	0	0
Putative transport proteins	146	0	0	0	0	0	**1**	0	0	0	0
Regulatory function	45	0	0	0	0	0	0	0	0	0	0
Transcription, RNA processing, and degradation	55	0	**3**	0	0	**1**	0	0	0	0	0
Translation and posttranslational modification	182	0	**5**	0	0	**5**	0	0	0	0	0
Transport and binding proteins	281	**3**	0	0	0	0	**1**	0	0	**8**	0
Others	25	0	0	0	0	0	0	0	0	0	0

Table 3.2 Summary of the data from the standard analysis of the five expression profile comparisons shown in Figure 3.5. Genes are classified according to their present classification at www.geneprotec.mbl.edu

Functional Group		Number of genes with significantly altered transcript levels									
		Column 1 A vs B		Column 2 B vs C		Column 3 B vs D		Column 4 C vs D		Column 5 A vs D	
	Total	Higher in A	Higher in B	Higher in B	Higher in C	Higher in B	Higher in D	Higher in C	Higher in D	Higher in A	Higher in D
Known *E. coli* genes	4,671	**12**	**71**	**1**	**0**	**55**	**9**	**11**	**3**	**25**	**12**
Enzyme	1086	**3**	**6**	0	0	**3**	**4**	0	**1**	**5**	0
Putative enzyme	535	0	**2**	0	0	**2**	0	**1**	0	**1**	0
Structure	91	**2**	**2**	0	0	**2**	**1**	0	**1**	**3**	0
Putative structure	36	0	**1**	0	0	**1**	0	0	0	**1**	0
Regulator	238	0	**7**	0	0	**4**	0	**2**	0	0	**3**
Putative regulator	154	0	**5**	0	0	**4**	0	0	0	**1**	**1**
Transport	399	**3**	0	0	0	0	**2**	0	0	**8**	0
Putative transport	332	0	**1**	0	0	**1**	0	0	0	0	0
Factor	118	0	**1**	0	0	0	0	0	0	0	**1**
Putative factor	30	0	**2**	0	0	0	0	0	0	0	0
Membrane	55	0	**2**	0	0	**1**	0	0	0	0	0
Putative membrane	210	0	**2**	0	0	0	0	0	0	0	0
Phenotype only	97	**1**	**2**	0	0	**1**	0	0	0	0	**1**
Leader	12	**1**	0	0	0	0	**1**	0	0	0	0
IS, Phage, Tn	318	0	**9**	**1**	0	**10**	0	**3**	0	0	**1**
RNA	114	0	0	0	0	0	0	0	0	0	0
Carrier	31	0	0	0	0	0	0	0	0	0	0
ORF	815	**2**	**29**	0	0	**26**	**1**	**5**	**1**	**6**	**5**

Table 3.3 Genes of known function found to have significantly altered expression profiles in a standard analysis of Comparison 1 (PHMB amended compared with PHMB unamended at same culture age).

Gene	Blattner number	Gene product description	Fold change
Nucleic acid associated			
stpA	b2669	DNA-binding protein StpA	+9.1
recA	b2699	DNA strand exchange and recombination protein	+6.6
dnaK	b0014	Chaperone-heat shock protein 70	+7.4
xseA	b2509	Exodeoxyribonuclease large subunit	+6.8
mcrA	b1159	5-Methylcytosine-specific restriction enzyme A	+21.9
evgS	b2370	Sensory histidine kinase regulating multidrug resistance	+6.9
Translation			
rplY	b2185	50S Ribosomal protein L25	+9.8
hha	b0460	Haemolysin expression modulator	+16.0
rpsP	b2609	30S Ribosomal subunit protein S16	+14.0
miaA	b4171	tRNA Δ-2-Isopentenylpyrophosphate (IPP) transferase	+7.7
ybcM	b0546	DLP12 prophage, putative transcriptional regulator	+17.2
Amino acid metabolism			
tnaL	b3707	Tryptophanase leader peptide	-33.9
tnaA	b3708	Tryptophan deaminase	-29.3
tdcR	b3119	tdcABC Operon (threonine dehydratase) transcriptional activator	+7.1
cysB	b1275	Transcriptional regulator of cysteine biosynthesis and regulator of sulphur assimilation	+8.2
cbl	b1987	Transcriptional regulator of cysteine biosynthesis	+12.2
aspA	b4139	Aspartate ammonia-lyase (aspartase)	-21.9
Energy metabolism			
pflB	b0903	Pyruvate formate lyase I, induced anaerobically	-18.5
glpD	b3426	Glycerol-3-phosphate dehydrogenase (aerobic)	+9.6

Table 3.3 Continued

Gene	Blattner number	Gene product description	Fold change
Transport and binding			
lamB	b4036	High affinity receptor for maltose and maltoseoligosaccharides, phage lambda receptor	-10.1
rbsD	b3748	Membrane associated component of high affinity ribose transport system	-12.1
gatB	b2093	Phosphotransferase system, galactitol-specific IIB component	-19.0
Fimbriae and flagella			
flgE	b1076	Flagella hook protein FlgE	-9.8
flgJ	b1081	Flagella protein FlgJ	-11.3
yadC	b0135	Putative fimbrial-like protein	+31.9
yehC	b2110	Putative periplasmic fimbrial chaperone	+7.2
Surface and outer membrane associated			
osmB	b1283	Osmotically inducible lipoprotein B precursor	+7.4
vacJ	b2346	VacJ lipoprotein precursor	+15.3
rfaL	b3622	O-Antigen ligase	+11.6
yefI	b2032	Putative transferase	+26.0
rfc	b2035	O-Antigen polymerase	+5.6
rfbX	b2037	Putative O-antigen transporter	+16.9
Others			
uspA	b3495	Universal stress protein A	-37.2
intB	b4271	Prophage P4 integrase	+5.5
cpxP	b3913	Periplasmic repressor of Cpx regulon	+17.5
b3914	b3914	Putative periplasmic protein	+26.1

Table 3.4 Genes of unknown function found to have significantly altered expression profiles in a standard analysis of Comparison 1 (PHMB amended compared with PHMB unamended at same culture age).

Gene	Blattner number	Gene product description	Fold change
rhs associated			
rhsD	b0497	RhsD protein precursor	+40.0
b0499	b0499	Conserved protein	+8.9
ybbD	b0500	Conserved hypothetical protein	+13.9
b0501	b0501	Unknown CDS	+6.6
rhsE	b1456	RhsE protein	+16.1
ydcD	b1457	Unknown CDS	+14.0
ybfD	b0706	H-repeat associated protein	+9.3
rhsB	b3482	RhsB core protein with unique extension	+14.3
yhhH	b3483	Unknown CDS	+19.8
yhiJ	b3488	Conserved hypothetical protein	+69.4
yhiK	b3489	Hypothetical protein	+11.6
yhiL	b3490	Hypothetical protein	+27.4
yibJ	b3595	Putative rhs protein	+21.5
Other unknowns			
hdeB	b3509	Protein HDEB precursor	-14.8
yeaC	b1777	Conserved hypothetical protein	-6.4
yaiN	b0357	Conserved hypothetical protein	+41.0
	b1202	Putative membrane protein	+5.2
ydjF	b1770	Putative transcriptional regulator	+8.1
ygiG	b3046	Putative outer membrane usher protein	+5.5
	b0299	Putative IS transposase	+5.9
ymgD	b1171	Unknown CDS	+15.2
	b1172	Conserved hypothetical protein	+22.5
yhiW	b3515	Putative transcriptional regulator	+6.6
yhiX	b3516	Putative transcriptional regulator	+19.0
yahA	b0315	Putative transcriptional repressor	+12.7
ycgW	b1160	Conserved hypothetical protein	+39.6
b1228	b1228	Unknown CDS	+44.2
ydhA	b1639	Conserved hypothetical protein	+10.0
ychF	b1203	Putative GTP binding protein	+8.4
yefG	b2034	Unknown CDS	+23.1
ybaJ	b0461	Conserved hypothetical protein	+7.6
htrL	b3618	Lipopolysaccharide biosynthesis	+8.6
YiiG	b3896	Conserved protein	+10.0
ydjO	b1730	Putative enzyme	+6.7
yfjW	b2642	CP4-57 prophage	+13.9
b1721	b1721	Putative regulator	+8.4
yhaB	b3120	Conserved protein	+48.8
b2854	b2854	Conserved protein, lysozyme like	+19.7

Table 3.4 Continued

Gene	Blattner number	Gene product description	Fold change
yjbM	b4048	Conserved hypothetical protein	+8.4
yebG	b1848	DNA damage inducible gene in SOS regulon, dependent on cAMP, H-NS	+13.8
yjcF	b4066	Conserved protein	+10.2
b1527	b1527	Conserved protein	+6.0
yedM	b1935	Unknown CDS	+5.5
yrhB	b3446	Unknown CDS	+6.8
b2863	b2863	Unknown CDS	+37.5
yeeN	b1983	Conserved protein	+9.9
b1963	b1963	Unknown CDS	+13.2

Table 3.5 Genes significantly altered in a standard analysis of Comparison 2 (onset of bacteriostatic compared with end of bacteriostatic).

Gene	Blattner number	Gene product description	Fold change
flxA	b1566	Qin prophage	-6.9

Table 3.6 Genes of known function significantly altered in a standard analysis of Comparison 3 (onset of bacteriostatic compared with fully recovered cells).

Gene	Blattner number	Gene product description	Fold change
O-antigen associated			
yefI	b2032	Putative transferase	-14.2
yefG	b2034	Unknown CDS	-10.4
rfc	b2035	O-Antigen polymerase	-5.8
rfbX	b2037	Putative O-antigen transporter	-8.8
Amino acid metabolism			
cbl	b1987	Transcriptional regulator of cysteine biosynthesis	-7.6
cysK	b2414	Subunit of cysteine synthase A and O-acetylserine sulph-hydrolase	+7.7
tnaL	b3707	Tryptophanase leader peptide	+7.7
tnaA	b3708	Tryptophan deaminase	+9.1
tdcR	b3119	Transcriptional activator for threonine dehydratase	-4.9
yhaB	b3120	Conserved protein	-8.5
Ribosomal			
rpsU	b3065	30S ribosomal subunit protein S21	-4.8
rplY	b2185	50S ribosomal protein l25	-5.5
Various			
osmB	b1283	Osmotically inducible membrane lipoprotein B	-4.6
flgE	b1076	Flagellar hook protein FlgE	+6.0
hha	b0460	Haemolysin expression modulator	-8.6
ybaJ	b0461	Conserved hypothetical protein	-5.6
sucA	b0726	2-Oxoglutarate decarboxylase, 2-oxoglutarate dehydrogenase E1 component	+6.0
mcrA	b1159	5-Methylcytosine-specific restriction enzyme A	-8.9
oraA	b2698	Regulatory protein RecX	-4.4
rbsD	b3748	High affinity ribose membrane transport protein RbsD	+12.1
miaA	b4171	Δ-2-Isopentenylpyrophosphate (IPP) tRNA-adenosine transferase	-5.5
intB	b4271	KpLE2 phage-like element; P4-like integrase	-4.9
aceA	b4015	Isocitrate lyase	+7.3

Table 3.7 Genes of unknown function significantly altered in a standard analysis of Comparison 3 (onset of bacteriostatic compared with fully recovered cells).

Gene	Blattner number	Gene product description	Fold change
rhs associated			
ybfD	b0706	Conserved protein	-5.7
rhsE	b1456	RhsE protein	-7.3
ydcD	b1457	Unknown CDS	-6.4
rhsB	b3482	RhsB core protein	-7.4
yhhH	b3483	Unknown CDS	-8.9
yhiJ	b3488	Conserved hypothetical protein	-10.4
yhiK	b3489	Unknown CDS	-8.2
yhiL	b3490	Unknown CDS	-9.0
yibJ	b3595	Putative rhs protein	-6.8
rhsD	b0497	RhsD core protein	-9.8
ybbD	b0500	Conserved hypothetical protein	-7.9
Other unknowns			
b3914	b3914	Putative periplasmic protein	-6.8
ymgD	b1171	Unknown CDS	-5.5
b1172	b1172	Conserved hypothetical protein	-7.5
yadC	b0135	Putative fimbrial-like protein	-11.6
yahA	b0315	Putative transcriptional repressor	-4.4
yaiN	b0357	Conserved hypothetical protein	-5.6
ybcM	b0546	DLP12 prophage; Putative transcriptional regulator	-8.3
ybcV	b0558	DLP12 prophage; putative envelope protein	-5.1
ybdO	b0603	Putative transcriptional regulator	-5.0
ycdK	b1010	Conserved protein	7.2
ycgW	b1160	Conserved hypothetical protein	-7.2
b1228	b1228	Unknown CDS	-4.9
flxA	b1566	Qin prophage	-7.6
ydjF	b1770	Putative transcriptional regulator	-6.5
b1963	b1963	Unknown CDS	-6.3
yeeN	b1983	Conserved protein	-8.5
yeiR	b2173	Putative enzyme	-8.1
yfaE	b2236	Conserved hypothetical protein, 2Fe-2S ferredoxin related	-4.2
yfjW	b2642	CP4-57 prophage	-10.7
b2854	b2854	Conserved protein, lysozyme like	-8.1
b2863	b2863	Unknown CDS	-5.6
yhbC	b3170	Conserved hypothetical protein	-5.5
glpD	b3426	Aerobic glycerol-3-phosphate dehydrogenase	-10.1
yrhB	b3446	Unknown CDS	-5.1
yhiW	b3515	Putative transcriptional regulator	-4.8
htrL	b3618	Involved in Lipopolysaccharide biosynthesis	-4.8

Table 3.7 Continued

Gene	Blattner number	Gene product description	Fold change
yicJ	b3657	Putative glycoside Pentoside Hexauronide (GPH) Transporter	+6.7
yiiG	b3896	Conserved protein	-5.6
yjbM	b4048	Conserved hypothetical protein	-5.1
yjcF	b4066	Hypothetcial protein	-4.3

Table 3.8 Genes significantly altered in a standard analysis of Comparison 4 (End of bacteriostatic compared with fully recovered cells).

Gene	Blattner number	Gene product description	Fold change
Amino acid metabolism			
cysK	b2414	Subunit of cysteine synthase A and O-acetylserine sulfhydrase A	+9.1
cbl	b1987	Transcriptional regulator of cysteine biosynthesis	-5.9
Flagellar			
flgE	b1076	Flagellar hook protein FlgE	+3.6
DNA associated			
hns	b1237	Transcriptional regulator, DNA-binding protein HLP-II, increases DNA thermal stability,	-6.2
tra5_2	b0541	DLP12 prophage, putative transposase for insertion sequence IS3	-4.4
rhs associated			
rhsB	b3482	RhsB core protein with unique extension	-6.0
yhiJ	b3488	Conserved hypothetical protein	-6.1
yhhH	b3483	Unknown CDS	-4.2
Other unknowns			
ycdK	b1010	Conserved protein	+6.2
ybcQ	b0551	DLP12 prophage; putative antitermination protein Q	-6.7
yefI	b2032	Putative transferase	-6.9
yhaB	b3120	Conserved protein	-6.1
yhhZ	b3442	Conserved protein	-4.7
b2863	b2863	Unknown CDS	-4.6

Table 3.9 Genes of known function significantly altered in a standard analysis of Comparison 5 (PHMB unexposed compared with fully recovered PHMB exposed).

Gene	Blattner number	Gene product description	Fold change
Flagellar associated			
fliC	b1923	Flagellar biosythesis, flagellin	-3.8
fliD	b1924	Flagellar biosynthesis	-3.7
flgJ	b1081	Flagellar biosynthesis	-2.9
Transport and binding			
lldP	b3603	l-Lactate permease	-6.0
gatA	b2094	Phosphotransferase system, galactitol-specific IIA component	-13.8
gatB	b2093	Phosphotransferase system, galactitol-specific IIB component	-11.9
gatC	b2092	Phosphotransferase system, galactitol specific IIC component	-8.6
malK	b4035	Maltose transport, repressor of mal operon	-3.7
lamB	b4036	Maltoporin, high affinity maltose and maltoseoligosaccharide, phage lambda receptor	-6.9
manX	b1817	Phosphotransferase enzyme II, AB component, mannose specific	-3.4
mglB	b2150	Galactose transport protein, periplasmic binding	-8.7
Other known			
cadA	b4131	Lysine decarboxylase	-3.5
tnaA	b3708	Tryptophan deaminase	-3.2
aspA	b4139	Aspartate ammonia-lyase (aspartase)	-6.4
fbaA	b2925	fructose 1,6-bisphosphate aldolase	-3.0
glgS	b3049	glycogen biosynthesis, rpoS dependent	-5.3
pflB	b0903	Pyruvate formate lyase 1, induced anaerobically	-3.8
cysB	b1275	Transcriptional regulator of cysteine biosynthesis and sulphur assimilation	+3.7
stpA	b2669	DNA-binding protein with chaperone activity	+3.4
fis	b3261	DNA binding protein for site-specific recombination and inversion, transcription of rRNA and tRNA operons, and DNA replication	+3.3
cpxP	b3913	Periplasmic repressor of cpx regulon	+5.3
b3914	b3914	Putative periplasmic protein	+3.9

Table 3.10 Genes of unknown function significantly altered in a standard analysis of Comparison 5 (PHMB unexposed compared with fully recovered PHMB exposed).

Gene	Blattner number	Gene product description	Fold change
yaiN	b0357	Conserved hypothetical protein	+7.0
yahA	b0315	Putative transcriptional regulator	+2.9
yebG	b1848	DNA damage inducible gene of SOS regulon	+4.2
rhsD	b0497	RhsD protein precursor	+4.1
ycgW	b1160	Conserved hypothetical protein	+5.5
b2863	b2863	Unknown CDS	+6.5
b1228	b1228	Unknown CDS	+9.9
yjdA	b4109	Conserved hypothetical protein	-3.1
b1502	b1502	Putative adhesin, similar to FimH	-3.6
ygeV	b2869	Putative transcriptional regulator	-4.5
yniA	b1725	Conserved protein, protein kinase-like	-2.7
yeeI	b1976	Conserved hypothetical protein	-7.4
hdeB	b3509	Conserved hypothetical protein	-27.6
hdeA	b3510	Conserved periplasmic protein	-8.5
yjfN	b4188	Conserved hypothetical protein	-3.7

```
RhsA            ----------------------------------------AAATATCAGTATGG--CC 16
RhsC            ------TTGCTGAGTTTTGAATACATTGAAAACCGCCAGCGGAAACGCTGGCGGTT--TT 52
RhsB            ----------------------------------------------------TTG--GT 5
RhsD            TCATAAATCATATGCGTTGAATGGATATTATCCATATAGTGAATTTGTTGATGATGAATT 60
RhsE            ---------------------------GACATTATGAATATTGGTTTGACTTTGCAT 30

RhsA            TGGTGATGGGAAAAAGGGCAGAA-AATGT-TGATGGTTTGTTAC-TTCCAAATAAATCAC 73
RhsC            TTGTATCAGGAAGATGTTCAGAGGAATAT-TTAGACATCGTTA--TACCAATCGAAACAT 109
RhsB            GAACGATCAACAAAGGGCCACTAGCACACCTGATTCCTCTCTAAATACCCTATCCGAACC 65
RhsD            CATCTGTGCTAAAAATGTTAGTTTAATAAAATATTGAAAGTGACCTGTAATAACAGTTGT 120
RhsE            ATCAGGTTTTTTATTACTCGCTGTAATGTACGA--GCCTGTCGTGAATAACGTCGAGAAC 88
                                          *      *

RhsA            ATATTTATCATGGTGATATAAATATTTTCCTAATTATTTCACTCTGATGGATATCTCACT 133
RhsC            ATATTCATGAAATATATATAAATATTTTCCTAATTGTTCTTATCTGACAGATATCTCACT 169
RhsB            TTCTCTTTTGTAACGTTCTAAATATATTCCTAAAAATCTTCAATTCATTGTGACCACA-- 123
RhsD            -TGTTGATTGAGAACAAATAAGTTTATGTGAAAAATATATAAATACATTAGCTGGTCTTG 179
RhsE            CCATCCGTTGCG--TTTGTGATTATTTTGTTGA---CTAAACAGACACCCGTTTCTCTGA 143
                   *   *           * * * * *     *              *         *

RhsA            TCAGGCTTTCTT-ATAAATCTGTAGGGTTTCGCCTGTCAGCAGACAAATAACCCGATAAA 192
RhsC            TAAGGCTTTCTT-ATAAATCTGTAGGGTTTCGCCTGTCAGCAGACAAATAACCCGATAAA 228
RhsB            --AGTTTTTCTTCGCTTTTTCGTATGAAGATAC-TGTCATTA---AAATAAT------AG 171
RhsD            TGTGTCATTTTATTTTTTTTTGT-TGCTAACAC-AGGGATATGAACAATAACT-----AA 232
RhsE            AGTAAAATCCCAGACTAAATCATCACATAAC-CATGACATTTTTCTGATATTCC------ 296
                       *               *          *  *  *        ***

RhsA            ACAAGGATGAG-AAATGAGCGGAAAACCGGC--------- 222
RhsC            ACAAGGATGAGCAGATGAGCGGAAAACCGGCGGCGCGTCA 268
RhsB            AAAAGGATTTTACGATGAGCGGAAAACCGG---------- 201
RhsD            AAGGGCACTTT-ATATGAGCGGAAAACCAGCGGCGCGTC- 270
RhsE            CCGGTAACGCC-AGATGT-CGACTCGCTTAACCACC---- 230
                          ***  **      *
```

Figure 3.6 Clustal W. sequence alignment of the rhs element promoters. * Denotes bases conserved in all five rhs element promoters.

3.4 Discussion

3.4.1 Statistical Issues

The purpose of gene arrays is to detect differentially expressed genes. However, there is no standard method for determining which genes are significantly altered in expression. Array analysis methodology appears to be performed on a basis of personal preference with no one method favoured over another. For example, even articles published in the same journal (in the following case The Journal of Bacteriology) over a short period of time (18 months), using arrays to monitor the expression profile of the same bacterium (*E. coli*), used a variety of statistical methods. Criteria for determining what is and is not significant include:-

- a basic four-fold induction (Zheng *et al.*, 2001, Phadtare *et al.*, 2002)
- a $\log_{10}$[IR] greater than 2.5 standard deviations from the mean $\log_{10}$[IR] (Pomposiello *et al.*, 2001)
- a SAM analysis (Weber *et al.*, 2002)
- a $\log_{10}$[IR] greater than 3 standard deviations from the mean $\log_{10}$[IR] **and** a students t test value P>0.05 (Tucker *et al.*, 2002).

It seems almost ironic that there is such a large variation in statistical methods for analysing such similar data. This variation in statistical methods employed, and the very fact that statistical analysis is needed, serve to highlight the limitations of gene array analysis. Array data are merely tools to steer the experimenter towards genes or clusters likely to be responding to the challenge imposed. They are never conclusive and the validation of array results through further biological characterisation is the most important aspect of any study involving gene arrays.

3.4.2 SAM Analysis

The lists of genes shown in Tables 3.3 to 3.10 refer only to the results of the standard statistical analysis. In this study, SAM typically identified around four times as many genes as the standard analysis. Comparisons 1 to 5 identified 83, 1, 64, 14 and 37 significant gene changes using a standard analysis, but a staggering 399, 9, 241, 37 and 18 changes in the corresponding SAM analysis. These results can be seen in Appendix A (Tables A.2, A.4, A.6, A.8 and A.10). This created a huge volume of data

for further biological characterisation. A comparison of the two methods showed a high proportion of overlap i.e. most genes in the standard analysis were deemed significantly altered in the SAM analysis (See Appendix B, Table B.1). However, the case is further complicated by closer inspection of the 'missing' genes from the standard analysis. For example, 62 of the 83 genes identified in the standard comparison 1 were found altered in the SAM comparison 1. Of the 21 genes 'missing' from the SAM analysis, 9 of them were found directly next to genes that SAM did find significantly altered e.g. *b1721* (a putative regulator) was deemed up-regulated in a standard analysis but not by SAM analysis. However, SAM considered the co-transcribed *b1720* to be up regulated (and also *b1722*, a downstream ORF found on the complementary strand). Therefore, due to the large volume of data generated by the SAM analysis and their large overlap with the data from the standard analysis, it was decided to concentrate only on the 'relatively few' genes identified by the standard analysis.

3.4.3 Analysis of gene paralogues

Analysis of the *E. coli* paralogue database reveals a few candidates for cross hybridisation creating false positives in these results. For example, due to the high sequence similarity found between the rhs elements it is possible that many false positives were created (Tables 3.4, 3.7, 3.8 and 3.10). However, the downstream ORFs associated with the rhs elements are highly variable in sequence (Hill *et al.*, 1994) and were found to be co-induced with the core rhs genes. For example, *b0499* was found to be significantly induced (~9 fold) in Comparison 1. It is possible that this is a false positive created by cross hybridisation of signals from *rhsE* (induced ~16 fold), *rhsD* (induced ~40 fold) and *yibJ* (induced ~22 fold). However, *b0499* is part of an operon in which *rhsD* (*b0497*), *ybbD* (*b0500*) and *b0501* all showed significant increases in expression in this comparison. Therefore, although it is a possible candidate for cross hybridisation, it is likely that, in this case, b0499 is actually showing an increase in expression level.

The *cpxP* (*b3913*) gene and the partially overlapping 'unknown function' *b3914* are candidates for cross hybridisation. Both *b3913* and *b3914* were found to be induced in Comparison 1 and Comparison 5. However, since the original first draft of the *E. coli*

K12 sequence was published (Blattner *et al.*, 1997), this region has been found to be incorrectly sequenced (creating an apparent frame shift) (Sean Philips, Pers. Com). *b3913* and *b3914* have since been recognised as one ORF, commonly referred to as *cpxP*. These 'two' apparent ORFs, therefore, provide a useful double-check for this gene and should always behave the same within the experiment.

3.4.4 Change in Expression Profile by PHMB treatment

A complete list of genes whose expression was deemed to be significantly altered by PHMB treatment can be seen in Tables 3.3 to 3.10. Genes have been classified according to the latest functional assignment (www.geneprotec.mbl.edu). Genes whose function are unknown are grouped into separate tables for each comparison. Genes in known operons or possible operons are grouped in tables together regardless of functional category. Among genes assigned a known function, nearly all functional groups are represented (Tables 3.1 and 3.2). These include genes involved in nucleic acid metabolism, translation, putative fimbriae genes, surface and outer membrane structure. Approximately half of genes induced immediately upon exposure to PHMB have no known function. Putative functions assigned to their protein products include transcriptional regulators and membrane proteins.

Upon exposure to PHMB 71 genes were induced and 12 repressed significantly. However, this change in expression profile was not maintained during the recovery process. Of the 71 genes induced, 46 genes subsequently became significantly repressed during the recovery as the cells regained their normal functional capacity and resumed growth. Only 7 of the genes that were induced immediately remained up regulated (i.e. ~10% of the initial 71 genes) in the recovered state, albeit at a lower level (Comparison 5). Unfortunately, none of these genes has any known function assigned to it. They can, however, be placed into two groups according to their induction profile: *yaiN, ycgW, b1228, b2863* and *rhsD* were all induced approximately 40-fold upon exposure and were found induced 4 to 10-fold in recovered cells; *yahA* and *yebG* were induced 12 to 14-fold upon exposure to PHMB and 3 to 4-fold in recovered cells.

Half of the 12 genes repressed immediately were maintained in a down-regulated state in recovered cells. These are *tnaA, aspA, pflB, lamB, flgJ* and *gatB*.

3.4.5 Genes Associated with the Outer Membrane

PHMB interacts primarily with LPS and is thought to have a negligible interaction with proteins in the outer membrane (Gilbert *et al.*, 1990a). Certain types of outer membrane protein could however be susceptible to interaction with PHMB and others could be affected indirectly by PHMB due to changes in the properties of LPS in the membrane, thereby affecting their function. Whereas genes involved directly in LPS biosynthesis and maintenance do not appear to have been altered in response to PHMB, many genes associated with the outer membrane were changed, such as those associated with the flagella, fimbriae, O-antigen and rhs elements.

3.4.5.1 O-Antigen

Five genes involved in the synthesis of the O-antigen were induced immediately after exposure to PHMB. These five genes are part of two clusters of genes (rfa [LPS core] and rfb [O-antigen]) involved in the synthesis and processing of the O-antigen. The genes *htrL* (*b3618*, a gene of unknown function which may be involved in LPS synthesis) and *rfaL* (encoding an O-antigen ligase) were induced ~9- and ~11-fold respectively. *yefI* (*b2032*, encoding a transferase), *rfc* (*b2035*, encoding an O-antigen polymerase) and *rfbX* (O-antigen transporter) were induced ~26-, ~5- and ~17-fold respectively. It is likely that the entire rfb cluster was induced, but the other ORFs were missed in this analysis. These genes returned to normal levels in recovered cells (Comparison 5), and were shown to be down regulated at some time after 5 hours. Comparison 3 (Table 3.6) shows that *yefI, yefG, rfc, rfbX* and *htrL* were down regulated approximately 14-, 10-, 6-, 9- and 5-fold respectively.

The O-antigen is a polysaccharide attached to the lipid core component of outer membrane LPS (which is synthesised separately) (Liu and Reeves, 1994). It consists of 10-30 repeating oligosaccharide 'O units' (generally composed of 3-6 sugars) (Liu *et al.*, 1996). O-antigen synthesis starts with the assembly of O units in the cytoplasm by sequential transfer of sugars onto a lipid carrier (undecaprenol

phosphate [UndP]). The completed O units are transferred to the periplasm (via RfbX) and then polymerised from the reducing end by an O-antigen polymerase (RfC). The O-antigen chains are then ligated to the core lipid A by an O-antigen ligase (RfaL) and the completed LPS translocated to the outer membrane (Liu *et al.*, 1996). However, K12 strains of *E. coli* do not have a functional O-antigen processing pathway (Liu and Reeves, 1994). Although it has the complete LPS core synthesis pathway, it has no O-antigen because of a defective O-antigen gene cluster (rfb). A mutation arising from the insertion of the *IS5* element has deleted a 1.1 kb section containing two glycosyltranferases-encoding genes (Liu and Reeves, 1994).

The biological relevance of the induction (and then repression) of these O-antigen genes by exposure to PHMB is not clear. However, a possible link can be made with the induction of the rhs elements (see below). In a K12 strain expressing a plasmid-encoded O-antigen (the O7 antigen), a *tolA* mutant was found to have reduced expression of O7 LPS (Gaspar *et al.*, 2000). *tolA* was shown to play a crucial role in the processing of the O-antigen. Although The *tolQRA* operon showed no sign of any significant transcriptional changes in expression throughout this experiment there is a case for a link to rhs expression which is discussed below.

3.4.5.2 rhs Elements

A rapid and strong induction of the rhs and rhs-associated genes was observed upon exposure to PHMB (see Table 3.3). *rhsB, rhsD* and *rhsE* showed mean induction ratios of approximately 14, 40 and 16 respectively. These genes are very similar (80-96 % identity) at the sequence level and their apparent induction could be caused by cross hybridisation. However, since genes of little or no homology which are associated with each rhs gene were also induced significantly, it is likely that each element (B, D and E) was actually and specifically induced. A closer inspection of the remaining two rhs elements (A and C) revealed that, for both , ORFs with which they were associated were up-regulated. The ORF downstream of *rhsA*, known as *yibJ*, was up-regulated ~20-fold and the ORF downstream of *rhsC*, known as *ybfD*, was up-regulated ~9-fold. *rhsC* was significantly up-regulated in two of the three experiments. It seems likely therefore that all five rhs elements were induced in response to the presence of PHMB.

The function of the rhs elements (so named as they were first identified as being recombinational hotspots) has remained elusive since their discovery (Lin *et al.*, 1984). This is mainly because they have never been shown to be expressed under normal laboratory conditions. This in itself makes the discovery of induction, by exposure to PHMB, both remarkable and a potentially important step in the elucidation of gene function. They are not essential but are conserved among *E. coli* strains. The rhs core ORF (~3.7 kb in size) contains a peptide motif xxGxxRYxYDxxGRL(IorT)xxxx that is repeated 28 times. It is likely to produce a cell surface protein which may have a macromolecular binding function (Hill *et al.*, 1994). This hypothesis has been suggested because of the similarity with the *Bacillus* WapA (wall associated protein) gene sequence (Foster, 1993). Unfortunately, no function has been assigned to WapA either. The downstream ORFs associated with the rhs core ORFs may play a role in the transport and processing of the core ORFs (which do not contain an obvious signal sequence). For example, YbbC (part of the *rhsD* element), has the codons for Leu-Phe-Ala-Cys starting at codon 15 (Hill *et al.*, 1994). This matches the sequence proposed to signal the cleavage and processing of lipoproteins (Wu and Tokunaga, 1986; von Heijne, 1989). Although other small rhs ORFs at positions like *ybbC* have apparent signal peptides, only *ybbC* has a proposed lipoprotein signal. Interestingly, of all the rhs elements induced, it is only *rhsD* (encoding a possible lipoprotein) that is maintained in an up-regulated state in recovered cells.

The initial massive induction of the rhs genes was reversed during the subsequent recovery from PHMB. *rhsE, rhsB, rhsD, yibJ* and 6 genes located downstream became down-regulated approximately 8-fold in a comparison between recovered cells and cells just exposed to PHMB (Comparison 3). This repression is likely to have occurred sometime around the five-hour point in Figure 3.5. The expression of the rhs elements then appeared to return to normal levels (i.e. very little or no expression) in recovered cells, except for *rhsD* the expression of which was maintained 4-fold above normal.

Previous promoter studies on the rhs elements have revealed little information. A mini Tn10 insertion into the *tolQRAB* operon results in an increased expression of

rhsA, *rhsB* and *rhsC*, but only when grown on solid media (Hill *et al.*, 1994). The biological relevance of this is unknown. No expression of rhs has previously been found in liquid cultures. The tol operon has recently been shown to play a role in O-antigen processing (see Section 3.4.5.1; Gaspar *et al.*, 2000). However, since the *tol* operon does not appear to change in expression during exposure to, or recovery from, PHMB, it is possible that some other unidentified factor is regulating both the O-antigen and rhs genes.

An alignment of the rhs promoter regions provided no insights to their regulation (See Figure 3.6). *rhsA* and *rhsC* have the most similar promoters (greater than 70% similarity), but were both weakly induced in comparison to the *rhsB, rhsD* and *rhsE* elements. The promoters for the *rhsB, rhsD* and *rhsE* elements are very different from each other and from the promoters for *rhsA* and *rhsC*. However, *rhsB*, *rhsD* and *rhsE* appeared to be strongly induced. Clearly, they must be regulated by some factor(s), but so far the nature of this regulation remains unknown. It is possible that one or more of the unknown genes found to have had their expression profile altered in response to PHMB may play crucial roles in this regulation.

The induction of these large putative cell-surface proteins could provide a physical barrier to the action of PHMB by preventing it from interacting directly with LPS in the outer membrane. Alternatively, these proteins could play a role in cellular aggregation and thus prevent PHMB from interacting with the cell wall by shielding cells and reducing the amount of exposed cell surface. However, no evidence for increased aggregation was observed in the growth experiments. Thus, these proteins may even be released into the growth medium in an attempt to bind PHMB. It has been suggested that WapA may actually be clipped from the membrane and secreted (Foster, 1993). The H-repeat is also an interesting component of the rhs elements that resembles an insertion sequence and is found in *rhsB, rhsC* and *rhsE* elements (Zhao *et al.*, 1993). The H-repeat (short for Hinc repeat) is named after the presence of a characteristic *Hinc* II site. H-repeats have been found in *Vibrio cholerae* and *Salmonella enterica* rfb clusters and are known to be involved in O-antigen variation. This provides another link to the expression of O-antigen associated genes, as discussed earlier. The rhs elements, their role during PHMB induced stress and their relationship with the O-antigen clearly deserves further study.

3.4.5.3 Flagella, Fimbriae and Pili

PHMB caused the down regulation of some of the flagella-associated genes: in Comparison 1, *flgE* and *flgJ* become down regulated approximately 10- and 11-fold respectively. The flagella are used for propulsion and each is driven by a motor embedded in the cell envelope at the base of the flagellum. The down-regulation of flagella would make sense if PHMB did cause cells to aggregate. SAM analysis of Comparison 1 showed *flgA, flgF, flgG and flgH*, as well as *flgJ*, to be significantly down regulated. It is interesting to note that FlgJ is a periplasmic flagella-specific muramidase, which hydrolyses the peptidoglycan layer to allow assembly of the rod structure to proceed through the periplasmic space (Hirano *et al.*, 2001). The breakdown of peptidoglycan in the periplasm could allow PHMB easier access to the cytoplasmic membrane, so down-regulation of the process might protect against PHMB.

In contrast to the flagella genes, it appears that some hypothetical fimbriae genes were induced when the flagella gene expression was being switched off, immediately after exposure to PHMB. *yadC* (encoding a putative fimbrial like protein) and *yehC* (encoding a putative periplasmic fimbrial chaperone) were induced approximately 32- and 7-fold respectively. Fimbriae are used as adherence factors in the colonisation of eukaryotic surfaces (Low *et al.*, 1987). This induction did not appear to be maintained following recovery, as *yadC* was found to be down regulated as the cells recovered (~12-fold in Comparison 3). The flagella genes did, however, appear to be kept in a down regulated state in recovered cells: *fliC, fliD* and *flgJ* were approximately 4-, 4- and 3-fold repressed in recovered cells. In addition to the previously mentioned hypothetical fimbriae genes, a putative pilus gene, *b1202*, was induced ~5-fold immediately after exposure to PHMB along with the downstream ORF, *b1203* (induced ~8-fold), which is thought to be a GTP-binding pilus chaperone (Colibase).

3.4.5.4 Others

The gene encoding a high-affinity maltose porin, LamB, was down-regulated approximately 10-fold upon exposure to PHMB and remained down regulated even in recovered cells (~7-fold). LamB is an integral outer-membrane protein that contains a hydrophobic path (known as a greasy slide) through which maltooligosaccharides, and therefore possibly PHMB, can cross into the periplasm (van Gelder *et al.*, 2002). Two lipid-anchored outer-membrane proteins, OsmB and VacJ, had their genes induced approximately 7- and ~15-fold respectively. These two genes did not appear to be maintained in an up-regulated manner in recovered cells (neither gene was up-regulated in comparison 5, but *osmB* was down-regulated ~5 fold in Comparison 3).

3.4.6 Genes Associated with the Periplasm

PHMB needs to pass through the periplasmic space (once the outer membrane has been breached) before it encounters the cytoplasmic membrane. However, relatively few periplasmic protein-encoding genes were changed in response to the presence of PHMB. Upon exposure to PHMB, the periplasmic repressor of the Cpx regulon (CpxP) became up-regulated approximately 20-fold. *cpxP* is itself a member of the Cpx regulon, but no other members appeared to be induced. CpxP has been shown to bind to misfolded pilus subunits and it could be induced directly (by some unknown mechanism) in response to damage to pili caused by PHMB (see Section 1.6.1). As mentioned previously, a pilus chaperone was also induced at the same time. Following recovery from PHMB stress, *cpxP* was maintained in an up-regulated state.

The *hdeAB* operon was down regulated immediately after exposure to PHMB. *hdeB* was down-regulated around 15-fold (*hdeA* was down-regulated approximately 8-fold initially, but only significantly in 2 of 3 experiments). These genes are linked to acid-resistance and may act as chaperones by preventing the aggregation of denatured proteins. The down-regulated state of the *hdeAB* operon was maintained even in recovered cells. The biological reasons for this are not clear. The repression of the *hdeAB* operon may be an indirect effect of a global H-NS regulated repression (see Section 3.4.9). The *hdeAB* operon is so-called because of its **H**-NS **d**ependent **e**xpression) (Gajiwala and Burley, 2000).

The only other known periplasmic-protein-encoding gene sharing altered expression in this investigation was *mglB*. The gene product of *mglB* is involved in galactose transport and chemotaxis and its expression was found to be down-regulated in recovered cells (Comparison 5).

The change in periplasmic-protein-encoding genes would be expected to be much greater if extensive damage to the contents of the periplasm was occurring. The up-regulation of DNA binding proteins and cytoplasmic-membrane-proteins (see below) suggests that the cytoplasmic membrane function has been disrupted. The lack of change in 'periplasmic genes' would, therefore, suggest that the route to the cytoplasmic membrane by PHMB is direct and that it has very little interaction with the contents of the periplasm.

3.4.7 Genes Associated with the Cytoplasmic Membrane

Many genes involved in the function of the cytoplasmic membrane were altered in response to PHMB, in particular those involved in sugar transport. Exposure to PHMB caused *rbsD* to become down-regulated immediately 12-fold. RbsD is involved in the high-affinity transport of ribose. The entire *gat* operon, involved in the galactitol-specific phosphotransferase system, was also probably down-regulated (at least the first 3 members; *gatA*, *gatB* and *gatC*) at this point since *gatB* was found to be down-regulated in Comparison 1 and all three genes, *gatABC,* were all found to be down regulated in recovered cells (Comparison 5). These genes were joined by *lldP* (encoding a lactate permease) which was down-regulated approximately 6-fold. Other sugar transport associated genes down-regulated in recovered cells were *malK* (maltose transport), *mglB* (galactose transport) and *manX* (encoding a mannose specific phosphotransferase enzyme).

3.4.8 Genes Associated with General Metabolism

The interruption of exponential growth phase by PHMB, led to some major changes in the expression of genes involved in general cell metabolism. Tryptophan, cysteine and aspartate metabolism all appeared to be altered. The tryptophanase leader peptide

(encoded by *tnaL*), tryptophan deaminase *(tnaA)* and aspartase (*aspA*) were down-regulated approximately 34-, 29- and 22-fold respectively in Comparison 1. The transcriptional regulators of cysteine biosynthesis, *cysB* and *cbl*, were induced approximately 8- and 12-fold respectively. These changes could represent a change as an indirect response to PHMB damage to proteins, or as a response to the change in growth rate. The transcriptional activator of threonine dehydratase was also induced 7-fold. This expression profile appeared to be reversed and over-compensated during the recovery. In Comparison 3, *cbl* was found to be repressed ~8-fold (along with *tdcR*, ~5-fold), whereas *tnaL* and *tnaA* were both induced (~8- and ~9-fold respectively). *cbl* was found to be repressed in Comparison 4, implying that the change in profile, for *cbl* at least, occurred late on during the recovery. Following the recovery, *tnaA* and *aspA* were found to be down-regulated (3- and 6-fold, respectively), while *cysB* was found to be up-regulated approximately 4-fold.

3.4.9 Genes Associated with DNA-Binding-Proteins

Many DNA-binding transcriptional regulators were induced/repressed in response to PHMB stress. This would be expected since transcriptional changes occur through alterations in DNA-binding proteins (i.e. transcription factors). However, the alteration in expression of a number of other genes encoding DNA-binding proteins implied that there could be damage to DNA. For example, transcripts for RecA (a DNA strand exchange and recombination protein), XseA (an exodeoxyribonuclease) and StpA (may play a DNA structural role) were all induced approximately 7-, 7- and 9- fold, respectively, immediately after exposure to PHMB. Whereas *recA* and *xseA* transcripts appeared to revert to normal levels in recovered cells, *stpA* was maintained in an up-regulated manner in recovered cells. Also up-regulated in recovered cells was *fis*, whose gene product is known to be involved in recombination and inversion of DNA.

Of particular note is the repression of *hns* that appears to occur in Comparison 4. H-NS exerts most of its transcriptional effects by negatively regulating genes; consequently, this repression of *hns* may be the key to resuming growth at near to normal levels by allowing the expression of certain genes. H-NS regulates the synthesis of a large number of gene products involved in a diverse range of biological

processes (Atlung and Ingmer, 1997). In the majority of cases, H-NS exerts its function by inhibiting gene transcription. This has been shown using both 2-D protein gels and array-based methods (Altung and Ingmer, 1997; Hommais *et al.*, 2001). Many of the target genes regulated by H-NS are also induced by changing environmental conditions such as osmolarity (e.g. *osmC* at high osmolarity), temperature (*fimE* at high temperature), pH (e.g. *hdeA* and *cadA* with acid) or growth phase (*hdeA* in stationary phase).

Approximately half of the proteins encoded by genes controlled by H-NS are associated with the membrane or periplasmic space. It is therefore not surprising that many genes found to have their transcript levels altered in response to PHMB have previously been identified as being positively or negatively regulated by H-NS. Indeed, as stated previously, the levels of H-NS were found to be significantly down-regulated in Comparison 4 between the cells that had fully recovered and those just about to recover. This implies that H-NS may be exerting its effect early in the response to PHMB stress. However, genes known to be repressed by H-NS include *rfaL, pflB, hha, stpA, yhiW, yhiX, hdeB, ybaJ, ydcD* and *yeeN*. These were all up-regulated upon exposure to PHMB, except for *pflB and hdeB*, which were strongly repressed. This might suggest that H-NS levels were low in cells during recovery, in contrast to the apparent up-regulation of the *hns* gene itself. In the final comparison between normal cells and recovered cells, H-NS negatively-regulated genes were found to be repressed (e.g. *cadA, pflB, hdeA, hdeB*) except for *stpA* which was up-regulated. This might suggest that H-NS levels were higher in exposed cells and that the apparent 'down-regulation' in Comparison 4 was due to an increase early on in the response. It is also possible that cross hybridisation of the *hns* signal to the *stpA* spot has masked a real increase in expression of *hns* that has not been identified. Alternatively, it is possible that an initial decrease in *hns* transcriptional levels led to the transcriptional activation of many of the normally repressed genes that allowed recovery from stress caused by PHMB. Once the stress was relieved, up-regulation of *hns* expression allowed normal growth to resume.

This is a very simple picture of expression and repression control by H-NS. It is clear that other regulators would also be involved in this response. H-NS is just one of the many transcriptional regulators that have been identified. It is also possible that

some of the putative regulators identified as having altered expression profiles are also exerting an effect. The highly integrated regulatory networks in *E. coli* are unlikely to rely solely on one regulator for a response to stress. H-NS may also have a more direct role to play in protection against PHMB. Since H-NS (and its paralogue StpA) binds DNA tightly, it may function as a protective protein , altering DNA structure, preventing it interacting with any PHMB present in the cytoplasm.

3.4.10 SAM Results

The results from the SAM analysis were not analysed extensively. However, there are some striking changes in expression identified in comparison 1 that are hard to ignore. The full results from this analysis can be seen in the Appendix (Appendix A, Table A.2). Most striking of all was the initial induction of members of the σ^{32}, SOS response and many more DNA metabolism associated genes. Of the main heat shock response, the genes encoding DnaK, DnaJ, ClpB, DegP and HtpX were all strongly induced (the gene encoding DnaK was the only one of this set found to be induced in the standard analysis, see Table 3.3). Furthermore, the gene encoding σ^{32}, the master regulator of the heat shock response (encode by *rpoH*), and the gene encoding an additional factor needed for transcriptional activation from σ^{32} promoters (HtgA) were also found to be strongly induced.

Eight known members of the SOS DNA damage response were found to be significantly up-regulated in the SAM analysis. These were *dinI*, *dinP*, *ydjM*, *ruvA*, *yebG*, *recN*, *recX* and *recA* (*yebG* and *recA* were also identified in the standard analysis). The SOS response is induced by DNA damage and includes around 30 unlinked genes involved in DNA damage tolerance and repair (Khil and Camerini-Otero, 2002). Regulation and control of the system is not fully understood and can occur through the LexA repressor, through RecA or sometimes independently of both RecA and LexA (Khil and Camerini-Otero, 2002). De-repression of the SOS response controlled via LexA occurs when RecA binds to single stranded DNA (Courcelle *et al.*, 2001). This produces a conformational change in RecA causing it to act as a protease capable of cleaving LexA (Sassanfar and Roberts, 1990). In addition to the SOS response, *recO*, *exoX* and *xseA* were also found to be up-regulated in the SAM

analysis. The products of these genes are also involved in DNA repair and maintenance.

3.4.11 Summary

Hitherto, PHMB was believed to cause bacteriostatic and bactericidal effects by disrupting the outer membrane and cytoplasmic membrane, respectively. The transcriptional profiling of the response to, and recovery from, PHMB stress in *E. coli* indicates that the levels of PHMB used in this investigation caused a brief bacteriostatic effect through interacting with the functioning of both the outer membrane and cytoplasmic membrane. However, the induction of genes involved in the repair of cytoplasmic constituents (in particular the σ^{32} and SOS response) further indicate that a substantial disruption of the cytoplasmic membrane occurred, and that PHMB may be causing substantial damage to proteins and DNA inside the cell even at bacteriostatic levels. PHMB has previously been shown to interact with and precipitate cytoplasmic contents, but these were thought to be secondary events in the bactericidal action.

It must be noted that the growth curves shown in Figure 3.2 may contain many different sub-populations of cells, each with different characteristics. Consequently, after exposure to PHMB, the recovery from the bacteriostatic phase may be led by one or many of the distinct sub-populations of cells with each having its own distinctive transcriptional profile. Therefore, the transcriptional profile generated in these experiments may actually be an amalgamation of many different sub-populations having separate and distinct responses to PHMB.

Chapter 4: Proteomic analysis

4.1 Introduction

The transcriptional analysis performed in Chapter 3 revealed changes in the expression of many cytoplasmic, periplasmic and membrane-protein encoding genes. However, one of the main drawbacks of any transcript-based method of expression profiling is that it does not take into account any post-transcriptional regulation or modification (Futcher *et al.*, 1999; Gygi *et al.*, 1999). These can have significant influences upon how a cell adapts to its environment. Proteomic analysis is the term given to the study of the **proteins** expressed by the **genome** of an organism. This is commonly accomplished by a combination of two-dimensional gel electrophoresis to separate and visualise proteins, and mass spectrometry for protein identification (Molloy *et al.*, 2000).

The aim of this chapter is to characterise further, at the protein level, the response of *E. coli* to PHMB using proteomic methods.

4.2 Methods

The experiment described in Section 3.3.2 was duplicated, but protein was extracted instead of RNA. Samples were taken at points equivalent to A (t = 3.75 h, unexposed cells) and D (t = 9 h, exposed cells) in Figure 3.2 i.e. a proteomic repeat of Comparison 5. The periplasmic proteins were isolated (Section 2.4.1), the concentration determined (Section 2.4.2) and then subjected to 2-dimensional gel electrophoresis (section 2.4.3). After visualisation (Section 2.4.3.4), gel images were scanned, the 'spots' aligned by eye and then quantified using computer based software (PDQuest, BioRad, Hemel Hempstead, UK). A statistical analysis identical to that performed upon macroarray and microarray data in a 'standard analysis' was attempted i.e. each spot value was normalised, the induction ratio calculated and significantly altered spots identified (Section 2.3.4). Spots of interest were sent to Dr Len Packman at Cambridge Protein and Nucleic Acid Chemistry Facility for identification (see Section 2.4.3.5).

4.3 Results

The periplasmic proteins of unexposed and PHMB exposed *E. coli* cells were subjected to 2-dimensional gel electrophoresis. Typical examples of scanned 2-dimensional gel images can be seen in Figure 4.1. Digital composite images were generated from triplicate scanned gel images (Figure 4.2). 'Landmark' spots were identified and used to align all the spots in the two sets of digital images (Figure 4.2). Spots were quantified (the full list of spots identified in each gel image, together with fold-change values when a matching spot in the corresponding gel image could be found, can be seen in Appendix C, Table C.1), those deemed altered were extracted and sent for identification by MALDI-TOF or electrospray MS/MS (Figure 4.3). The results of the spot identification can be seen in Table 4.1.

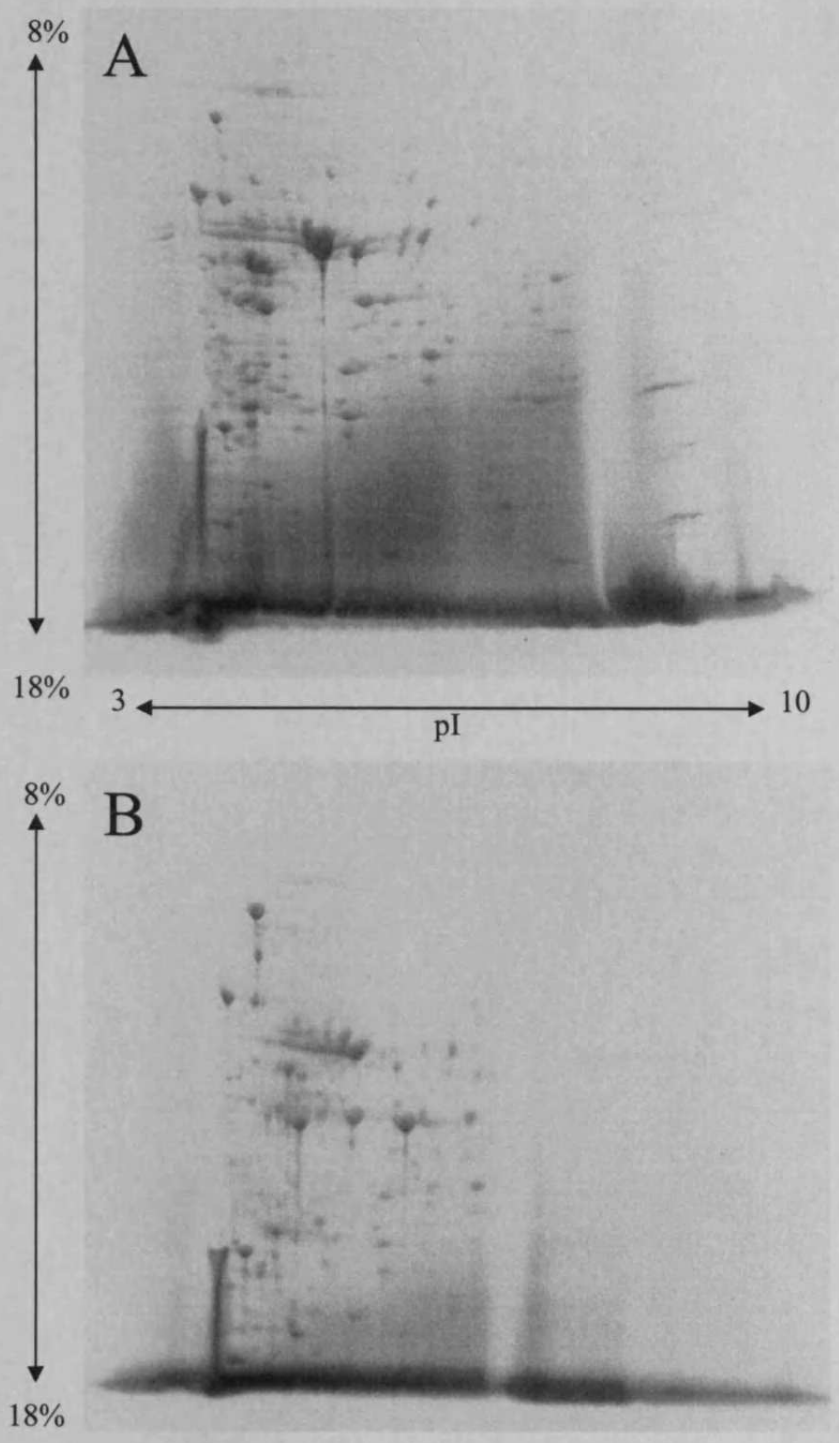


Figure 4.1 Examples of typical 2-dimensional gel images of periplasmic-proteins extracted from non-exposed cells (A) and PHMB recovered cells (B).

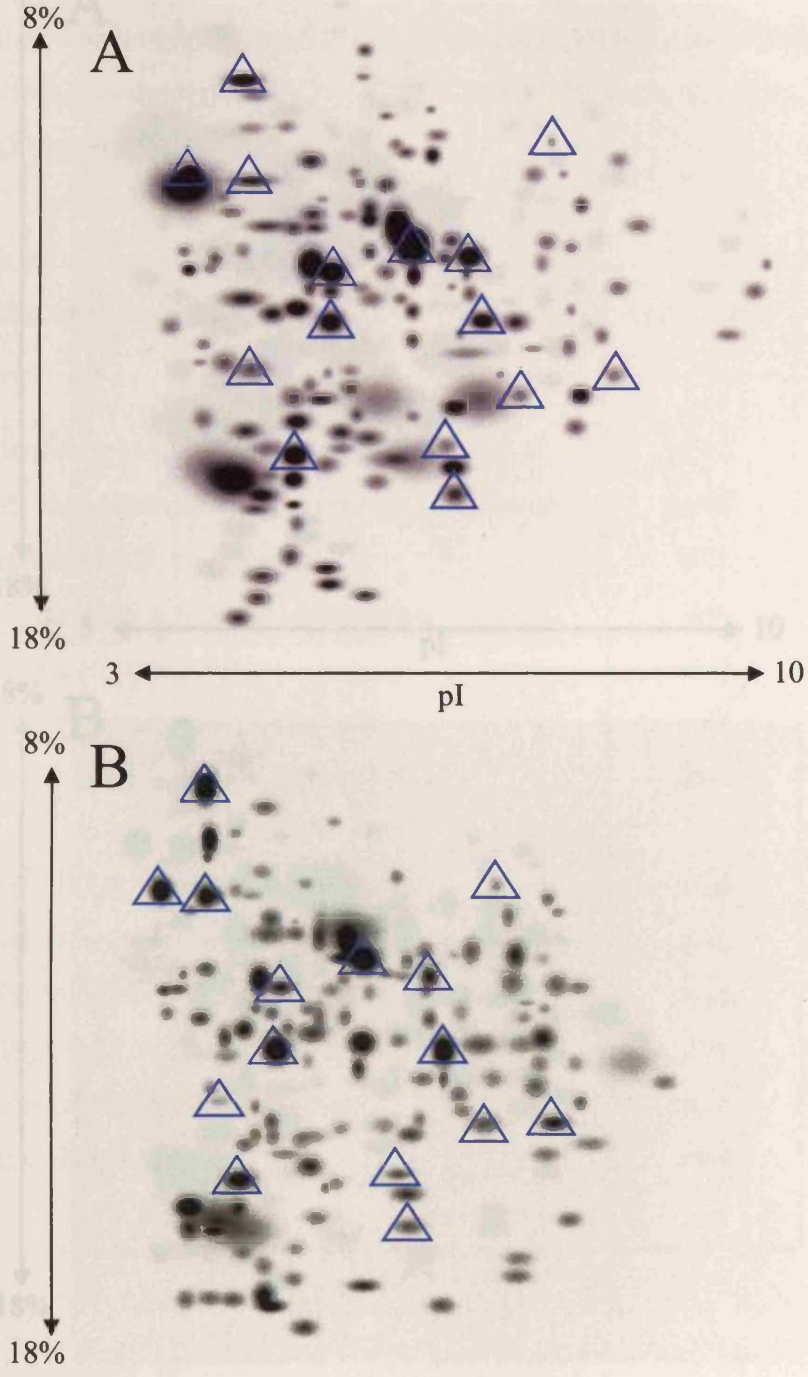


Figure 4.2 Digital composite images generated from triplicate periplasmic-protein gel images from non-exposed cells (A) and triplicate periplasmic-protein gel images from PHMB recovered cells (B). Some landmark spots are are indicated by blue triangles.

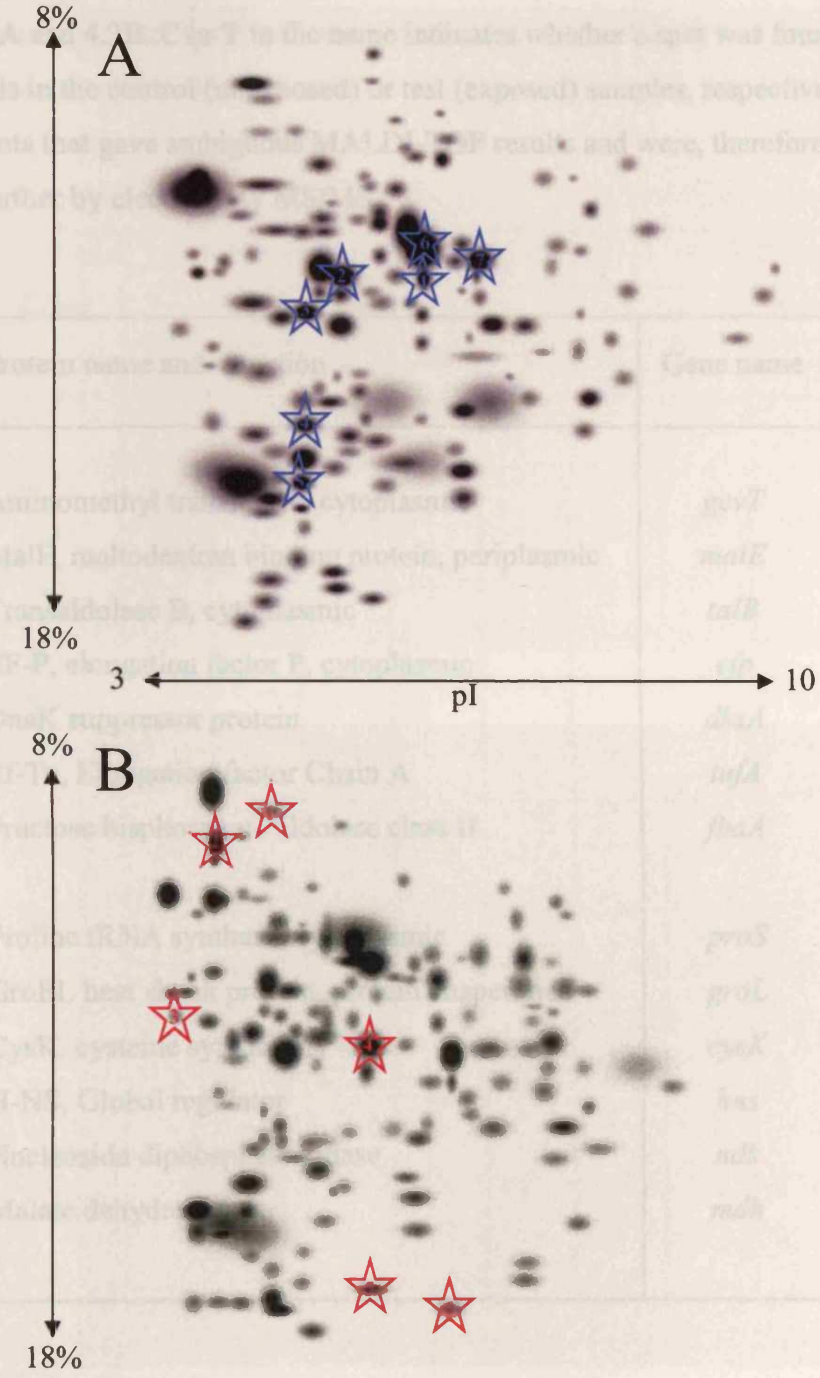


Figure 4.3 Spots deemed to be significantly altered whose proteins were subsequently identified by mass spectrometry. Blue stars indicate proteins found at higher levels in unexposed cells (A), red stars indicate proteins found at higher levels in recovered cells. See Table 4.1 for protein identification.

Table 4.1 Identity of proteins (by MALDI-TOF) deemed to have altered levels from Figures 4.3A and 4.3B. C or T in the name indicates whether a spot was found at higher levels in the control (unexposed) or test (exposed) samples, respectively. # Denotes spots that gave ambiguous MALDI-TOF results and were, therefore, analysed further by electrospray MS/MS.

Spot	Protein name and function	Gene name	Blattner Number
C1	Aminomethyl transferase, cytoplasmic	*gcvT*	b2905
C2	MalE, maltodextran binding protein, periplasmic	*malE*	b4034
C3	Transaldolase B, cytoplasmic	*talB*	b0008
C4#	EF-P, elongation factor P, cytoplasmic	*efp*	b4147
C5	DnaK suppressor protein	*dksA*	b0145
C6	Ef-Tu, Elongation factor Chain A	*tufA*	b3339
C7#	Fructose bisphosphate aldolase class II	*fbaA*	b2925
T1	Proline tRNA synthase, cytoplasmic	*proS*	b0194
T2	GroEL heat shock protein, protein chaperone	*groL*	b4143
T3	CysK, cysteine synthase A	*cysK*	b2414
T4	H-NS, Global regulator	*hns*	b1237
T5#	Nucleoside diphosphate kinase	*ndk*	b2518
T6	Malate dehydrogenase	*mdh*	b3236

Table 4.2 Average fold induction values for genes whose protein product was deemed to be found at significantly altered levels. Values are shown only if they were found to be significantly altered in expression level in Comparisons 1, 2, 3, 4 and 5 in Chapter 3.

Spot	Gene	Standard Analysis					SAM analysis				
		Comparison					Comparison				
		1	2	3	4	5	1	2	3	4	5
C1	*gcvT*										
C2	*malE*										
C3	*talB*										
C4	*efp*										
C5	*dksA*										
C6	*tufA*										
C7	*fbaA*					-3.0					
T1	*proS*										
T2	*groL*										
T3	*cysK*			+7.7	+9.1				+7.5	+8.9	
T4	*hns*				-6.2						
T5	*ndk*										
T6	*mdh*										

4.4 Discussion

2-Dimensional gel electrophoresis is a highly sensitive procedure. Slight fluctuations in salt levels and the presence of any residual surfactants can have major effects on the isoelectric focusing points of samples. This makes analysis of the response to a surfactant (i.e. PHMB) difficult to perform accurately using this method. Despite this, as can be seen in Figure 4.1, a reproducible pattern of protein spots was observed from a periplasmic protein preparation extracted from cells unexposed to PHMB (A) and from cells exposed to PHMB (B).

Analysis of 2-dimensional protein gels is complex and prone to error. Replicate gel images must have all corresponding spots aligned and matched together in order to form a composite 'digital' image. Composite images must then be aligned with another composite digital image in order to determine any differences. Spot recognition and matching has to be done by eye which is highly time consuming and prone to human error. Determining what is and what is not a spot can be difficult and is subject to different interpretation. This can be seen by the very different looking digital images seen in Figure 4.2 (in comparison with the original gel images seen in Figure 4.1). The largest spots (i.e. the most abundant) are easier to align than the smaller spots (which are less abundant), but it is often the smaller spots that show the greatest changes in abundance.

After alignment of the digital composite images, spot density quantification (by PD-Quest) was followed by an attempt at statistical analysis (identical to that performed in the standard analysis in Chapter 3). However, the validity of this method was immediately called into question as it became apparent that there was a large number of unmatched spots. Since the standard analysis method works by identifying significantly altered induction ratios, any spot that has no matched 'partner' would have an 'infinitely large' induction ratio. This clearly causes problems in a statistical analysis. Therefore, spots (which could be matched in the corresponding gels) were deemed to be 'significantly' altered if their $\log_{10}$[IR] was greater than 2 standard deviations from the mean of all **matched** spot $\log_{10}$[IR] values in the three replicate experiments OR, in the case of unmatched spots, if the unmatched spots could **clearly** be identified as being absent in the corresponding gels. These "mixed criteria" clearly

show the limitations of this method for identifying spots of significantly altered abundance.

The spots deemed altered in abundance were re-checked by eye on the composite images and on the original gel images. A selection of spots that could be confidently identified as present on 3 original gel images and absent in the corresponding 3, or which were deemed significantly altered in the statistical analysis, were chosen for identification by MALDI-TOF analysis (See Figure 4.3). The results from this analysis can be seen in Table 4.1. The majority of the protein spots were identified by MALDI-TOF (10 out of 13). In these cases, mass spectra showed two or more peptides that were matched to a SWISS-PROT *E. coli* database homologue, thus establishing identity. In the remaining 3 cases protein identification was compromised by unreliable peptide spectra (either too few peptides or poor spectra) and needed further characterisation following MALDI-TOF to confirm identification. This was achieved by electrospray MS/MS.

The results of the protein identification were disappointing. Whereas all 13 spots were positively identified as being of *E. coli* K12 origin, only one of the proteins identified was periplasmic (C2: MalE, a maltodextran binding protein). The remaining 12 were all of cytoplasmic origin. This indicates that the method used to extract only the periplasmic proteins was fundamentally flawed. Unfortunately, this only emerged at the final step of protein identification. The presence of cytoplasmic proteins in the PHMB exposed samples could represent disruption of the cytoplasmic membrane and therefore leakage of the cytoplasmic contents into the periplasm. However, the identification of cytoplasmic proteins (as well as the periplasmic MalE) in the unexposed samples indicates that the extraction method itself was failing to exclude cytoplasmic proteins.

The method of protein preparation used here was unsuccessful at isolating only periplasmic proteins. However, although a dedicated specifically-cytoplasmic protein profile needs to be performed, some information about the alteration in cytoplasmic protein profile may be gleaned from these results. This information validates the array results obtained in Chapter 3.

Both H-NS and CysK are found at higher levels in cells exposed to PHMB. The transcripts for the genes encoding these proteins (*hns* and *cysK*) were identified as being significantly altered in Chapter 3. Furthermore, GroEL was found at higher levels in the exposed cells, providing further evidence that damage to cytoplasmic proteins is occurring and cytoplasmic protein repair mechanisms have been activated. In addition, the DnaK suppressor protein DksA was found at higher levels in the unexposed cells. Since *dnaK* was found to be significantly up-regulated upon exposure to PHMB, it is possible that this up-regulation occurred due to the decreased levels of DksA. DksA is a putative zinc-binding protein of 151 amino acids (Bass *et al.*, 1996). Although the basis of its suppression of *dnaK* is unknown, it has been suggested that some stress response factors may be involved (Bass *et al.*, 1996).

Although these results are clearly not conclusive, these findings do provide further evidence that the changes observed in Chapter 3 are transmitted to the protein level.

Chapter 5: Phenotypic Characterisation

5.1 Introduction

The emergence of array technology has allowed the rapid collection of large volumes of data (Merlin *et al.*, 2002). However, the power of genomic sequencing (to date more than 60 bacterial genomes are known in their entirety) has out stripped the capacity for functional characterisation of the ORFs identified. Computer analysis of the *E. coli* genome has revealed that even after 50 years of extensive research almost 50% of the predicted ORFs still remain to be characterised. This poses serious problems for the interpretation of array data.

In order for the value and meaning of array data to reach its full potential, genes of both known and, just as importantly, unknown function must be further analysed in the context of the original study. This functional characterisation can be carried out by a phenotypic study in both the gene deletant system and in the corresponding over-expression system. Two complimentary resources were exploited in this work in order to do a phenotypic study on as many *E. coli* ORFs as possible.

The Genome Analysis Project Japan (http://ecoli.aist-nara.ac.jp/) has isolated clones of each ORF predicted from the genome sequence of *E. coli* W3110 (Mori *et al.*, 2000). Every ORF has been cloned into a plasmid (known as pCA24N) containing the IPTG inducible promoter pT5/*lac*, an N-terminus histidine tag of the target ORF and an in-frame fusion of green fluorescent protein (GFP) at the C-terminus of the target. A *cis* coded *lacIq* is present to allow strict repression of the expression from the pT5/*lac* promoter. The structure of IPTG resembles the structure of lactose thus it can bind to the LacI repressor, modifying it so that it loses its affinity for the promoter (Ptashne, 1986). The structure of pCA24N can be seen in Figure 5.1.

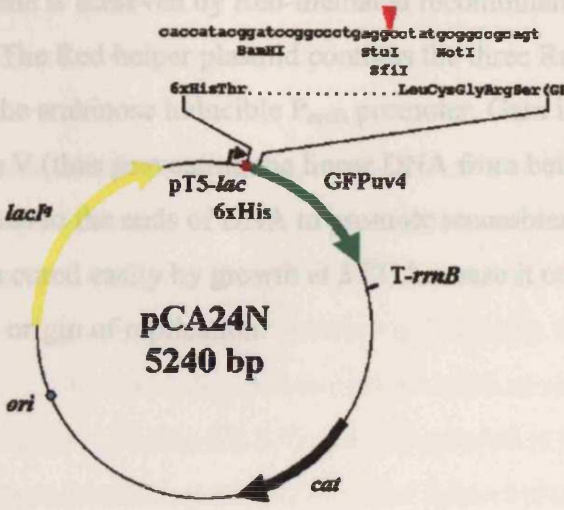


Figure 5.1 The structure of pCA24N. Diagram taken from The Genome Analysis Project Japan website (http://ecoli.aist-nara.ac.jp/)

Following on from the sequencing of *E. coli* MG1655 the University of Wisconsin *E. coli* Genome Project (www.genome.wisc.edu) has undertaken to construct a set of mutant strains in *E. coli* MG1655. To date, just over 1000 ORFs have been deleted according to the method devised by Datsenko and Wanner (2000) (which is discussed in greater depth below). Although the set of mutants is far from complete, approximately 30 deletants of interest to this study have been constructed.

A further 5 genes (1 of known function and 4 of unknown function) were selected as candidates for the production of mutant stains. These genes (*CpxP, b3914, ygeQ, ycgW, yaiN*) were chosen as they were found to be maintained in an up-regulated manner even after cells had recovered from PHMB-induced stress. The method chosen for the production of these mutants was based upon that of Datsenko and Wanner. Although initially these mutants were constructed in *E. coli* W3110, to allow a direct comparison with those mutants made by the U.W. *E. coli* Genome Project, the deletion was later transferred into MG1655.

The basic strategy is to replace a chromosomal sequence with a sequence containing a selectable antibiotic resistance gene. The sequence containing the insert is generated in a PCR reaction by using primers that also contain 36 nt extensions

homologous to regions adjacent to the gene to be inactivated. The exchange of wild-type for inactivated gene is achieved by Red-mediated recombination in the flanking homologous regions. The Red helper plasmid contains the three Red genes (γ, β and *exo*) expressed from the arabinose inducible P_{araB} promoter. Gam inhibits the host RecBCD exonuclease V (thus preventing the linear DNA from being degraded), while Bet and Exo gain access to the ends of DNA to promote recombination. The Red helper plasmid can be cured easily by growth at 37°C because it contains a temperature-sensitive origin of replication.

5.2 Methods

5.2.1 Gene Disruption

PCR products were generated from the template plasmid pKD4. The kanamycin resistance gene was amplified using primers which also contained 5'extensions (~36 nt) homologous to regions adjacent to the gene to be inactivated. The primers and their sequences used in these reactions can be found in Table 2.3. PCR products were gel purified (Section 2.2.5.4), *Dpn* I digested to remove residual plasmid DNA (Section 2.2.5.1), re-purified (Section 2.2.5.5) and resuspended in MilliQ water. *E. coli* W3110 transformants harbouring the pKD46 (Red helper) plasmid were induced with arabinose (0.5%) at 30°C with ampicillin to D_{600} ~0.5 and made electrocompetent (Section 2.2.3.3). Electroporation was performed using 20 ng of PCR product (Section 2.2.3.5). After primary selection on kanamycin containing agar, mutants were grown overnight at 42°C in plain media, re-plated on kanamycin plates and tested for Amp sensitivity to confirm the loss of pKD46.

5.2.2 PCR Verification

Four PCRs were used to verify that mutants had the correct structure. Three PCR screens using locus specific primers and the respective common test primer (K_1/K_T or K_2) were used to test for both new junction fragments (See Table 2.3 for primers and their sequences). A third PCR was carried out with both flanking locus specific primers to verify the loss of the parental (nonmutant) fragment and gain of the new mutant specific fragment. A diagrammatic view of the PCR reactions can be seen in Figure 5.2.

5.2.3 Transfer into *E. coli* MG1655

Genomic DNA (100 ng) (purified as described in Section 2.2.4.3) from verified mutants was sonicated in a water bath for 2 minutes before electroporation (Section 2.2.3.5) into *E. coli* MG1655. Following electroporation, shocked cells were incubated in 1ml SOC at 37°C for 4 hours before plating onto LB agar containing

kanamycin. The resulting mutants were verified using the PCR reactions described above.

5.2.4 MIC Assay

Assays were performed to determine the MIC for PHMB in the knock-out strains, knock-out strains over-expressing the knocked-out gene and on the wild-type strains over-expressing the gene of interest, as described in Section 2.5.

5.3 Results

5.3.1 Construction of knockout Strains

Five linear 1.5 kb DNA fragments for knocking out each of the five genes *cpxP, b1228, yaiN, ycgW* and *ygeQ* were successfully produced from pKD4 template in standard PCR reactions. Transformation of 'Knock-out DNA' into W3110 containing the pKD46 recombination plasmid typically yielded 10-20 kanamycin-resistant colonies for *cpxP* and *b1228*. A colony was chosen from each plate and screened for the correct insert structure. Both colonies gave the expected fragments in four PCR test reactions (See Figure 5.2 and 5.3 for the results of the PCR screen on the C*pxP* knock-out) showing that these mutants have the correct structures. The remaining three genes were transformed a further 10 times. Single, weak growing colonies occasionally grew, but when screened showed no sign of having the correct insert structure. These mutants were, therefore, unable to be created. To allow a direct comparison with MG1655 knock-out strains, gDNA from the verified knock-out strains in W3110 (*cpxP* and *b1228*) was sonicated for 5 minutes and used to transform MG1655 cells. This was a highly inefficient method, but the *cpxP* knockout yielded a single kanamycin resistant colony that was shown to contain the correct knock-out structure.

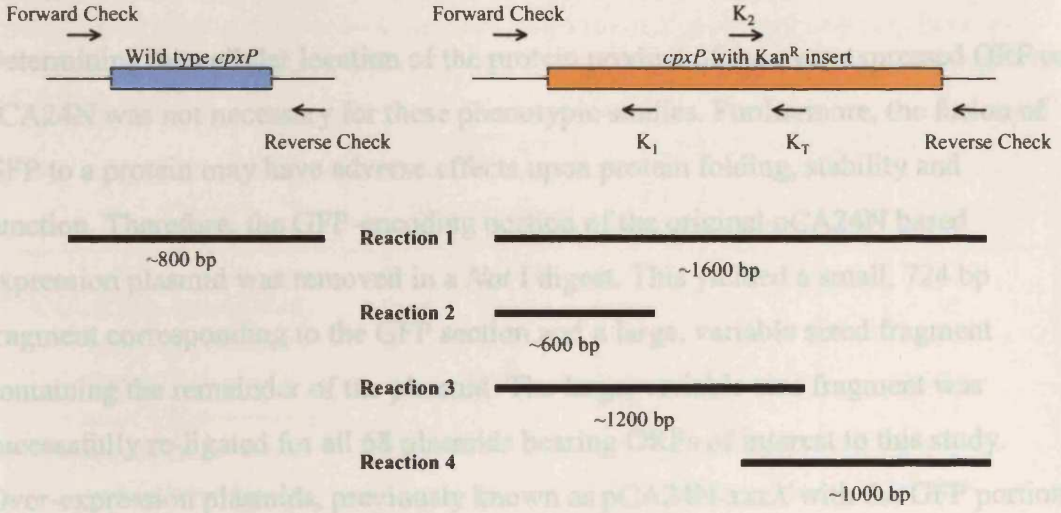


Figure 5.2 Primer pairings used in PCR verification of knock-out strains. Band sizes shown are approximate sizes for the wild-type (shown in blue, left) and *cpxP* knock-out (shown in orange, right) strains. Reaction 1: Forward Check + Reverse check; Reaction 2: Forward Check + K_1; Reaction 3: Forward Check + K_T; Reaction 4: K_2 + Reverse Check

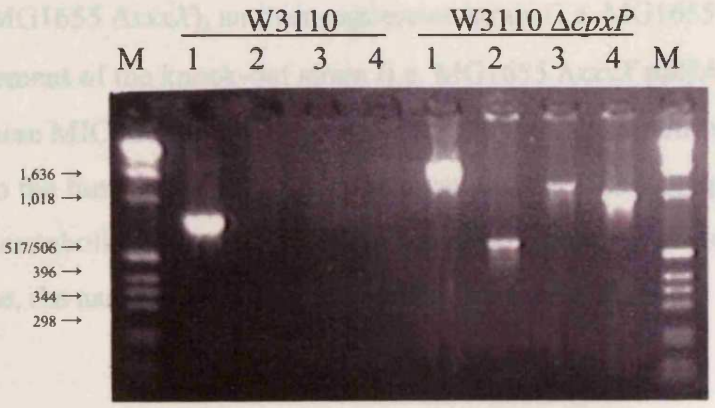


Figure 5.3 PCR verification of knock-out structure of W3110 *cpxP* mutant. 1, 2, 3 and 4 Correspond to PCR Reactions 1-4 described above in Figure 5.2

5.3.2 Modification of Over-expression Plasmids

Determining the cellular location of the protein product of the over-expressed ORF on pCA24N was not necessary for these phenotypic studies. Furthermore, the fusion of GFP to a protein may have adverse effects upon protein folding, stability and function. Therefore, the GFP-encoding portion of the original pCA24N based expression plasmid was removed in a *Not* I digest. This yielded a small, 724 bp fragment corresponding to the GFP section and a large, variable sized fragment containing the remainder of the plasmid. The large, variable size fragment was successfully re-ligated for all 68 plasmids bearing ORFs of interest to this study. Over-expression plasmids, previously known as pCA24N-*xxxX* with the GFP portion removed were subsequently referred to as pMJA-*xxxX*.

5.3.4 Phenotypic Characterisation

Strains bearing knock-outs for genes that had their transcriptional profile significantly altered in any of the comparisons performed in Chapter 3, were analysed in an MIC assay. When a knock-out strain was available, MICs were performed on the knock-out strain (i.e. MG1655 Δ*xxxX*), an over-expression strain (i.e. MG1655 pMJA-*xxxX*) and on a complement of the knock-out strain (i.e. MG1655 Δ*xxxX* pMJA-*xxxX*). The results of these MIC measurements are shown in Table 5.1. Strains were categorised according to the function of the ORF of interest into the following groups: DNA/RNA associated, metabolism, stress, inner membrane, periplasm, outer membrane, flagella and fimbriae, rhs associated and those of unknown function.

Table 5.1 Phenotypic characterisation of strains based upon MIC assay. See footnote for details of scaling. Symbols in brackets indicate the transcriptional profile in the standard analysis i.e. (↑↔↔↔↔↓) denotes expression significantly up regulated in Comparison 1, no change in Comparisons 2, 3 and 4, and significantly down regulated in Comparison 5.

	MIC[a] for strain in which indicated gene was:		
Gene	Knocked out	Knock-out Complemented	Over-expressed
DNA/RNA-associated			
hns (↔↔↔↔↓↔)	− − −	− −	+
stpA (↑↔↔↔↔↑)	n/a[b]	n/a	− − −
xseA (↑↔↔↔↔↔)	−	− −	− −
recA (↑↔↔↔↔↔)	−	+	+ + + +
recX (↔↔↓↔↔)	=	− −	+ + +
fis (↔↔↔↔↔↑)	n/a	n/a	=
intB (↑↔↓↔↔)	+ + +	+ +	+ + +
yebG (↑↔↔↔↑)	=	+ + +	+
Metabolism			
cysB (↑↔↔↔↑)	+ +	− − − −	− − −
cysK (↔↔↑↑↔)	−	− − −	− − −
cbl (↑↔↓↓↔)	− −	+ + + +	+
tnaA (↓↔↑↔↓)	n/a	n/a	=
tnaL (↓↔↑↔↔)	n/a	n/a	− −
aspA (↓↔↔↔↓)	+ + + +	+ +	− − − −
glpD (↑↔↓↔↔)	+ + + + +	− −	+ +
tdcR (↑↔↓↔↔)	=	−	− −
aceA (↔↔↑↔↔)	=	− − −	− − − −
Stress-associated			
dnaK (↑↔↔↔↔)	n/a	n/a	− − −
uspA (↓↔↔↔↔)	n/a	n/a	− −
gadX (↑↔↔↔↔)	+ + + + +	+ + + + +	+
gadW (↑↔↓↔↔)	−	+ + +	− −
Inner membrane-associated			
mglB (↔↔↔↔↓)	+	+	+
lldP (↔↔↔↔↓)	=	+ +	+
malK (↔↔↔↔↓)	+	− −	+ +
manX (↔↔↔↔↓)	+	−	+ + + +
gatC (↔↔↔↔↓)	=	+ +	− −
yicJ (↔↔↑↔↔)	−	−	− −
rbsD (↓↔↑↔↔)	n/a	n/a	− − −
evgS (↑↔↔↔↔)	+ +	− − − −	− −
Periplasm-associated			
cpxP (↑↔↓↔↑)	−	−	− −
hdeA (↔↔↔↔↓)	− −	− − −	+ +
hdeB (↓↔↔↔↓)	n/a	n/a	+ +
ydhA (↑↔↔↔↔)	=	− −	+

Table 5.1 Continued.

Gene	MIC[a] of strain in which indicated gene was:		
	Knocked out	Knockout-Complemented	Over-expressed
Outer membrane			
htrL (↑↔↓↔↔)	=	− − − −	+
rfbX (↑↔↓↔↔)	=	− −	+
osmB (↑↔↓↔↔)	+	+ + + +	+ + + + +
vacJ (↑↔↔↔↔)	n/a	n/a	− − − −
rfaL (↑↔↔↔↔)	n/a	n/a	+
yefI (↑↔↓↓↔)	n/a	n/a	+ +
rfc (↑↔↓↔↔)	n/a	n/a	+
Flagella and fimbriae			
flgJ (↓↔↔↔↓)	+	− − − −	=
flgE (↓↔↑↑↔)	n/a	n/a	=
fliC (↔↔↔↔↓)	− −	+	+
yadC (↑↔↓↔↔)	n/a	n/a	+
rhs element-associated			
Element A			
rhsA (↔↔↔↔↔)	n/a	n/a	+
yibA (↔↔↔↔↔)	n/a	n/a	− −
yibJ (↑↔↓↔↔)	n/a	n/a	=
yibG (↔↔↔↔↔)	n/a	n/a	+
Element B			
rhsB (↑↔↓↓↔)	n/a	n/a	+ + +
yhhH (↑↔↓↓↔)	n/a	n/a	−
yhhI (↔↔↔↔↔)	n/a	n/a	− −
Element C			
rhsC (↔↔↔↔↔)	n/a	n/a	=
b0702 (↔↔↔↔↔)	n/a	n/a	+ +
b0703 (↔↔↔↔↔)	n/a	n/a	−
ybfC (↔↔↔↔↔)	n/a	n/a	+ + +
ybfD (↑↔↓↔↔)	n/a	n/a	− − − −
Element D			
rhsD (↑↔↓↔↑)	n/a	n/a	+ +
ybbC (↔↔↔↔↔)	n/a	n/a	− −
b0499(↑↔↔↔↔)	n/a	n/a	+ +
ybbD (↑↔↓↔↔)	n/a	n/a	=
Element E			
rhsE (↑↔↓↔↔)	−	− − −	− −
ydcD (↑↔↓↔↔)	n/a	n/a	+
b1458 (↔↔↔↔↔)	n/a	n/a	=
b1459 (↔↔↔↔↔)	n/a	n/a	+ +
ydcC (↔↔↔↔↔)	n/a	n/a	− − − −
ydcE (↔↔↔↔↔)	n/a	n/a	+

Table 5.1 Continued

Gene	MIC of strain in whihc indicated gene was:		
	Knocked out	Knockout-Complemented	Over-expressed
Unknown function			
yaiN (↑↔↓↔↑)	n/a	n/a	=
ycgW (↑↔↓↔↑)	n/a	n/a	− − −
ygeV (↔↔↔↔↓)	n/a	n/a	−
b2863 (↑↔↔↓↑)	n/a	n/a	+ +
yhaB (↔↔↔↓↔)	n/a	n/a	=

[a] The MIC of MG1655 was 3.5 mg/l. The scale for the average MIC in three replicate experiments is as follows:

	MIC	≤1.25 mg/l	− − − − −
1.25 mg/l <	MIC	≤1.75 mg/l	− − − −
1.75 mg/l <	MIC	≤2.25 mg/l	− − −
2.25 mg/l <	MIC	≤2.75 mg/l	− −
2.75 mg/l <	MIC	< 3.25 mg/l	−
3.25 mg/l ≤	MIC	≤3.75 mg/l	=
3.75 mg/l <	MIC	< 4.50 mg/l	+
4.50 mg/l ≤	MIC	< 5.25 mg/l	+ +
5.25 mg/l ≤	MIC	< 6.00 mg/l	+ + +
6.00 mg/l ≤	MIC	< 6.75 mg/l	+ + + +
6.75 mg/l ≥	MIC		+ + + + +

[b] Knock-out strains that were not available for analysis are indicated by n/a.

5.4 Discussion

5.4.1 Potential Essential Genes: *yaiN, ycgW* and *ygeQ*

For three of the desired mutant strains, attempts to construct them were unsuccessful. The production of knock-out strains using the Datsenko and Wanner method had highly variable results. Arguably, the most difficult step was the production of a batch of cells highly competent for electrotransformation. This however, was achieved on three occasions (typically one in ten batches of cells were highly competent). Whereas the *cpxP* knock-out was constructed using the first batch of competent cells, all efforts to construct *yaiN, ycgW* and *ygeQ* knock-outs failed. This was continued through two more batches of highly competent cells (tested for competency using the *cpxP* knock-out DNA). Although not conclusive evidence, this suggests that *yaiN, ycgW* and *ygeQ* may be essential for cellular function and thus refractory to inactivation.

Indeed, since the completion of this experimental work, *yaiN* and *ycgW* have been designated in a different study as likely to be essential genes (Gerdes *et al.*, 2003). In this comprehensive screen of MG1655 (3,746 genes, ~87% of the total genes), genes were assessed on their 'essentiality for robust, aerobic growth in a rich tryptone-based medium'. The results of this genome-wide screen are generally consistent with previously published data. However, there are discrepancies between these results and those obtained by the Wisconsin Genome Project (e.g. 12% of the 1,379 mutants currently listed at the Wisconsin Genome Project are listed as essential) so further work is necessary to confirm essentiality (Gerdes *et al.*, 2003). Unfortunately, *ygeQ* could not be screened in this work due to 'technical difficulties' (Gerdes *et al.*, 2003). However, *b1228* has been designated as essential (Gerdes *et al.*, 2003), despite the isolation of a strain bearing this knock-out in W3110 in this study. Despite these discrepancies, the work of Gerdes *et al.* does lend some credence to theory that *yaiN* and *ycgW* (and possibly *ygeQ* too) are essential genes in *E. coli*.

5.4.2 Phenotypic Characterisation

Phenotypic characterisation of strains based on an MIC assay is a relatively easy method to determine if altering the expression of an ORF has an effect on

susceptibility to PHMB. However, as can be seen in Table 5.1 the results do not always follow a clear and defined pattern. Over-expression of an ORF in a sensitive knock-out strain does not always lead to an increase in resistance. Similarly, over-expression in a more resistant knock-out strain does not always lead to sensitivity. Altering the expression of a gene could have varying effects upon the phenotype observed depending on not only its function but also on its normal and altered expression levels. A normally functioning cell contains a highly integrated network of regulatory, metabolic and structural pathways. In a knock-out system, where there is absolutely no expression of the gene of interest, it is possible that other genes could compensate for the function of the missing gene product. However, applying the same logic, due to the nature of this highly integrated network, altering the expression of just one gene could cause a variety of indirect effects in many of the connecting pathways. Furthermore, normal regulatory controls in a cell allow optimum expression levels under normal conditions for an ORF. Once this regulation has been removed (by uncontrolled expression from an expression plasmid) over-expression of an ORF could prove to be detrimental to the cell when expressed to too high a level. In order to address this problem it would have been worthwhile to include a set of strains over-expressing ORFs which were shown to not be PHMB regulated in this study. This would determine whether over-expression of an ORF *per se* is enough to impose a burden that PHMB may add to to give lethality.

These considerations create problems for the interpretation of these results. However, what is important is noting whether an alteration in expression level of a gene has an effect on the phenotype, in some way, regardless of whether it follows the 'expected' pattern. If there is a phenotypic difference (i.e. a change from the normal) there is clearly something important about the function of the gene of interest in relation to PHMB susceptibility.

What is immediately apparent upon inspection of the results in Table 5.1 is that altering the expression of the ORFs tested here leads to a change in susceptibility to PHMB in the majority of cases. Not one of the 'sets' of knockout, complement and over-expression strains for an ORF is completely unchanged in susceptibility. Of the 40 strains over-expressing ORFs for which no knock-out strain was available, only 9 showed no change in susceptibility to PHMB (it is possible that their knock-outs and

their complement strains, if available, would show a phenotypic difference). This has implications for the results obtained in Chapter 3. Since altering the expression of the majority of ORFs has some effect, it is likely that the majority of the changes observed in Chapter 3 are physiologically relevant to the recovery from PHMB stress and are not indirect consequences of, for example, a slow down in growth phase or entry into a lag phase. Most genes appear to play some role in susceptibility to or tolerance of PHMB, hence their change in expression in cells exposed to PHMB. Since in these experiments the expression of only 1 gene is altered at a time, in an exaggerated and artificial way, the alteration of a large number of genes (as observed in Chapter 3), that either aid recovery from or alter susceptibility to PHMB, would lead to the rapid recovery from PHMB stress.

The results in Table 5.1, grouped according to cellular function, are now discussed in relation to the array data from Chapter 3. Further discussion on significance of collective results is deferred to Chapter 8.

5.4.2.1 General Metabolism

The strains associated with genes involved in cysteine metabolism present an interesting puzzle. CysB was induced immediately after exposure to PHMB (Comparison 1, Table 3.3). The *cysB* knock-out would, therefore, be expected to be sensitive to PHMB, and the complement and over-expressing strains resistant. The reverse appears to be true however. CysB is the master regulator for sulphur assimilation in *E. coli* (cysteine is the preferred source of sulphur) (Van der Ploeg *et al.*, 2001). It is a positive regulator of the cys regulon, which includes both *cbl* and *cysK* (two other cysteine associated genes changed in transcriptional profile in Chapter 3). The **CysB** like protein, encoded by *cbl*, is involved in an accessory regulatory circuit within the cys regulon (Iwanicka-Nowicka and Hryniewicz, 1995). The *cbl* knock-out strain was sensitive to PHMB, whereas the over-expression and complement strains were resistant. This would be the expected phenotype of strains expressing a gene associated with PHMB resistance, since *cbl* was induced immediately in response to the presence of PHMB (Section 3.4.8).

The *cysK* knock-out strain appeared to have no noticeable difference in sensitivity compared with the over-expression and complemented strains. All three strains gave MIC values below the normal wild-type levels. These lower MIC values could be indicative of interference with the final step in the cysteine biosynthesis pathway. *E. coli* can exploit two alternative metabolic pathways for cysteine biosynthesis in response to varying environmental conditions (Iwanicka-Nowicka and Hryniewicz, 1995). O-Acetyl serine (thiol) lyase (OASL) is present in two isoforms, A and B, which catalyse the synthesis of cysteine from OAS and sulphide. OASL-B also has the additional activity to use thiosulphate to synthesise S-sulphocysteine, which can then be reduced to cysteine(Van der Ploeg *et al.*, 1997). OASL-A is encoded by *cysK*, while OASL-B is encoded by *cysM*. The three genes, *cysK, cysB* and *cbl*, are clearly linked. CysB positively regulates both *cysK* and *cbl* (as well as negatively regulating its own gene, *cysB*). Cbl regulates both OASL gene isoforms, *cysK* and *cysM* (Van der Ploeg *et al.*, 1997). Since over-expression of *cbl* leads to resistance, and the over-expression of *cysK* caused sensitivity, it is possible that resistance mediated through *cbl* is occurring via an increase in *cysM* expression (whose expression profile does not alter in the array experiments). However, the increased sensitivity of the *cysB* over-expression strain (and the resistance shown by the *cysB* knock-out strain) complicates this idea. It is possible that Cbl and CysB could be regulating other genes involved in resistance and susceptibility to PHMB. Both have been shown to bind to the *tauABCD* and *ssuEADCB* operon promoters, but at different regions (Van der Ploeg *et al.*, 2001), and it is possible that these genes somehow provide a resistance mechanism to PHMB.

The *tdcR* gene encodes the positive regulator of the *tdc* operon (which encodes threonine dehydratase). Gene knock-out, complementation and over-expression made no noticeable differences to PHMB MIC values. All three strains were slightly more sensitive to PHMB than the wild-type.

The *glpD* gene encodes an aerobic, cytoplasmic, membrane-associated glycerol-3-phosphate dehydrogenase. The MICs of the knock-out, complemented and over-expression strains showed a similar pattern to that of the corresponding *cysB* strains (except the over-expression of *glpD* in the wild-type makes the strain more

resistant to PHMB). This, like the *cysB* strain phenotypes, is the reverse of what would be expected for a gene induced initially in response to PHMB (Table 3.3).

The *aceA* gene encodes isocitrate lyase which catalyses the conversion of isocitrate into succinate and glyoxylate. The *aceA* knock-out strain showed no difference in sensitivity to PHMB than the wild-type strain. However, when *aceA* was over-expressed in both the knock-out and wild-type strain, cells become more sensitive to PHMB. This is consistent with the suppression of *aceA* transcription in cells made bacteriostatic by adding PHMB, compared with fully recovered cells (Table 3.6).

The *aspA* knock-out was resistant to PHMB, and the over-expression of *aspA* in this knock-out led to a decrease in resistance (i.e. increased sensitivity). Over-expression of *aspA* in the wild-type strain caused cells to become very sensitive to PHMB. The *aspA* gene encodes an aspartate ammonia lyase, known as aspartase, which converts L-aspartate into fumarate and ammonia. The significance of the *aspA* response, which has been noted in other stressful conditions and often negeclted, is discussed more fully in Chapter 8.

5.4.2.2 DNA-associated

The *recA* knock-out was sensitive to PHMB and the over-expression of *recA* in this knock-out strain led to PHMB resistance. Over-expression of *recA* in the wild-type strain led to a stronger resistance to PHMB. RecA is involved in DNA replication, repair, restriction and modification, and it is a global regulator of the SOS regulon (Stohl *et al.*, 2002; Kim and Levin, 1990). The RecA protein triggers de-repression of the SOS regulon by the inactivation, by cleavage, of the LexA repressor protein (Volkert and Landini, 2001). The SOS response to DNA damage involves at least 31 genes (including *yebG*) encoding proteins involved in several types of DNA repair mechanism such as nucleotide excision repair, recombinational repair and polymerases involved in the mutagenic bypass replication of lesions (Volkert and Landini, 2001). The expression of *recA* is influenced by topological alterations in DNA (Urios *et al.*, 1990). If PHMB is disrupting the cytoplasmic membrane and gaining access to the cytoplasm it could interact with genomic DNA (see Chapter 6)

and cause damage. The induction of RecA could, therefore, help repair damage caused by PHMB or its activity in signalling damage may be altering the apparent phenotype.

Strains associated with the *recX* gene, encoding the regulator of RecA, were also tested for sensitivity to PHMB. RecX was known as OraA at the outset of this study, but has recently been renamed as RecX (Stohl *et al.*, 2002; Pages *et al.*, 2003). *recX* (*oraA*) was also identified as being significantly altered in Chapter 3 (See Comparison 3, Table 3.6). All three strains (knock-out, complement and over-expression) were sensitive to PHMB. The knock-out was slightly sensitive and the over-expression and complement more so. The regulation of *recA*/RecA by RecX is complex. Over-expression of *recX* leads to reduced RecA levels (Stohl *et al.*, 2002). Therefore, it is no surprise that the over-expression of (and complement by) *recX* leads to increased sensitivity to PHMB. Following on from this, it would be expected that the *recX* knock-out would be more resistant to PHMB (due to increased RecA levels) if regulation of *recA*/RecA by RecX were this simple. However, it has been suggested that at low levels RecX may actually enhance the stabilisation of the *recA* transcript or the RecA protein or even its activity (Stohl *et al.*, 2002) . The complete absence of RecX (as in the case of the knock-out strain) could therefore lead to an increased sensitivity to PHMB due to decreased RecA levels or activity.

The *yebG* knock-out strain showed no difference to the wild-type in sensitivity to PHMB. When *yebG* was over-expressed in this strain however, the cells became slightly resistant. When *yebG* was over-expressed in the wild-type strain, the cells became even more resistant. Expression of *yebG* has been shown to occur as part of the SOS response to damaged DNA and as part of the entry into stationary phase (Lomba *et al.*, 1997; Oh and Kim, 1999). The induction of *yebG* following exposure to PHMB may indicate, therefore, that damage had occurred to DNA or it may be induced as part of the transcriptional changes that occur as cells enter stationary phase (albeit an enforced stationary phase). Whereas the transcriptional control of *yebG* is partly characterised (it has also been shown to be dependent on H-NS and cAMP, but does not require RpoS), no biological function has been found, to date, for YebG (Oh and Kim, 1999). The relevance of its induction by PHMB, therefore, remains to be elucidated.

There were no differences among MIC values *xseA* knock-out, over-expression and complement strains. All 3 strains had MICs slightly below that of the wild-type. The *xseA* gene encodes an exonuclease VII which functions by degrading ssDNA. The *intB* knock-out, complement and over-expression strains were all more resistant to PHMB. IntB is a prophage integrase.

The *hns* knock-out strain was sensitive to PHMB, while the compliment was more resistant (but still more sensitive than the wild-type). However, the over-expression strain was more resistant to PHMB than the wild-type. H-NS (**h**istone-like **n**ucleoid **s**tructuring protein) is a global regulator and thus inactivating it would seriously perturb the general cellular regulation (Lammi *et al.*, 1984). Because H-NS is involved in the control of global expression, these results must be put into context. The apparent sensitivity in the *hns* knock-out strain could be caused by a de-regulation of H-NS regulated genes, but it could also be caused by a change in DNA structure. H-NS is a histone like protein that causes the condensation of DNA (hence the reason why it represses the transcription of most of the genes under its control). H-NS could itself have a direct protective role against PHMB, if indeed the latter is interacting with and damaging DNA (see Chapter 6). The resistance shown in the over-expression strain could be due to increased protection of gDNA caused by it having a more compact and super-coiled structure (thus, reducing the size of the target for PHMB). Since many of the target genes repressed by H-NS are induced by environmental stresses such as high osmolarity, extremes of temperature and pH, it would be thought that a *hns* knock-out could be more resistant to stresses (since these genes would be under less stringent repression). However in this case the reverse is true, adding further credibility to the idea that the direct effect of H-NS on gDNA structure could be a basis for resistance to PHMB.

The over-expression of *dnaK* and *stpA* caused increased sensitivity to PHMB. StpA is a homologue of H-NS, but whereas *hns* over-expression leads to resistance, *stpA* over-expression led to sensitivity. This provides further evidence for the difference in biological role of the *stpA* and *hns* gene products.

5.4.2.3 Flagella

The *flgJ* knock-out strain was slightly more resistant to PHMB in comparison with the wild-type strain. Over-expression of *flgJ* in this knock-out led to PHMB sensitivity. Interestingly, over-expression of *flgJ* in the wild-type strain gave no difference in sensitivity to PHMB. The *flgJ* gene is a member of a flagella-associated operon which includes *flgBCDEFGHIJKL*. It is likely that most, if not all, of the operon has had a decrease in expression after exposure to PHMB, but that only *flgE* and *flgJ* have been identified as being significantly altered. Reasons for this could include variation in transcript stability, primer efficiency in the reverse transcription step and hybridisation. Unfortunately, no other flagella-associated genes from this operon were available from the Wisconsin Genome Project at the time this study was performed.

The flagellum consists of a basal body, a hook and a filament (Nambu *et al.*, 1999). The basal body consists of a rod attached to an inner membrane associated MS ring (composed of FliF subunits) and two outer L and P rings (composed of FlgI and FlgH subunits, respectively). The rod (composed of FlgB, FlgC, FlgF and FlgG) crosses the inner membrane, periplasm and outer membrane. In addition to the structural genes required for the rod subunit, other genes are vital for correct rod formation; *flgJ* is one such gene. Whereas FliF and FlgH create the holes necessary for the rod to pass through the inner and outer membrane respectively, the muramidase FlgJ is essential for rod assembly through the periplasmic space. *flgJ* mutants have been shown to produce an MS ring lacking the rod and other flagellar substructures.

It is thought that FlgJ is responsible for the hydrolysis, and thus breakdown, of the tough peptidoglycan in the periplasm allowing space for rod assembly to occur (Nambu *et al.*, 1999). Presumably, over-expression of *flgJ* would lead to an increase in the breakdown of peptidoglycan in the periplasm. This could allow PHMB a more direct and easier route to the cytoplasmic membrane. The resistance of the *flgJ* knock-out strain could therefore arise due to a decrease in the breakdown of peptidoglycan. Alternatively, the lack of a flagellar structure could lead to an increased ability of cells to aggregate, and thus protect themselves, although no evidence for this has been observed in this investigation.

The filament (the major component of the flagellar in terms of mass) is made up of around 20,000 subunits of flagellin,a 51.2 kDa protein, encoded by *fliC* (Auvray *et al.*, 2001). The *fliC* knock-out was sensitive to PHMB, while the complement and over-expression strains were more resistant to PHMB than the wild-type. This is the opposite of the pattern expected from the array data (Table 3.9) which showed this gene suppressed in Comparison 5. However, *fliC* is not part of the main flagellar operon (it is transcribed separately) and was found to be down-regulated in only one of the triplicate experiments in Comparison 1, although *fliC* is down-regulated significantly in recovered cells.

It is possible that the down-regulation of *fliC* is not a direct response to provide resistance to PHMB, but as a way of conserving valuable resources. If the flagella are not being produced (because, for example, the hydrolysis of the peptidoglycan in the periplasm leads to an increased susceptibility to attack from PHMB), the creation of large amounts of flagellin would be a huge waste of energy and resources. It is easy to place a bias on flagella mediated resistance on the effects of FlgJ. However, it must be noted that the entire flagella operon may be down regulated. Switching off flagella-gene expression may have some other role in PHMB resistance.

5.4.2.4 Outer Membrane Associated

The *osmB* knock-out strain was slightly more resistant to PHMB than the wild-type strain. However, over-expression of *osmB* in this knock-out strain and in the wild-type led to strong resistance. Indeed, over-expression of *osmB* in the wild-type gave an MIC value more than twice that of the wild-type. OsmB is an 'osmotically inducible lipoprotein' whose function has remained unknown since its discovery (Gutierrez *et al.*, 1987). It is structurally similar to the major *E. coli* lipoprotein encoded by *lpp*. Expression of *osmB* has been shown to be dependent upon osmolarity and growth phase (Jung *et al.*, 1990). The expression triggered by entry into stationary phase is independent of the osmotic environment and is thought to caused by an entirely σ^S-dependent mechanism. The altered expression of *osmB* observed after exposure to PHMB could be caused by an alteration in osmotic status caused by loss of ions (e.g.

displacement of Mg^{2+} from the membrane) or by entry into the enforced stationary phase. Alternatively, since the over-expression of *osmB* leads to PHMB resistance, OsmB could be helping to stabilise the membrane in response to disruption caused by PHMB.

The *rfbX* knock-out strain had a similar MIC value to that of the wild-type strain. Over-expression of *rfbX* gave slight resistance in the wild-type, whereas over-expression of *rfbX* in the knock-out gave slight sensitivity to PHMB. This pattern was repeated for the equivalent *htrL* strains (except the sensitivity of the *htrL* complement strain was greater). Since *rfbX* and *htrL* are involved in O-antigen biosynthesis and *E. coli* does not express a functional O-antigen, the biological significance of these results remains elusive.

Five genes associated with membrane composition, whose expression was induced immediately upon exposure to PHMB were over-expressed in the wild-type strain. Four of these five over-expression strains were slightly more resistant to PHMB than the wild-type. These included three genes associated with the production of the O-antigen (*rfaL, yefI* and *rfc*). The fourth gene, *yadC*, is predicted to encode a fimbrial like protein. The over-expression of the fifth gene, *vacJ*, caused PHMB sensitivity. VacJ is a lipoprotein and it is possible that high levels of it could be disrupting membrane function. The role of the O-antigen in PHMB resistance clearly warrants further investigation.

5.4.2.5 Inner Membrane Associated

The *evgS* knock-out strain was more resistant to PHMB than the wild-type strain. Over-expression of *evgS* in the knock-out strain caused PHMB sensitivity. Over-expression of *evgS* in the wild-type caused PHMB sensitivity but to a lesser extent. These phenotypes, like the equivalent *cysB* strains, appear to be the opposite of what would be expected for an 'induced' gene in Comparison 1 (Table 3.3). The *evgS* gene is co-transcribed with *evgA* (which is found significantly induced in 2 of the three experiments in comparison 1, but unfortunately no knock-out was available in MG1655) (Aiso and Ohki, 2003). EvgS is an integral membrane histidine kinase. It is thought to respond to environmental signals, such as sulphate ions and temperature,

which cause its autophosphorylation (Utsumi *et al.*, 1994; Beier *et al.*, 1995). Autophosphorylation of EvgS allows the phosphorylation, and activation, of EvgA. The presence of EvgS is not necessary for phosphorylation of EvgA, and *evgA* over-expression in an *evgS* knock-out strain can still induce multiple drug resistance (Nishino and Yamaguchi, 2002). This could account for the lack of sensitivity shown by the *evgS* knock-out strain.

EvgA is a transcriptional regulator that has, so far, been shown to regulate gadABC, hdeAB, *yhiUV*, *emrKY* (multidrug transport systems) and 21 other genes of unknown function (Masuda and Church, 2002). The *emrKY* operon showed no sign of a change in expression profile in the standard analysis (Tables 3.3 to 3.10). However, *emrY* appeared to be up-regulated according to the SAM analysis of Comparison 1 (Appendix A, Table A.2). The *yhiV* gene was induced in two of three experiments in Comparison 1 (standard analysis). However, there was no sign of change in *yhiU* expression. Interestingly, *yhiUV* is found directly next to *yhiW* and *yhiX* (also known as *gadW* and *gadX*) which were both induced in Comparison 1. Clearly, this system has some form of involvement in the responses to PHMB and deserves further investigation.

5.4.2.6 Periplasm-Associated

The *cpxP* knock-out, complement and over-expression strains all show near identical sensitivity to PHMB which is higher than that of the wild-type. It would appear that too high or too low expression levels of CpxP is detrimental to a cell's resistance to PHMB. Since CpxP is a repressor of the Cpx response, its induction (as part of the Cpx response) would lead to a switching off of the system. It could be that in this case, a lack of CpxP could lead to detrimental effects caused by the permanent switching on of the Cpx response (hence the reason for the auto-regulation provided by CpxP) and when too much CpxP is present its repression of the Cpx response inhibits the expression of genes that are needed to combat the effects of PHMB.

The *ydhA* knock-out strain had similar sensitivity to PHMB as the wild-type strain. The over-expression strain was slightly resistant, whereas the complement was

slightly sensitive. YdhA is a conserved hypothetical protein whose function remains completely unknown.

The *hdeA* knock-out strain was sensitive to PHMB, while the over-expression strain was resistant. This would be the expected phenotypes if *hdeA* was induced in response to PHMB. However, *hdeA* and *hdeB* were repressed in response to PHMB (Tables 3.4 and 3.10). Over-expression of *hdeA* in the *hdeA* knock-out did not restore PHMB resistance. It appears that disruption of the chromosomal copy of *hdeA* had a detrimental and irreversible (by complementation) effect on cells. This could be due to the disruption of *hdeB* expression (which is co-transcribed with *hdeA*). The increased resistance to PHMB caused by over-expression of *hdeB* in the wild-type strain is consistent with this notion.

HdeA knock-outs have previously been shown to be acid sensitive and expression of *hdeA* in the knock-out strains restores acid resistance (Waterman and Small, 1996; Gajiwala and Burley, 2000). It is an interesting paradox that *hdeAB* were repressed in recovered cells, when *hdeAB* over-expression clearly provides a mechanism for resistance. Expression of *hdeAB* is repressed by H-NS (hence their name, **H**-NS **d**ependent **e**xpression). Since over-expression of *hdeA* and *hdeB* (except in the *hdeA* knock-out strain) led to PHMB resistance, their down-regulation could be an indirect effect caused by increased levels of H-NS in the cell. The protective effects of H-NS either structurally or at the transcriptional control level could outweigh the protective effects of having HdeA and HdeB present in the periplasm. The functions of HdeA and its structural homologue HdeB are not clear. HdeA has been suggested to act as a pH sensor that activates acid responsive systems or as a component of a proton pump (Gajiwala and Burley, 2000). A chaperone-like activity has also been shown by preventing aggregation of denatured proteins. It is currently thought that HdeA may bind to acid-denatured proteins in the periplasmic space and prevents irreversible aggregation (Gajiwala and Burley, 2000). Once the acid stress is relieved, HdeA could then allow the proteins to refold. However, HdeA was shown to have no influence on protein aggregation at neutral pH. These models do not account for the resistance to PHMB shown by *hdeA* and *hdeB* over-expression, since acidic conditions are not present.

The *gadX* (previously known as *yhiX*) knock-out and complement strains were strongly resistant to PHMB. The *gadX* over-expression strain was only slightly more resistant than the wild-type to PHMB. It would appear that the disruption of the chromosomal copy of *gadX* could have led to increased PHMB resistance. The over-expression of *gadX* has been shown to be detrimental to *E. coli*, which might explain why the over-expression strain is not as resistant as the complement (Tramonti *et al.*, 2002). The complement may express *gadX* to a biologically advantageous level, whereas the over-expression in the wild-type may express to a biologically disadvantageous level. This does not, however, explain why the *gadX* knock-out was resistant. GadX functions as a postitive transcriptional regulator of, among others, *gadA* and *gadBC* (part of the acid resistance response) (Masuda and Church, 2003). The control of regulation appears to be highly complex and is integrally linked to GadW. Recently a non-coding RNA has been found to be located directly between the *gadX* and *gadW* genes (Storz, Pers. Com.). It is possible that it is in fact this non-coding RNA that is important in the response to PHMB. Non-coding RNAs have been shown to have a wide range of functions including transcriptional and translational control (Gottesman *et al.*, 2001; Storz, 2000).

The *gadW* (previously known as *yhiW*) knock-out and over-expression strains were slightly more sensitive to PHMB than the wild-type. Over-expression of *gadW* in the *gadW* knock-out strain led to PHMB resistance. As mentioned above, *gadX* and *gadW* are found directly next to each other (*gadX* is also next to *gadA*). They were thought to be transcribed separately (Ma *et al.*, 2002). However, recently it has been suggested this may not be the case, and that *gadX* and *gadW* may be co-transcribed (Tucker *et al.*, 2003). GadW inhibits the activation of *gadA* and *gadBC* by GadX, but can induce *gadA* and *gadBC* itself when GadX is absent. Both proteins are capable of binding the same promoters. Whereas GadX and GadW form homodimers in vivo, they can also form heterodimers (although the interaction is much weaker). The physiological relevance of these heterodimers is unknown.

It appears that there is a fine balance in the levels of GadX and GadW needed to coordinate expression. Too much or too little of either GadX or GadW appears to lead to very different resistance phenotypes to PHMB. Both *gadX* and *gadW* expression levels appear to be important for PHMB resistance in *E. coli*. GadX has

been shown to regulate the expression of *ybaS, ybaT, gadBC, gadA, hdeAB, hdeD, yhiD, yhiE, slp, yhiF, yhiN, yhiM* (Tucker *et al.*, 2003). The case is further complicated by the discovery of the non-coding RNA found sandwiched between *gadX* and *gadW*. Whereas no function has been assigned to this RNA, it is likely to play some form of role in the regulation of the *gad* system.

5.4.2.7 Unknowns

Four genes whose functions were completely unknown were also over-expressed. Three of these (*yaiN, ycgW* and *ygeQ*) were genes that resisted conversion to knock-outs in the previous section (i.e. potentially essential and thus important genes). *yhaB* and *yaiN* over-expression gave no difference in sensitivity to PHMB, whereas over-expression of *ycgW* and *ygeQ* caused increased PHMB sensitivity and resistance, respectively.

5.4.2.8 Sugar Transport

The six sets of strains associated with genes involved in sugar transport showed varying and contrasting phenotypes. This is surprising since five of these genes (*mglB, manX, gatC, lldP* and *malK*) have their transcriptional profile significantly down-regulated in Comparison 5 (Table 3.9).
Both *yicJ* and *mglB* associated strains showed similar but opposite phenotypes (the *mglB* knock-out, complement and over-expression strain were all slightly resistant, whereas the corresponding *yicJ* strains were slightly sensitive). The remaining four sets of strains showed conflicting phenotypes. The *manX* knock-out was slightly resistant, the over-expression strain very resistant, but the complement slightly sensitive. The *gatB* knock-out had normal resistance, the over-expression strain was sensitive, whereas the complement was resistant. The *lldP* knock-out had normal resistance, the over expresser slight resistance and the complement stronger resistance. The *malK* knock-out was slightly resistant, the over expresser more resistant, whereas the complement was very sensitive. These variable results imply that the alterations in the expression of genes involved in sugar transport can have highly variable effects on PHMB resistance.

5.4.2.9 The rhs Elements

A typical rhs element contains a 3.7 kb GC-rich 'core' that extends a further 500 bp (approximately) into an AT-rich region known as the core extension (Wang *et al.*, 1998). This large, single ORF is known as the core ORF and can potentially encode a protein in excess of 150 kDa (Hill *et al.*, 1995). The core sequence encodes a predicted protein containing 28 repetitions of the peptide motif xxGxxRYxYDxxGRL(IorT)xxxx (Zhao et al., 1995). This motif is similar to the one repeated 31 times in the *Bacillus subtilis* wall associated protein, WAPA; Gxxxx(Y,F)xYDxxG (Foster, 1993). Each core ORF is followed immediately, or even over lapped by, a shorter, highly variable ORF, known as the downstream ORF (dsORF) (Wang et al., 1998). These dsORFs are AT-rich and most appear to have signal peptides, for export from the cytoplasm, in their N termini (Hill *et al.*, 1995). Several rhs elements also contain one or more partial repetitions of the core, located downstream of the main core. These fragments vary in size, but always include the variable 3' end of the core. These fragments are not always part of intact reading frames.

E. coli K12 contains 5 rhs elements (A to E) containing a total of nine core extensions. Five core extensions are associated with the five primary rhs cores and the remaining four are found in secondary 3' core fragments. All nine are distinctly different in sequence (Hill *et al.*, 1995). A great deal of confusion surrounds the annotation of the rhs elements and the ORFs associated with them. The rhs elements are named according to their physical location on the *E. coli* chromosome. However, there is variation among *E. coli* strains over presence, number and location of rhs elements. There is also variation in the core extensions associated with the main cores in different strains. Further complications arise by the lack of consistency in the naming of the same ORF in the same strain. For example, the 840 bp ORF found down stream of the *rhsA* core is referred to as ORF-2 (Feulner *et al.*, 1990), ORF-A2 (Zhao *et al.*, 1993) and as dsORF-A1 (Vlazny and Hill, 1995). Interestingly, these three papers have the same corresponding author. The sequencing of the *E. coli* chromosome (by two competing groups) and their subsequent computerised annotation of predicted ORFs has complicated matters further by predicting ORFs of

different sizes and at different locations within the rhs elements and by providing alternative names for all these ORFs.

Over-expression of the ORFs associated with the rhs elements provided no clear patterns for resistance to PHMB. Over-expression of the core ORFs led to PHMB resistance with *rhsB* and *rhsD*, no effect with *rhsA* and *rhsC*, and slight sensitivity with *rhsE*. *rhsE* is the only rhs element thought to be non-functional in *E. coli* K12 (Sadosky *et al.*, 1991; Wang *et al.*, 1998). Over-expression of the down stream ORFs also had variable results. For the *rhsA* element, over-expression of *yibA* led to slight sensitivity, whereas over-expression of the further downstream *yibJ* and *yibG* had no effect. As noted above, over-expression of the *rhsB* core provides resistance to PHMB, however, over-expression of the down stream *yhhH* and *yhhI* led to slight sensitivity. The reverse appeared to happen in the *rhsC* element. Over-expression of the core ORF had no effect, but over-expression of two of the three down stream ORFs (*ybfB* and *ybfC*) led to strong PHMB resistance. Over-expression of *b0499*, one of the three downstream ORFs associated with *rhsD*, led to PHMB resistance. Of the other two ORFs over-expression of *ybbC* gave PHMB sensitivity while *ybbD* had no effect. Over-expression of *ydcD*, *b1459* and *ydcE* of the *rhsE* element gave PHMB resistance, whereas over-expression of *ydcC* caused a strong sensitivity to PHMB.

There is no obvious pattern to the PHMB resistance shown by over-expression of the various rhs element ORFs. Since no expression has been seen before and no phenotype has previously been seen in knock-out strains, the function of the rhs elements and the ORFs contained within them remains a mystery.

The confusion surrounding the rhs elements creates problems when performing phenotypic studies on over-expression strains. Since the down stream ORFs may play an essential role in signalling and processing of the rhs core ORFs, it seems unlikely that over-expression of the core ORF *per se* would have any effect since it would be unable to get to its target destination. However, there is clearly an increase in resistance to PHMB when the r*hsB* and *rhsD* core ORFs are over expressed. The variation in susceptibility of strains over expressing the down steam ORFs complicates the case further. Since there also appears to be variation in which

dsORFs are associated with which core (see above) the case becomes even more complicated. It is possible that different dsORFs play roles in targeting different core proteins to different locations within the cell. This could account for the huge variations found in MIC values associated with these rhs element ORFs.

Chapter 6: PHMB-DNA interaction

6.1 Introduction

The lethal action of PHMB is considered to involve interaction at the cytoplasmic membrane to cause non-specific alterations in membrane permeability (see Section 1.3). However, the results from array experiments (Chapter 3) implied that, not only is the cytoplasmic membrane disrupted at bacteriostatic levels, but also a possible interaction between PHMB and the cytoplasmic contents could be occurring. The induction of genes associated with the SOS response would suggest that damage to DNA could be occurring. Moreover, the phenotypes observed by DNA/RNA associated gene knock-out, complement and over-expression strains in Chapter 5 indicate that alterations in nucleic acid metabolism can have large effects on susceptibility to PHMB. Furthermore, previous work has identified clusters of densely stained precipitates in *Acanthamoeba castellanii* treated with high concentrations of PHMB, and also noted that PHMB treatment produced increased amounts of phosphorus inside the cells compared with untreated controls, and that these accumulations were often confined to cell walls and nuclei (Khunkitti *et al.*, 1998a; Khunkitti *et al.*, 1998b; Khunkitti *et al.*, 1999). Reduced membrane permeability causing retention of phosphorus, coagulation of proteins, and aggregation of phospholipids, have been considered as possible causes of elevated phosphorus but the possibility of association between PHMB and nucleic acids has not been not considered hitherto, despite rapid growth in literature in the last 20 years on the interaction of DNA with the natural polyamines, putrescine, spermine and spermidine (Wallace, 2003; Cohen, 1998). Furthermore, a variety of other synthetic polycationic compounds are currently being developed as vehicles for non-viral transfection of DNA into cells for therapeutic purposes (Vijayanathan *et al.*, 2002).

This chapter describes experiments aimed at revealing the *in vitro* interactions between PHMB and various nucleic acids.

6.2.1 Materials and Methods

Nucleic acids used in PHMB-nucleic acid interaction experiments are described in Section 2.6. PHMB (Vantocil) and fluorescent PHMB are described in Section 2.1.2.

6.2.2 Nucleic Acid Precipitation

Precipitation experiments were performed as described in Section 2.6.1. Extinction coefficients of nucleic acid and PHMB were determined at 236 nm and 260 and the concentrations of nucleic acid and PHMB remaining in solution were estimated using simultaneous equations.

6.2.3 Fluorescence Polarisation

Fluorescence polarisation experiments were performed as described in Section 2.6.2. Polarisation index was calculated as $p = (F_v - F_h)/(F_v + F_h)$ where F_v and F_h were the fluorescence intensities in the vertical (parallel to the excitation plane) and horizontal planes respectively. The polarisation index of free fluorescent-PHMB, p_f, and of fully bound fluorescent-PHMB, p_b, were determined for fluorescent-PHMB in the absence of nucleic acid, and in the presence of excess nucleic acid, respectively. The respective molar fluorescence constants Q_f and Q_b (where $Q = (F_v + F_h)/$[molar concn.]) of free and bound forms were determined in a similar way. These constants and the observed value of p for a given mixture enabled the calculation (Dandliker *et al.*, 1981) of the ratio, R, of bound to free fluorescent PHMB as

$$R = \frac{Q_f}{Q_b}\left(\frac{p - p_f}{p_b - p}\right)$$

Finally, since $R = L_b / L_f$, and $L_b + L_f$ = the total concentration of fluorescent-PHMB present ($[PHMB]_{tot}$), L_b and L_f were calculated using Equations 1 and 2 respectively.

Equation 1:

$$L_b = [PHMB]_{tot}\left(\frac{R}{1+R}\right)$$

Equation 2:

$$L_f = [PHMB]_{tot}\left(\frac{1}{1+R}\right)$$

6.3 Results

6.3.1 Nucleic Acid Precipitation with PHMB

Titrating dsDNA (100 bp, 2.5 μM strands, i.e. 0.25 mM nucleotide phosphate residues) with increasing concentrations of PHMB caused the concentrations of free PHMB (measured as dissolved monomer) and of dissolved DNA to change as shown in Figure 6.1a. At the equivalence point (i.e. addition of 0.25 mM PHMB biguanide units to 0.25 mM nucleotide phosphate residues), residual concentrations of DNA and PHMB were about 5% and 10% of totals added, respectively, indicating extensive precipitation of each from solution. The straight line of unit slope shows the expected concentration of dissolved PHMB, assuming no loss by precipitation. The extent to which measured values fell below this line during the titration shows the loss of PHMB by precipitation, which increased modestly as PHMB was added. At the same low concentrations of added PHMB, the concentration of dissolved DNA was similarly reduced by precipitation. However, as concentrations of added PHMB approached 0.2 mM further equal increments in PHMB produced marked reductions in both dissolved DNA and dissolved PHMB. These results indicated a strong complex formation between DNA and PHMB that appears virtually complete when the total PHMB concentration (as monomer) was 0.25 mM. Figure 6.1b shows concentrations of the DNA-PHMB complex measured as either precipitated-DNA (initial concentration minus measured dissolved concentration) or precipitated-PHMB (total PHMB added minus measured dissolved concentration), each as a function of added PHMB. Close correspondence between the curves in Figure 6.1b supports the notion of complex formation between DNA and PHMB. The plot of precipitated-PHMB concentration against precipitated-DNA concentration (not shown) was linear with a slope of 106, indicating 106 PHMB biguanide units associated per 100 bp, i.e. equivalence between nucleotide units and biguanide units in the complex. Moreover, the sigmoidal shape of the curves indicated a cooperativity in the binding of PHMB to DNA, i.e. initial binding of PHMB to DNA promoted stronger binding of further PHMB.

Similar curves were obtained when double-stranded DNA was replaced by single-stranded DNA of the same length and at the same concentration of strands (Figure 6.2a). As with dsDNA, there was a close correspondence between precipitated-DNA and precipitated-PHMB, with 98 biguanides bound per 100-base strand, i.e. 0.98 biguanides per nucleotide residue.

To determine whether DNA fragment-size was critical in this phenomenon, a mixture of fragments of dsDNA derived by *Hin*d III digestion of Lambda DNA was used. The sample (a DNA molecular size marker for calibrating electrophoresis gels) contained an equimolar mixture of eight fragments ranging in size from 125 to 23130 bp (see Section 6.2.1). For the titration with PHMB, the concentration of digested DNA was adjusted so that the total nucleotide concentration (0.2mM, corresponding to 2.1 nM of each double-stranded fragment) was similar to that used in the previous experiments (0.25 mM). Titration curves (not shown) were very similar to those shown in Figure 6.1a, and with the addition of PHMB equivalent to a total of 0.25 mM monomer, the residual soluble concentrations of PHMB and DNA were 10% and <5% respectively. The corresponding complex-formation curves (Figure 6.2b) showed the now-familiar close equivalence between precipitated-DNA and precipitated-PHMB with 1.18 biguanides bound per nucleotide residue, and the sigmoidal form characteristic of cooperativity. These results indicated that complex formation via a cooperative binding of PHMB to DNA occurred with a range of sizes of DNA.

Addition of PHMB concentrations in the range equivalent to 0-0.075 mM biguanides to tRNA produced no reduction in soluble tRNA, and no disappearance of any added PHMB (Figure 6.3a), indicating no precipitation under these conditions. In the added PHMB concentration range 0.1-0.2 mM, there was a slight reduction in dissolved concentration, but at higher concentrations (around 0.2 mM PHMB monomer, 80% of the total nucleotide concentration) there was a very marked precipitation of all tRNA over a narrow range of PHMB concentration (0.2-0.25 mM). Complex formation curves (Figure 6.3b) measured as either precipitated-tRNA or precipitated-PHMB were very similar (with 0.95 biguanide units bound per nucleotide residue, assuming the length of tRNA as the modal value of 72, see Section 6.2.1), and showed very strong cooperativity.

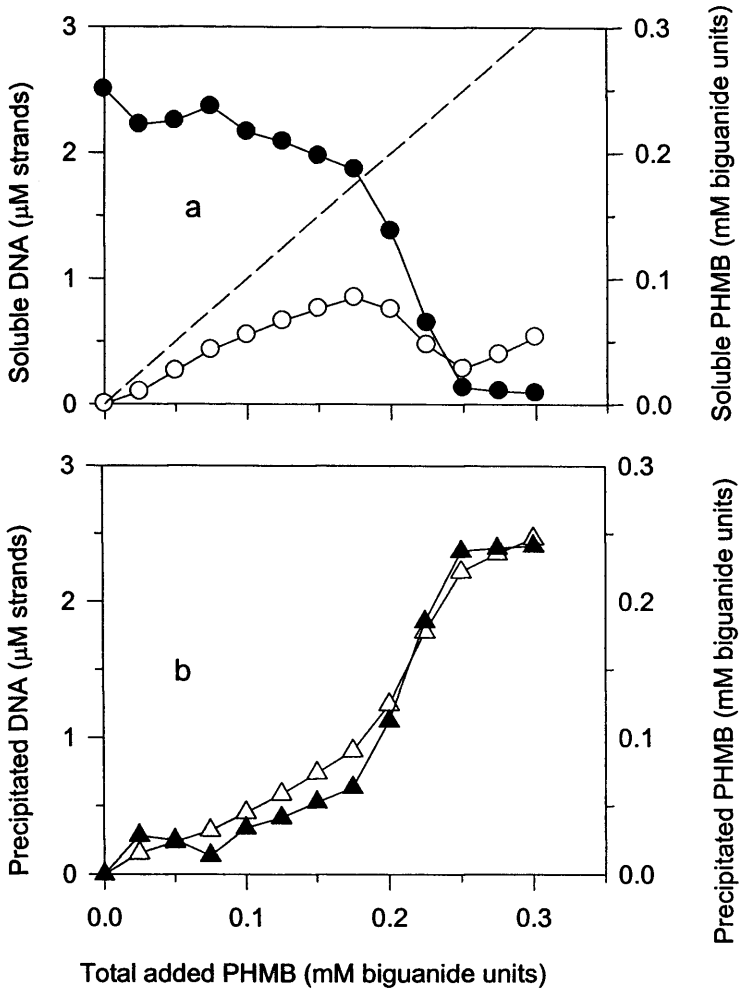


Figure 6.1 Titration of dsDNA with increasing concentrations of PHMB to form insoluble complexes. Initial total concentration of dsDNA (100 bp) was 2.5 μM strands (i.e. 0.25 mM nucleotide phosphate residues). PHMB concentration is expressed as molarity of biguanide units. (a) Concentrations in solution: filled circles, free (soluble) dsDNA; open circles, free (soluble) PHMB. Broken line, line of unit slope, i.e. expected soluble PHMB in the absence of any precipitation. (b) Concentrations lost by precipitation: filled triangles, complexed-DNA calculated as [total initial DNA minus measured soluble DNA]; open triangles , complexed-PHMB calculated as [added PHMB minus measured soluble PHMB].

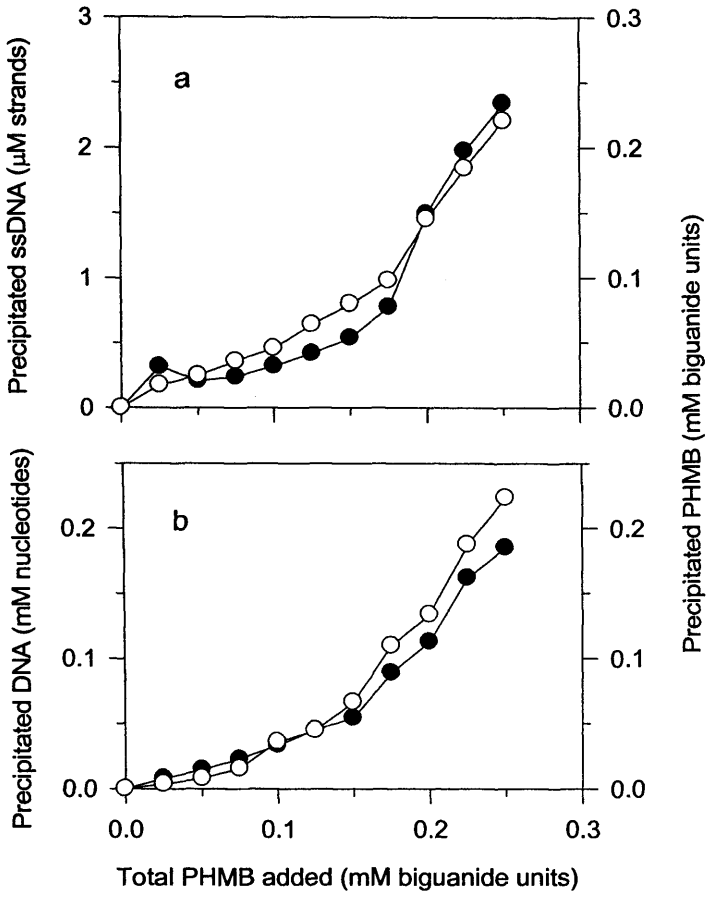


Figure 6.2 PHMB-nucleic acid complex formation during titration of (a) ssDNA, (b) HinD III-digested lambda DNA, with PHMB. Filled symbols, concentration of complexed nucleic acid calculated as the decrease in soluble DNA; open symbols, complexed PHMB calculated as difference between added PHMB and measured soluble PHMB.

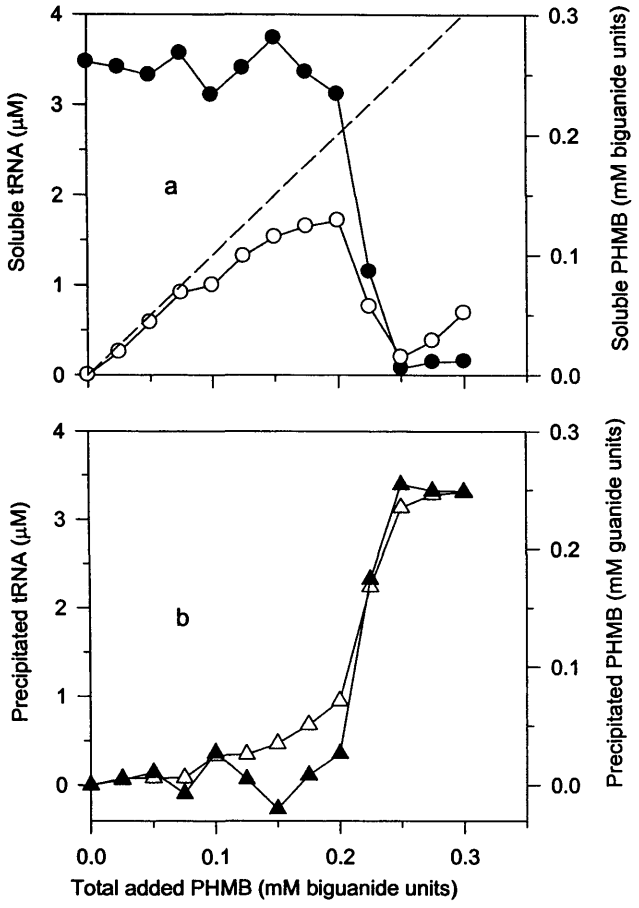


Figure 6.3 Titration of tRNA with increasing concentrations of PHMB to form insoluble complexes. Initial total concentration of yeast tRNA was equivalent to 0.25 mM nucleotide phosphate residues). (a) Concentrations in solution: filled circles, free (soluble) tRNA; open circles, free (soluble) PHMB. Dotted line, soluble PHMB in the absence of tRNA. (b) Concentrations lost by precipitation: filled triangles, complexed-tRNA calculated as [total initial tRNA minus measured soluble tRNA]; open circles, complexed-PHMB calculated as [added PHMB minus measured soluble PHMB].

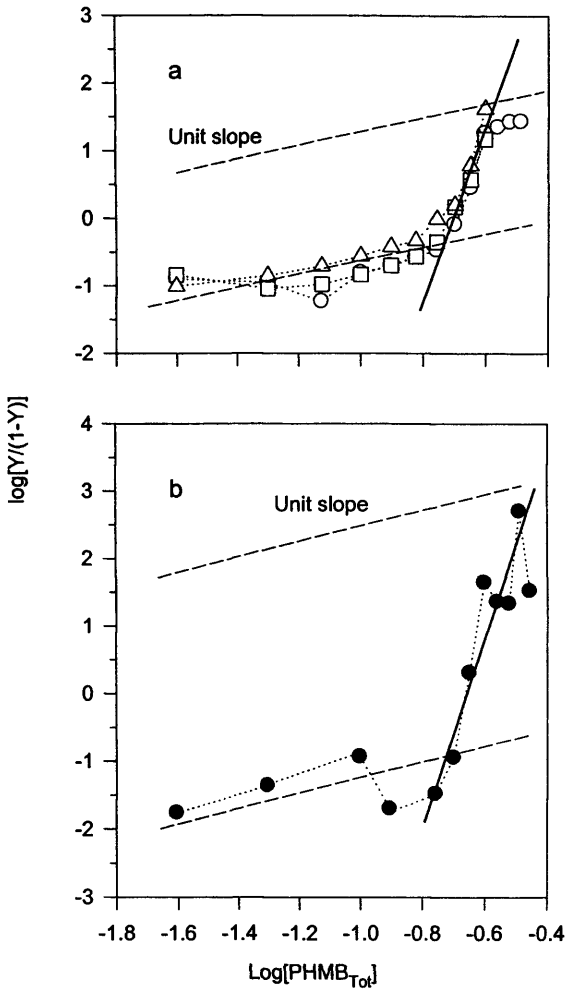


Figure 6.4 Hill-type plots for binding of PHMB to (a) dsDNA (open circles), ssDNA (open triangles) and Hind III-digested Lambda DNA (open squares), and (b) tRNA. Data are based on results in Figures 6.1b, 6.2 and 6.3 and equation 3. Dashed straight lines are lines of unit slope; solid straight lines are lines of linear regression through points in the central section of the plots where $\log[PHMB]_{Tot}$ has values between -0.7 and -0.6. Slopes of these lines correspond to *h* in equation 3. Values of *h* are shown in Table 6.1. See text for details.

6.3.2 Fluorescence Polarisation of DNA by PHMB

When small fluors are excited with plane-polarised light, rapid rotation of the molecules prior to emission results in fluorescence emission at all angles of rotation. When the fluor becomes attached to larger molecules that rotate more slowly, the plane of emission becomes progressively confined to the plane of excitation. Polarisation of fluorescence into this plane allows an assessment of binding to large molecules.

In order to investigate interactions between PHMB and DNA at lower concentrations at which precipitation did not occur, polarisation of the fluorescence from fluorescent-PHMB during its titration with nucleic acids, was measured (Figure 6.5a). Addition of increasing concentrations of either DNA or tRNA produced increasing polarisation, indicating progressive restriction of rotation of fluorescent-PHMB molecules, consistent with its binding to the much larger nucleic acid molecules. Control experiments using the same procedure but in which (a) water replaced the stock solutions of nucleic acid, and (b) free fluor (1,8-naphthalimide) replaced fluorescent-PHMB, produced no increase in the baseline polarisation (Figure 6.5a).

Figure 6.5b shows values of concentration of bound-PHMB (calculated from values of p as described in the Experimental section, and expressed as biguanide unit equivalents) plotted against total nucleic acid added (expressed as nucleotide equivalents added). For DNA, the titration curve closely followed the line of unit slope up to 10 μM DNA nucleotides, i.e. one biguanide unit was bound for every nucleotide unit added. At higher concentrations of added DNA, the number of bound biguanide residues exceeded the number of nucleotide residues added, so that almost all 25 μM biguanide units present were in complex form when only about 18 μM nucleotides had been added.

For tRNA, at concentrations of added tRNA nucleotides up to 17.5 μM, the amount of complexed PHMB fell below the line of unit slope, showing that only about 60-70% of nucleotides were complexed with biguanide units. However,

between 17.5 μM and 22.5 μM nucleotides, the binding of biguanides strongly recovered to the same point as with DNA, i.e. all PHMB biguanide units bound to less than an equivalent amount of nucleotides.

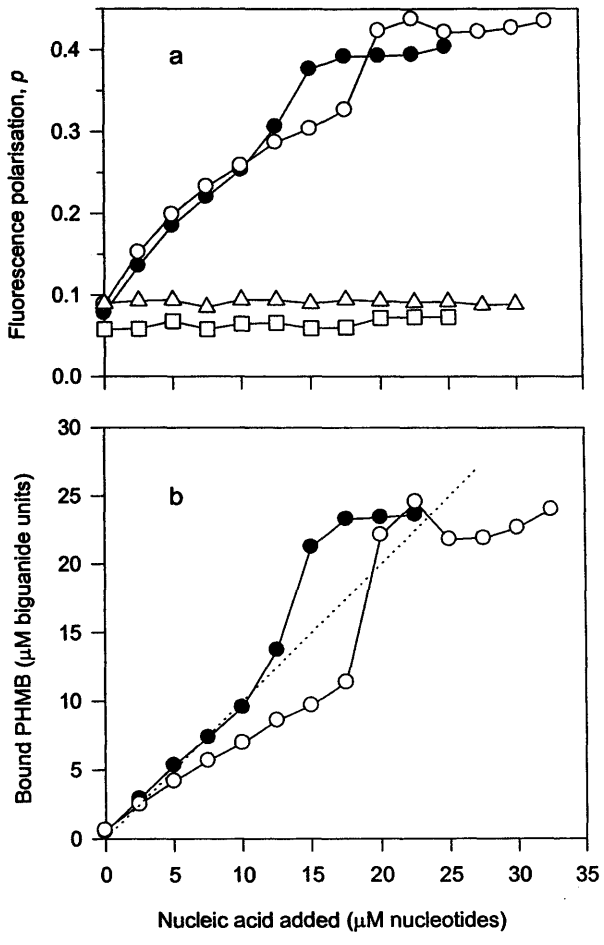


Figure 6.5 Fluorescence polarisation of fluorescent-PHMB during titration with DNA (filled circles) or tRNA (open circles). (a) Polarisation index was calculated as $p = (F_v - F_h)/(F_v + F_h)$ where F_v and F_h were the fluorescence intensities in the vertical (parallel to the excitation plane) and horizontal planes respectively. Aliquots (3 μl) of either DNA or tRNA stock solution were added to fluorescent-PHMB, and fluorescence measurements made at 30°C between additions (see text for details). In the control experiments, either the stock nucleic acid solutions were replaced by the same aliquots of water (open squares), or the fluorescent-PHMB was replaced with 1,8-naphthalimide at the equivalent concentration (open triangles). See text for details. (b) Bound PHMB was calculated from p as described in the text. The dotted line is the line of unit slope.

Table 6.1 Measures of cooperativity for binding of PHMB to nucleic acids to form a precipitated complex. [a] Value for non-cooperative binding is 1; values increase as cooperativity increases.

Nucleic acid		Apparent Hill coefficient [a]
Type	Size, characteristics	
DNA	100 bp duplex	14.8
DNA	100 bases, single strand	10.3
DNA	Mixed sizes, 125-23130 bp	14.6
tRNA	Mixture from yeast, modal length is 72 bases	14.5

6. 4 Discussion

At micromolar concentrations of polymeric PHMB and nucleic acid molecules, single-stranded DNA, double-stranded DNA of different lengths, and tRNA all interacted with PHMB to remove a proportion of both species (nucleic acid and PHMB) from solution as a precipitate. Given the polyanionic nature of the (deoxy)ribose phosphate backbones, and the polycationic nature of PHMB (see Figure 1.3), electrostatic interaction is likely to be the dominant factor in complex formation. This is consistent with the parallel precipitation of biguanides and nucleotide phosphates, invariably on a 1:1 basis throughout the titrations (Figure 6.1-6.3).

Recent work has shown that the optimum ratio of polyethyleneimine/DNA needed to maximise gene translocation into cells corresponded to a N/P ratio of 4.5 (Lee *et al.*, 2003). The spacing between positive charges on polyethyleneimine is 3 bonds, making for a spacing between phosphate-bound nitrogens of $4.5 \times 3 = 13.5$ bonds on average. Thus for this polymeric cation, 13.5-bond loops, on average, between phosphate-bound cationic groups optimize translocation. Examination of the structure of PHMB (Figure 1.3) shows that there are in fact 13 C-N or C-C bonds between successive positively charged imino groups. Thus it may be that the "intra-cation spacing" in PHMB is already optimized for DNA binding.

The sigmoidal form of the curves in Figures 6.1b, 6.2 and 6.3b shows that initial binding of PHMB promotes the stronger binding of further PHMB. The occurrence of this cooperativity with ssDNA shows it does not arise simply from increased PHMB binding following separation of DNA strands. In enzyme kinetics, cooperativity is assessed using the Hill equation, of which one form is given by Equation 3:

$$\log (Y/(1-Y)) = h \log [L]_T + h \log K$$

where Y is the fraction of occupied sites, $[L]_T$ is the total concentration of substrate (ligand) added, K is a constant and h is the Hill coefficient, a measure of cooperativity. Applying this to the present results, Y/(1-Y) was calculated as $([NA]_T - [NA]_f) / [NA]_f$ where $[NA]_f$ was the concentration of free nucleic acid in

solution, and $[NA]_T$ was the total initial nucleic acid concentration in solution. Hill plots of log (Y/(1-Y)) against log [substrate] are typically sigmoidal with the central part approximating to the linear Equation 1. Figure 6.4 shows the plots of log (Y/(1-Y)) against log $[PHMB]_T$ for the four types of nucleic acid used in the precipitation study, and the apparent Hill coefficients, measured as the slopes of the central regions of these plots are in Table 6.1. These values are particularly high; for enzymes, values are always well below 10 and typically in the range 1-3. Thus, the precipitation arising from binding of PHMB to DNA is strongly cooperative.

Although the Hill equation is based loosely on deviations from a hyperbolic saturation curve, it is essentially empirical and constants derived using it should not be interpreted in terms of any specific mechanism of cooperativity (Cornish-Bowden, 1995). Nevertheless, values of the Hill coefficient may be taken as an upper limit on the number of binding sites. For the defined length DNA (100 bp), 10.3-14.8 PHMB binding sites per 100 bp molecule (Table 6.1) equates, for fully complexed DNA, to an average PHMB length of 9.7-6.7 biguanides, which maps within the known range of 2-15 but is a little higher than the mean of 5.5. For tRNA with 14.5 binding sites per 72 base molecule (Table 6.1) complete occupancy of sites requires, on average, 4.96 biguanides per PHMB molecule, very close to the known average length. This analysis strengthens the hypothesis that PHMB-nucleic acid binding is based on a 1:1 association between biguanide and phosphate residues.

A possible mechanism that could account for cooperativity is illustrated in Figure 6.6. At low concentration of PHMB (Figure 6.6a), available sites on the nucleic acid significantly outnumber the biguanides added, and there is little precipitation. As PHMB concentration increases (Figure 6.6b), the net charge on complexed nucleic acids molecules progressively decreases allowing nucleic acid molecules to associate, causing some precipitation. At higher concentrations of PHMB, available nucleotide sites become limited and PHMB molecules find sufficient sites only by bridging between nucleic acid molecules (Figure 6.6c). As the cross-linked complexes become more extensive, their ability to bind further PHMB is enhanced, making the complex progressively less soluble.

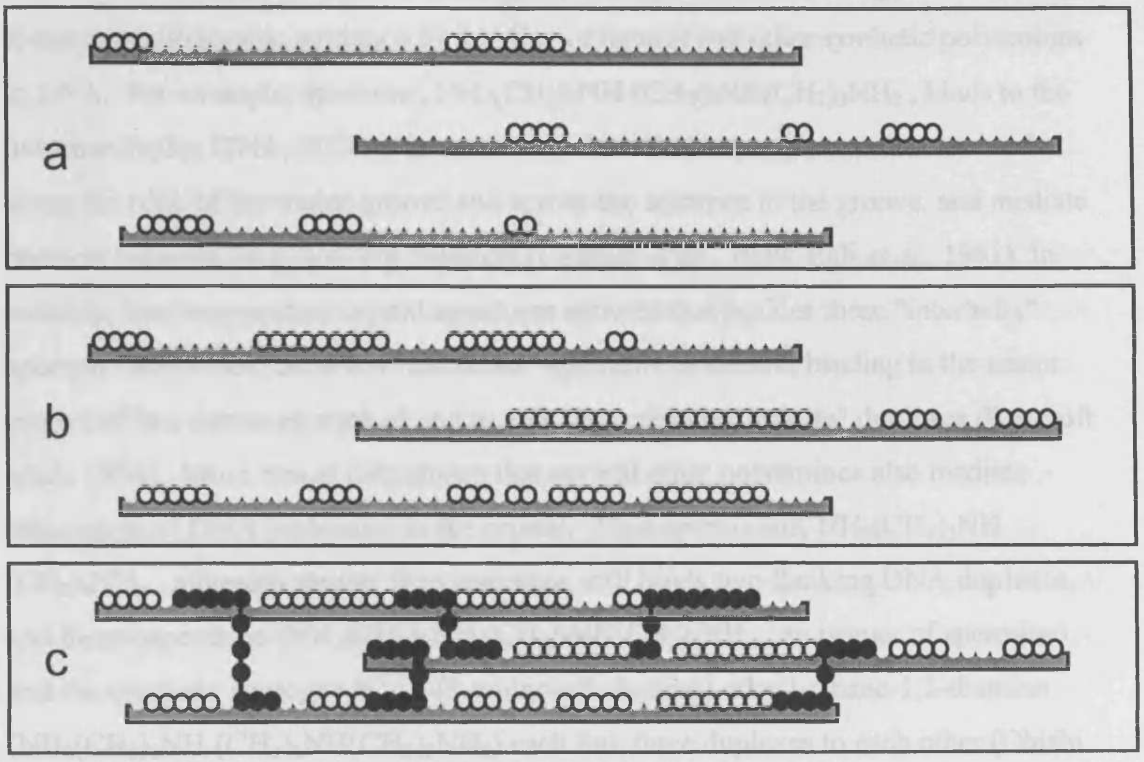


Figure 6.6 Proposed mechanism for precipitation of nucleic acid by addition of PHMB. Long grey bars are nucleic acid molecules, each cup-shape representing a nucleotide phosphate residue capable of binding positively charged biguanide groups (circles) that occur in oligomers from $n = 2$ to 15. Filled circles represent PHMB molecules that bridge between nucleic acid molecules. (a) PHMB at low concentration occupies a small fraction of the available sites on nucleic acid. At intermediate concentrations (b), PHMB causes significant decrease in net charge and promotes association of nucleic acid molecules. At higher concentration of PHMB (c), PHMB molecules find sufficient binding sites only by bridging across nucleic acids. See text for details.

The notion that PHMB can bridge between DNA strands is supported by X-ray crystallographic evidence for binding of natural and other synthetic polyamines to DNA. For example, spermine, $NH_2(CH_2)_3NH\ (CH_2)_4NH(CH_2)_3NH_2$, binds to the hexamer duplex DNA, $d(CG)_3$, in several modes. Some spermine molecules bind along the edge of the major groove and across the entrance to the groove, and mediate contacts between neighbouring duplexes (Gessner *et al.*, 1989; Egli *et al.*, 1991). In addition, low temperature crystal structures showed that besides these "interhelix" spermine molecules, there are "intrahelix" spermine molecules binding in the minor groove of two duplexes stacked end to end, thus creating "infinite" duplexes (Bancroft *et al.*, 1994). More recent data shows that several other polyamines also mediate association of DNA molecules in the crystal. Thus spermidine, $NH_2(CH_2)_3NH\ (CH_2)_4NH_2$, although shorter than spermine, still binds two flanking DNA duplexes, and thermospermine ($NH_2(CH_2)_3NH\ (CH_2)_3NH(CH_2)_5NH_2$, an isomer of spermine) and the synthetic analogue N^1-[2-(2-amino-ethylamino)-ethyl]-ethane-1,2-diamine ($NH_2(CH_2)_2NH\ (CH_2)_2NH(CH_2)_2NH_2$) each link three duplexes to each other (Ohishi *et al.*, 1996a; Ohishi *et al.*, 1996b; Ohishi *et al.*, 2002a; Ohishi *et al.*, 2002b). If these comparatively short polyamines can bridge between duplexes then there is every possibility that the long PHMB molecules, with on average more cationic groups per molecule, will do so equally, if not more, effectively.

Fluorescence polarisation experiments in which fluorescent-PHMB was titrated with DNA showed that, approaching the mid-titration point, the amount of biguanide converted to the complex form was equal to the total amount of nucleotides added (Figure. 6.5b). This implies that every nucleotide site on the added DNA molecules was associated with a biguanide residue, and is consistent with a strong association between DNA and PHMB and with the 1:1 binding ratio deduced from the precipitation experiments. Remarkably, further increments of DNA were able to bind more than an equivalent amount of biguanides. It is possible that in this later stage, all nucleotides sites are fully occupied and some molecules of PHMB are bound to DNA through some but not all of the biguanides present in the molecule, i.e. PHMB molecules are bound only along a part of their structure, with "tails" into the medium. This would effectively immobilise more biguanides than there are nucleotides present. This is analogous to the situation in Figure 6.6 where only part of some PHMB molecules is attached to a given DNA molecule.

For the first part of the titration of PHMB with tRNA, the biguanides bound fell short of the nucleotides added, by about 30%. This implies that not all nucleotides in tRNA are readily accessible to the PHMB and reflects the more extensive secondary structure in tRNA compared with DNA, and in particular the occlusion of some phosphate groups. The partially complexed tRNA molecules must carry a residual negative charge and this presumably accounts for the absence of precipitation of tRNA by even quite high concentrations of PHMB (Figure 6.3). However at about 17.5 μM nucleotides (equivalent to about 70% of biguanides present), the binding of PHMB suddenly strengthened (Figure 6.5b) and PHMB was fully complexed before the full equivalent of nucleotides was added. This extreme cooperative behaviour paralleled the results of the precipitation experiments (Figure 6.3) and it may arise from sudden degradation of the tRNA secondary structure caused by the progressively stronger interactions with PHMB, thus making more phosphate accessible for PHMB-binding.

Hitherto, mechanisms for the biocidal action of PHMB have focussed primarily on the readily observable changes occurring at the cell envelope, including loss of LPS, changes in membrane integrity and loss of function of membrane proteins (see Section 1.3). A number of cytoplasmic changes have also been observed (Khunkitti *et al.*, 1998a), but these have not yet been incorporated into a mechanistic theory. Nevertheless, they clearly imply ingress of PHMB to the cytoplasm, and this, together with the results presented here, that PHMB can bind tightly and cooperatively to DNA and RNA, opens a clear possibility that the binding of PHMB to nucleic acids inside the cell may be an important contributor to the bacteriostatic as well as the biocidal action of this compound. This prospect is lent credence by the well-known interaction of DNA with natural polyamines (spermine, spermidine) that achieves significant compaction of DNA and as a result has important implications in all aspects of cell biology, including cell growth and division, replication and repair of DNA, synthesis and processing of RNA, and protein synthesis (Cohen, 1998). Moreover, in enterobacteria numerous nucleoid-associated proteins also bind very tightly to DNA and contribute to its compaction and organisation. PHMB-nucleic acid complex formation might also account for the observed aggregation of

phosphorus near the cell wall and nuclei in PHMB-treated *Acanthamoeba* (Khunkitti *et al.*, 1999).

Given that transcriptional regulation is paramount in enabling prokaryotic adaptation, growth and survival in the face of ever-changing and potentially hostile environments, the possibilities for PHMB to interfere in cell-function by binding to both DNA and RNA, are hard to ignore. For example, if PHMB-nucleic acid binding is indeed an important feature of its biocidal action, systems involved in DNA compaction and gene-regulation such as H-NS (histone-like nucleoid structuring protein), the SOS response and DNA protection and repair are highly likely to be affected by the presence of PHMB.

Furthermore, the dosage dependent effects of PHMB observed previously (Figure 3.1; Davies *et al.*, 1968; Broxton *et al.*, 1983) may even be caused by the highly cooperative nature of the PHMB-DNA binding interaction. The change from bacteriostatic to bactericidal effects by small increments in PHMB concentration may be caused due to a major increase in nucleic acid precipitation thus leading to cell death.

Chapter 7: Comparison of Microarrays and Macroarrays

7.1 Introduction

Array technology has been shown in this study to be a powerful tool for measuring differences between populations of cells. The commercial availability of both microarrays and macroarrays has allowed an extensive study of the transcriptional profile of *E. coli* to a variety of environmental stimuli. An increasing volume of literature is focusing on the statistical issues associated with the analysis of the large volumes of data produced by such experiments (Nadon and Shoemaker, 2002; Brody *et al.*, 2002; Zhang, 2002). Other key issues such as image quantification, reproducibility, data transformation, data normalisation and interpretation have also been addressed (Coombes *et al.*, 2002).

In order to address these problems, the Microarray Gene Expression Data Society (MGED) was formed in 1999 (Brazma *et al.*, 2000). The aims of MGED were to record how array analysis is performed to allow the assessing of methodology and for others to reproduce the results. For this reason the MIAME system was created (**m**inimum **i**nformation **a**bout **m**icroarray **e**xperiments) to outline the minimum information required to interpret unambiguously array data and to allow subsequent verification by other groups. This system addresses the variation in results produced by different groups using different systems on different samples. However, no study has been performed into the variation caused directly by different systems i.e. the same group, with the same samples, but using different methods. Exploitation of the ExGen Project (School of Biosciences, Birmingham University) has allowed such a comparison to be performed in this study.

Naturally, the basic principles behind array technology are common to both microarray and macroarray procedures. RNA is isolated from samples, often reverse transcribed into labelled cDNA, which is hybridised to immobilised DNA. However, there is a huge scope for variation in how this can be achieved which is represented well in a comparison between the commercially available Sigma-Genosys Panorama *E. coli* Arrays (as used in Chapter 3) and the ExGen *E. coli* Arrays. Below is a description of the two techniques. A diagrammatic summary can be seen in Figures 7.1 and 7.2.

7.1.1 Macroarrays

Sigma-Genosys Panorama *E. coli* gene arrays (an example of a macroarray) comprise 4,290 PCR products spotted onto a positively charged 12 × 24 cm nylon hybridisation membrane. Each PCR product is spotted in duplicate at equal mass per spot (10 ng) and cross linked onto the membrane with UV light. Spots are 1.5mm apart (centre to centre) and arranged into 3 fields. Each field consists of 'boxes' in 16 rows (A to P) of 24 columns. Each 'box' has the capacity to contain 8 spots.

The majority of PCR products are amplified from start to stop codon. ^{33}P-labelled cDNA is generated from 2 μg of template RNA using 4,290 *E. coli* ORF specific labelling primers (Sigma-Genosys, Poole, UK) (see Section 2.3.1). Each cDNA sample is hybridised to a separate membrane and the image visualised using phosphoimaging. Membranes can be re-used up to 10 times.

7.1.2 Microarrays

The microarrays produced by the ExGen Project of Birmingham University are based upon the 'Array Ready Oligo Set' supplied by Operon (Qiagen, UK). 5,978 70mer oligonucleotides representing three strains of *E. coli* (K12, O157:H7 EDL933 and O157:H7 Sakai) are spotted onto a poly-lysine coated glass slide.

The set of oligos is based upon the K12 strain (4,289 ORFs). Alignment of the three strains (by BLAST analysis) reveals an additional 1,416 unique ORFs when comparing O157:H7 EDL933 with K12 and a further 273 unique ORFs for O157:H7 Sakai when compared with K12 and EDL933. However, since the strain used in this study is W3110 (a K12 derivative) only the data from the K12 associated spots will be considered.

Oligonucleotide design is strict and follows stringent rules. Briefly, all oligos have a melting temperature of 75°C +/- 5°C, do not occur within 40 bases of the 3' end, contain no single nucleotide base repeats longer than 8 bases and contain no hairpin structures with stem length longer than 9 bases. More information on oligo

design can be found at www.operon.com. Each oligo is spotted in duplicate at equal mass per spot and cross-linked onto the glass slide by irradiation with UV light. Aminoallyl modified cDNA is generated from 20 μg of template RNA using random hexamer primers (see Section 2.3.2). Control and test cDNA are distinguished by labelling with CyDye (either Cy3 or Cy5) and hybridised onto the same slide before visualisation by laser scanning.

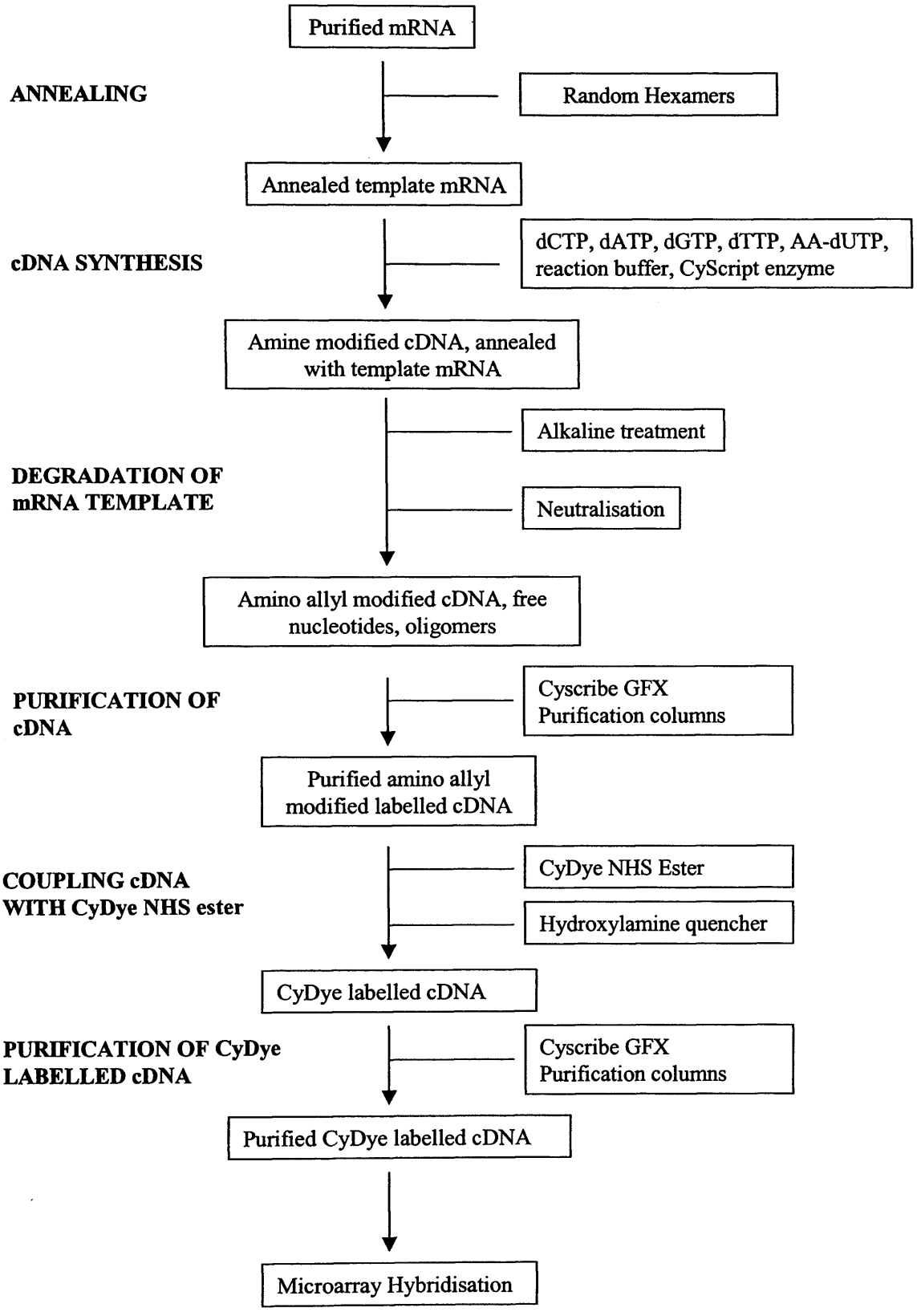


Figure 7.1. Preparation of CyDye labelled cDNA for microarray analysis

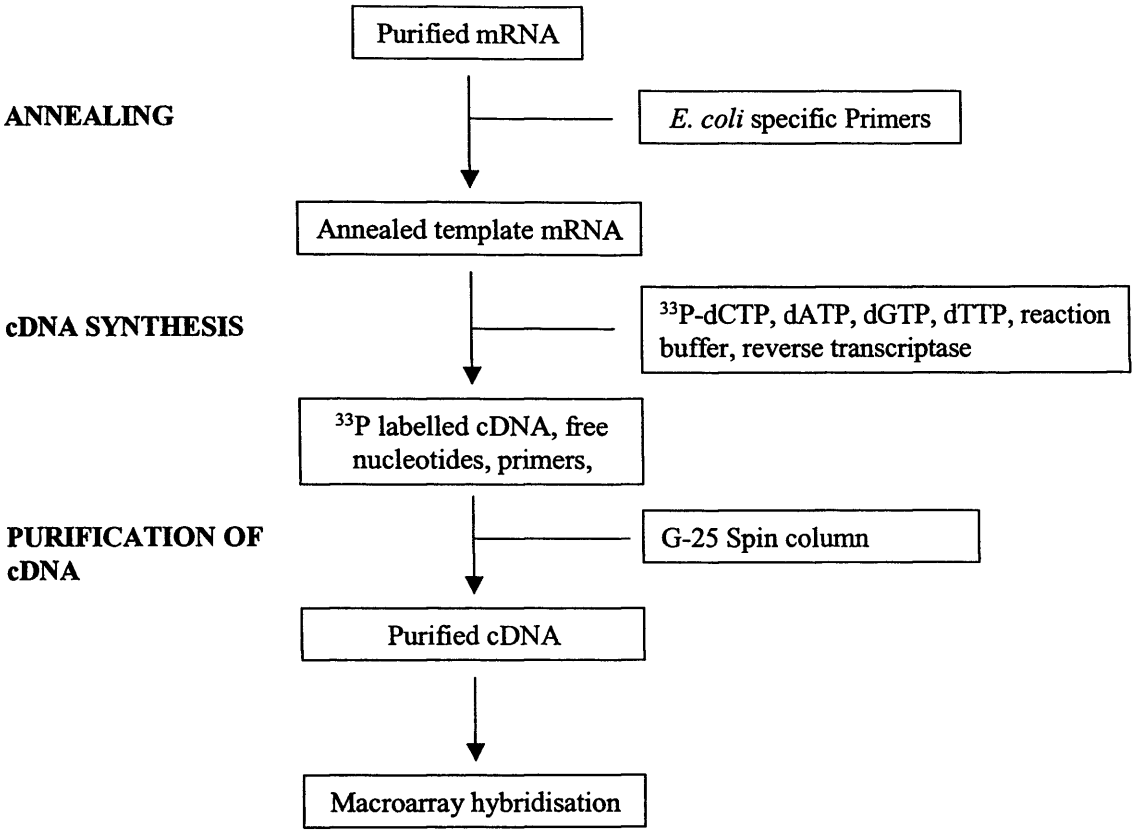


Figure 7.2 Preparation of ^{33}P labelled cDNA for macroarray analysis

7.3 Results

7.3.1 Lab-on-a-chip Analysis of RNA

The original macroarray analysis was performed approximately 12 months before the microarray analysis carried out in this chapter. During this period RNA samples were stored at –80°C. Prior to performing the microarrays, RNA was analysed using the Agilent 'lab-on-a-chip' system. Figures 7.3 and 7.4 clearly show that no degradation of RNA had occurred during storage. Also, a near perfect ratio of 2:1 was found for the amount of 23S:16S rRNA indicating a high quality RNA prep. If any degradation had occurred, a whole series of peaks would be seen early on in the spectrograph, particularly between 19 and 39 seconds.

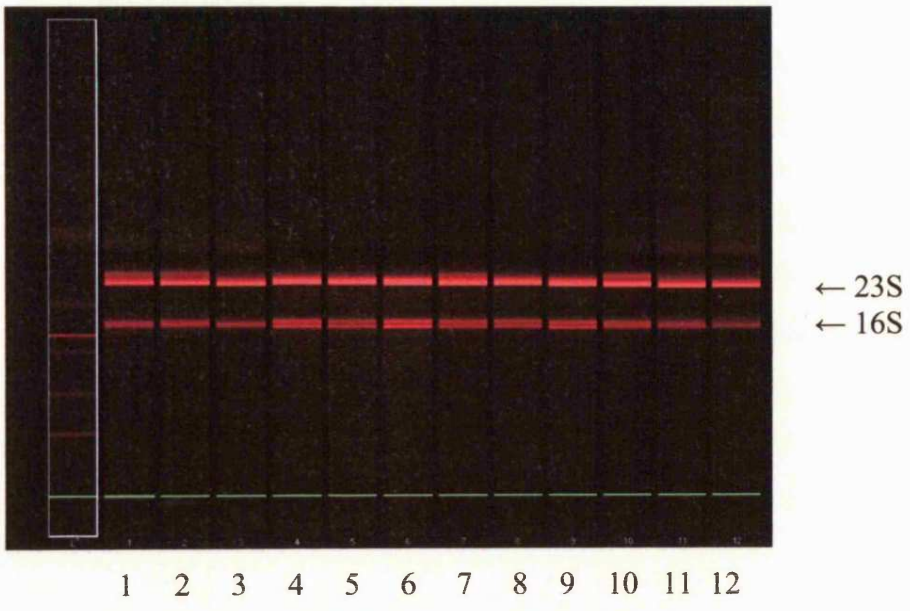


Figure 7.3 Lab-on-a-chip gel analysis of RNA samples used for macroarrays in Chapter 3. Lanes 1-3, 4-6, 7-9, 10-12 correspond to the RNA samples isolated from points A, B, C and D, respectively) in Chapter 3, Figure 3.5.

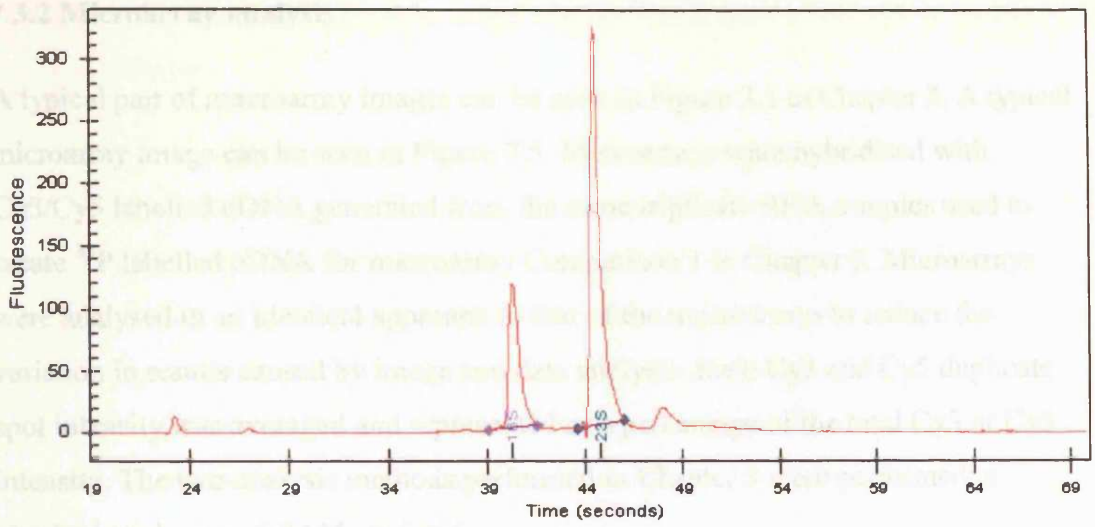


Figure 7.4 Lab-on-a-chip spectrograph analysis of sample 1 from Figure 7.3. Samples gave 23S:16S rRNA ratios of 2:1.

7.3.2 Microarray analysis

A typical pair of macroarray images can be seen in Figure 3.1 in Chapter 3. A typical microarray image can be seen in Figure 7.5. Microarrays were hybridised with Cy3/Cy5 labelled cDNA generated from the same triplicate RNA samples used to create ^{33}P labelled cDNA for macroarray Comparison 1 in Chapter 3. Microarrays were analysed in an identical approach to that of the macroarrays to reduce the variation in results caused by image and data analysis. Each Cy3 and Cy5 duplicate spot intensity was averaged and represented as a percentage of the total Cy3 or Cy5 intensity. The two analysis methods performed in Chapter 3 were performed: a standard analysis and SAM analysis.

A complete list of the results obtained from the microarray analysis can be seen in Appendix E (SAM analysis in Table E.2; Standard analysis in Table E.3). The data obtained from each microarray analysis were aligned against the corresponding macroarray analysis. A comparison (in alignment form) of microarray replicate 1 standard analysis results versus macroarray replicate 1 standard analysis results, in Blattner number order, can be seen in Table 7.1. Further alignments (microarray replicate 2 standard analysis versus macroarray replicate 2 standard analysis; microarray replicate 3 standard analysis versus macroarray replicate 3 standard analysis; microarray triplicate standard analysis versus macroarray triplicate standard analysis; microarray SAM analysis versus macroarray SAM analysis) can be seen in Appendix F (Tables F.1, F.2, F.3 and F.4 respectively). An overall summary of these comparisons can be seen in Table 7.2 (Standard analysis) and Table 7.3 (SAM analysis).

The 'common' changes provide an estimate of the percentage identity (directly) between results from a microarray and a macroarray experiment i.e. single genes whose change in expression is 'confirmed' by being altered in both microarray and macroarray procedures. 'Regional' changes provide an estimate of the overall percentage identity between the results obtained from microarray and macroarray procedures i.e. groups of genes whose change in expression can be 'confirmed' by the alternative array procedure. For example, the flagellar operon in Table 7.1: *flgB*, *flgC* and *flgG* were altered in microarrays, while *flgE* and *flgJ* were altered in macroarrays.

Whereas this gave 0% similarity by common changes, it gave 100% similarity by regional changes.

Regional changes were defined as significant changes that occur with 3 ORFs of any significantly changed ORF i.e. if *b1000* was identified, regional changes would include *b0997-b0999* and *b1001-b1003* in either array. It is important to note that these only provide an insight into regional changes in gene expression and do not represent operons, but merely clusters of genes in the same region that are regulated in a similar manner. This was done for ease of comparison and to avoid issues such as what is or is not an operon. Furthermore, a comparison of the fold-induction values obtained for genes deemed to significantly altered in triplicate in both microarray and macroarray experiments can be seen in Table 7.4.

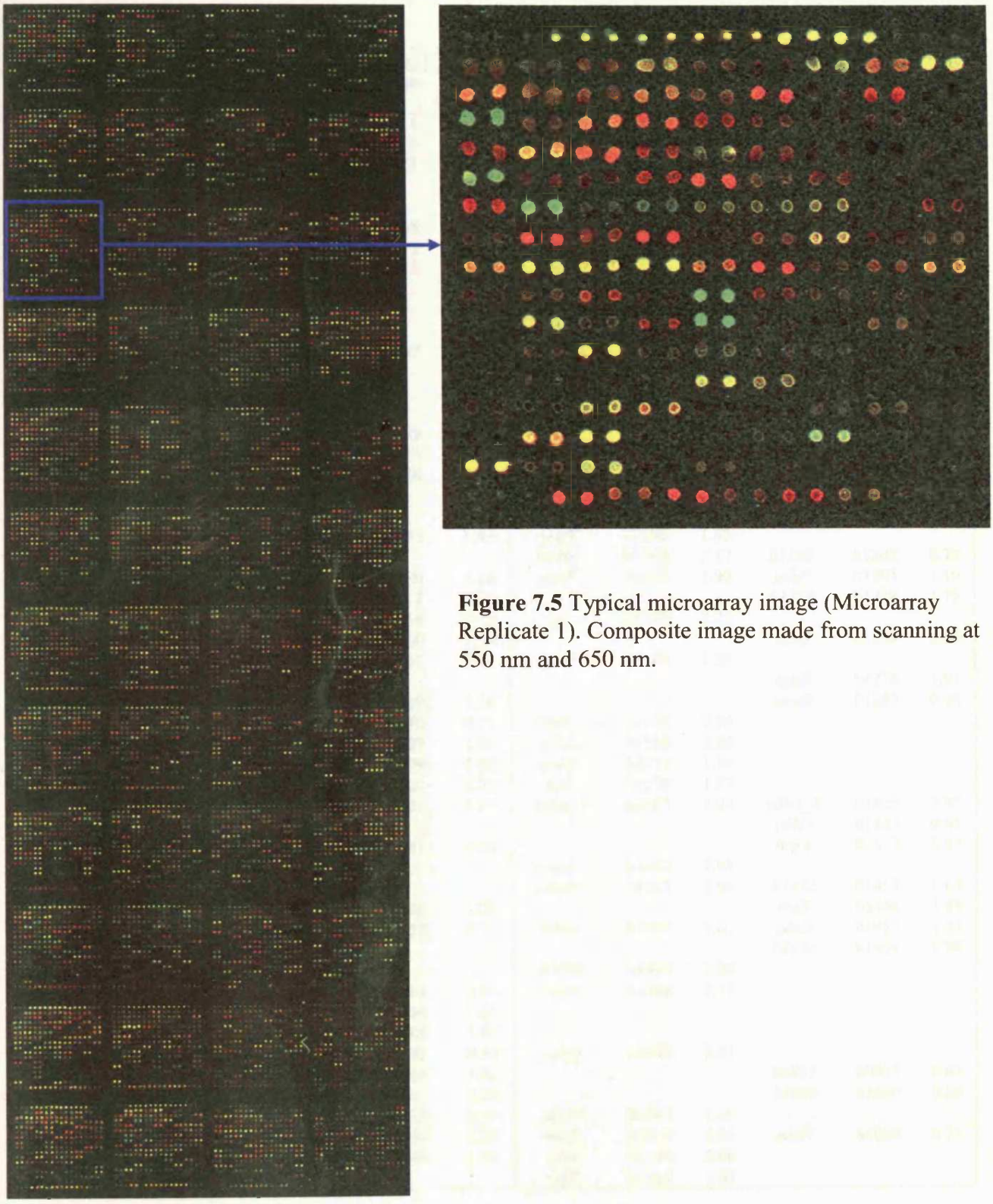


Figure 7.5 Typical microarray image (Microarray Replicate 1). Composite image made from scanning at 550 nm and 650 nm.

Table 7.1 Gene changes and alignment, in Blattner number order, for microarray replicate 1 standard analysis results and macroarray replicate 1 standard analysis results.

Microarray 1			Macroarray 1			Microarray 1			Macroarray 1		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	IR	Gene name	Blattner No.	Log IR
			dnaK	b0014	0.88	*ycaL*	b0909	2.25			
			dnaJ	b0015	0.82				*cmk*	b0910	0.73
			insB_1	b0021	0.76	*ycbQ*	b0938	2.01			
hepA	b0059	1.96							*b0955*	b0955	0.67
yadE	b0130	1.64							*b1030*	b1030	0.86
			yadC	b0135	1.44	*ymdD*	b1044	1.69			
			b0212	b0212	0.80	*mviN*	b1069	1.81	*mviN*	b1069	0.73
yafT	b0217	2.48	*b0217*	b0217	1.28	*ycfJ*	b1110	2.08			
ykfE	b0220	1.71				*ycfR*	b1112	1.65			
dinJ	b0226	1.71				*ycfB*	b1133	1.98			
gpt	b0238	1.77							*b1138*	b1138	0.86
			b0247	b0247	0.72	*b1145*	b1145	1.79			
insB_3	b0274	1.68				*mcrA*	b1159	2.18	*mcrA*	b1159	1.42
yagY	b0292	1.85				*ycgW*	b1160	1.61	*b1160*	b1160	1.65
yagZ	b0293	1.87				*ycgX*	b1161	1.85	*b1161*	b1161	0.98
			b0299	b0299	0.69	*ycgE*	b1162	1.77			
yahA	b0315	2.86	*yahA*	b0315	1.12				*b1164*	b1164	1.05
			b0316	b0316	0.97	*ymgA*	b1165	1.67			
yahM	b0327	1.76							*b1171*	b1171	1.26
adhC	b0356	1.84							*b1172*	b1172	1.40
			b0357	b0357	1.34	*ycgY*	b1196	1.85			
yaiS	b0364	2.22				*b1202*	b1202	2.17	*b1202*	b1202	0.72
			b0370	b0370	1.12	*ychF*	b1203	1.99	*ychF*	b1203	1.10
			yajI	b0412	0.76				*b1228*	b1228	1.75
			ybaA	b0456	0.72	*tpr*	b1229	2.21			
hha	b0460	1.83	*hha*	b0460	1.30				*ychG*	b1239	0.66
			ybaJ	b0461	0.90	*yciL*	b1269	1.93			
ybaN	b0468	1.70							*cysB*	b1275	1.01
			apt	b0469	1.14				*osmB*	b1283	0.98
			dnaX	b0470	0.71	*ycjF*	b1322	2.05			
rhsD	b0497	1.61	*rhsD*	b0497	1.81	*ydaL*	b1340	2.05			
			b0499	b0499	0.92	*ynaE*	b1375	1.74			
			ybbD	b0500	1.24	*hslJ*	b1379	1.77			
			b0501	b0501	0.77	*ydbA_2*	b1405	2.01	*ydbA_2*	b1405	0.97
b0538	b0538	2.06							*ydbD*	b1407	0.88
			tra5_2	b0541	0.72				*hrpA*	b1413	0.67
emrE	b0543	2.09				*rimL*	b1427	1.63			
ybcK	b0544	2.51				*ansP*	b1453	1.96	*b1453*	b1453	0.68
			b0546	b0546	1.28				*rhsE*	b1456	1.34
			b0557	b0557	0.77	*ydcD*	b1457	1.61	*ydcD*	b1457	1.27
ybcY	b0562	2.08							*b1458*	b1458	0.90
ylcE	b0563	2.07				*b1459*	b1459	2.09			
			b0603	b0603	0.76	*narZ*	b1468	2.55			
			b0648	b0648	1.05						
			b0685	b0685	1.47						
ybfL	b0705	2.12	*b0705*	b0705	0.80	*ycbQ*	b0938	2.01			
			ybfD	b0706	1.21				*b0955*	b0955	0.67
			ybgF	b0742	0.69				*b1030*	b1030	0.86
			b0833	b0833	0.67	*ymdD*	b1044	1.69			
			grxA	b0849	0.70	*mviN*	b1069	1.81	*mviN*	b1069	0.73
infA	b0884	2.18	*infA*	b0884	1.08	*ycfJ*	b1110	2.08			
ycaD	b0898	1.93				*ycfR*	b1112	1.65			

Table 7.1 Continued

Microarray 1			Macroarray 1			Microarray 1			Macroarray 1		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			b1472	b1472	0.66				*b1983*	b1983	1.04
			b1527	b1527	0.79				*cbl*	b1987	1.05
marR	b1530	1.72							*yefI*	b2032	1.45
marA	b1531	2.43	*marA*	b1531	1.01	*wbbJ*	b2033	2.19			
			ydeH	b1535	0.82	*wbbI*	b2034	2.09	*yefG*	b2034	1.43
ydfO	b1549	1.79	*b1549*	b1549	1.20				*rfc*	b2035	0.84
cspI	b1552	1.73				*rfbX*	b2037	1.91	*rfbX*	b2037	1.06
cspF	b1558	1.91							*rfbC*	b2038	0.72
relF	b1562	1.80							*rfbA*	b2039	0.87
relE	b1563	2.09							*rfbD*	b2040	0.92
relB	b1564	2.08				*wza*	b2062	1.64			
			b1568	b1568	0.70	*b2070*	b2070	2.18			
ydfA	b1571	1.76				*b2071*	b2071	2.34			
ydfB	b1572	1.68							*b2081*	b2081	0.68
			ydfC	b1573	0.71	*b2084*	b2084	1.61			
dicB	b1575	1.89				*b2088*	b2088	1.76			
			ydhA	b1639	1.03	*yehA*	b2108	1.74			
			b1643	b1643	1.00				*yehC*	b2110	0.82
			b1644	b1644	0.87				*b2145*	b2145	0.74
			b1648	b1648	0.90				*b2174*	b2174	0.90
b1649	b1649	1.83				*rsuA*	b2183	1.66			
			lhr	b1653	0.71				*rplY*	b2185	1.04
ynhC	b1681	1.78							*yfaE*	b2236	0.74
b1706	b1706	1.68				*b2250*	b2250	1.93			
infC	b1718	1.74				*elaD*	b2269	2.10	*b2269*	b2269	0.69
			b1721	b1721	0.88	*cvpA*	b2313	1.74			
b1729	b1729	2.08				*truA*	b2318	1.85	*truA*	b2318	0.71
			b1730	b1730	0.74	*usg*	b2319	1.61			
			b1743	b1743	0.69				*vacJ*	b2346	1.19
			b1770	b1770	0.85				*evgA*	b2369	0.72
			b1785	b1785	0.82				*evgS*	b2370	0.85
			b1786	b1786	0.80	*eutI*	b2458	1.80			
yeaZ	b1807	2.07							*b2504*	b2504	0.68
			b1815	b1815	0.72	*b2506*	b2506	2.53			
b1821	b1821	1.65	*b1821*	b1821	0.68				*xseA*	b2509	0.95
b1825	b1825	1.64				*yfhL*	b2562	1.71			
			b1826	b1826	1.18				*lepA*	b2569	0.71
			holE	b1842	0.93	*yfiM*	b2586	1.66			
			yebE	b1846	0.72	*aroF*	b2601	1.81			
			yebG	b1848	1.19				*rplS*	b2606	1.06
ruvB	b1860	1.69							*trmD*	b2607	1.26
bisZ	b1872	1.69							*yfjA*	b2608	0.78
			insB_5	b1893	0.67				*rpsP*	b2609	1.40
			b1933	b1933	0.70	*yfjW*	b2642	2.30	*b2642*	b2642	1.08
			b1935	b1935	0.76				*b2649*	b2649	0.76
b1936	b1936	2.54				*b2667*	b2667	2.12			
yedJ	b1962	2.62	*yedJ*	b1962	1.14	*stpA*	b2669	2.07	*stpA*	b2669	1.03
b1963	b1963	2.79	*b1963*	b1963	1.18	*proV*	b2677	1.93			
yedV	b1968	2.19				*emrA*	b2685	1.72			
			b1969	b1969	0.72	*b2689*	b2689	1.73			
			b1974	b1974	0.71	*yqaB*	b2690	1.91			
			b1979	b1979	1.38				*recX*	b2698	0.76

Table 7.1 Continued

Microarray 1			Macroarray 1			Microarray 1			Macroarray 1		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			recA	b2699	0.91	*cspA*	b3556	2.02			
			recC	b2822	1.11				*yiaA*	b3562	0.68
			b2845	b2845	0.71	*yibJ*	b3595	3.02	*yibJ*	b3595	1.43
yqeH	b2846	2.46							*yibI*	b3598	0.77
yqeJ	b2848	1.77				*htrL*	b3618	1.79	*b3618*	b3618	0.93
			b2849	b2849	0.70				*rfaL*	b3622	0.99
			b2851	b2851	0.78	*rfaZ*	b3624	1.78	*rfaZ*	b3624	0.87
			b2854	b2854	1.31	*rfaI*	b3627	1.84	*rfaI*	b3627	0.70
			b2863	b2863	1.50				*rfaS*	b3629	0.81
b2862	b2862	2.02							*rpmB*	b3637	0.83
			pepP	b2908	0.69				*rep*	b3778	0.68
ygfA	b2912	1.75				*rfe*	b3784	1.74			
sprT	b2944	1.75							*yigF*	b3817	0.80
			b2974	b2974	0.76	*yiiG*	b3896	1.83	*yiiG*	b3896	1.19
yghS	b2985	2.20				*b3913*	b3913	1.78	*b3913*	b3913	1.12
b3022	b3022	2.04	*b3022*	b3022	0.86	*b3914*	b3914	1.85	*b3914*	b3914	1.55
			b3046	b3046	0.79	*yiiX*	b3937	1.91			
ygiG	b3058	1.79							*yijP*	b3955	0.69
ygjD	b3064	1.64				*trmA*	b3965	1.81			
			rpsU	b3065	0.84	*btuB*	b3966	1.60			
ygjK	b3080	1.61				*birA*	b3973	1.81			
ygjN	b3083	2.67	*ygjN*	b3083	0.75				*rplJ*	b3985	0.88
			tdcR	b3119	0.96				*rplL*	b3986	0.71
			yhaB	b3120	1.88				*dinF*	b4044	0.78
			yraH	b3142	1.11				*yjbM*	b4048	0.91
deaD	b3162	2.72							*yjcF*	b4066	0.99
			rpsO	b3165	1.01				*yjeQ*	b4161	0.82
			infB	b3168	0.82				*miaA*	b4171	0.98
			yhbC	b3170	0.80				*yjeB*	b4178	0.78
			yhbX	b3173	0.74				*yjfI*	b4181	0.78
			yhbZ	b3183	0.72	*sgaU*	b4197	1.60			
yhcA	b3215	2.43							*priB*	b4201	0.67
			b3238	b3238	0.69				*yjfZ*	b4204	0.68
envR	b3264	1.75				*ytfP*	b4222	1.76			
			yhdV	b3267	0.72				*b4257*	b4257	1.00
			rplX	b3309	0.69				*intB*	b4271	0.75
			rpsJ	b3321	0.89	*yi21*	b4272	1.97			
			rpsG	b3341	0.69	*yjhA*	b4311	1.73			
			yheL	b3343	0.78	*uxuR*	b4324	1.85			
mrcA	b3396	1.74				*yjjA*	b4360	1.96			
			glpD	b3426	1.03				*yjjY*	b4402	0.66
			yhhZ	b3442	1.39						
			yrhB	b3446	0.87						
			rhsB	b3482	1.17						
			yhhH	b3483	1.34						
			yhiJ	b3488	1.94						
yhiK	b3489	1.96	*yhiK*	b3489	1.05						
yhiL	b3490	1.62	*yhiL*	b3490	1.57						
yhiM	b3491	2.59							*acnB*	b0118	-0.76
			yhiV	b3514	0.86				*insA_2*	b0265	-1.53
			b3515	b3515	0.93				*insA_3*	b0275	-1.42
yhiX	b3516	1.74	*yhiX*	b3516	1.45				*cynR*	b0338	-1.09

Table 7.1 Continued

Microarray 1			Macroarray 1			Microarray 1			Macroarray 1		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
ybeK	b0651	-1.07							*xerD*	b2894	-0.68
sdhD	b0722	-1.15							*fba*	b2925	-0.71
sdhA	b0723	-1.22							*cmtB*	b2934	-1.42
b0725	b0725	-1.43							*yggZ*	b2963	-0.96
sucA	b0726	-1.28	*sucA*	b0726	-0.73				*ribB*	b3041	-0.99
			modF	b0760	-0.69				*glgS*	b3049	-0.70
glnH	b0811	-1.11	*glnH*	b0811	-1.28				*yhcE*	b3217	-0.72
			dps	b0812	-1.68				*b3254*	b3254	-0.68
cspD	b0880	-1.18							*ugpB*	b3453	-0.68
			pflB	b0903	-1.00	*uspA*	b3495	-1.05	*uspA*	b3495	-0.90
flgB	b1073	-1.07							*yhiS*	b3504	-1.11
flgC	b1074	-1.07							*hdeB*	b3509	-1.17
			flgE	b1076	-0.90	*hdeA*	b3510	-1.10	*hdeA*	b3510	-1.09
flgG	b1078	-1.37				*dctA*	b3528	-1.29			
			flgJ	b1081	-0.98	*lldP*	b3603	-1.41	*lldP*	b3603	-0.86
			b1140	b1140	-0.79	*lldR*	b3604	-1.24			
			b1194	b1194	-0.75	*tnaL*	b3707	-1.76	*tnaL*	b3707	-1.69
			sapD	b1291	-0.78	*tnaA*	b3708	-1.84	*tnaA*	b3708	-1.51
			b1332	b1332	-0.77	*yieC*	b3720	-1.01			
			ydaC	b1347	-0.66	*atpC*	b3731	-1.19			
ynaF	b1376	-1.48	*b1376*	b1376	-0.70	*atpD*	b3732	-1.03			
aldA	b1415	-1.28				*atpG*	b3733	-1.08			
			ydfB	b1572	-0.85	*atpH*	b3735	-0.96			
			fumA	b1612	-0.78	*atpF*	b3736	-0.97			
sodB	b1656	-0.93							*yieN*	b3746	-0.93
			b1777	b1777	-0.70	*rbsD*	b3748	-0.96	*rbsD*	b3748	-0.85
			b1871	b1871	-0.98	*yifA*	b3762/3	-1.15			
fliD	b1924	-1.30	*fliD*	b1924	-0.67				*fadB*	b3846	-0.68
fliS	b1925	-1.19				*glnA*	b3870	-1.23			
fliT	b1926	-1.06				*fdoG*	b3894	-1.11			
yedF	b1930	-0.98							*trmA*	b3965	-0.83
			gatC	b2092	-1.27	*aceB*	b4014	-1.58			
			gatB	b2093	-1.46	*aceA*	b4015	-0.95			
gatA	b2094	-1.39	*gatA*	b2094	-1.12	*malG*	b4032	-1.24			
gatZ	b2095	-1.63				*malF*	b4033	-1.63			
gatY	b2096	-1.40				*malE*	b4034	-1.46			
mglA	b2149	-1.57	*mglA*	b2149	-0.77	*malK*	b4035	-1.09	*malK*	b4035	-0.87
mglB	b2150	-1.26	*mglB*	b2150	-1.20	*lamB*	b4036	-1.51	*lamB*	b4036	-1.06
galS	b2151	-0.97				*malM*	b4037	-1.71			
nuoI	b2281	-1.23				*acs*	b4069	-1.39			
nuoG	b2283	-1.14							*phnD*	b4105	-1.30
			nuoE	b2285	-0.75				*b4115*	b4115	-0.71
nuoC	b2286	-1.10				*melR*	b4118	-1.24			
			pta	b2297	-0.75	*melA*	b4119	-1.02			
			b2363	b2363	-0.85	*aspA*	b4139	-1.10	*aspA*	b4139	-0.98
			acrD	b2470	-0.81				*b4144*	b4144	-0.96
sseA	b2521	-0.97							*aidB*	b4187	-0.77
hcaR	b2537	-1.16							*yjfO*	b4189	-0.70
yfiQ	b2584	-1.13				*treC*	b4239	-1.30			
srlE	b2703	-1.37				*treB*	b4240	-1.28			
ygcN	b2766	-1.01							*yjgF*	b4243	-0.90
recJ	b2892	-0.96							*yjjX*	b4394	-0.80

Table 7.2 Summary of results from standard analyses of microarrays and macroarrays. Percentage values are given to the nearest 1%.

	Number of significant changes			
Array results	Total	Common		Regional
MICROARRAY 1	205	54	(26%)	116 (57%)
MACROARRAY 1	249		(22%)	117 (47%)
MICROARRAY 2	180	34	(19%)	85 (47%)
MACROARRAY 2	244		(14%)	87 (36%)
MICROARRAY 3	209	29	(14%)	88 (42%)
MACROARRAY 3	266		(11%)	96 (36%)
TRIPLICATE MICROARRAY	65	10	(15%)	26 (40%)
TRIPLICATE MACROARRAY	82		(12%)	21 (26%)

Table 7.3 Summary of results from SAM analyses of microarrays and macroarrays. Percentage values are given to the nearest 1%. * Indicates a SAM analysis with Δ=3.72, false significant number of 0.69 (see Table E.2). # Indicates a SAM analysis with Δ=3.48, false significant number of 0.97 (see Table A.1).

	Number of significant changes			
Array results	Total	Common		Regional
MICROARRAY	76*	16	(21%)	58 (76%)
MACROARRAY	399#		(4%)	91 (23%)

Table 7.4 Fold-change values of genes significantly altered in triplicate in both microarray and macroarray experiments.

	Microarray Replicates					Macroarray Replicates				
Gene	1	2	3	Average	S.D.	1	2	3	Average	S.D.
ycgW	40.3	67.3	19.7	42.4	23.9	45.0	28.2	45.4	39.6	9.8
b1963	619.7	241.9	83.0	314.9	275.7	15.1	9.6	14.8	13.2	3.1
yhiL	41.4	953.0	14.0	336.1	534.4	37.5	13.6	31.0	27.4	12.4
yibJ	1036.3	736.0	248.9	673.7	397.4	27.2	15.2	22.3	21.5	6.0
htrL	61.7	64.9	45.8	57.5	10.2	8.5	8.1	9.2	8.6	0.6
tnaL	56.9	46.8	44.3	49.3	6.7	48.8	18.2	34.7	33.9	15.3
tnaA	69.7	45.2	47.7	54.2	13.4	32.6	30.4	24.8	29.3	4.0
b3914	70.5	132.7	296.9	166.7	117.0	35.1	17.6	25.6	26.1	8.8
lamB	32.2	50.5	15.1	32.6	17.7	11.5	6.8	12.0	10.1	2.9
aspA	12.5	13.6	23.8	16.6	6.2	9.6	44.6	11.5	21.9	19.7

7.4 Discussion

7.4.1 Assessment of RNA quality

The lack of any visible degradation and the perfect 2:1 ratio of 23S:16S rRNA suggests that any variations found between the results from the microarray and the macroarray analyses were likely to be caused by experimental differences and not by biological differences (i.e. changes to the RNA profile caused by degradation).

However, it is interesting to see the presence of a doublet 16S rRNA band in the RNA samples corresponding to samples 4-9. When the sample was heated to 90°C for 5 minutes prior to loading on the lab-on-a-chip this doublet band disappears (data not shown). The increased sensitivity of the lab-on-a-chip system has allowed this band to be visualised (it could not be seen on a standard 1.5% agarose gel, see Figure 3.1). This doublet band could be caused by the association of PHMB with RNA molecules providing possible evidence for the *in vivo* interaction between PHMB and cytoplasmic nucleic acids although in this case, it is not clear why only the 16S rRNA is affected and why a distinct doublet is produced rather than a smear representing different levels of binding to individual rRNA molecules.

7.4.2 Macroarrays versus Microarrays

Macroarrays identified more genes with significantly altered expression than were identified with microarrays. This can clearly be seen in Tables 7.2 and 7.3. The standard analysis results from macroarray and microarray replicates 1, 2, 3, the combined triplicate results and the SAM analyses all showed that macroarrays identify more changes. Typically, this corresponded to approximately 50 extra genes per replicate experiment, which contributed to the identification of an extra 15 significant genes in triplicate using the standard analysis. SAM analysis of the microarray and macroarray data proved this case further. SAM analysis of microarrays created a list of 76 genes with a false significant number (FSN) of 0.69 (i.e. of the 76, '0.69' are likely to be false positives), whereas a list of 399 genes was created for macroarrays (with a FSN value of 0.97). This would imply that macroarrays were more sensitive to detecting changes in transcriptional profile than microarrays.

Closer inspection of the microarray and macroarray results provided some interesting insights. There appeared to be a large variation in the similarity between the results from the same RNA samples. For example, comparing microarray 1 with macroarray 1 showed that there were 54 common changes identified (approximately 26% and 22% of the genes identified respectively). When regional changes were taken into consideration, these arrays showed 57% and 47% similarity (reflecting 116 and 117 gene changes in regions verified by both arrays). However, this 'high' similarity was not found in the comparisons between microarray 2 and macroarray 2 (only 34 common changes, 88 and 96 regional changes respectively) or in the comparison of microarray 3 and macroarray 3 (only 29 common changes, 88 and 96 regional changes respectively). Overall, in the triplicate comparison, just 10 genes were identified as significantly altered in both arrays (a further 27 and 24 genes identified in regional changes).

The direct overlap in results from microarray and macroarray experiments was much lower than expected. However, this overlap did increase when regional changes were taken into account. Although a crude analysis, it does provide an insight into the localisation of the changes that occurred. Close inspection of the results revealed that in some regions there is an excellent match between identified gene changes. Some regions showed areas of perfect matching, some showed matching and some reciprocation, while other regions showed no matches but perfect reciprocation (that is genes identified as being transcriptionally altered appeared to alternate back-and-forth between the two arrays being compared e.g. for a series of genes ABCDE, one array could identify genes A, C and D while the alternate array could identify genes B and E for reciprocation of results to be said to be occurring.). However, certain regions show no matching or reciprocation at all. Examples of these types of matching and reciprocation can be found in each individual replicate comparison, triplicate change comparisons and in the SAM analysis. To illustrate this point the following results have been taken from Table 7.1. Figure 7.6 shows examples of perfect overlapping regions (A and B), some regions of matching with reciprocation (C and D), some regions of matching with extensive overlapping (E), a region of reciprocation with no matching (F) and two regions where no matching, overlap or reciprocation occurs (G and H). Only Blattner numbers are shown.

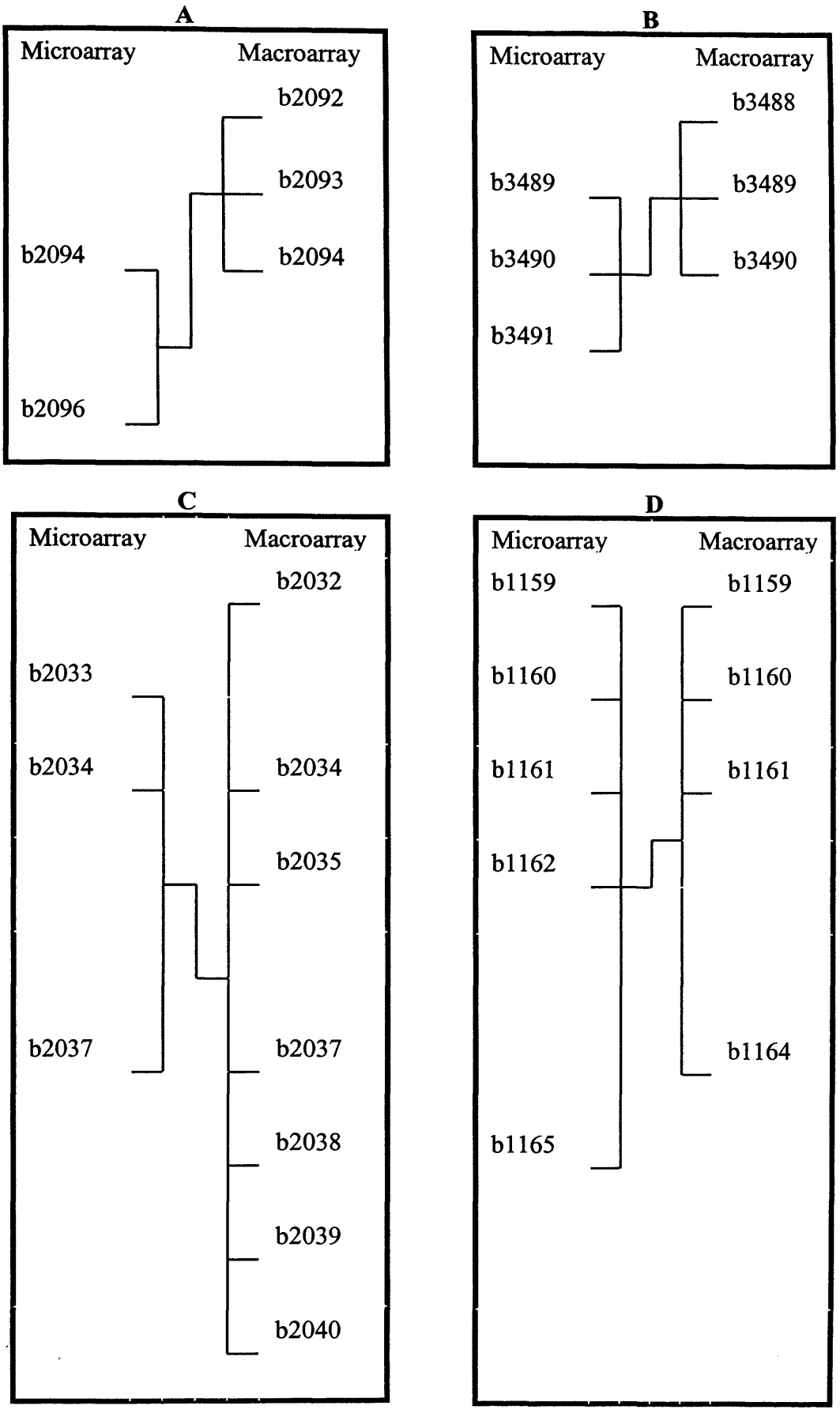


Figure 7.6 Examples of some regional and specific alignments of the results from microarray replicate 1 and macroarray replicate 1.

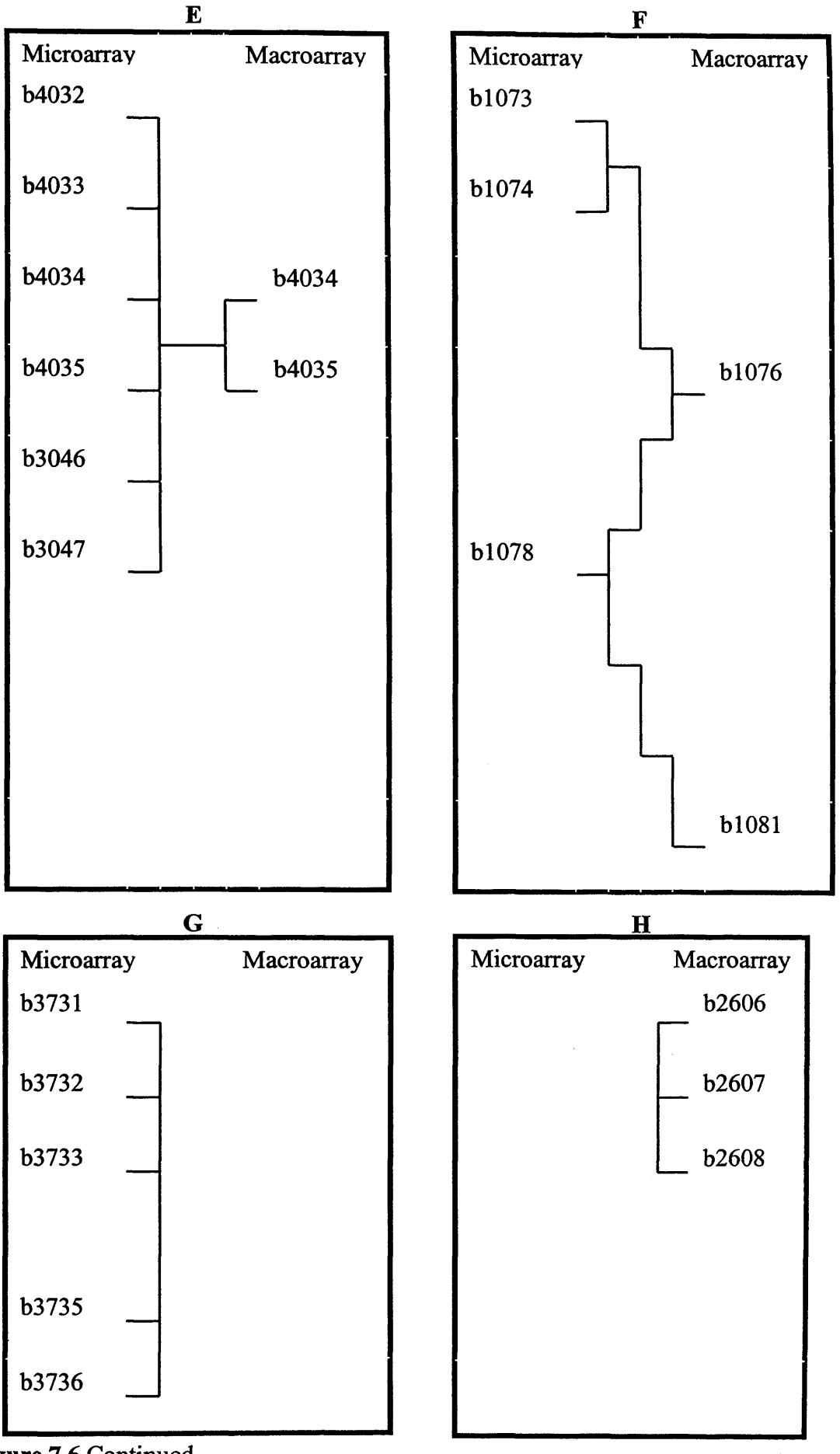


Figure 7.6 Continued

7.4.3 Quantitative Comparison

Table 7.4 shows the fold-change values in each replicate experiment for the ten genes that were significantly altered in all three replicates in both microarrays and macroarrays. Since these are the most reproducible gene changes, the genes in this set are likely to have large induction ratios that are more reproducible than the other genes, hence their identification in both microarrays and macroarrays.

Based on these ten genes, macroarray replicates appear to be more consistently reproducible than microarray replicates. Fold-change (FC) values in macroarrays were typically smaller than those in the corresponding microarray replicates (and their standard deviations smaller). Microarrays appeared to have a much larger variation in FCs for some genes, leading to larger average FCs and standard deviations. This is represented well by *yhiL*, which has FCs of 41.4, 953.0, 14.0 giving an average of 336.1 and a standard deviation of 534.4. However, not all genes in microarrays show this pattern of large variation in FC e.g. *tnaL* had FCs of 56.9, 46.8 and 44.3 giving an average of 49.3 with a standard deviation of only 6.7.

The large variation in replicate FCs observed in microarrays could account for the difference in results provided in a SAM analysis. Whereas macroarrays identified 399 genes with a FSN of 0.97, microarrays identified a mere 76 genes with a FSN of 0.69 using this statistical method. SAM analysis (See Chapter 3 for a description of the technique) is based upon comparing the differences between the signals of control replicates with those of test replicates. Large variations in the intensity of replicate signals (as illustrated with the 10 most reproducible changes above) could prevent many genes being identified as significantly altered. Therefore, the SAM analysis would appear to be more effective at identifying significantly altered genes using the data obtained by macroarray analysis.

7.4.4 Summary

There were clearly some big differences between the results obtained from microarrays and macroarrays. However, the above examples show that upon closer inspection, there was, to a certain extent, a large amount of reciprocation between the results obtained using the two methods.

An important implication obtained from these results is that when one or more gene(s) from the same operon is identified as being significantly altered, and other members are not, that the 'missing' gene changes may actually occur (and are represented in the sample) and that they can even be confirmed using a different type of array analysis or technique on the same sample. Array results need to be interpreted correctly to glean as much information as possible. The comparison between microarrays and macroarrays undertaken here, shows that when looking at gene changes the surrounding genes must also be taken into consideration. Array results are often thought to be unsound when only one or two members of an operon are found altered, while the remaining members of the operon appear to have no change in expression profile. These results show that when this occurs it is likely that the 'missing genes' are altered in expression but for some reason do not appear to be so. Possible reasons for this occurring are differences in annealing temperatures of primers, size of the ORF, primer efficiency and mRNA stability.

However, the results obtained from whole genome arrays must be placed into the correct context. The most important aspect of arrays is that they can only point to regions of the chromosome that are changed in expression. Further phenotypic or biochemical characterisation is needed to validate any results obtained. Even when this characterisation is performed, the lack of identified change of fellow operon members in the array should not be used as a reason to ignore them.

Furthermore, these experiments showed that arrays will only identify certain parts of a large network of responses. The 'complete' response is unlikely to be identified in its entirety. For example, one of the striking results from the macroarrays was the massive induction of many members of the rhs elements. This massive induction was almost completely missed by the microarray analysis, and if microarrays were performed instead of macroarrays (in Chapter 3) their induction

would, most likely, have been overlooked and missed. Similarly, the microarrays showed a massive repression of the *atp* operon. This repression was not seen in the corresponding macroarray experiments. The biological implications of this repression are discussed below in Section 7.4.5.

7.4.5 Biological implications

The original purpose of this study was to investigate the transcriptional response of *E. coli* to PHMB. The large overlap in the results obtained from microarray and macroarray analysis, coupled with the large number of genes of unknown function identified, allowed little extra biological information to be gleaned from this new microarray data. However, one potentially important group of genes were found to be repressed in the microarrays that the macroarrays did not identify. These genes encode some of the subunits of the membrane bound ATP synthase. Figure 7.7 shows the structure of the *atp* operon and which genes were found to be significantly altered in each array comparison.

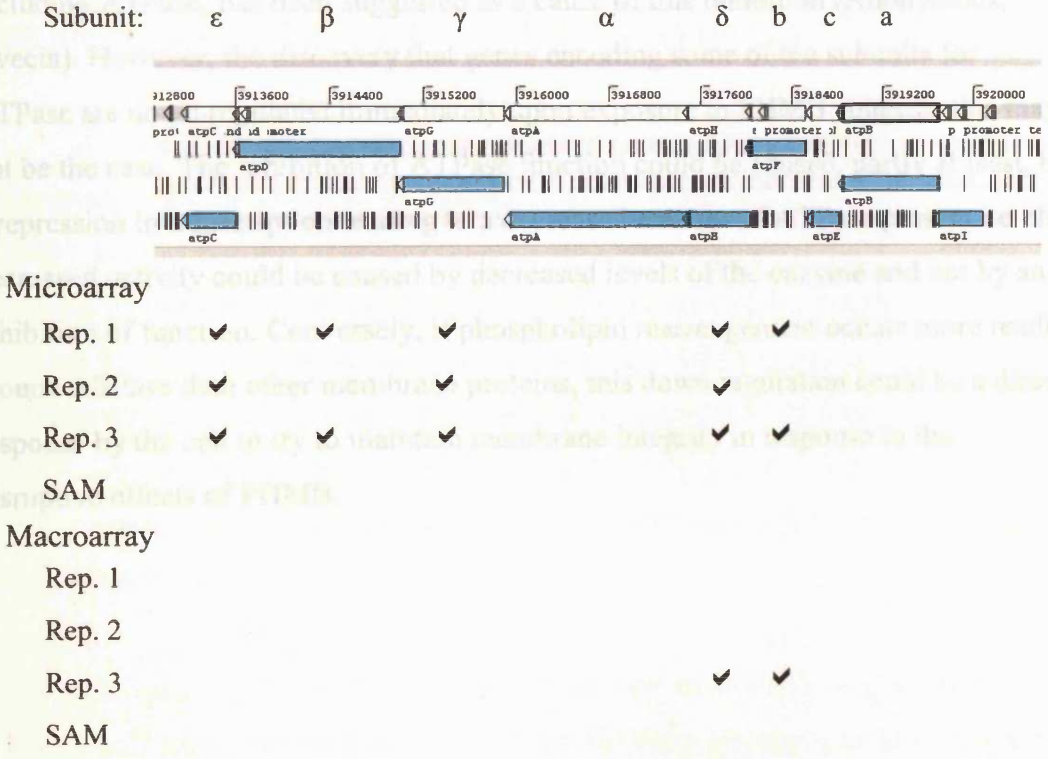


Figure 7.7 Structure of the *atp* operon (taken from the Artemis program). Ticks denote genes that are deemed to be significantly altered.

The membrane bound ATP synthase (ATPase) couples ATP synthesis/ hydrolysis with proton translocation (Groth, 2000). ATPase is a multi-subunit complex comprised of a peripheral catalytic F_1 domain and a membrane integrated F_0 domain (Deckers-Hebestreit *et al.*, 2000). The F_1 part is comprised of 5 types of subunit arranged in the stoichiometry $\alpha_3\beta_3\gamma\delta\epsilon$. ATP synthesis/hydrolysis occurs in the β subunits. The F_0 part is comprised of 3 types of subunit, arranged in the stoichiometry ab_2c_{12}, and functions as a proton channel. Proton translocation is mediated by subunit c. Four protons are translocated per ATP synthesis/hydrolysed i.e. in a complete cycle 3 ATP molecules are synthesised/hydrolysed, while 12 protons are translocated.

The inhibition of ATPase function has been postulated to occur at sub-lethal concentrations of PHMB (Broxton *et al.*, 1983). This was thought to be caused by a change in the ATPase boundary phospholipid environment. Phospholipid rearrangement into the hexagonal phase, concentrated around membrane proteins including ATPase, has been suggested as a cause of this inhibition (Anonymous, Avecia). However, the discovery that genes encoding some of the subunits for ATPase are down regulated immediately upon exposure to PHMB suggests this may not be the case. The inhibition of ATPase function could be caused, partly at least, by a repression in transcription leading to a decreased amount of ATPase present i.e. the decreased activity could be caused by decreased levels of the enzyme and not by an inhibition of function. Conversely, if phospholipid rearrangement occurs more readily around ATPase than other membrane proteins, this down regulation could be a direct response by the cell to try to maintain membrane integrity in response to the disruptive effects of PHMB.

Chapter 8: General Discussion

Hitherto, mechanisms for the biocidal action of PHMB have focused primarily on the readily observable changes occurring at the cell envelope, including loss of LPS, changes in membrane integrity and loss of function of membrane (see Section 1.3). A number of cytoplasmic changes have been observed but these have not been incorporated into mechanistic theory (Khunkitti *et al.*, 1998a). At the outset of this study, little was known about the response of bacteria to the presence of PHMB.

8.1 The response of *E. coli* to PHMB

The transcriptional analysis undertaken in this study (Chapter 3) indicated that *E. coli* responds to bacteriostatic levels of PHMB by altering the expression of a large number of genes. The large number of genes identified by the SAM analysis illustrates this point well. Indeed, the sheer volume of data generated using this method made analysis so difficult that it was not possible to give these results the close inspection that was afforded to those from the standard analysis.

As was noted in Chapter 3, the transcriptional profile generated in these experiments may comprise an amalgamation of different transcriptional responses by separate sub-populations of cells. Further work utilising flow cytometry would provide a useful insight into whether this is the case. Indeed, the transcriptional response to PHMB may not be truly adaptive. Since PHMB is a man-made chemical it is unlikely that *E. coli* would have a designated response to its presence. Therefore, when *E. coli* responds to the presence of PHMB it may actually alter the expression of genes in a way that is deleterious to itself. This could explain some of the unexpected results obtained in Chapter 5.

The majority of the alterations observed were reversed during the subsequent recovery indicating that once the initial stress had been overcome, cellular function returned to normal. PHMB was shown to cause changes in the expression of genes involved in outer membrane, periplasmic, inner membrane and cytoplasmic function. Since cells are dynamic and highly integrated systems, the disruption of any one region or aspect of metabolism of a cell was likely to have detrimental effects on the

function of other regions. Therefore, the alteration in the expression of genes encoding cytoplasmic-proteins came as no great surprise. However, the nature of some of the genes induced was particular interesting with regards to the current hypothesis of PHMB mechanism of action.

Genes associated with stresses including acid resistance, alkali resistance and osmotic shock were altered in expression alongside the controlling factors of the Evg and Cpx response systems. These systems are all involve with sensing and responding to environmental insults that affect the outer membrane, periplasm and cytoplasmic membrane. Since PHMB interacts initially with the outer membrane, periplasm and cytoplasmic membrane the induction of members of these systems would appear to be a logical method of reacting to any damage caused.

However, the induction of the σ^H (heat shock) response, SOS (DNA damage) response and other DNA metabolism associated genes implied that significant damage is occurring to the cytoplasmic contents. Although this would be almost predictable at bactericidal levels of PHMB (when inner membrane disruption is thought to occur), these experiments were performed at bacteriostatic levels of PHMB from which the cells, shortly after, recover. These findings led to expansion of the hypothesis for PHMB mechanism of action discussed below (Section 8.2).

Furthermore, although not part of the official σ^S 'general' stress response, a set of genes were, again, identified in array experiments as being altered at the transcriptional level in response to environmental stress. These are *tnaA, tnaL, cysB, cysK* and *aspA*. For example, many array experiments show genes such as *tnaA, tnaL, cysB, cysK* and *aspA* being altered in response to various stresses such as the presence of alkali, superoxide, hydrogen peroxide, acetate, 4,5-dihydroxy-2-cyclopenten-1-one and growth arrest (Bordi *et al.*, 2003; Phadtare *et al.*, 2002; Polen *et al.*, 2003; Zheng *et al.*, 2001; Chang *et al.*, 2002). These are sometimes dismissed as erroneous results caused by, for example, a change in growth rate due to the stress being investigated. It is possible that they may in fact be part of another broadly acting general stress response that is under the control of some other factor(s). However, as Chapter 5 clearly shows, these genes may have a crucial role in the cell's response to the stress and may not be the mere innocent bystanders they are often mistaken for. Further

characterisation of the roles of these 'usual suspects' in response to environmental stresses is needed.

Whereas the induction of genes associated with nucleic acid metabolism can be understood in relation to damage to DNA, the majority of the *E. coli* response could not be interpreted as clearly. Lipopolysacharride is thought to be the primary target for PHMB, but relatively few genes associated with LPS metabolism were transcriptionally altered. Those that were affected are mostly involved in the synthesis of the O-antigen (a polysaccharide attached to the lipid core of a LPS). However, the O-antigen is not thought to be functional in K12 strains of *E. coli.* Furthermore, the physiological reasons behind the strong induction of the rhs elements still remain to be elucidated. Since their discovery, the function the rhs elements has remained elusive particularly because no conditions have previously been found that lead to their induction.

The proteomic analysis performed in chapter 4 was largely ineffective as the periplasmic profile sought was clearly contaminated with cytoplasmic proteins. Furthermore, the difficulties that arose during the alignment and matching of the 2-D PAGE images only served to further hinder the analysis. As noted previously, any transcript-based method needs further validation. However, in this case, the brief proteomic analysis performed by 2-D PAGE was not productive. This was partly due to problems with the periplasmic protein preparation and partly due to limitations of the software available to analyse the images. Further proteomic profiling of the membrane proteins would be invaluable in assessing alterations in membrane composition, but this would need to be done with a more reliable system for analysis.

In contrast to the proteomic profiling performed in Chapter 4, the phenotypic characterisation by MIC assay performed in Chapter 5 was highly productive and validated much of the gene expression data obtained in Chapter 3. Phenotypic characterisation of knock out strains, over-expressing strains and complementation strains revealed that the large majority of genes, previously identified in Chapter 3, which were tested in this manner did in fact have an effect on susceptibility to PHMB. Thus it appears that most of the transcriptional alterations observed played some role in the recovery from PHMB exposure.

8.2 A new hypothesis for the PHMB mechanism of action

A combination of the supposed damage to genomic DNA observed in Chapter 3, the nature of PHMB itself (repeating positive biguanide units) and previous work in this field led to a study of the direct interaction between PHMB and nucleic acids (Chapter 6). The finding of a strong cooperativity in binding between PHMB and the co-precipitation observed led to the derivation of the model seen in Figure 6.6. Furthermore, since the results from Chapter 3 predict that disruption of the inner membrane does occur at bacteriostatic levels, the difference between bacteriostatic and bactericidal levels of PHMB may not be based upon whether or not disruption of the inner membrane occurs (as previously thought), but on the level of interaction between PHMB and genomic DNA. At lower concentrations (bacteriostatic levels), the damage caused by the interaction between PHMB and DNA could be tolerable and even repairable. Since binding is highly cooperative, small increases in PHMB concentrations could cause a massive increase in damage and/or precipitation, thus leading to cell death. This could account for the dosage dependent effects observed by Broxton *et al.* (1983) and Davies *et al* (1968), the aggregation of phosphorus near the cell wall in PHMB-treated *Acanthamoeba* (Khunkitti *et al.*, 1998a) and the fine line observed between bacteriostatic and bactericidal levels in the growth curves performed in Chapter 3 (see Figure 3.1). This hypothesis can also account for the large changes in minimum inhibitory concentrations observed between prokaryotic and eukaryotic cells and the compounds low mammalian toxicity (See Table 1.2). Eukaryotes tested tend to have high MICs whereas prokaryotes tend to have lower MICs. This could be due to the compartmentalisation of genomic DNA and the presence of internal structures blocking the direct interaction between PHMB and DNA.

8.3 Additional: Array comparison

During the course of this work a unique opportunity arose to compare and contrast the results obtained by different gene array techniques. Data analysis in Chapter 3 had already shown that the results obtained from the raw array data were dependent upon

the method used for analysis. However, by working with the ExGen Project at Birmingham University a comparison of array techniques (i.e. macroarray versus microarray) was performed (Chapter 7). This was extremely fruitful and the results clearly showed that there was a large overlap in the results obtained from the two techniques. The overlap in results obtained, coupled with the large number of genes of unknown function in *E. coli*, allowed very little extra biological information about the response of *E. coli* to PHMB to be obtained from the additional microarray experiments. This is reassuring for groups working with arrays in regard to which technique is 'best'. However, the direct matching, overlap and reciprocation of genes deemed to be significantly altered in the comparison of the two techniques does raise some interesting questions as to where this variation arises.

Factors such as the use of random or ORF specific primers, the target probe (PCR product or oligonucleotide), the biological variation between the replicates, could all effect the results obtained. For this reason, further work has been planned and is currently underway to perform a more extensive comparison of array techniques. PCR-product based macroarrays (produced by Sigma-Genosys, as used in chapter 3), oligonucleotide based microarrays (produced by the ExGen Project, Birmingham University, as used in Chapter 7) and PCR-product based microarrays (produced by the Mori lab, Japan) are currently being used to determine where this variation arises. Whereas the original experiment has been kept the same (i.e. exposure of *E. coli* to PHMB), biological variation has been eliminated by the removal of triplicate replicates (there will be just one control and one test sample) and each reverse transcriptase reaction will be performed with random primers, ORF specific primers and a random primer/ORF specific primer mix.

Although the comparison of the two array techniques allowed the validation of some of the data, it is important to note that further independent verification needs to be performed. Quantitative real time PCR or Northern blotting could be used to verify the transcript levels of genes of interest.

Chapter 9 References

Aiso, T. and R. Ohki (2003). Instability of Sensory Histidine Kinase mRNAs in *Escherichia Coli*. Genes Cells 8(2): 179-187.

Atlung, T. and H. Ingmer (1997). H-NS: A Modulator of Environmentally Regulated Gene Expression. Mol Microbiol 24(1): 7-17.

Auvray, F., J. Thomas, G. M. Fraser and C. Hughes (2001). Flagellin Polymerisation Control by a Cytosolic Export Chaperone. J Mol Biol 308(2): 221-229.

Bancroft, D., L. D. Williams, A. Rich and M. Egli (1994). The Low-Temperature Crystal Structure of the Pure-Spermine Form of Z-DNA Reveals Binding of a Spermine Molecule in the Minor Groove. Biochemistry 33(5): 1073-1086.

Bass, S., Q. Gu and A. Christen (1996). Multicopy Suppressors of *prc* Mutant *Escherichia Coli* Include Two HtrA (DegP) Protease Homologs (HhoAB), DksA, and a Truncated R1pA. J Bacteriol 178(4): 1154-1161.

Becker, W.M., J.B. Reece, M.F. Poenie and W.F. Poenie. The World of the Cell. 3rd Edition. The Benjamin Cummings Publishing Company. Menlo Park, CA.

Beier, D., B. Schwarz, T. M. Fuchs and R. Gross (1995). In Vivo Characterization of the Unorthodox BvgS Two-Component Sensor Protein of *Bordetella Pertussis*. J Mol Biol 248(3): 596-610.

Ben-Zvi, A. P. and P. Goloubinoff (2001). Review: Mechanisms of Disaggregation and Refolding of Stable Protein Aggregates by Molecular Chaperones. J Struct Biol 135(2): 84-93.

Bertani, D., A.B. Oppenheim and F. Narberhaus (2001). FEBS Lett. 2001 Mar 23; 493(1):17-20.

Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao (1997). The Complete Genome Sequence of *Escherichia Coli* K-12. Science 277(5331): 1453-1474.

Bloomgarden, Z. T. (1996). Conference Report: Renal Disease, Metformin, and the Adipocyte. Diabetes Care 19(9): 1038-1041.

Bordi, C., L. Theraulaz, V. Mejean and C. Jourlin-Castelli (2003). Anticipating an Alkaline Stress Through the Tor Phosphorelay System. Mol Microbiol 48(1): 211-223.

Brazma, A., A. Robinson, G. Cameron and M. Ashburner (2000). One-Stop Shop for Microarray Data. Nature 403(6771): 699-700.

Bremer, E. and R. Kramer (2000). Coping with Osmotic Challenges: Osmoregulation through Accumulation and Release of Compatible Solutes in Bacteria. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 79-98.

Brocklehurst, K. R. and A. P. Morby (2000). Metal-Ion Tolerance in *Escherichia Coli*: Analysis of Transcriptional Profiles by Gene-Array Technology. Microbiology 146 (Pt 9): 2277-2282.

Brody, J. P., B. A. Williams, B. J. Wold and S. R. Quake (2002). Significance and Statistical Errors in the Analysis of DNA Microarray Data. Proc Natl Acad Sci USA 99(20): 12975-12978.

Broxton, P., P. M. Woodcock and P. Gilbert (1983). A Study of the Antibacterial Activity of Some Polyhexamethylene Biguanides Towards *Escherichia Coli* ATCC 8739. J Appl Bacteriol 54(3): 345-353.

Broxton, P., P. M. Woodcock and P. Gilbert (1984a). Binding of Some Polyhexamethylene Biguanides to the Cell Envelope of *Escherichia Coli* ATCC 8739. Microbios 41(163): 15-22.

Broxton, P., P. M. Woodcock and P. Gilbert (1984b). Injury and Recovery of *Escherichia Coli* ATCC 8739 from Treatment with Some Polyhexamethylene Biguanides. Microbios 40(161-162): 187-193.

Broxton, P., P. M. Woodcock, F. Heatley and P. Gilbert (1984c). Interaction of Some Polyhexamethylene Biguanides and Membrane Phospholipids in *Escherichia Coli*. J Appl Bacteriol 57(1): 115-124.

Bullitt, E. and L. Makowski (1995). Structural Polymorphism of Bacterial Adhesion Pili. Nature 373(6510): 164-167.

Chang, D. E., D.J. Smalley and T. Conway (2002). Gene Expression Profiling of *Escherichia coli* Growth Transitions: an Extended Stringent Response Model. Mol Microbiol 45(2): 289-306.

Chung, C. T., S. L. Niemela and R. H. Miller (1989). One-Step Preparation of Competent *Escherichia Coli*: Transformation and Storage of Bacterial Cells in the Same Solution. Proc Natl Acad Sci USA 86(7): 2172-2175.

Cohen, S. S. (1998). A Guide to Polyamines. Oxford, Oxford University.

Collinet, B., H. Yuzawa, T. Chen, C. Herrera and D. Missiakas (2000). RseB Binding to the Periplasmic Domain of RseA Modulates the RseA:SigmaE Interaction in the Cytoplasm and the Availability of SigmaE.RNA-Polymerase. J Biol Chem 275(43): 33898-33904.

Coombes, K. R., W. E. Highsmith, T. A. Krogmann, K. A. Baggerly, D. N. Stivers and L. V. Abruzzo (2002). Identifying and Quantifying Sources of Variation in Microarray Data Using High-Density cDNA Membrane Arrays. J Comput Biol 9(4): 655-669.

Cornish-Bowden (1995). Fundamentals of Enzyme Kinetics. London, Portland Press.

Courcelle, J., A. Khodursky, B. Peter, P. O. Brown and P. C. Hanawalt (2001). Comparative Gene Expression Profiles Following UV Exposure in Wild-Type and SOS-Deficient *Escherichia Coli*. Genetics 158(1): 41-64.

Cox, N. A., M. E. Berrang, R. J. Buhr and J. S. Bailey (1999). Bactericidal Treatment of Hatching Eggs II. Use of Chemical Disinfectants with Vacuum to Reduce *Salmonella*. J Appl Poultry Res 8: 321-326.

Cronan, J. E., Jr. and C. O. Rock (1987). Biosynthesis of Membrane Lipids. *Escherichia Coli* and *Salmonella*. Cellular and Molecular Biology. F. C. Neidhardt. Washington, D.C., ASM Press.

Curd, F. H. S. and F. L. Rose (1946). Synthetic Antimalarials. Part X-Some Aryl Diguanide Derivatives. J Chem Soc Part II: 729-737.

Dartigalongue, C., D. Missiakas and S. Raina (2001). Characterization of the *Escherichia Coli* Sigma E Regulon. J Biol Chem 276(24): 20866-20875.

Datsenko, K. A. and B. L. Wanner (2000). One-Step Inactivation of Chromosomal Genes in *Escherichia Coli* K-12 Using PCR Products. Proc Natl Acad Sci USA 97(12): 6640-6645.

Davies, A., M. Bentley and B.S. Field (1968). Comparison of the Action of 'Vantocil', Cetrimide and Chlorhexidine on *Escherichia Coli* and its Spheroplasts and the Protoplasts of Gram Positive Bacteria. Journal of Applied Bacteriology 31: 448-461.

Deckers-Hebestreit, G., J. Greie, W. Stalz and K. Altendorf (2000). The ATP Synthase of *Escherichia Coli*: Structure and Function of F(0) Subunits. Biochim Biophys Acta 1458(2-3): 364-373.

Egli, M., L. D. Williams, Q. Gao and A. Rich (1991). Structure of the Pure-Spermine Form of Z-DNA (Magnesium Free) at 1-A Resolution. Biochemistry 30(48): 11388-11402.

Elbein, A. D., Y. T. Pan, I. Pastuszak and D. Carroll (2003). New Insights on Trehalose: A Multifunctional Molecule. Glycobiology 13(4): 17-27.

Fang, L., B. Xia and M. Inouye (1999). Transcription of *cspA*, the Gene for the Major Cold-Shock Protein of *Escherichia Coli*, Is Negatively Regulated at 37 Degrees C by the 5'-Untranslated Region of its mRNA. FEMS Microbiol Lett 176(1): 39-43.

Feulner, G., J. A. Gray, J. A. Kirschman, A. F. Lehner, A. B. Sadosky, D. A. Vlazny, J. Zhang, S. Zhao and C. W. Hill (1990). Structure of the rhsA Locus from *Escherichia Coli* K-12 and Comparison of rhsA with Other Members of the rhs Multigene Family. J Bacteriol 172(1): 446-456.

Foster, J. W. (2000). Microbial Responses to Acid Stress. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 99-116.

Foster, S. J. (1993). Molecular Analysis of Three Major Wall-Associated Proteins of *Bacillus Subtilis* 168: Evidence for Processing of the Product of a Gene Encoding a 258 kDa Precursor Two-Domain Ligand-Binding Protein. Mol Microbiol 8(2): 299-310.

Futcher, B., G. I. Latter, P. Monardo, C. S. McLaughlin and J. I. Garrels (1999). A Sampling of the Yeast Proteome. Mol Cell Biol 19(11): 7357-7368.

Gajiwala, K. S. and S. K. Burley (2000). HDEA, a Periplasmic Protein That Supports Acid Resistance in Pathogenic Enteric Bacteria. J Mol Biol 295(3): 605-612.

Garwin, J. L., A. L. Klages and J. E. Cronan, Jr. (1980). Structural, Enzymatic, and Genetic Studies of Beta-ketoacyl-acyl Carrier Protein Synthases I and II of *Escherichia Coli*. J Biol Chem 255(24): 11949-11956.

Gaspar, J. A., J. A. Thomas, C. L. Marolda and M. A. Valvano (2000). Surface Expression of O-Specific Lipopolysaccharide in *Escherichia Coli* Requires the Function of the TolA Protein. Mol Microbiol 38(2): 262-275.

Gerdes, S. Y., M. D. Scholle, J. W. Campbell, G. Balazsi, E. Ravasz, M. D. Daugherty, A. L. Somera, N. C. Kyrpides, I. Anderson, M. S. Gelfand, A. Bhattacharya, V. Kapatral, M. D'Souza, M. V. Baev, Y. Grechkin, F. Mseeh, M. Y. Fonstein, R. Overbeek, A. L. Barabasi, Z. N. Oltvai and A. L. Osterman (2003). Experimental Determination and System Level Analysis of Essential Genes in *Escherichia Coli* MG1655. J Bacteriol 185(19): 5673-5684.

Gessner, R. V., C. A. Frederick, G. J. Quigley, A. Rich and A. H. Wang (1989). The Molecular Structure of the Left-Handed Z-DNA Double Helix at 1.0-A Atomic Resolution. Geometry, Conformation, and Ionic Interactions of D(CGCGCG). J Biol Chem 264(14): 7921-7935.

Gilbert, P., D. Pemberton and D. E. Wilkinson (1990a). Barrier Properties of the Gram-Negative Cell Envelope Towards High Molecular Weight Polyhexamethylene Biguanides. J Appl Bacteriol 69(4): 585-592.

Gilbert, P., D. Pemberton and D. E. Wilkinson (1990b). Synergism within Polyhexamethylene Biguanide Biocide Formulations. J Appl Bacteriol 69(4): 593-598.

Gottesman, S., G. Storz, C. Rosenow, N. Majdalani, F. Repoila and K. M. Wassarman (2001). Small RNA Regulators of Translation: Mechanisms of Action and Approaches for Identifying New Small RNAs. Cold Spring Harb Symp Quant Biol 66: 353-362.

Grallert, H. and J. Buchner (2001). Review: A Structural View of the GroE Chaperone Cycle. J Struct Biol 135(2): 95-103.

Grant, R. A., D. J. Filman, S. E. Finkel, R. Kolter and J. M. Hogle (1998). The Crystal Structure of Dps, a Ferritin Homolog That Binds and Protects DNA. Nat Struct Biol 5(4): 294-303.

Groth, G. (2000). Molecular Models of the Structural Arrangement of Subunits and the Mechanism of Proton Translocation in the Membrane Domain of F(1)F(0) ATP Synthase. Biochim Biophys Acta 1458(2-3): 417-427.

Gutierrez, C., J. Barondess, C. Manoil and J. Beckwith (1987). The Use of Transposon TnphoA to Detect Genes for Cell Envelope Proteins Subject to a Common Regulatory Stimulus. Analysis of Osmotically Regulated Genes in *Escherichia Coli*. J Mol Biol 195(2): 289-297.

Gygi, S. P., Y. Rochon, B. R. Franza and R. Aebersold (1999). Correlation between Protein and mRNA Abundance in Yeast. Mol Cell Biol 19(3): 1720-1730.

Hagenmaier, S., Y. D. Stierhof and U. Henning (1997). A New Periplasmic Protein of *Escherichia Coli* Which is Synthesized in Spheroplasts but not in Intact Cells. J Bacteriol 179(6): 2073-2076.

Han, W. and P. Christen (2003). Mechanism of the Targeting Action of DnaJ in the DnaK Molecular Chaperone System. J Biol Chem 278(21): 19038-19043.

Hanahan, D. (1983). Studies on Transformation of *Escherichia Coli* with Plasmids. J Mol Biol 166(4): 557-580.

Hauthal, H. G. (1992). Trends in Surfactants. Chemica*oggi* May: 9-13.

Hengge-Aronis, R. (2000a). A Role for the Sigma S Subunit of RNA Polymerase in the Regulation of Bacterial Virulence. Adv Exp Med Biol 485: 85-93.

Hengge-Aronis, R. (2000b). General Stress Response in *Escherichia coli*. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 161-178.

Hengge-Aronis, R. (2002a). Recent Insights into the General Stress Response Regulatory Network in *Escherichia Coli*. J Mol Microbiol Biotechnol 4(3): 341-346.

Hengge-Aronis, R. (2002b). Signal Transduction and Regulatory Mechanisms Involved in Control of the Sigma(S) (RpoS) Subunit of RNA Polymerase. Microbiol Mol Biol Rev 66(3): 373-395.

Hengge-Aronis, R. (2002c). Stationary Phase Gene Regulation: What Makes an *Escherichia Coli* Promoter SigmaS-Selective? Curr Opin Microbiol 5(6): 591-595.

Hengge-Aronis, R., W. Klein, R. Lange, M. Rimmele and W. Boos (1991). Trehalose Synthesis Genes Are Controlled by the Putative Sigma Factor Encoded by *rpoS* and are Involved in Stationary-Phase Thermotolerance in *Escherichia Coli*. J Bacteriol 173(24): 7918-7924.

Hill, C. W., G. Feulner, M. S. Brody, S. Zhao, A. B. Sadosky and C. H. Sandt (1995). Correlation of Rhs Elements with *Escherichia Coli* Population Structure. Genetics 141(1): 15-24.

Hill, C. W., C. H. Sandt and D. A. Vlazny (1994). Rhs Elements of *Escherichia Coli*: A Family of Genetic Composites Each Encoding a Large Mosaic Protein. Mol Microbiol 12(6): 865-871.

Hirano, T., T. Minamino and R. M. Macnab (2001). The Role in Flagellar Rod Assembly of the N-Terminal Domain of Salmonella FlgJ, a Flagellum-Specific Muramidase. J Mol Biol 312(2): 359-369.

Hommais, F., E. Krin, C. Laurent-Winter, O. Soutourina, A. Malpertuy, J. P. Le Caer, A. Danchin and P. Bertin (2001). Large-Scale Monitoring of Pleiotropic Regulation of Gene Expression by the Prokaryotic Nucleoid-Associated Protein, H-NS. Mol Microbiol 40(1): 20-36.

Houry, W. A. (2001). Chaperone-Assisted Protein Folding in the Cell Cytoplasm. Curr Protein Pept Sci 2(3): 227-244.

Hugo, W. B. and A.R. Longworth (1964). Some Aspects of the Mode of Action of Chlorhexidine. J Pharm Pharmacol 16: 655-662.

Hultgren, S. J., S. Normark and S. N. Abraham (1991). Chaperone-Assisted Assembly and Molecular Architecture of Adhesive Pili. Annu Rev Microbiol 45: 383-415.

Hung, D. L., T. L. Raivio, C. H. Jones, T. J. Silhavy and S. J. Hultgren (2001). Cpx Signaling Pathway Monitors Biogenesis and Affects Assembly and Expression of P Pili. Embo J 20(7): 1508-1518.

Ikeda, T., A. Ledwith, C. H. Bamford and R. A. Hann (1984). Interaction of a Polymeric Biguanide Biocide with Phospholipid Membranes. Biochim Biophys Acta 769(1): 57-66.

Ikeda, T., S. Tazuke and C. H. Bamford (1985). Interaction of Membrane-Active Biguanides with Negatively Charged Species. A Model for Their Interaction with Target Sites in Microbial Membranes. J Chem Research (S): 180-181.

Iwanicka-Nowicka, R. and M. M. Hryniewicz (1995). A New Gene, *cbl*, Encoding a Member of the LysR Family of Transcriptional Regulators Belongs to *Escherichia Coli* Cys Regulon. Gene 166(1): 11-7.

Jones, D. S. (1995). Preliminary Investigations Concerning the Anti-Adherence Properties of Polyhexamethylenebiguanide (Vantocil IB™). International Journal of Pharmaceutics 117: 123-126.

Jones, F. T., D. V. Rives and J. B. Carey (1995). *Salmonella* Contamination in Commercial Eggs and an Egg Production Facility. Poult Sci 74(4): 753-757.

Jung, J. U., C. Gutierrez, F. Martin, M. Ardourel and M. Villarejo (1990). Transcription of *osmB*, a Gene Encoding an *Escherichia Coli* Lipoprotein, is Regulated by Dual Signals. Osmotic Stress and Stationary Phase. J Biol Chem 265(18): 10574-10581.

Jung, J. U., C. Gutierrez and M. R. Villarejo (1989). Sequence of an Osmotically Inducible Lipoprotein Gene. J Bacteriol 171(1): 511-520.

Kaasen, I., J. McDougall and A. R. Strom (1994). Analysis of the *otsBA* Operon for Osmoregulatory Trehalose Synthesis in *Escherichia Coli* and Homology of the OtsA and OtsB Proteins to the Yeast Trehalose-6-Phosphate Synthase/Phosphatase Complex. Gene 145(1): 9-15.

Kachur, A. V., C. J. Koch and J. E. Biaglow (1998). Mechanism of Copper-Catalyzed Oxidation of Glutathionine. Free Radical Res 28: 259-269.

Kandror, O., A. DeLeon and A. L. Goldberg (2002). Trehalose Synthesis Is Induced Upon Exposure of *Escherichia Coli* to Cold and is Essential for Viability at Low Temperatures. Proc Natl Acad Sci USA 99(15): 9727-9732.

Khil, P.P. and R.D. Camerini-Otero (2002). Over 1000 Genes are Involved in the DNA Damage Response of *Escherichia coli*. Mol Microbiol 44(1): 89-105.

Khunkitti, W., A. C. Hann, D. Lloyd, J. R. Furr and A. D. Russell (1998a). Biguanide-Induced Changes in *Acanthamoeba Castellanii*: An Electron Microscopic Study. Journal of Applied Microbiology 84: 53-62.

Khunkitti, W., D. Lloyd, J. R. Furr and A. D. Russell (1998b). *Acanthamoeba Castellanii*: Growth, Encystment, Excystment and Biocide Susceptibility. J Infect 36(1): 43-48.

Khunkitti, W., A. C. Hann, D. Lloyd, J. R. Furr and A. D. Russell (1999). X-Ray Microanalysis of Chlorine and Phosphorus Content in Biguanide-Treated *Acanthamoeba Castellanii*. J Appl Microbiol 86(3): 453-459.

Kim, J. and R. E. Levin (1990). Influence of Caffeine on the Induction of SOS Functions RecA and UmuC by Mitomycin C in *Escherichia Coli*. Microbios 64(260-261): 185-195.

Knight, D. and M. Cooke (2002). Biocides: Big Changes Are Afoot in the Biocide Industry. Chemistry in Britain May: 26-29.

Kusnetsov, J. M., A. I. Tulkki, H. E. Ahonen and P. J. Martikainen (1997). Efficacy of Three Prevention Strategies against *Legionella* in Cooling Water Systems. J Appl Microbiol 82(6): 763-768.

Lammi, M., M. Paci, C. L. Pon, M. A. Losso, A. Miano, R. T. Pawlik, G. L. Gianfranceschi and C. O. Gualerzi (1984). Proteins from the Prokaryotic Nucleoid: Biochemical and 1H NMR Studies on Three Bacterial Histone-Like Proteins. Adv Exp Med Biol 179: 467-477.

Lee, C. H., Y. H. Ni, C. C. Chen, C. Chou and F. H. Chang (2003). Synergistic Effect of Polyethylenimine and Cationic Liposomes in Nucleic Acid Delivery to Human Cancer Cells. Biochim Biophys Acta 1611(1-2): 55-62.

Lin, R. J., M. Capage and C. W. Hill (1984). A Repetitive DNA Sequence, Rhs, Responsible for Duplications within the *Escherichia Coli* K-12 Chromosome. J Mol Biol 177(1): 1-18.

Liu, D., R. A. Cole and P. R. Reeves (1996). An O-Antigen Processing Function for Wzx (RfbX): A Promising Candidate for O-Unit Flippase. J Bacteriol 178(7): 2102-2107.

Liu, D. and P. R. Reeves (1994). *Escherichia Coli* K12 Regains its O-Antigen. Microbiology 140 (Pt 1): 49-57.

Lomba, M. R., A. T. Vasconcelos, A. B. Pacheco and D. F. de Almeida (1997). Identification of *yebG* as a DNA Damage-Inducible *Escherichia Coli* Gene. FEMS Microbiol Lett 156(1): 119-122.

Low, D., B. Braaten and v. d. Woude (1987). Fimbriae. *Escherichia Coli* and *Salmonella*. Cellular and Molecular Biology. F. C. Neidhardt. Washington, D.C., ASM Press.

Ma, Z., H. Richard, D. L. Tucker, T. Conway and J. W. Foster (2002). Collaborative Regulation of *Escherichia Coli* Glutamate-Dependent Acid Resistance by Two AraC-Like Regulators, GadX and GadW (YhiW). J Bacteriol 184(24): 7001-7012.

Mann, C. and J. Crabbe (1998). Bacteria & Antibacterial Agents, Oxford University Press.

Masuda, N. and G. M. Church (2002). *Escherichia coli* gene expression responsive to levels of the response regulator EvgA. J Bacteriol. 184(22): 6225-34.

Masuda, N. and G. M. Church (2003). Regulatory Network of Acid Resistance Genes in *Escherichia Coli*. Mol Microbiol 48(3): 699-712.

Mehnert, H. (2001). Metformin, the Rebirth of a Biguanide: Mechanism of Action and Place in the Prevention and Treatment of Insulin Resistance. Exp Clin Endocrinol Diabetes 109 Suppl 2: S259-264.

Merlin, C., S. McAteer and M. Masters (2002). Tools for Characterization of *Escherichia Coli* Genes of Unknown Function. J Bacteriol 184(16): 4573-4581.

Molloy, M. P., B. R. Herbert, M. B. Slade, T. Rabilloud, A. S. Nouwens, K. L. Williams and A. A. Gooley (2000). Proteomic Analysis of the *Escherichia Coli* Outer Membrane. Eur J Biochem 267(10): 2871-2881.

Morein, S., A. Andersson, L. Rilfors and G. Lindblom (1996). Wild-Type *Escherichia Coli* Cells Regulate the Membrane Lipid Composition in a Window between Gel and Non-Lamellar Structures. J Biol Chem 271(12): 6801-6809.

Mori, H., K. Isono, T. Horiuchi and T. Miki (2000). Functional Genomics of *Escherichia Coli* in Japan. Res Microbiol 151(2): 121-128.

Morita, M. T., Y. Tanaka, T. S. Kodama, Y. Kyogoku, H. Yanagi and T. Yura (1999). Translational Induction of Heat Shock Transcription Factor Sigma32: Evidence for a Built-in RNA Thermosensor. Genes Dev 13(6): 655-665.

Nadon, R. and J. Shoemaker (2002). Statistical Issues with Microarrays: Processing and Analysis. Trends Genet 18(5): 265-271.

Nambu, T., T. Minamino, R. M. MacNab and K. Kutsukake (1999). Peptidoglycan-Hydrolyzing Activity of the FlgJ Protein, Essential for Flagellar Rod Formation in *Salmonella Typhimurium*. J Bacteriol 181(5): 1555-1561.

Neidhardt, F. C. and M. A. Savageau (1987). Regulation Beyond the Operon. *Escherichia Coli* and *Salmonella*. Cellular and Molecular Biology. F. C. Neidhardt. Washington, D.C., ASM Press.

Nies, D. H. (1999). Microbial Heavy-Metal Resistance. Appl Microbiol Biotechnol 51(6): 730-750.

Nies, D. H. and S. Silver (1995). Ion Efflux Systems Involved in Bacterial Metal Resistance. J Indust Microbiol 14: 186-199.

Nikaido, H. (1987). Outer Membrane. *Escherichia Coli* and *Salmonella*. Cellular and Molecular Biology. F. C. Neidhardt. Washington, D.C., ASM Press.

Nishino, K. and A. Yamaguchi (2002). EvgA of the Two-Component Signal Transduction System Modulates Production of the YhiUV Multidrug Transporter in *Escherichia Coli*. J Bacteriol 184(8): 2319-2323.

Nunoshiba, T., E. Hidalgo, C. F. Amabile Cuevas and B. Demple (1992). Two-Stage Control of an Oxidative Stress Regulon: The *Escherichia Coli* SoxR Protein Triggers Redox-Inducible Expression of the *soxS* Regulatory Gene. J Bacteriol 174(19): 6054-6060.

Oh, T. J. and I. G. Kim (1999). Identification of Genetic Factors Altering the SOS Induction of DNA Damage-Inducible *yebG* Gene in *Escherichia Coli*. FEMS Microbiol Lett 177(2): 271-277.

Ohishi, H., I. Nakanishi, K. Inubushi, G. van der Marel, J. H. van Boom, A. Rich, A. H. Wang, T. Hakoshima and K. Tomita (1996a). Interaction between the Left-Handed Z-DNA and Polyamine-2. The Crystal Structure of the D(CG)3 and Spermidine Complex. FEBS Lett 391(1-2): 153-156.

Ohishi, H., N. Terasoma, I. Nakanishi, G. van der Marel, J. H. van Boom, A. Rich, A. H. Wang, T. Hakoshima and K. Tomita (1996b). Interaction between Left-Handed Z-DNA and Polyamine - 3. The Crystal Structure of the D(CG)3 and Thermospermine Complex. FEBS Lett 398(2-3): 291-296.

Ohishi, H., K. Suzuki, K. Grzeskowiak, K. Fukuyama and T. Ishida (2002a). The X-Ray Crystallographic Study of Long Chain Left Handed Z-DNA and Polyamine Complex and Structural Chemistry Study of Effection by Polyamine for the Z-DNA. Nucleic Acids Res Suppl(2): 53-54.

Ohishi, H., K. Suzuki, M. Ohtsuchi, T. Hakoshima and A. Rich (2002b). The Crystal Structure of N(1)-[2-(2-Amino-Ethylamino)-Ethyl]-Ethane-1,2-Diamine (Polyamines) Binding to the Minor Groove of D(CGCGCG)(2), Hexamer at Room Temperature. FEBS Lett 523(1-3): 29-34.

O'Malley, L. P. (2000). Biodegradation of Polyhexamethylene Biguanide and Other Cationic Compounds. Cardiff School of Biosciences, Cardiff University.

Pages, V., N. Koffel-Schwartz and R. P. Fuchs (2003). *recX*, a New SOS Gene That is Co-Transcribed with the *recA* Gene in *Escherichia Coli*. DNA Repair (Amst) 2(3): 273-84.

Phadtare, S., I. Kato and M. Inouye (2002). DNA Microarray Analysis of the Expression Profile of *Escherichia Coli* in Response to Treatment with 4,5-Dihydroxy-2-Cyclopenten-1-one. J Bacteriol 184(23): 6725-6729.

Phadtare, S., K. Yamanaka and M. Inouye (2000). The Cold Shock Response. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 33-46.

Polen, T., D. Rittmann, V.F. Wendisch and H. Sahm (2003). DNA Microarray Analyses of the Long-Term adaptive response of Escherichia coli to Acetate and Propionate. Appl Environ Micro 69(3): 1759-1774.

Pomposiello, P. J., M. H. Bennik and B. Demple (2001). Genome-Wide Transcriptional Profiling of the *Escherichia Coli* Responses to Superoxide Stress and Sodium Salicylate. J Bacteriol 183(13): 3890-3902.

Ptashne, M. (1986). A Genetic Switch: Phage Lambda and Higher Organisms. Cambridge, MA., Cell Press & Blackwell Scientific Publications.

Raffa, R. G. and T. L. Raivio (2002). A Third Envelope Stress Signal Transduction Pathway in *Escherichia Coli*. Mol Microbiol 45(6): 1599-1611.

Raivio, T. L., M. W. Laird, J. C. Joly and T. J. Silhavy (2000). Tethering of CpxP to the Inner Membrane Prevents Spheroplast Induction of the Cpx Envelope Stress Response. Mol Microbiol 37(5): 1186-97.

Raivio, T. L., D. L. Popkin and T. J. Silhavy (1999). The Cpx Envelope Stress Response Is Controlled by Amplification and Feedback Inhibition. J Bacteriol 181(17): 5263-5272.

Raivio, T. L. and T. J. Silhavy (1997). Transduction of Envelope Stress in *Escherichia Coli* by the Cpx Two-Component System. J Bacteriol 179(24): 7724-7733.

Raivio, T. L. and T. J. Silhavy (1999). The SigmaE and Cpx Regulatory Pathways: Overlapping but Distinct Envelope Stress Responses. Curr Opin Microbiol 2(2): 159-165.

Raivio, T. L. and T. J. Silhavy (2000). Sensing and Responding to Envelope Stress. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 19-32.

Raivio, T. L. and T. J. Silhavy (2001). Periplasmic Stress and ECF Sigma Factors. Annu Rev Microbiol 55: 591-624.

Ramos, J. L., M. T. Gallegos, S. Marques, M. I. Ramos-Gonzalez, M. Espinosa-Urgel and A. Segura (2001). Responses of Gram-Negative Bacteria to Certain Environmental Stressors. Curr Opin Microbiol 4(2): 166-171.

Rasmussen, K. and D. Kappes (2000). Implementation of the Biocidal Products Directive: Technical Notes for Guidance in Support of Directive 98/8/EC. Biocides 20(12).

Rezuchova, B., H. Miticka, D. Homerova, M. Roberts and J. Kormanec (2003). New Members of the *Escherichia Coli* SigmaE Regulon Identified by a Two-Plasmid System. FEMS Microbiol Lett 225(1): 1-7.

Rose, F. L. and G. Swain (1956). Bisbiguanides Having Antibacterial Activity. Journal of the Chemical Society Part IV: 4422-4425.

Rouviere, P.E., A. De Las Penas, J. Mecsas, C.Z. Lu, K.E. Rudd and C.A. Gross (1995). *rpoE*, the Gene Encoding the Second Sigma Factor, SigmaE, in *Escherichia coli*. EMBO J 14(5): 1032-1042.

Sadosky, A. B., J. A. Gray and C. W. Hill (1991). The RhsD-E Subfamily of *Escherichia Coli* K-12. Nucleic Acids Res 19(25): 7177-7183.

Sassanfar, M. and J.W. Roberts (1990). Nature of the SOS-Inducing Signal in *Escherichia coli*. The Involvement of DNA Replication. J Mol Biol 212: 79-96.

Schweder, T., K. H. Lee, O. Lomovskaya and A. Matin (1996). Regulation of *Escherichia Coli* Starvation Sigma Factor (Sigma S) by ClpXP Protease. J Bacteriol 178(2): 470-476.

Sinensky, M. (1974). Homeoviscous Adaptation: a Homeostatic Process That Regulates the Viscosity of Membrane Lipids in *Escherichia Coli*. Proc Natl Acad Sci USA 71(2): 522-525.

Smirnova, G. V., N. G. Muzyka and O. N. Oktyabrsky (2000). The Role of Antioxidant Enzymes in Response of *Escherichia Coli* to Osmotic Upshift. FEMS Microbiol Lett 186(2): 209-213.

Stancik, L. M., D. M. Stancik, B. Schmidt, D. M. Barnhart, Y. N. Yoncheva and J. L. Slonczewski (2002). pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia Coli*. J Bacteriol 184(15): 4246-4258.

Stohl, E. A., L. Blount and H. S. Seifert (2002). Differential Cross-Complementation Patterns of *Escherichia Coli* and *Neisseria Gonorrhoeae* RecA Proteins. Microbiology 148(Pt 6): 1821-1831.

Storz, G. (2002). An Expanding Universe of Noncoding RNAs. Science 296(5571): 1260-1263.

Storz, G. and R. Hengge-Aronis (2000). Bacterial Stress Responses. Washington, D.C., ASM Press: Preface.

Storz, G. and M. Zheng (2000). Oxidative Stress. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 47-60.

Tao, H., C. Bausch, C. Richmond, F. R. Blattner and T. Conway (1999). Functional Genomics: Expression Analysis of *Escherichia Coli* Growing on Minimal and Rich Media. J Bacteriol 181(20): 6425-6440.

Thieringer, H. A., P. G. Jones and M. Inouye (1998). Cold Shock and Adaptation. Bioessays 20(1): 49-57.

Tramonti, A., P. Visca, M. De Canio, M. Falconi and D. De Biase (2002). Functional Characterization and Regulation of *gadX*, a Gene Encoding an AraC/XylS-Like Transcriptional Activator of the *Escherichia Coli* Glutamic Acid Decarboxylase System. J Bacteriol 184(10): 2603-2613.

Tsaneva, I. R. and B. Weiss (1990). soxR, a Locus Governing a Superoxide Response Regulon in *Escherichia Coli* K-12. J Bacteriol 172(8): 4197-4205.

Tucker, D. L., N. Tucker and T. Conway (2002). Gene Expression Profiling of the pH Response in *Escherichia Coli*. J Bacteriol 184(23): 6551-6558.

Tucker, D. L., N. Tucker, Z. Ma, J. W. Foster, R. L. Miranda, P. S. Cohen and T. Conway (2003). Genes of the GadX-GadW Regulon in *Escherichia Coli*. J Bacteriol 185(10): 3190-3201.

Tusher, V. G., R. Tibshirani and G. Chu (2001). Significance Analysis of Microarrays Applied to the Ionizing Radiation Response. Proc Natl Acad Sci USA 98(9): 5116-5121.

Urios, A., G. Herrera, V. Aleixandre and M. Blanco (1990). Expression of the *recA* Gene is Reduced in *Escherichia Coli* Topoisomerase I Mutants. Mutat Res 243(4): 267-272.

Utsumi, R., S. Katayama, M. Taniguchi, T. Horie, M. Ikeda, S. Igaki, H. Nakagawa, A. Miwa, H. Tanabe and M. Noda (1994). Newly Identified Genes Involved in the Signal Transduction of *Escherichia Coli* K-12. Gene 140(1): 73-77.

van der Ploeg, J. R., E. Eichhorn and T. Leisinger (2001). Sulfonate-Sulfur Metabolism and its Regulation in *Escherichia Coli*. Arch Microbiol 176(1-2): 1-8.

van der Ploeg, J. R., R. Iwanicka-Nowicka, M. A. Kertesz, T. Leisinger and M. M. Hryniewicz (1997). Involvement of CysB and Cbl Regulatory Proteins in Expression of the *tauABCD* Operon and Other Sulfate Starvation-Inducible Genes in *Escherichia Coli*. J Bacteriol 179(24): 7671-7678.

van Gelder, P., F. Dumas, I. Bartoldus, N. Saint, A. Prilipov, M. Winterhalter, Y. Wang, A. Philippsen, J.P. Rosenbusch and T. Schirmer (2002). Sugar Transport through Maltoporin of *Escherichia coli*: Role of the Greasy Slide. J Bacteriol 184(11): 2994-2999

Vijayanathan, V., T. Thomas and T. J. Thomas (2002). DNA Nanoparticles and Development of DNA Delivery Vehicles for Gene Therapy. Biochemistry 41(48): 14085-14094.

Vlazny, D. A. and C. W. Hill (1995). A Stationary-Phase-Dependent Viability Block Governed by Two Different Polypeptides from the RhsA Genetic Element of *Escherichia Coli* K-12. J Bacteriol 177(8): 2209-2213.

Voet, D. and J.G. Voet (1995). Biochemistry. 2nd Edition. John Wiley and Sons, Inc. New York.

Volkert, M. R. and P. Landini (2001). Transcriptional Responses to DNA Damage. Curr Opin Microbiol 4(2): 178-185.

von Heijne, G. (1989). The Structure of Signal Peptides from Bacterial Lipoproteins. Protein Eng 2(7): 531-534.

Wallace, H. M. (2003). Polyamines and Their Role in Human Disease- an Introduction. Biochem Soc Trans 31: 354-355.

Walter, S. (2002). Structure and Function of the GroE Chaperone. Cell Mol Life Sci 59(10): 1589-1597.

Wang, Y. D., S. Zhao and C. W. Hill (1998). Rhs Elements Comprise Three Subfamilies Which Diverged Prior to Acquisition by *Escherichia Coli*. J Bacteriol 180(16): 4102-4110.

Waterman, S. R. and P. L. Small (1996). Characterization of the Acid Resistance Phenotype and *rpoS* Alleles of Shiga-Like Toxin-Producing *Escherichia Coli*. Infect Immun 64(7): 2808-2811.

Weber, A. and K. Jung (2002). Profiling Early Osmostress-Dependent Gene Expression in *Escherichia Coli* Using DNA Macroarrays. J Bacteriol 184(19): 5502-5507.

Welch, W. J. and C. R. Brown (1996). Influence of Molecular and Chemical Chaperones on Protein Folding. Cell Stress Chaperones 1(2): 109-115.

Wild, J., P. Rossmeissl, W. A. Walter and C. A. Gross (1996). Involvement of the DnaK-DnaJ-GrpE Chaperone Team in Protein Secretion in *Escherichia Coli*. J Bacteriol 178(12): 3608-3613.

Wu, H. C. and M. Tokunaga (1986). Biogenesis of Lipoproteins in Bacteria. Curr Top Microbiol Immunol 125: 127-157.

Wu, J. and B. Weiss (1992). Two-Stage Induction of the *soxRS* (Superoxide Response) Regulon of *Escherichia Coli*. J Bacteriol 174(12): 3915-3920.

Yasuda, K., C. Ohmizo and T. Katsu (2003). Potassium and Tetraphenylphosphonium Ion-Selective Electrodes for Monitoring Changes in the Permeability of Bacterial Outer and Cytoplasmic Membranes. J Microbiol Methods 54(1): 111-115.

Yeates, T. (2002). PHMB for Use in Household and Institutional Cleaning Products. Biocides 20(11).

Yura, T., M. Kanemori and M. T. Morita (2000). The Heat Shock Response: Regulation and Function. Bacterial Stress Responses. G. Storz and R. Hengge-Aronis. Washington, D.C., ASM Press: 3-18.

Yura, T., H. Nagai and H. Mori (1993). Regulation of the Heat-Shock Response in Bacteria. Annu Rev Microbiol 47: 321-350.

Yura, T. and K. Nakahigashi (1999). Regulation of the Heat-Shock Response. Curr Opin Microbiol 2(2): 153-158.

Zhang, A., S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis and G. Storz (1998). The OxyS Regulatory RNA Represses RpoS Translation and Binds the Hfq (HF-I) Protein. Embo J 17(20): 6061-6068.

Zhao, S. and C. W. Hill (1995). Reshuffling of Rhs Components to Create a New Element. J Bacteriol 177(5): 1393-1398.

Zhao, S., C. H. Sandt, G. Feulner, D. A. Vlazny, J. A. Gray and C. W. Hill (1993). Rhs Elements of *Escherichia Coli* K-12: Complex Composites of Shared and Unique Components That Have Different Evolutionary Histories. J Bacteriol 175(10): 2799-2808.

Zheng, M., B. Doan, T. D. Schneider and G. Storz (1999). OxyR and SoxRS Regulation of Fur. J Bacteriol 181(15): 4639-4643.

Zheng, M., X. Wang, L. J. Templeton, D. R. Smulski, R. A. LaRossa and G. Storz (2001). DNA Microarray-Mediated Transcriptional Profiling of the *Escherichia Coli* Response to Hydrogen Peroxide. J Bacteriol 183(15): 4562-4570.

Zhou, Y., S. Gottesman, J. R. Hoskins, M. R. Maurizi and S. Wickner (2001). The RssB Response Regulator Directly Targets Sigma(S) for Degradation by ClpXP. Genes Dev 15(5): 627-637.

Appendix A

Table A.1 Input parameters and computed quantities for the SAM performed on data from macroarray comparison 1. Results from this analysis can be seen in Table A.2.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	3.48957
(Upper Cutoff, Lower Cutoff)	(5.01844, -6.36448)
Computed Quantities	
Computed Exchangeability Factor S0	1.93847E-06
S0 percentile	0.02
False Significant Number (Median, 90 percentile)	(0.97484, 6.58015)
False Discovery Rate (Median, 90 percentile)	(0.24432, 1.64916)
Pi0Hat	0.48742

Table A.2 Results of SAM performed on data from macroarray comparison 1.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
b0357	b0357	15.02	1.67E-03	1.11E-04	29.97647	0.13
b0499	b0499	12.67	6.43E-04	5.08E-05	8.79768	0.13
b0501	b0501	13.40	7.92E-05	5.91E-06	6.16018	0.13
b0543	b0543	12.17	1.22E-04	9.99E-06	2.55021	0.13
b1228	b1228	17.54	2.33E-03	1.33E-04	39.49966	0.13
b1527	b1527	22.91	2.80E-04	1.22E-05	5.85378	0.13
b1730	b1730	11.96	1.23E-04	1.03E-05	6.51246	0.13
b1762	b1762	16.28	1.15E-04	7.09E-06	3.28589	0.13
b1970	b1970	12.38	2.04E-04	1.65E-05	3.48136	0.13
b1983	b1983	15.09	7.62E-04	5.05E-05	8.99523	0.13
b2081	b2081	16.07	1.26E-04	7.87E-06	4.86681	0.13
b2339	b2339	17.73	1.85E-04	1.04E-05	3.15656	0.13
b2668	b2668	14.36	2.88E-04	2.01E-05	2.07178	0.13
b3046	b3046	21.87	1.03E-04	4.69E-06	5.31716	0.13
b3618	b3618	33.28	1.11E-03	3.32E-05	8.42859	0.13
b3913	b3913	14.71	6.87E-04	4.67E-05	15.63903	0.13
cysB	b1275	15.87	4.59E-04	2.89E-05	7.76034	0.13
dcm	b1961	13.49	3.44E-04	2.55E-05	3.44738	0.13
fabA	b0954	22.12	2.66E-04	1.20E-05	2.72065	0.13
greA	b3181	14.22	1.00E-04	7.06E-06	2.83392	0.13
guaB	b2508	13.60	2.52E-04	1.85E-05	2.48297	0.13
osmB	b1283	28.80	3.29E-04	1.14E-05	6.98740	0.13
pabC	b1096	21.03	9.89E-05	4.70E-06	3.20646	0.13
recA	b2699	12.25	2.09E-03	1.71E-04	6.45208	0.13
rhsB	b3482	13.51	4.96E-04	3.67E-05	12.70366	0.13
rpsP	b2609	15.29	4.45E-04	2.91E-05	8.73843	0.13
sfhB	b2594	12.67	7.57E-05	5.98E-06	4.11438	0.13
slp	b3506	15.91	2.13E-04	1.34E-05	2.15474	0.13
yafO	b0233	13.29	1.10E-04	8.31E-06	1.56537	0.13
yahA	b0315	14.74	9.95E-04	6.75E-05	12.36745	0.13
ybbD	b0500	13.06	5.35E-05	4.09E-06	12.65143	0.13
ybfC	b0704	12.00	1.92E-04	1.60E-05	2.71860	0.13
ybhD	b0768	13.15	1.12E-04	8.48E-06	2.26898	0.13
yceB	b1063	13.37	9.01E-05	6.74E-06	1.77201	0.13
ychF	b1203	18.24	1.66E-04	9.11E-06	7.38813	0.13
ydcD	b1457	17.66	7.11E-04	4.03E-05	13.11922	0.13
yehA	b2108	13.40	1.10E-04	8.20E-06	3.35004	0.13
yhiK	b3489	14.47	1.70E-04	1.17E-05	10.25855	0.13
yi41	b4278	15.48	8.81E-05	5.69E-06	3.46497	0.13
yibJ	b3595	25.01	9.86E-04	3.94E-05	19.93377	0.13
yjbM	b4048	31.52	1.46E-04	4.63E-06	8.19692	0.13
yjcF	b4066	15.54	5.39E-05	3.47E-06	9.66237	0.13
yjhC	b4280	13.58	1.20E-04	8.84E-06	3.55163	0.13

Table A.2 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
ychN	b1219	11.69	9.22E-05	7.89E-06	2.04203	0.13
intB	b4271	11.65	1.19E-04	1.02E-05	5.42857	0.13
b1524	b1524	11.62	2.01E-04	1.73E-05	3.04153	0.13
aroF	b2601	11.50	1.73E-04	1.50E-05	1.75543	0.13
b2623	b2623	11.42	1.94E-04	1.70E-05	2.39234	0.13
tdcB	b3117	11.34	9.43E-05	8.31E-06	1.77147	0.13
yebG	b1848	11.30	7.96E-04	7.05E-05	13.54232	0.13
yhhH	b3483	11.26	2.48E-04	2.20E-05	17.03423	0.13
b2340	b2340	11.10	8.50E-05	7.66E-06	3.15953	0.13
yhhS	b3473	10.95	5.81E-05	5.31E-06	2.35666	0.13
rpsO	b3165	10.88	8.20E-04	7.53E-05	5.80654	0.13
yjhB	b4279	10.84	1.91E-04	1.77E-05	3.65332	0.13
asnC	b3743	10.78	5.62E-05	5.21E-06	2.73586	0.13
trmD	b2607	10.72	2.33E-03	2.18E-04	7.94067	0.13
b0482	b0482	10.63	2.24E-04	2.11E-05	1.61791	0.13
miaA	b4171	10.62	2.51E-04	2.36E-05	7.34083	0.13
b0600	b0600	10.61	2.41E-04	2.27E-05	1.84780	0.13
b1171	b1171	10.54	9.45E-04	8.96E-05	14.83005	0.13
b0546	b0546	10.43	1.38E-04	1.33E-05	16.86423	0.13
b0819	b0819	10.41	2.46E-04	2.36E-05	1.47213	0.13
b0847	b0847	10.37	1.71E-04	1.65E-05	1.40588	0.13
b1963	b1963	10.33	4.59E-04	4.44E-05	12.74716	0.13
ydjM	b1728	10.27	5.70E-05	5.55E-06	3.02981	0.13
ygcA	b2785	10.23	8.17E-05	7.99E-06	2.16135	0.13
b1060	b1060	10.12	6.40E-05	6.32E-06	2.64812	0.13
hemM	b1209	10.10	4.59E-05	4.55E-06	1.57894	0.13
ftsL	b0083	10.09	3.11E-04	3.08E-05	2.42261	0.13
frdA	b4154	10.06	9.47E-05	9.41E-06	1.96744	0.13
def	b3287	9.89	8.31E-05	8.40E-06	1.59494	0.13
recN	b2616	9.81	6.84E-04	6.97E-05	2.43145	0.13
yiiM	b3910	9.76	3.84E-04	3.93E-05	2.11164	0.13
ydhA	b1639	9.75	8.56E-05	8.77E-06	8.74988	0.13
b2854	b2854	9.72	5.58E-05	5.75E-06	19.29425	0.13
b0859	b0859	9.71	9.94E-05	1.02E-05	1.98702	0.13
pepP	b2908	9.60	1.75E-04	1.83E-05	3.13152	0.13
yciH	b1282	9.60	6.53E-05	6.81E-06	2.10765	0.13
b1604	b1604	9.49	1.29E-04	1.36E-05	3.46149	0.13
htgA	b0012	9.48	5.50E-05	5.81E-06	2.95871	0.13
ydfC	b1573	9.44	9.43E-05	9.99E-06	4.60318	0.13
htrA	b0161	9.37	2.70E-04	2.88E-05	3.85268	0.13
pncB	b0931	9.35	1.79E-04	1.91E-05	3.18017	0.13
b1596	b1596	9.33	1.32E-04	1.42E-05	3.51332	0.13
b1642	b1642	9.29	5.83E-05	6.28E-06	2.52034	0.13

Table A.2 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
tdcR	b3119	9.20	5.10E-05	5.54E-06	6.68139	0.13
b1722	b1722	9.13	6.95E-05	7.61E-06	1.95308	0.13
glpD	b3426	9.10	6.60E-05	7.25E-06	9.32251	0.13
rplD	b3319	9.03	4.04E-04	4.47E-05	2.91524	0.13
yhiJ	b3488	9.00	6.41E-04	7.12E-05	46.31167	0.13
adhC	b0356	8.96	1.52E-04	1.69E-05	2.62769	0.13
yejD	b2183	8.95	1.37E-04	1.53E-05	2.63851	0.13
stpA	b2669	8.95	3.61E-04	4.03E-05	8.75879	0.13
yjjI	b4380	8.90	7.47E-05	8.39E-06	1.65592	0.13
yfaE	b2236	8.88	2.01E-04	2.26E-05	5.01661	0.13
yidR	b3689	8.82	1.81E-04	2.05E-05	2.54557	0.13
fabF	b1095	8.82	7.15E-04	8.10E-05	3.14157	0.13
b0955	b0955	8.82	5.25E-05	5.95E-06	3.22664	0.13
ycfJ	b1110	8.79	1.05E-04	1.19E-05	2.53826	0.13
yedA	b1959	8.74	3.11E-04	3.56E-05	2.80875	0.13
ruvA	b1861	8.72	1.62E-04	1.86E-05	2.97390	0.13
b0557	b0557	8.71	1.08E-04	1.24E-05	3.99638	0.13
eco	b2209	8.70	9.00E-05	1.03E-05	2.68897	0.13
ygjO	b3084	8.61	1.11E-04	1.29E-05	2.50469	0.13
b2496	b2496	8.57	9.18E-05	1.07E-05	2.21986	0.13
b1122	b1122	8.54	1.02E-04	1.19E-05	2.06167	0.13
b0619	b0619	8.51	1.63E-04	1.92E-05	2.65679	0.13
ybbC	b0498	8.51	2.79E-04	3.29E-05	4.12261	0.13
b4257	b4257	8.49	7.61E-05	8.97E-06	5.62620	0.13
dinI	b1061	8.48	7.02E-05	8.28E-06	4.02906	0.13
rplA	b3984	8.47	4.74E-04	5.60E-05	3.34653	0.13
b2255	b2255	8.45	1.74E-04	2.05E-05	1.52873	0.13
b1935	b1935	8.43	3.54E-04	4.20E-05	5.46538	0.13
xseA	b2509	8.42	1.83E-04	2.18E-05	6.13022	0.13
pabB	b1812	8.42	1.02E-04	1.21E-05	1.80738	0.13
rep	b3778	8.37	4.43E-05	5.29E-06	3.66071	0.13
b1667	b1667	8.33	1.17E-04	1.40E-05	2.96597	0.13
b0458	b0458	8.26	2.45E-04	2.97E-05	2.28207	0.13
b0245	b0245	8.26	6.34E-05	7.67E-06	1.89682	0.13
b1566	b1566	8.26	3.85E-03	4.66E-04	3.47493	0.13
rplV	b3315	8.20	2.64E-04	3.22E-05	2.72695	0.13
proW	b2678	8.19	4.87E-04	5.94E-05	3.59438	0.13
purA	b4177	8.18	3.52E-04	4.31E-05	2.74480	0.13
aroB	b3389	8.12	7.97E-05	9.81E-06	1.23277	0.13
b0603	b0603	8.11	1.56E-04	1.92E-05	5.07521	0.13
dicB	b1575	8.07	1.42E-04	1.76E-05	3.08861	0.13
b1720	b1720	8.03	4.12E-04	5.13E-05	6.28488	0.13
rnd	b1804	7.94	5.83E-05	7.34E-06	1.69960	0.13

Table A.2 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
ybfB	b0702	7.87	1.47E-04	1.87E-05	1.88648	0.13
yafP	b0234	7.86	2.07E-05	2.64E-06	3.62056	0.13
b0807	b0807	7.84	3.35E-05	4.28E-06	2.49821	0.13
rplS	b2606	7.78	9.95E-04	1.28E-04	4.79490	0.13
b2856	b2856	7.78	5.03E-05	6.47E-06	3.36835	0.13
insB_2	b0264	7.76	3.23E-04	4.16E-05	4.98449	0.13
b2174	b2174	7.74	6.40E-05	8.27E-06	6.08673	0.13
recO	b2565	7.71	1.04E-04	1.35E-05	1.63118	0.13
rhsC	b0700	7.71	3.77E-04	4.89E-05	5.20994	0.13
b1679	b1679	7.70	5.50E-05	7.14E-06	1.65561	0.13
yhbZ	b3183	7.68	2.09E-04	2.72E-05	3.72080	0.13
insB_3	b0274	7.67	3.22E-04	4.19E-05	3.77709	0.13
leuO	b0076	7.63	4.15E-05	5.45E-06	2.06817	0.13
oraA	b2698	7.58	6.80E-05	8.97E-06	5.17557	0.13
b1648	b1648	7.56	3.12E-04	4.13E-05	5.24103	0.13
b2100	b2100	7.55	2.39E-04	3.17E-05	1.82324	0.13
b0235	b0235	7.53	4.41E-05	5.86E-06	1.37157	0.13
rpsU	b3065	7.51	5.59E-04	7.45E-05	5.09985	0.13
b0453	b0453	7.51	3.94E-05	5.24E-06	1.96037	0.13
b1821	b1821	7.50	1.92E-04	2.56E-05	3.94226	0.13
yjjG	b4374	7.49	4.47E-05	5.96E-06	1.97943	0.13
b1162	b1162	7.48	1.79E-04	2.39E-05	4.61672	0.13
rplB	b3317	7.48	3.78E-04	5.05E-05	2.03653	0.13
fimB	b4312	7.47	3.43E-04	4.60E-05	2.63975	0.13
basS	b4112	7.40	1.65E-04	2.23E-05	2.37498	0.13
b0545	b0545	7.38	1.09E-04	1.47E-05	2.26142	0.13
b1567	b1567	7.38	9.42E-05	1.28E-05	2.86725	0.13
rnpA	b3704	7.37	1.55E-04	2.10E-05	1.97652	0.13
rplU	b3186	7.32	3.46E-04	4.73E-05	2.86289	0.13
b1560	b1560	7.28	2.81E-05	3.86E-06	2.55289	0.13
b3048	b3048	7.24	2.56E-04	3.54E-05	2.89697	0.13
yaeL	b0176	7.24	4.68E-05	6.47E-06	3.58024	0.13
b1374	b1374	7.22	6.19E-05	8.58E-06	1.95826	0.13
b3524	b3524	7.13	1.45E-04	2.04E-05	2.88832	0.13
ycfH	b1100	7.12	1.33E-04	1.87E-05	1.51484	0.13
ybaB	b0471	7.12	1.31E-04	1.84E-05	1.89010	0.13
yjhO	b4305	7.11	3.07E-05	4.32E-06	2.81420	0.13
yehC	b2110	7.11	6.09E-04	8.56E-05	7.01396	0.13
b2667	b2667	7.09	1.17E-04	1.64E-05	1.98763	0.13
rplR	b3304	7.09	1.61E-04	2.26E-05	2.10457	0.13
b0556	b0556	7.08	2.79E-05	3.94E-06	3.34760	0.13
ybaJ	b0461	7.05	1.37E-04	1.94E-05	7.45593	0.13
metG	b2114	7.04	8.34E-05	1.18E-05	3.55341	0.13

Table A.2 Continued.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
cfa	b1661	7.04	2.82E-04	4.00E-05	2.02211	0.13
appY	b0564	7.03	2.39E-04	3.40E-05	1.91988	0.13
galF	b2042	7.02	6.57E-05	9.36E-06	2.05182	0.13
yceD	b1088	7.01	1.25E-04	1.79E-05	1.97578	0.13
b0989	b0989	6.97	8.91E-05	1.28E-05	2.96090	0.13
b1160	b1160	6.95	1.55E-03	2.23E-04	38.85820	0.13
yhbC	b3170	6.95	1.58E-04	2.28E-05	6.08688	0.13
yggN	b2958	6.95	5.85E-05	8.43E-06	2.50313	0.13
nagB	b0678	6.94	1.34E-04	1.93E-05	1.38959	0.13
emrY	b2367	6.92	5.43E-05	7.85E-06	1.22895	0.13
rhsE	b1456	6.83	3.19E-04	4.67E-05	14.26517	0.13
yejL	b2187	6.81	1.06E-04	1.56E-05	1.89912	0.13
ycfC	b1132	6.78	1.15E-04	1.69E-05	1.67922	0.13
dnaJ	b0015	6.77	2.51E-04	3.71E-05	4.94130	0.13
rpsN	b3307	6.77	2.13E-04	3.15E-05	2.82970	0.13
dnaK	b0014	6.76	6.65E-04	9.84E-05	7.29396	0.13
cbl	b1987	6.76	4.37E-04	6.46E-05	11.87252	0.13
ygjN	b3083	6.74	1.49E-04	2.21E-05	5.26311	0.13
b0365	b0365	6.73	2.81E-05	4.18E-06	2.89633	0.13
dinP	b0231	6.64	1.14E-04	1.72E-05	1.36974	0.13
rpsI	b3230	6.62	1.16E-04	1.75E-05	2.16934	0.13
dnaX	b0470	6.62	1.08E-04	1.64E-05	4.93858	0.13
rpsT	b0023	6.62	3.27E-04	4.94E-05	3.57390	0.13
glpR	b3423	6.61	2.96E-05	4.47E-06	2.74333	0.13
yfhC	b2559	6.61	1.14E-04	1.73E-05	1.31058	0.13
b0558	b0558	6.57	1.37E-04	2.09E-05	5.70155	0.13
yjeQ	b4161	6.52	1.09E-04	1.67E-05	5.17684	0.13
b0753	b0753	6.50	1.32E-04	2.02E-05	1.49143	0.13
ygfE	b2910	6.46	2.41E-04	3.73E-05	2.50872	0.13
b0570	b0570	6.45	7.47E-05	1.16E-05	1.60850	0.13
ykgB	b0301	6.45	1.05E-04	1.63E-05	1.73210	0.13
b2628	b2628	6.45	8.07E-05	1.25E-05	1.94000	0.13
ybaA	b0456	6.42	5.53E-05	8.61E-06	5.28533	0.13
ybcI	b0527	6.42	3.61E-05	5.62E-06	1.53265	0.13
rplQ	b3294	6.41	1.27E-04	1.99E-05	2.16669	0.13
b1169	b1169	6.40	5.36E-04	8.36E-05	3.10000	0.13
hha	b0460	6.40	2.56E-04	3.99E-05	15.66890	0.13
fliR	b1950	6.39	8.33E-05	1.30E-05	1.80175	0.13
yadC	b0135	6.34	1.65E-04	2.61E-05	31.58550	0.13
b2271	b2271	6.31	9.34E-05	1.48E-05	1.53527	0.13
b1031	b1031	6.31	7.32E-05	1.16E-05	3.14964	0.13
proP	b4111	6.30	6.21E-04	9.86E-05	3.27837	0.13
b0725	b0725	6.29	1.62E-04	2.57E-05	1.44584	0.13

Table A.2 Continued.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
rpoH	b3461	6.27	2.81E-04	4.48E-05	2.94891	0.13
yfjB	b2615	6.27	3.17E-05	5.05E-06	1.41378	0.13
b1172	b1172	6.27	7.39E-04	1.18E-04	22.05197	0.13
gltF	b3214	6.25	4.31E-04	6.90E-05	3.16475	0.13
htpX	b1829	6.24	9.41E-04	1.51E-04	4.73113	0.13
argR	b3237	6.20	1.06E-04	1.70E-05	1.70372	0.13
rpsE	b3303	6.20	1.80E-04	2.90E-05	1.82368	0.13
b0326	b0326	6.19	5.33E-05	8.61E-06	1.76179	0.13
b2295	b2295	6.19	1.18E-04	1.91E-05	2.46238	0.13
mreC	b3250	6.17	8.96E-05	1.45E-05	1.19987	0.13
ybgA	b0707	6.17	7.67E-05	1.24E-05	1.64070	0.13
b1826	b1826	6.17	2.99E-04	4.85E-05	6.22021	0.13
yheL	b3343	6.16	7.29E-05	1.18E-05	3.96594	0.13
relF	b1562	6.15	1.04E-04	1.69E-05	2.06612	0.13
b2345	b2345	6.15	2.10E-04	3.41E-05	3.92391	0.13
yfhD	b2558	6.14	8.93E-05	1.45E-05	2.48216	0.13
b0517	b0517	6.13	7.38E-05	1.20E-05	1.27051	0.13
gadA	b3517	6.12	4.54E-05	7.42E-06	2.29403	0.13
yfjA	b2608	6.11	3.75E-04	6.14E-05	5.53209	0.13
nusA	b3169	6.09	6.58E-04	1.08E-04	2.95618	0.13
gusC	b1615	6.08	1.49E-04	2.46E-05	2.18084	0.13
rplY	b2185	6.08	9.19E-04	1.51E-04	9.28235	0.13
lit	b1139	6.07	3.95E-05	6.51E-06	3.02177	0.13
glpG	b3424	6.06	6.01E-05	9.92E-06	3.56821	0.13
b1743	b1743	6.04	6.94E-05	1.15E-05	4.78737	0.13
cca	b3056	6.03	4.02E-05	6.67E-06	2.12480	0.13
b0544	b0544	5.99	1.71E-04	2.85E-05	1.71373	0.13
rpmA	b3185	5.96	7.37E-05	1.24E-05	1.98119	0.13
b1844	b1844	5.96	3.28E-05	5.50E-06	2.01163	0.13
ptr	b2821	5.94	3.83E-05	6.45E-06	2.08514	0.13
b2639	b2639	5.94	2.53E-04	4.27E-05	1.74194	0.13
b1111	b1111	5.91	6.18E-05	1.04E-05	1.77435	0.13
b2850	b2850	5.91	2.02E-04	3.42E-05	4.97247	0.13
b2851	b2851	5.88	1.53E-04	2.60E-05	4.68961	0.13
yafN	b0232	5.87	2.17E-05	3.69E-06	3.23917	0.13
rpmC	b3312	5.86	3.43E-04	5.85E-05	1.54036	0.13
gor	b3500	5.85	4.66E-05	7.97E-06	2.06749	0.13
b1025	b1025	5.85	2.70E-04	4.61E-05	2.26778	0.13
b1678	b1678	5.84	6.01E-05	1.03E-05	2.07474	0.13
ygfD	b2918	5.83	1.05E-04	1.79E-05	2.69960	0.13
rimI	b4373	5.83	4.15E-05	7.11E-06	1.90910	0.13
yhiF	b3507	5.82	5.75E-05	9.88E-06	2.55734	0.13
rfaL	b3622	5.78	2.14E-04	3.71E-05	8.61193	0.13

Table A.2 Continued.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
yiiG	b3896	5.78	6.24E-05	1.08E-05	8.68446	0.13
frr	b0172	5.72	2.06E-04	3.60E-05	2.43256	0.13
evgA	b2369	5.72	1.88E-04	3.29E-05	5.18239	0.13
rcsA	b1951	5.71	5.93E-05	1.04E-05	3.19165	0.13
b1933	b1933	5.71	3.23E-05	5.65E-06	4.05852	0.13
rpsG	b3341	5.68	2.20E-04	3.87E-05	3.45688	0.13
yhiM	b3491	5.68	2.55E-04	4.49E-05	4.17489	0.13
lysA	b2838	5.67	7.09E-05	1.25E-05	1.33475	0.13
avtA	b3572	5.67	5.32E-05	9.37E-06	1.26197	0.13
yi82_1	b0017	5.67	2.83E-05	4.99E-06	1.85027	0.13
b0302	b0302	5.67	3.22E-05	5.68E-06	2.34591	0.13
rplM	b3231	5.66	3.76E-04	6.63E-05	2.38135	0.13
yhhT	b3474	5.66	1.10E-04	1.94E-05	2.25732	0.13
b1980	b1980	5.66	7.12E-05	1.26E-05	2.10709	0.13
yefI	b2032	5.65	5.95E-04	1.05E-04	23.64828	0.13
b1504	b1504	5.59	4.66E-05	8.33E-06	2.24986	0.13
rplC	b3320	5.57	6.30E-04	1.13E-04	2.52040	0.13
yhbU	b3158	5.57	1.63E-04	2.93E-05	1.95078	0.13
b0362	b0362	5.57	5.07E-05	9.11E-06	1.60107	0.13
b2584	b2584	5.56	2.94E-04	5.28E-05	1.63582	0.13
b1809	b1809	5.55	2.17E-04	3.91E-05	3.33967	0.13
yebH	b1822	5.54	7.93E-05	1.43E-05	3.44584	0.13
yjaA	b4011	5.53	8.97E-05	1.62E-05	3.42367	0.13
ydiB	b1692	5.51	3.56E-05	6.47E-06	2.10447	0.13
lhr	b1653	5.51	1.02E-04	1.86E-05	4.39556	0.13
b1770	b1770	5.50	1.08E-04	1.96E-05	8.03465	0.13
arsB	b3502	5.49	4.86E-05	8.85E-06	2.96884	0.13
b1586	b1586	5.48	2.48E-04	4.53E-05	2.71514	0.13
b0770	b0770	5.46	4.28E-05	7.84E-06	1.81680	0.13
truA	b2318	5.45	1.22E-04	2.24E-05	4.20776	0.13
ydbA	b1405	5.45	1.46E-04	2.67E-05	5.78541	0.13
clpB	b2592	5.45	8.45E-05	1.55E-05	2.57568	0.13
b1968	b1968	5.44	1.93E-04	3.55E-05	3.16973	0.13
b1170	b1170	5.44	3.69E-04	6.78E-05	4.28987	0.13
b0833	b0833	5.44	7.37E-05	1.35E-05	3.61791	0.13
b2665	b2665	5.43	2.23E-05	4.10E-06	1.66893	0.13
yigG	b3818	5.42	2.97E-05	5.47E-06	3.96198	0.13
rfc	b2035	5.42	1.51E-04	2.79E-05	5.26565	0.13
yeeC	b2010	5.41	6.87E-05	1.27E-05	1.56923	0.13
pdxJ	b2564	5.41	3.04E-05	5.61E-06	1.50737	0.13
b1546	b1546	5.41	1.14E-04	2.10E-05	1.75782	0.13
criR	b0620	5.38	3.25E-04	6.04E-05	3.30963	0.13
b2760	b2760	5.38	1.96E-04	3.65E-05	1.86342	0.13

Table A.2 Continued.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
tyrA	b2600	5.36	6.15E-05	1.15E-05	1.43362	0.13
yeeF	b2014	5.36	3.29E-05	6.14E-06	2.62683	0.13
yjfZ	b4204	5.35	6.58E-05	1.23E-05	4.22271	0.13
b1811	b1811	5.35	7.30E-05	1.37E-05	2.32463	0.13
b0986	b0986	5.35	1.24E-04	2.31E-05	2.07655	0.13
yiiL	b3901	5.34	8.03E-05	1.50E-05	1.77841	0.13
b3472	b3472	5.33	1.36E-04	2.55E-05	2.50337	0.13
hflC	b4175	5.32	1.16E-04	2.17E-05	2.34148	0.13
b2862	b2862	5.31	3.41E-04	6.42E-05	3.64852	0.13
b1165	b1165	5.31	8.52E-05	1.60E-05	2.81386	0.13
ldcC	b0186	5.29	7.48E-05	1.42E-05	1.76383	0.13
b1047	b1047	5.28	2.20E-04	4.17E-05	2.11494	0.13
b2027	b2027	5.28	3.22E-04	6.10E-05	3.89073	0.13
yefG	b2034	5.27	4.20E-04	7.96E-05	21.26114	0.13
xylE	b4031	5.27	7.37E-05	1.40E-05	1.22954	0.13
yaeF	b0193	5.26	2.91E-05	5.54E-06	1.34688	0.13
yidS	b3690	5.23	6.42E-05	1.23E-05	1.24724	0.13
hflX	b4173	5.22	1.06E-04	2.04E-05	3.23429	0.13
b1668	b1668	5.21	9.95E-05	1.91E-05	2.82766	0.13
b2466	b2466	5.21	8.30E-05	1.59E-05	1.72034	0.13
yjfI	b4181	5.21	2.68E-05	5.14E-06	4.29320	0.13
b1202	b1202	5.20	4.01E-05	7.72E-06	5.07043	0.24
b1972	b1972	5.18	1.33E-04	2.56E-05	2.17229	0.24
b2974	b2974	5.18	6.38E-05	1.23E-05	4.38452	0.24
farR	b0730	5.18	2.97E-04	5.74E-05	2.07780	0.24
acrA	b0463	5.18	4.80E-05	9.28E-06	1.76202	0.24
yhiL	b3490	5.17	5.29E-04	1.02E-04	24.36293	0.24
b2224	b2224	5.16	1.15E-04	2.22E-05	1.74479	0.24
b2857	b2857	5.15	2.36E-05	4.58E-06	3.59014	0.24
lasT	b4403	5.15	2.04E-04	3.97E-05	2.13569	0.24
b1462	b1462	5.13	8.71E-05	1.70E-05	2.21369	0.24
yfiM	b2586	5.12	1.84E-04	3.59E-05	3.49948	0.24
yifB	b3765	5.11	4.00E-05	7.83E-06	2.35694	0.24
yhhJ	b3485	5.11	8.87E-05	1.74E-05	2.26218	0.24
sdaB	b2797	5.10	8.44E-05	1.65E-05	1.29158	0.24
yjgL	b4253	5.10	6.73E-04	1.32E-04	2.23602	0.24
b1645	b1645	5.10	7.88E-05	1.55E-05	2.63817	0.24
yijP	b3955	5.09	1.01E-04	1.98E-05	2.77396	0.24
b0459	b0459	5.08	1.32E-04	2.59E-05	1.79005	0.24
tra5_4	b2089	5.08	1.05E-04	2.07E-05	2.14424	0.24
yeiR	b2173	5.08	2.42E-04	4.76E-05	3.86060	0.24
b2863	b2863	5.06	1.74E-03	3.44E-04	34.36140	0.24
yafX	b0248	5.05	8.74E-05	1.73E-05	1.51721	0.24

Table A.2 Continued.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
moaA	b0781	5.05	4.71E-05	9.33E-06	1.94149	0.24
gltX	b2400	5.02	1.20E-04	2.39E-05	2.32351	0.24
b0609	b0609	5.02	2.30E-05	4.59E-06	1.56030	0.24
flgH	b1079	-94.26	-4.31E-04	4.58E-06	0.35606	0.13
flgA	b1072	-18.54	-4.35E-04	2.35E-05	0.30263	0.13
pflB	b0903	-15.35	-7.80E-04	5.08E-05	0.06887	0.13
b1420	b1420	-14.90	-1.11E-04	7.46E-06	0.33906	0.13
nadR	b4390	-12.73	-6.68E-04	5.25E-05	0.47351	0.13
tnaA	b3708	-12.11	-1.74E-03	1.43E-04	0.03360	0.13
b0881	b0881	-11.79	-9.32E-05	7.90E-06	0.33145	0.13
flgJ	b1081	-10.84	-8.95E-05	8.26E-06	0.08939	0.13
flhA	b1879	-10.77	-1.58E-03	1.46E-04	0.47549	0.13
malK	b4035	-10.60	-2.09E-04	1.97E-05	0.14073	0.13
pta	b2297	-10.33	-2.50E-04	2.42E-05	0.24359	0.13
eno	b2779	-9.65	-8.48E-04	8.78E-05	0.24851	0.13
flgG	b1078	-9.47	-7.63E-04	8.06E-05	0.25117	0.13
b2483	b2483	-9.39	-1.62E-04	1.72E-05	0.36458	0.13
b0518	b0518	-9.22	-1.04E-04	1.13E-05	0.37059	0.13
soxR	b4063	-9.13	-6.41E-05	7.02E-06	0.86096	0.13
ybeK	b0651	-9.13	-1.76E-04	1.93E-05	0.33647	0.13
flgF	b1077	-9.02	-2.93E-04	3.25E-05	0.28841	0.13
nuoM	b2277	-8.94	-1.42E-04	1.59E-05	0.46934	0.13
gatZ	b2095	-8.70	-5.34E-04	6.14E-05	0.30044	0.13
pckA	b3403	-8.47	-2.31E-04	2.73E-05	0.42362	0.13
pspB	b1305	-8.46	-7.77E-05	9.18E-06	0.58935	0.13
rimJ	b1066	-8.33	-6.56E-04	7.87E-05	0.46580	0.13
nuoK	b2279	-8.32	-9.16E-05	1.10E-05	0.37893	0.13
sufI	b3017	-8.27	-1.22E-04	1.47E-05	0.68089	0.13
trxB	b0888	-7.96	-2.35E-04	2.95E-05	0.27444	0.13
yfiD	b2579	-7.96	-6.24E-05	7.84E-06	0.30688	0.13
murB	b3972	-7.60	-1.96E-04	2.57E-05	0.66190	0.13
nuoH	b2282	-7.43	-9.77E-05	1.31E-05	0.55544	0.13
yhjY	b3548	-7.39	-8.40E-05	1.14E-05	0.68764	0.13
b1312	b1312	-7.32	-2.17E-04	2.96E-05	0.50939	0.13
hybG	b2990	-7.27	-5.68E-05	7.81E-06	0.45108	0.13
pgk	b2926	-7.24	-4.89E-04	6.76E-05	0.36636	0.13
b1483	b1483	-7.18	-1.12E-03	1.56E-04	0.41853	0.13
ychE	b1242	-7.12	-3.85E-05	5.41E-06	0.40053	0.13
minE	b1174	-7.07	-1.52E-03	2.15E-04	0.43510	0.13
uspA	b3495	-6.98	-2.58E-03	3.69E-04	0.10082	0.13
yicE	b3654	-6.98	-1.12E-04	1.60E-05	0.49395	0.13
ytfA	b4205	-6.95	-7.99E-04	1.15E-04	0.44577	0.13

Table A.2 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
tnaL	b3707	-6.95	-1.47E-03	2.12E-04	0.03134	0.13
b0701	b0701	-6.88	-1.05E-04	1.53E-05	0.27309	0.13
radC	b3638	-6.78	-2.89E-04	4.27E-05	0.26279	0.13
hycB	b2724	-6.65	-2.66E-04	3.99E-05	0.41062	0.13
hybB	b2995	-6.62	-4.63E-04	6.99E-05	0.22632	0.13
yhfQ	b3374	-6.60	-4.17E-04	6.32E-05	0.50876	0.13
fnr	b1334	-6.58	-7.79E-05	1.18E-05	0.75639	0.13
b2770	b2770	-6.55	-1.72E-04	2.63E-05	0.47176	0.24
hycF	b2720	-6.48	-1.96E-04	3.02E-05	0.50728	0.24
nikE	b3480	-6.48	-8.89E-04	1.37E-04	0.26445	0.24
glgS	b3049	-6.46	-1.32E-04	2.03E-05	0.23180	0.24
nuoI	b2281	-6.45	-6.64E-05	1.03E-05	0.36214	0.24
flgI	b1080	-6.36	-4.00E-04	6.28E-05	0.65291	0.24

Table A.3 Input parameters and computed quantities for the SAM performed on data from macroarray comparison 2. Results from this analysis can be seen in Table A.4.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	1.82970
(Upper Cutoff, Lower Cutoff)	(∞, -6.31473)
Computed Quantities	
Computed Exchangeability Factor S0	1.28493E-06
S0 percentile	0.01
False Significant Number (Median, 90 percentile)	(0.96319, 7.70550)
False Discovery Rate (Median, 90 percentile)	(10.70208, 85.61665)
Pi0Hat	0.96319

Table A.4 Results of SAM performed on data from macroarray comparison 2.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
b1566	b1566	-12.36	-4.35E-03	3.52E-04	0.19534	10.70
b1972	b1972	-7.90	-8.74E-05	1.11E-05	0.64440	10.70
b3618	b3618	-7.47	-6.86E-04	9.18E-05	0.45350	10.70
rnd	b1804	-7.17	-2.19E-05	3.05E-06	0.84552	10.70
b1527	b1527	-6.70	-1.34E-04	2.00E-05	0.60383	10.70
yrbK	b3199	-6.51	-5.18E-05	7.96E-06	0.41146	10.70
yhhH	b3483	-6.45	-1.39E-04	2.15E-05	0.47148	10.70
rpmC	b3312	-6.34	-1.29E-04	2.03E-05	0.86832	10.70
yjcF	b4066	-6.31	-3.68E-05	5.83E-06	0.38730	10.70

Table A.5 Input parameters and computed quantities for the SAM performed on data from macroarray comparison 3. Results from this analysis can be seen in Table A.6.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	4.29965
(Upper Cutoff, Lower Cutoff)	(7.92938, -6.27216)
Computed Quantities	
Computed Exchangeability Factor S0	1.13046E-06
S0 percentile	0.01
False Significant Number (Median, 90 percentile)	(0.53122, 2.12488)
False Discovery Rate (Median, 90 percentile)	(0.22042, 0.88169)
Pi0Hat	0.53122

Table A.6 Results of SAM performed on data from macroarray comparison 3.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
tnaA	b3708	62.66	4.93E-04	7.87E-06	9.16392	0.22
sucA	b0726	22.22	2.68E-04	1.21E-05	5.71450	0.22
kefB	b3350	17.01	2.13E-04	1.25E-05	1.80029	0.22
aceA	b4015	16.02	1.14E-04	7.13E-06	6.39762	0.22
b0375	b0375	13.93	1.15E-04	8.25E-06	2.63562	0.22
pspB	b1305	12.84	1.20E-04	9.35E-06	2.07725	0.22
gatY	b2096	12.66	1.39E-04	1.09E-05	1.75796	0.22
yajG	b0434	11.92	3.02E-04	2.53E-05	1.66703	0.22
flgA	b1072	11.71	3.20E-04	2.73E-05	2.69577	0.22
argH	b3960	11.07	4.26E-05	3.85E-06	1.32761	0.22
b1565	b1565	11.04	1.77E-04	1.61E-05	1.43317	0.22
flgJ	b1081	10.81	2.45E-05	2.27E-06	3.78816	0.22
trxB	b0888	10.70	1.94E-04	1.81E-05	3.17625	0.22
ykgC	b0304	10.37	5.72E-04	5.52E-05	1.47305	0.22
flgE	b1076	10.22	1.10E-03	1.07E-04	5.58404	0.22
ybeL	b0643	9.76	6.94E-05	7.12E-06	1.85677	0.22
cysK	b2414	9.41	5.09E-04	5.41E-05	7.53867	0.22
b1526	b1526	9.27	6.33E-04	6.83E-05	2.35106	0.22
nadR	b4390	9.24	6.35E-04	6.88E-05	2.05785	0.22
ydaH	b1336	9.00	6.32E-05	7.02E-06	2.52736	0.22
b2968	b2968	8.58	4.16E-03	4.85E-04	2.97166	0.22
trxA	b3781	8.56	1.15E-03	1.34E-04	4.44428	0.22
yabQ	b0057	8.50	3.62E-04	4.27E-05	3.57659	0.22
b1010	b1010	8.37	1.37E-03	1.64E-04	7.16763	0.22
phnO	b4093	8.10	1.19E-04	1.47E-05	2.52447	0.22
gltA	b0720	8.08	7.73E-04	9.57E-05	3.84800	0.22
atpF	b3736	7.93	1.12E-04	1.42E-05	3.43842	0.22
osmB	b1283	-40.36	-3.01E-04	7.46E-06	0.21690	0.22
rpsP	b2609	-32.75	-3.53E-04	1.08E-05	0.29755	0.22
b1524	b1524	-31.03	-1.85E-04	5.97E-06	0.38278	0.22
yjbM	b4048	-24.91	-1.32E-04	5.32E-06	0.20252	0.22
b1566	b1566	-24.83	-4.64E-03	1.87E-04	0.14167	0.22
b3618	b3618	-22.06	-9.88E-04	4.48E-05	0.21250	0.22
yibJ	b3595	-20.56	-8.79E-04	4.28E-05	0.15329	0.22
b2668	b2668	-19.51	-2.61E-04	1.34E-05	0.53113	0.22
guaB	b2508	-19.35	-2.03E-04	1.05E-05	0.51823	0.22
slp	b3506	-19.21	-2.19E-04	1.14E-05	0.44780	0.22
yjcF	b4066	-18.89	-4.62E-05	2.45E-06	0.23116	0.22
b2339	b2339	-17.33	-1.66E-04	9.57E-06	0.38720	0.22
b1527	b1527	-16.16	-2.31E-04	1.43E-05	0.31541	0.22
leuO	b0076	-16.08	-3.97E-05	2.47E-06	0.50664	0.22
ybbD	b0500	-16.02	-5.06E-05	3.16E-06	0.12881	0.22

Table A.6 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
yhiK	b3489	-15.61	-1.65E-04	1.06E-05	0.12497	0.22
b2639	b2639	-15.48	-3.58E-04	2.31E-05	0.39782	0.22
b1983	b1983	-15.29	-7.48E-04	4.89E-05	0.12753	0.22
b1970	b1970	-15.20	-2.03E-04	1.33E-05	0.29050	0.22
b3048	b3048	-15.10	-2.66E-04	1.76E-05	0.32088	0.22
def	b3287	-14.86	-8.46E-05	5.70E-06	0.62009	0.22
yjhC	b4280	-14.72	-1.08E-04	7.32E-06	0.35541	0.22
rpmC	b3312	-14.66	-4.02E-04	2.74E-05	0.58912	0.22
rep	b3778	-14.51	-3.71E-05	2.56E-06	0.39140	0.22
ydcD	b1457	-14.28	-6.42E-04	4.50E-05	0.16632	0.22
b0819	b0819	-14.11	-2.09E-04	1.48E-05	0.72680	0.22
emrY	b2367	-14.00	-5.36E-05	3.83E-06	0.81595	0.22
greA	b3181	-13.94	-9.88E-05	7.08E-06	0.36364	0.22
b0501	b0501	-13.79	-6.46E-05	4.69E-06	0.31618	0.22
tdcB	b3117	-13.34	-8.49E-05	6.37E-06	0.60757	0.22
b0543	b0543	-13.22	-1.02E-04	7.70E-06	0.49114	0.22
b0482	b0482	-13.20	-2.39E-04	1.81E-05	0.59233	0.22
rhsB	b3482	-12.90	-4.63E-04	3.58E-05	0.14132	0.22
fliR	b1950	-12.78	-9.46E-05	7.40E-06	0.49494	0.22
b0619	b0619	-12.73	-1.21E-04	9.50E-06	0.53739	0.22
rplS	b2606	-12.72	-8.21E-04	6.45E-05	0.34693	0.22
b1228	b1228	-12.60	-1.87E-03	1.48E-04	0.21687	0.22
yahA	b0315	-12.55	-8.36E-04	6.66E-05	0.22710	0.22
pabC	b1096	-12.50	-7.47E-05	5.98E-06	0.47991	0.22
b2340	b2340	-12.47	-7.51E-05	6.02E-06	0.39632	0.22
b0805	b0805	-12.43	-8.31E-05	6.69E-06	0.48847	0.22
yhhS	b3473	-12.37	-4.96E-05	4.01E-06	0.50868	0.22
dcm	b1961	-12.22	-2.98E-04	2.44E-05	0.38455	0.22
yi52_8	b2192	-12.20	-1.05E-03	8.64E-05	0.46257	0.22
b0600	b0600	-12.17	-2.47E-04	2.03E-05	0.53008	0.22
b3524	b3524	-11.91	-1.64E-04	1.38E-05	0.26232	0.22
intB	b4271	-11.88	-1.15E-04	9.72E-06	0.20873	0.22
yrbK	b3199	-11.83	-5.50E-05	4.65E-06	0.37537	0.22
appY	b0564	-11.67	-2.28E-04	1.96E-05	0.54210	0.22
b1972	b1972	-11.24	-1.46E-04	1.30E-05	0.40675	0.22
yhhH	b3483	-11.21	-2.33E-04	2.08E-05	0.11268	0.22
pepP	b2908	-11.12	-1.39E-04	1.25E-05	0.45926	0.22
b2623	b2623	-11.05	-1.57E-04	1.42E-05	0.53057	0.22
fabA	b0954	-10.84	-2.04E-04	1.88E-05	0.51585	0.22
b1762	b1762	-10.81	-1.08E-04	9.97E-06	0.35008	0.22
trmD	b2607	-10.73	-1.98E-03	1.84E-04	0.25951	0.22
tdcR	b3119	-10.69	-4.76E-05	4.45E-06	0.20697	0.22
b0245	b0245	-10.56	-5.37E-05	5.08E-06	0.59925	0.22

Table A.6 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
b0499	b0499	-10.53	-5.37E-04	5.10E-05	0.25977	0.22
b1980	b1980	-10.53	-6.89E-05	6.54E-06	0.49123	0.22
recA	b2699	-10.45	-1.76E-03	1.68E-04	0.29069	0.22
b3046	b3046	-10.45	-8.66E-05	8.28E-06	0.31524	0.22
sfhB	b2594	-10.41	-6.13E-05	5.89E-06	0.38683	0.22
b2854	b2854	-10.40	-5.16E-05	4.96E-06	0.12404	0.22
b0357	b0357	-10.35	-1.39E-03	1.34E-04	0.19326	0.22
b0453	b0453	-10.27	-3.68E-05	3.58E-06	0.54241	0.22
ybfC	b0704	-10.26	-1.69E-04	1.65E-05	0.44187	0.22
dut	b3640	-10.20	-5.95E-05	5.83E-06	0.46677	0.22
yfhC	b2559	-10.15	-1.09E-04	1.07E-05	0.77387	0.22
gadB	b1493	-10.07	-2.06E-04	2.05E-05	0.68305	0.22
yi52_7	b2030	-10.06	-1.10E-03	1.10E-04	0.37447	0.22
yjhB	b4279	-10.06	-1.85E-04	1.84E-05	0.29880	0.22
b2545	b2545	-10.04	-1.11E-04	1.11E-05	0.76397	0.22
b1060	b1060	-10.03	-5.08E-05	5.06E-06	0.50604	0.22
asnC	b3743	-10.02	-4.58E-05	4.57E-06	0.48224	0.22
b2542	b2542	-9.97	-1.27E-04	1.28E-05	0.77567	0.22
b0546	b0546	-9.95	-1.28E-04	1.29E-05	0.12954	0.22
b0235	b0235	-9.92	-3.36E-05	3.39E-06	0.79342	0.22
miaA	b4171	-9.91	-2.36E-04	2.38E-05	0.18859	0.22
fimB	b4312	-9.88	-3.51E-04	3.55E-05	0.36500	0.22
rplA	b3984	-9.84	-3.33E-04	3.39E-05	0.50706	0.22
glpD	b3426	-9.83	-6.57E-05	6.68E-06	0.11163	0.22
yehA	b2108	-9.83	-1.09E-04	1.11E-05	0.30566	0.22
ygcA	b2785	-9.78	-8.33E-05	8.51E-06	0.45240	0.22
rnpA	b3704	-9.67	-8.60E-05	8.90E-06	0.72542	0.22
yceB	b1063	-9.62	-7.47E-05	7.77E-06	0.63871	0.22
flgM	b1071	-9.60	-2.86E-04	2.98E-05	0.53549	0.22
ygiF	b3054	-9.57	-1.35E-04	1.41E-05	0.62655	0.22
b4257	b4257	-9.56	-7.31E-05	7.65E-06	0.21029	0.22
b2191	b2191	-9.49	-1.60E-04	1.68E-05	0.40829	0.22
b1963	b1963	-9.48	-4.17E-04	4.40E-05	0.16151	0.22
yafO	b0233	-9.41	-9.30E-05	9.89E-06	0.69563	0.22
b0556	b0556	-9.38	-2.69E-05	2.87E-06	0.32343	0.22
b1171	b1171	-9.33	-8.28E-04	8.88E-05	0.18237	0.22
b1730	b1730	-9.30	-1.11E-04	1.19E-05	0.23650	0.22
yafD	b0209	-9.18	-3.87E-05	4.22E-06	0.81140	0.22
tra5_3	b1026	-9.12	-1.54E-04	1.69E-05	0.55817	0.22
mreB	b3251	-9.09	-1.63E-04	1.80E-05	0.74514	0.22
b0955	b0955	-9.03	-4.73E-05	5.24E-06	0.37800	0.22
b3042	b3042	-9.01	-1.05E-04	1.16E-05	0.54484	0.22
rpsO	b3165	-8.93	-6.52E-04	7.31E-05	0.34138	0.22

Table A.6 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
ybfB	b0702	-8.91	-1.43E-04	1.61E-05	0.54220	0.22
argR	b3237	-8.87	-9.80E-05	1.11E-05	0.61621	0.22
nusA	b3169	-8.83	-4.13E-04	4.68E-05	0.58444	0.22
ybiF	b0813	-8.83	-5.82E-05	6.59E-06	0.64717	0.22
b2081	b2081	-8.82	-9.45E-05	1.07E-05	0.40585	0.22
rnd	b1804	-8.75	-4.71E-05	5.38E-06	0.66751	0.22
tdh	b3616	-8.59	-1.13E-04	1.31E-05	0.68578	0.22
b1560	b1560	-8.59	-2.66E-05	3.09E-06	0.42483	0.22
yafP	b0234	-8.54	-1.62E-05	1.90E-06	0.43438	0.22
flgN	b1070	-8.45	-4.02E-04	4.75E-05	0.63678	0.22
yfaE	b2236	-8.42	-1.90E-04	2.26E-05	0.24143	0.22
b2856	b2856	-8.42	-4.66E-05	5.54E-06	0.34868	0.22
yaeM	b0173	-8.33	-7.44E-05	8.93E-06	0.73391	0.22
b0603	b0603	-8.33	-1.55E-04	1.86E-05	0.20441	0.22
ruvA	b1861	-8.31	-1.27E-04	1.53E-05	0.47927	0.22
yhiJ	b3488	-8.31	-5.91E-04	7.11E-05	0.09907	0.22
yiaV	b3586	-8.29	-4.21E-05	5.08E-06	0.43507	0.22
b3913	b3913	-8.27	-4.96E-04	6.00E-05	0.32426	0.22
ychF	b1203	-8.26	-1.20E-04	1.45E-05	0.37577	0.22
yfgB	b2517	-8.26	-1.23E-04	1.49E-05	0.77093	0.22
aroF	b2601	-8.20	-1.57E-04	1.92E-05	0.60899	0.22
yafN	b0232	-8.19	-1.92E-05	2.35E-06	0.38735	0.22
rplD	b3319	-8.18	-3.45E-04	4.21E-05	0.43937	0.22
insB_3	b0274	-8.05	-3.26E-04	4.05E-05	0.25430	0.22
b0557	b0557	-8.04	-7.05E-05	8.76E-06	0.51200	0.22
recN	b2616	-8.02	-5.39E-04	6.71E-05	0.53623	0.22
hydH	b4003	-7.94	-8.62E-05	1.09E-05	0.60475	0.22
smpA	b2617	-7.94	-6.34E-05	7.98E-06	0.55477	0.22
ykgB	b0301	-7.93	-9.89E-05	1.25E-05	0.60211	0.22
b1935	b1935	-7.91	-3.34E-04	4.22E-05	0.23028	0.22
yfiA	b2597	-7.90	-1.37E-03	1.74E-04	0.37636	0.22
b0769	b0769	-7.88	-2.69E-04	3.42E-05	0.71234	0.22
rpsC	b3314	-7.88	-4.83E-04	6.14E-05	0.49105	0.22
b2496	b2496	-7.86	-6.06E-05	7.70E-06	0.63741	0.22
lit	b1139	-7.83	-4.41E-05	5.63E-06	0.25420	0.22
b0825	b0825	-7.81	-5.57E-05	7.13E-06	0.81038	0.22
aroB	b3389	-7.78	-7.30E-05	9.38E-06	0.82708	0.22
b0859	b0859	-7.77	-9.18E-05	1.18E-05	0.54091	0.22
eda	b1850	-7.75	-3.75E-05	4.84E-06	0.78622	0.22
b1808	b1808	-7.69	-1.43E-04	1.86E-05	0.70960	0.22
ycfC	b1132	-7.67	-8.12E-05	1.06E-05	0.71377	0.22
rpsU	b3065	-7.63	-5.49E-04	7.19E-05	0.21112	0.22
b2061	b2061	-7.59	-3.83E-05	5.04E-06	0.70104	0.22

Table A.6 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
ychN	b1219	-7.57	-7.56E-05	9.98E-06	0.58166	0.22
yebG	b1848	-7.57	-6.01E-04	7.94E-05	0.30068	0.22
sdaB	b2797	-7.55	-6.75E-05	8.94E-06	0.81931	0.22
flhC	b1891	-7.55	-3.82E-04	5.07E-05	0.54665	0.22
ycfJ	b1110	-7.53	-8.64E-05	1.15E-05	0.50020	0.22
b0024	b0024	-7.49	-3.59E-05	4.80E-06	0.37216	0.22
b0302	b0302	-7.47	-3.00E-05	4.01E-06	0.46603	0.22
insB_2	b0264	-7.41	-3.12E-04	4.22E-05	0.22718	0.22
oraA	b2698	-7.34	-6.45E-05	8.78E-06	0.23456	0.22
yjjG	b4374	-7.34	-2.75E-05	3.74E-06	0.69551	0.22
gadA	b3517	-7.34	-4.63E-05	6.31E-06	0.42443	0.22
frdA	b4154	-7.31	-8.68E-05	1.19E-05	0.54945	0.22
b0790	b0790	-7.28	-7.65E-05	1.05E-05	0.73222	0.22
nagB	b0678	-7.25	-1.18E-04	1.63E-05	0.75261	0.22
kbl	b3617	-7.24	-1.51E-04	2.09E-05	0.28308	0.22
b1933	b1933	-7.21	-3.09E-05	4.28E-06	0.27823	0.22
b0365	b0365	-7.21	-2.84E-05	3.94E-06	0.33874	0.22
rhsC	b0700	-7.18	-3.47E-04	4.83E-05	0.25628	0.22
yejF	b2180	-7.18	-4.80E-05	6.68E-06	0.44506	0.22
ybbC	b0498	-7.16	-2.33E-04	3.26E-05	0.36761	0.22
yedA	b1959	-7.15	-2.94E-04	4.10E-05	0.39253	0.22
dicB	b1575	-7.13	-1.18E-04	1.65E-05	0.43806	0.22
b2174	b2174	-7.08	-5.77E-05	8.15E-06	0.24625	0.22
yhiN	b3492	-7.08	-1.07E-04	1.52E-05	0.51072	0.22
ptr	b2821	-7.05	-3.80E-05	5.39E-06	0.48444	0.22
b0671	b0671	-7.03	-7.73E-05	1.10E-05	0.34755	0.22
cfa	b1661	-7.02	-2.55E-04	3.63E-05	0.54265	0.22
recO	b2565	-7.01	-8.38E-05	1.19E-05	0.68964	0.22
yeeC	b2010	-7.01	-7.35E-05	1.05E-05	0.61204	0.22
coaA	b3974	-6.98	-1.17E-04	1.68E-05	0.76844	0.22
b1643	b1643	-6.92	-2.38E-05	3.44E-06	0.30465	0.22
ybiJ	b0802	-6.90	-5.35E-05	7.76E-06	0.84143	0.22
ilvC	b3774	-6.89	-1.08E-04	1.56E-05	0.76268	0.22
ybaJ	b0461	-6.82	-1.28E-04	1.88E-05	0.18770	0.22
rpoE	b2573	-6.81	-4.51E-05	6.62E-06	0.30006	0.22
yidR	b3689	-6.80	-1.58E-04	2.33E-05	0.46825	0.22
ilvY	b3773	-6.80	-1.33E-04	1.96E-05	0.76879	0.22
rplB	b3317	-6.80	-3.41E-04	5.02E-05	0.54007	0.22
yhbC	b3170	-6.80	-1.53E-04	2.25E-05	0.19269	0.22
ygfD	b2918	-6.80	-1.25E-04	1.84E-05	0.24521	0.22
ydhA	b1639	-6.78	-7.12E-05	1.05E-05	0.26276	0.22
b0558	b0558	-6.76	-1.32E-04	1.96E-05	0.20360	0.22
b0545	b0545	-6.68	-9.24E-05	1.38E-05	0.52515	0.22

Table A.6 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
rhlB	b3780	-6.67	-4.80E-04	7.19E-05	0.69228	0.22
ribH	b0415	-6.67	-3.73E-05	5.59E-06	0.56091	0.22
ygfE	b2910	-6.64	-2.07E-04	3.12E-05	0.48304	0.22
b1604	b1604	-6.63	-1.28E-04	1.92E-05	0.29892	0.22
b2294	b2294	-6.59	-5.66E-05	8.58E-06	0.39812	0.22
b1162	b1162	-6.59	-1.68E-04	2.54E-05	0.26617	0.22
htgA	b0012	-6.57	-5.10E-05	7.77E-06	0.38625	0.22
b1742	b1742	-6.56	-1.30E-04	1.98E-05	0.44085	0.22
yeiR	b2173	-6.56	-2.84E-04	4.34E-05	0.12794	0.22
b1720	b1720	-6.56	-3.50E-04	5.33E-05	0.28685	0.22
pncB	b0931	-6.54	-1.18E-04	1.81E-05	0.54671	0.22
rpsI	b3230	-6.53	-7.41E-05	1.13E-05	0.65617	0.22
yijP	b3955	-6.48	-8.70E-05	1.34E-05	0.44674	0.22
xseA	b2509	-6.48	-1.43E-04	2.21E-05	0.34597	0.22
b3047	b3047	-6.47	-1.84E-04	2.84E-05	0.54756	0.22
stpA	b2669	-6.47	-2.56E-04	3.97E-05	0.37009	0.22
gltF	b3214	-6.45	-4.17E-04	6.47E-05	0.33857	0.22
b2100	b2100	-6.41	-1.98E-04	3.08E-05	0.62717	0.22
ybaA	b0456	-6.37	-4.98E-05	7.82E-06	0.26959	0.22
menA	b3930	-6.36	-8.00E-05	1.26E-05	0.74047	0.22
yheL	b3343	-6.36	-6.90E-05	1.09E-05	0.29170	0.22
cbl	b1987	-6.35	-4.09E-04	6.44E-05	0.14311	0.22
b0847	b0847	-6.32	-1.51E-04	2.38E-05	0.74659	0.22
b1728	b1728	-6.31	-5.17E-05	8.20E-06	0.39218	0.22
yaeL	b0176	-6.30	-4.05E-05	6.42E-06	0.37670	0.22
adhC	b0356	-6.29	-9.06E-05	1.44E-05	0.62982	0.22
yhbZ	b3183	-6.27	-1.55E-04	2.47E-05	0.45829	0.22

Table A.7 Input parameters and computed quantities for the SAM performed on data from macroarray comparison 4.
Results from this analysis can be seen in Table A.8.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	4.45362
(Upper Cutoff, Lower Cutoff)	(9.12068, -7.67986)
Computed Quantities	
Computed Exchangeability Factor S0	1.31E-06
S0 percentile	0.01
False Significant Number (Median, 90 percentile)	(0.63653, 3.81920)
False Discovery Rate (Median, 90 percentile)	(1.72036, 10.32216)
Pi0Hat	0.63653

Table A.8 Results of S.A.M. performed on data from macroarray comparison 4.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
trxA	b3781	14.62	1.10E-03	7.51E-05	3.86647	1.72
sucA	b0726	11.98	1.69E-04	1.41E-05	2.08473	1.72
aceA	b4015	11.95	9.59E-05	8.03E-06	3.42435	1.72
trxB	b0888	10.31	1.85E-04	1.80E-05	2.90453	1.72
cysK	b2414	9.79	5.21E-04	5.32E-05	8.92494	1.72
accB	b3255	9.21	8.44E-05	9.17E-06	1.67666	1.72
yeeE	b2013	9.12	8.49E-05	9.30E-06	2.86196	1.72
yhhH	b3483	-15.31	-9.44E-05	6.16E-06	0.23898	1.72
tdh	b3616	-13.25	-1.00E-04	7.55E-06	0.71124	1.72
rfbA	b2039	-12.15	-6.14E-05	5.05E-06	0.29019	1.72
miaA	b4171	-11.61	-1.43E-04	1.23E-05	0.27732	1.72
yahA	b0315	-11.59	-5.68E-04	4.90E-05	0.30195	1.72
b2345	b2345	-11.12	-1.19E-04	1.07E-05	0.43772	1.72
b2863	b2863	-10.72	-1.15E-03	1.08E-04	0.22154	1.72
mcrA	b1159	-10.69	-4.42E-05	4.13E-06	0.24837	1.72
b2851	b2851	-10.06	-1.16E-04	1.15E-05	0.35458	1.72
rpmC	b3312	-9.95	-2.73E-04	2.74E-05	0.67846	1.72
b0834	b0834	-9.90	-6.84E-05	6.91E-06	0.53977	1.72
rplR	b3304	-9.52	-8.69E-05	9.13E-06	0.69150	1.72
b0619	b0619	-9.42	-1.43E-04	1.51E-05	0.49634	1.72
yjgT	b4265	-9.11	-1.24E-04	1.36E-05	0.66428	1.72
yhiJ	b3488	-9.06	-3.09E-04	3.41E-05	0.17350	1.72
eutH	b2452	-8.74	-1.59E-04	1.81E-05	0.55903	1.72
rfc	b2035	-8.42	-9.56E-05	1.14E-05	0.27038	1.72
b2849	b2849	-8.42	-3.34E-05	3.97E-06	0.39337	1.72
pdxH	b1638	-8.36	-3.02E-05	3.61E-06	0.71318	1.72
rfaG	b3631	-8.33	-1.46E-04	1.75E-05	0.72897	1.72
rpmA	b3185	-8.29	-3.93E-05	4.75E-06	0.71763	1.72
mviN	b1069	-8.28	-1.66E-04	2.00E-05	0.31942	1.72
rpsI	b3230	-8.25	-6.45E-05	7.83E-06	0.68663	1.72
kbl	b3617	-8.20	-8.78E-05	1.07E-05	0.40527	1.72
ybfC	b0704	-8.19	-1.31E-04	1.61E-05	0.50493	1.72
avtA	b3572	-8.17	-4.36E-05	5.34E-06	0.82918	1.72
b2848	b2848	-8.14	-3.76E-05	4.62E-06	0.51814	1.72
yjeB	b4178	-7.95	-3.02E-05	3.80E-06	0.33546	1.72
b1549	b1549	-7.92	-3.09E-05	3.90E-06	0.37306	1.72
rpsS	b3316	-7.68	-2.57E-04	3.35E-05	0.61165	1.72

Table A.9 Input parameters and computed quantities for the SAM performed on data from macroarray comparison 5. Results from this analysis can be seen in Table A.10.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	4.89542
(Upper Cutoff, Lower Cutoff)	(8.81482, -∞)
Computed Quantities	
Computed Exchangeability Factor S0	1.12824E-06
S0 percentile	0.01
False Significant Number (Median, 90 percentile)	(0.80196, 4.89194)
False Discovery Rate (Median, 90 percentile)	(4.45532, 27.17744)
Pi0Hat	0.80196

Table A.10 Results of SAM performed on data from macroarray comparison 5.

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
b1160	b1160	42.71	1.81E-04	4.23E-06	5.41337	4.46
b1171	b1171	17.72	1.16E-04	6.57E-06	2.70459	4.46
acpP	b1094	16.13	3.01E-04	1.87E-05	1.77169	4.46
yahA	b0315	14.61	1.58E-04	1.08E-05	2.80863	4.46
b2863	b2863	13.76	2.76E-04	2.01E-05	6.28593	4.46
pncB	b0931	12.13	6.05E-05	4.99E-06	1.73863	4.46
stpA	b2669	11.64	1.04E-04	8.95E-06	3.24150	4.46
recA	b2699	10.74	3.36E-04	3.13E-05	1.87556	4.46
yjdJ	b4127	10.24	4.84E-05	4.73E-06	2.09085	4.46
fabF	b1095	9.83	3.91E-04	3.98E-05	2.17055	4.46
sodA	b3908	9.03	1.98E-04	2.19E-05	1.73196	4.46
ubiG	b2232	9.00	3.92E-05	4.36E-06	1.63792	4.46
b1172	b1172	8.99	6.77E-05	7.53E-06	2.92971	4.46
hslV	b3932	8.97	5.10E-05	5.68E-06	1.83179	4.46
b1122	b1122	8.96	5.14E-05	5.74E-06	1.53639	4.46
b3914	b3914	8.83	1.46E-04	1.66E-05	3.69161	4.46
b0499	b0499	8.83	1.06E-04	1.20E-05	2.28540	4.46
b0217	b0217	8.81	7.79E-05	8.84E-06	2.68403	4.46

Appendix B

Table B.1 Alignment by Blattner number order of the results of standard analysis and S.A.M from macroarray experiments. Parameters for SAM analysis can be seen in Table A.1.

Standard Analysis			SAM Analysis			Standard Analysis			SAM Analysis		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
			htgA	b0012	3.0				*b0557*	b0557	4.0
dnaK	b0014	7.4	*dnaK*	b0014	7.3				*b0558*	b0558	5.7
			dnaJ	b0015	4.9				*appY*	b0564	1.9
			yi82_1	b0017	1.9				*b0570*	b0570	1.6
			rpsT	b0023	3.6				*b0600*	b0600	1.9
			leuO	b0076	2.1				*b0603*	b0603	5.1
			ftsL	b0083	2.4				*b0609*	b0609	1.6
yadC	b0135	31.9	*yadC*	b0135	31.6				*b0619*	b0619	2.7
			htrA	b0161	3.9				*criR*	b0620	3.3
			frr	b0172	2.4				*nagB*	b0678	1.4
			yaeL	b0176	3.6				*rhsC*	b0700	5.2
			ldcC	b0186	1.8				*ybfB*	b0702	1.9
			yaeF	b0193	1.4				*ybfC*	b0704	2.7
			dinP	b0231	1.4	*ybfD*	b0706	9.3			
			yafN	b0232	3.2				*ybgA*	b0707	1.6
			yafO	b0233	1.6				*b0725*	b0725	1.4
			yafP	b0234	3.6				*farR*	b0730	2.1
			b0235	b0235	1.4				*b0753*	b0753	1.5
			b0245	b0245	1.9				*ybhD*	b0768	2.3
			yafX	b0248	1.5				*b0770*	b0770	1.8
			insB_2	b0264	5.0				*moaA*	b0781	1.9
			insB_3	b0274	3.8				*b0807*	b0807	2.5
b0299	b0299	5.9							*b0819*	b0819	1.5
			ykgB	b0301	1.7				*b0833*	b0833	3.6
			b0302	b0302	2.4				*b0847*	b0847	1.4
yahA	b0315	12.7	*yahA*	b0315	12.4				*b0859*	b0859	2.0
			b0326	b0326	1.8				*pncB*	b0931	3.2
			adhC	b0356	2.6				*fabA*	b0954	2.7
b0357	b0357	41.0	*b0357*	b0357	30.0				*b0955*	b0955	3.2
			b0362	b0362	1.6				*b0986*	b0986	2.1
			b0365	b0365	2.9				*b0989*	b0989	3.0
			b0453	b0453	2.0				*b1025*	b1025	2.3
			ybaA	b0456	5.3				*b1031*	b1031	3.1
			b0458	b0458	2.3				*b1047*	b1047	2.1
			b0459	b0459	1.8				*b1060*	b1060	2.6
hha	b0460	16.0	*hha*	b0460	15.7				*dinI*	b1061	4.0
ybaJ	b0461	7.6	*ybaJ*	b0461	7.5				*yceB*	b1063	1.8
			acrA	b0463	1.8				*yceD*	b1088	2.0
			dnaX	b0470	4.9				*fabF*	b1095	3.1
			ybaB	b0471	1.9				*pabC*	b1096	3.2
			b0482	b0482	1.6				*ycfH*	b1100	1.5
rhsD	b0497	40.0							*ycfJ*	b1110	2.5
			ybbC	b0498	4.1				*b1111*	b1111	1.8
b0499	b0499	8.9	*b0499*	b0499	8.8				*b1122*	b1122	2.1
ybbD	b0500	13.9	*ybbD*	b0500	12.7				*ycfC*	b1132	1.7
b0501	b0501	6.6	*b0501*	b0501	6.2				*lit*	b1139	3.0
			b0517	b0517	1.3	*mcrA*	b1159	21.9			
			ybcI	b0527	1.5	*b1160*	b1160	39.6	*b1160*	b1160	38.9
			b0543	b0543	2.6				*b1162*	b1162	4.6
			b0544	b0544	1.7				*b1165*	b1165	2.8
			b0545	b0545	2.3				*b1169*	b1169	3.1
b0546	b0546	17.2	*b0546*	b0546	16.9				*b1170*	b1170	4.3

Table B.1 Continued

Standard Analysis			SAM Analysis			Standard Analysis			SAM Analysis		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
b1171	b1171	15.2	*b1171*	b1171	14.8				*b1821*	b1821	3.9
b1172	b1172	22.5	*b1172*	b1172	22.1				*yebH*	b1822	3.4
b1202	b1202	5.2	*b1202*	b1202	5.1				*b1826*	b1826	6.2
ychF	b1203	8.4	*ychF*	b1203	3.1				*htpX*	b1829	4.7
			hemM	b1209	4.3				*b1844*	b1844	2.0
			ychN	b1219	14.8	*yebG*	b1848	13.8	*yebG*	b1848	13.5
b1228	b1228	44.2	*b1228*	b1228	22.1				*ruvA*	b1861	3.0
cysB	b1275	8.2	*cysB*	b1275	5.1				*b1933*	b1933	4.1
			yciH	b1282	3.1	*b1935*	b1935	5.5	*b1935*	b1935	5.5
osmB	b1283	7.4	*osmB*	b1283	4.3				*fliR*	b1950	1.8
			b1374	b1374	14.8				*rcsA*	b1951	3.2
			ydbA_2	b1405	22.1				*yedA*	b1959	2.8
rhsE	b1456	16.1	*rhsE*	b1456	5.1				*dcm*	b1961	3.4
ydcD	b1457	14.0	*ydcD*	b1457	3.1	*b1963*	b1963	13.2	*b1963*	b1963	12.7
			b1462	b1462	4.3				*b1968*	b1968	3.2
			b1504	b1504	14.8				*b1970*	b1970	3.5
			b1524	b1524	22.1				*b1972*	b1972	2.2
b1527	b1527	6.0	*b1527*	b1527	5.1				*b1980*	b1980	2.1
			b1546	b1546	3.1	*b1983*	b1983	9.9	*b1983*	b1983	9.0
			b1560	b1560	4.3	*cbl*	b1987	12.2	*cbl*	b1987	11.9
			relF	b1562	14.8				*yeeC*	b2010	1.6
			b1566	b1566	22.1				*yeeF*	b2014	2.6
			b1567	b1567	5.1				*b2027*	b2027	3.9
			ydfC	b1573	3.1	*yefI*	b2032	26.0	*yefI*	b2032	23.6
			dicB	b1575	4.3	*yefG*	b2034	23.1	*yefG*	b2034	21.3
			b1586	b1586	14.8	*rfc*	b2035	5.6	*rfc*	b2035	5.3
			b1596	b1596	22.1	*rfbX*	b2037	16.9			
			b1604	b1604	5.1				*galF*	b2042	2.1
			gusC	b1615	3.1				*b2081*	b2081	4.9
ydhA	b1639	10.0	*ydhA*	b1639	4.3				*tra5_4*	b2089	2.1
			b1642	b1642	14.8				*b2100*	b2100	1.8
			b1645	b1645	22.1				*yehA*	b2108	3.4
			b1648	b1648	5.1	*yehC*	b2110	7.2	*yehC*	b2110	7.0
			lhr	b1653	3.1				*metG*	b2114	3.6
			cfa	b1661	4.3				*yeiR*	b2173	3.9
			b1667	b1667	14.8				*b2174*	b2174	6.1
			b1668	b1668	22.1				*yejD*	b2183	2.6
			b1678	b1678	5.1	*rplY*	b2185	9.8	*rplY*	b2185	9.3
			b1679	b1679	3.1				*yejL*	b2187	1.9
			ydiB	b1692	4.3				*eco*	b2209	2.7
			b1720	b1720	14.8				*b2224*	b2224	1.7
b1721	b1721	8.4							*yfaE*	b2236	5.0
			b1722	b1722	2.0				*b2255*	b2255	1.5
			b1728	b1728	3.0				*b2271*	b2271	1.5
b1730	b1730	6.7	*b1730*	b1730	6.5				*b2295*	b2295	2.5
			b1743	b1743	4.8				*truA*	b2318	4.2
			b1762	b1762	3.3				*b2339*	b2339	3.2
b1770	b1770	8.1	*b1770*	b1770	8.0				*b2340*	b2340	3.2
			rnd	b1804	1.7				*b2345*	b2345	3.9
			b1809	b1809	3.3	*vacJ*	b2346	15.2			
			b1811	b1811	2.3				*emrY*	b2367	1.2
			pabB	b1812	1.8				*evgA*	b2369	5.2

Table B.1 Continued

Standard Analysis			SAM Analysis			Standard Analysis			SAM Analysis		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
evgS	b2370	6.9							*cca*	b3056	2.1
			gltX	b2400	2.3				*rpsU*	b3065	5.1
			b2466	b2466	1.7				*ygjN*	b3083	5.3
			b2496	b2496	2.2				*ygjO*	b3084	2.5
			guaB	b2508	2.5				*tdcB*	b3117	1.8
xseA	b2509	6.8	*xseA*	b2509	6.1	*tdcR*	b3119	7.1	*tdcR*	b3119	6.7
			yfhD	b2558	2.5	*yhaB*	b3120	48.8			
			yfhC	b2559	1.3				*yhbU*	b3158	2.0
			pdxJ	b2564	1.5				*rpsO*	b3165	5.8
			recO	b2565	1.6				*nusA*	b3169	3.0
			b2584	b2584	1.6				*yhbC*	b3170	6.1
			yfiM	b2586	3.5				*greA*	b3181	2.8
			clpB	b2592	2.6				*yhbZ*	b3183	3.7
			sfhB	b2594	4.1				*rpmA*	b3185	2.0
			tyrA	b2600	1.4				*rplU*	b3186	2.9
			aroF	b2601	1.8				*gltF*	b3214	3.2
			rplS	b2606	4.8				*rpsI*	b3230	2.2
			trmD	b2607	7.9				*rplM*	b3231	2.4
			yfjA	b2608	5.5				*argR*	b3237	1.7
rpsP	b2609	14.0	*rpsP*	b2609	8.7				*mreC*	b3250	1.2
			yfjB	b2615	1.4				*def*	b3287	1.6
			recN	b2616	2.4				*rplQ*	b3294	2.2
			b2623	b2623	2.4				*rpsE*	b3303	1.8
			b2628	b2628	1.9				*rplR*	b3304	2.1
			b2639	b2639	1.7				*rpsN*	b3307	2.8
b2642	b2642	13.9							*rpmC*	b3312	1.5
			b2665	b2665	1.7				*rplV*	b3315	2.7
			b2667	b2667	2.0				*rplB*	b3317	2.0
			b2668	b2668	2.1				*rplD*	b3319	2.9
stpA	b2669	9.1	*stpA*	b2669	8.8				*rplC*	b3320	2.5
			proW	b2678	3.6				*rpsG*	b3341	3.5
			oraA	b2698	5.2				*yheL*	b3343	4.0
recA	b2699	6.6	*recA*	b2699	6.5				*aroB*	b3389	1.2
			b2760	b2760	1.9				*glpR*	b3423	2.7
			ygcA	b2785	2.2				*glpG*	b3424	3.6
			sdaB	b2797	1.3	*glpD*	b3426	9.6	*glpD*	b3426	9.3
			ptr	b2821	2.1	*yrhB*	b3446	6.8			
			lysA	b2838	1.3				*rpoH*	b3461	2.9
			b2850	b2850	5.0				*b3472*	b3472	2.5
			b2851	b2851	4.7				*yhhS*	b3473	2.4
b2854	b2854	19.7	*b2854*	b2854	19.3				*yhhT*	b3474	2.3
			b2856	b2856	3.4	*rhsB*	b3482	14.3	*rhsB*	b3482	12.7
			b2857	b2857	3.6	*yhhH*	b3483	19.8	*yhhH*	b3483	17.0
			b2862	b2862	3.6				*yhhJ*	b3485	2.3
b2863	b2863	37.5	*b2863*	b2863	34.4	*yhiJ*	b3488	69.3	*yhiJ*	b3488	46.3
			pepP	b2908	3.1	*yhiK*	b3489	11.6	*yhiK*	b3489	10.3
			ygfE	b2910	2.5	*yhiL*	b3490	27.4	*yhiL*	b3490	24.4
			ygfD	b2918	2.7				*yhiM*	b3491	4.2
			yggN	b2958	2.5				*gor*	b3500	2.1
			b2974	b2974	4.4				*arsB*	b3502	3.0
b3046	b3046	5.5	*b3046*	b3046	5.3				*slp*	b3506	2.2
			b3048	b3048	2.9				*yhiF*	b3507	2.6

Table B.1 Continued

Standard Analysis			SAM Analysis			Standard Analysis			SAM Analysis		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
b3515	b3515	6.6							*b0518*	b0518	-2.7
yhiX	b3516	19.0							*ybeK*	b0651	-3.0
			gadA	b3517	2.3				*b0701*	b0701	-3.7
			b3524	b3524	2.9				*b0881*	b0881	-3.0
			avtA	b3572	1.3				*trxB*	b0888	-3.6
yibJ	b3595	21.5	*yibJ*	b3595	19.9	*pflB*	b0903	-18.5	*pflB*	b0903	-14.5
b3618	b3618	8.6	*b3618*	b3618	8.4				*rimJ*	b1066	-2.1
rfaL	b3622	11.6	*rfaL*	b3622	8.6				*flgA*	b1072	-3.3
			yidR	b3689	2.5	*flgE*	b1076	-9.8			
			yidS	b3690	1.2				*flgF*	b1077	-3.5
			rnpA	b3704	2.0				*flgG*	b1078	-4.0
			asnC	b3743	2.7				*flgH*	b1079	-2.8
			yifB	b3765	2.4				*flgI*	b1080	-1.5
			rep	b3778	3.7	*flgJ*	b1081	-11.3	*flgJ*	b1081	-11.2
			yigG	b3818	4.0				*minE*	b1174	-2.3
yiiG	b3896	10.0	*yiiG*	b3896	8.7				*ychE*	b1242	-2.5
			yiiL	b3901	1.8				*pspB*	b1305	-1.7
			yiiM	b3910	2.1				*b1312*	b1312	-2.0
b3913	b3913	17.5	*b3913*	b3913	15.6				*fnr*	b1334	-1.3
b3914	b3914	26.1							*b1420*	b1420	-2.9
			yijP	b3955	2.8				*b1483*	b1483	-2.4
			rplA	b3984	3.3	*b1777*	b1777	-6.4			
			yjaA	b4011	3.4				*flhA*	b1879	-2.1
			xylE	b4031	1.2	*gatB*	b2093	-19.0			
yjbM	b4048	8.4	*yjbM*	b4048	8.2				*gatZ*	b2095	-3.3
yjcF	b4066	10.2	*yjcF*	b4066	9.7				*nuoM*	b2277	-2.1
			proP	b4111	3.3				*nuoK*	b2279	-2.6
			basS	b4112	2.4				*nuoI*	b2281	-2.8
			frdA	b4154	2.0				*nuoH*	b2282	-1.8
			yjeQ	b4161	5.2				*pta*	b2297	-4.1
miaA	b4171	7.7	*miaA*	b4171	7.3				*b2483*	b2483	-2.7
			hflX	b4173	3.2				*yfiD*	b2579	-3.3
			hflC	b4175	2.3				*hycF*	b2720	-2.0
			purA	b4177	2.7				*hycB*	b2724	-2.4
			yjfI	b4181	4.3				*b2770*	b2770	-2.1
			yjfZ	b4204	4.2				*eno*	b2779	-4.0
			yjgL	b4253	2.2				*pgk*	b2926	-2.7
			b4257	b4257	5.6				*hybG*	b2990	-2.2
intB	b4271	5.5	*intB*	b4271	5.4				*hybB*	b2995	-4.4
			yi41	b4278	3.5				*sufI*	b3017	-1.5
			yjhB	b4279	3.7				*glgS*	b3049	-4.3
			yjhC	b4280	3.6				*yhfQ*	b3374	-2.0
			yjhO	b4305	2.8				*pckA*	b3403	-2.4
			fimB	b4312	2.6				*nikE*	b3480	-3.8
			rimI	b4373	1.9	*uspA*	b3495	-37.2	*uspA*	b3495	-9.9
			yjjG	b4374	2.0	*hdeB*	b3509	-14.8			
			yjjI	b4380	1.7						
			lasT	b4403	2.1						
									yhjY	b3548	-1.5
									radC	b3638	-3.8
									yicE	b3654	-2.0
						tnaL	b3707	-33.9	*tnaL*	b3707	-31.9

Table B.1 Continued

Standard Analysis			SAM Analysis		
Gene name	Blattner No.	IR	Gene name	Blattner No.	IR
tnaA	b3708	-29.3	*tnaA*	b3708	-29.8
rbsD	b3748	-12.1			
			murB	b3972	-1.5
			malK	b4035	-7.1
lamB	b4036	-10.1			
			soxR	b4063	-1.2
aspA	b4139	-21.9			
			ytfA	b4205	-2.2
			nadR	b4390	-2.1

Appendix C

Table C.1 Results of 2-Dimensional PAGE. A blank space indicates that no match was found for the that 'spot'. Spot numbers in bold denote spots that were sent for identification by mass spectrometry.

Spot ID	Gel set 1			Gel set 2			Gel set 3		
	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC
202	15526			4408			3611		
301	2329			3709			2544		
401	2062			2822			2066		
402	1322			2886			1259		
403	612			1306			9061		
501		**2482**			**51033**			**12694**	
602	4951	3898	-1.27	6306	2403	-2.62	1651	2904	1.76
603	2198	2840	1.29	3487	1483	-2.35	3816	876	-4.35
604	1335			1783			2360		
605	1200	2975	2.48	1823	12669	6.95	320	1357	4.25
803	9608			4199			2164		
804	27718	42920	1.55	39157	52527	1.34	22816	35935	1.57
1003	4803			6802			8490		
1101		3985			1190			21040	
1102		8465			743			4112	
1203	74916	36255	-2.07	20017	24987	1.25	18461	19351	1.05
1302	5549			5045			2306		
1402	4178			2670			3658		
1403	4988	1275	-3.91	1304	2021	1.55		1781	
1502	9445	5774	-1.64	7092	2868	-2.47	5467	5570	1.02
1602	1015	5324	5.25	1795	3759	2.09	3125	3989	1.28
1701	874	993	1.14	2925	943	-3.10		874	
1702	1765			2531					
1801		18861			4875			7657	
1802		2188			10681			1248	
1902	20268	56181	2.77	115840	38063	-3.04	10292	120663	11.72
2002	5384			4564			18489		
2104	8822	8654	-1.02	12604	5816	-2.17	7523	5954	-1.26
2105	3900	2663	-1.46	3404	29059	8.54	2368	4850	2.05
2106	5276	4346	-1.21	2399	3691	1.54	2967	2805	-1.06
2107		4595			2685			1808	
2202	1063			1078					
2203	2474			4540					
2401		795			1354			1715	
2402		7103			2610			2302	
2503	7732			7676	7710	1.00	4422	4559	1.03
2702	5976			9450			28818		
2802	6961	3304	-2.11	5227	8072	1.54	9021	32024	3.55
2803		1009			1843			1527	
2804		1009			2735			986	
2901	1565			406			1906		
2902	**1232**	**24228**	**19.67**	**2222**			**1896**	**98467**	**51.92**
3102	2547	23648	9.28	5144	1060	-4.85	3085	1639	-1.88
3103	1586	2143	1.35	2611	1988	-1.31	2202	6670	3.03
3104	4186			2216			4334		
3105		3341			3292			21676	
3203	26165	24720	-1.06	28911	32095	1.11	27979	22442	-1.25
3204	**13240**	**6205**	**-2.13**	**15268**	**4705**	**-3.25**	**14394**	**4442**	**-3.24**
3303	**12750**	**3097**	**-4.12**	**11919**	**2891**	**-4.12**	**38463**	**2916**	**-13.19**

Table C.1 Continued

Spot ID	Gel set 1			Gel set 2			Gel set 3		
	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC
3306	4634			3309			2619		
3307		2365			1187			2333	
3402	1300	1370	1.05	761	1885	2.48	1601	929	-1.72
3501	**19906**	**11432**	**-1.74**	**17340**	**8896**	**-1.95**	**18005**	**7070**	**-2.55**
3503	2592	2204	-1.18	819	10906	13.32	3285	4697	1.43
3602	54149	30036	-1.80	36281	25085	-1.45	79178	22250	-3.56
3705	5908			5345			13158		
3706	4661			39864			27777		
3707	4761	8533	1.79	2286	16244	7.11	2225	1393	-1.60
3802	5834	1598	-3.65	5944	711	-8.36	4989	3069	-1.63
3901	982	772	-1.27	985	937	-1.05	579	370	-1.56
3902		**4856**			**5844**			**4480**	
4103	7373	18221	2.47	12891	3446	-3.74	6554	9591	1.46
4104	6060	15218	2.51	5839	65412	11.20	17499	11755	-1.49
4105		2614			914			1727	
4106		1470			1471			973	
4202	2844			3416			1916		
4203	1282			1676			825		
4304	5492	2736	-2.01	5264	2748	-1.92	3376	3226	-1.05
4305	2318	2800	1.21	2145	1370	-1.57	1919	2444	1.27
4306	6122	4288	-1.43	4572	1861	-2.46	3057	2373	-1.29
4307		2492			3119			3856	
4308		1362			2187			1913	
4401	21560	45403	2.11	14221	43774	3.08	26425	45190	1.71
4403	3454			3462			3016		
4404	395			2717			1708		
4405	872						2830		
4406		1062			12085			1836	
4502	5099	2819	-1.81	4281	2210	-1.94	1661	1541	-1.08
4503	834	2615	3.14		2624		1631	4855	2.98
4504		9973			4213			6360	
4505		2856			1483			3300	
4601	**39445**	**8245**	**-4.78**	**30360**	**5836**	**-5.20**	**7876**	**4179**	**-1.88**
4701	427	5428	12.71	674	2686	3.99	1710	3382	1.98
4702	1291	3850	2.98	1011	1396	1.38	2268	1402	-1.62
4703					349			650	
4801	1250						897		
4802	803	888	1.11	1135	706	-1.61	1505	1044	-1.44
4902	931						1022		
5003	4550	6888	1.51	4084	3289	-1.24	3861	6906	1.79
5203	5448	10589	1.94	2705	5268	1.95	3390	7730	2.28
5204	9357	1794	-5.21	3997	1596	-2.51	3042	784	-3.88
5205	2655	3683	1.39	3233	5325	1.65	3590	1559	-2.30
5206		5549			3186			5604	
5301	9343	1558	-6.00	846	831	-1.02	3653	723	-5.05
5501	1303	2973	2.28	2384	1626	-1.47	758	3399	4.48
5502	2518	4616	1.83	1441	1215	-1.19	1298	2160	1.66
5605	3079			2718			1948		
5606	3032	1599	-1.90	2684	983	-2.73	1050	1681	1.60
5607		1463			1288			2774	
5703	65278	16565	-3.94	75838	50457	-1.50	72869	48692	-1.50

Table C.1 Continued

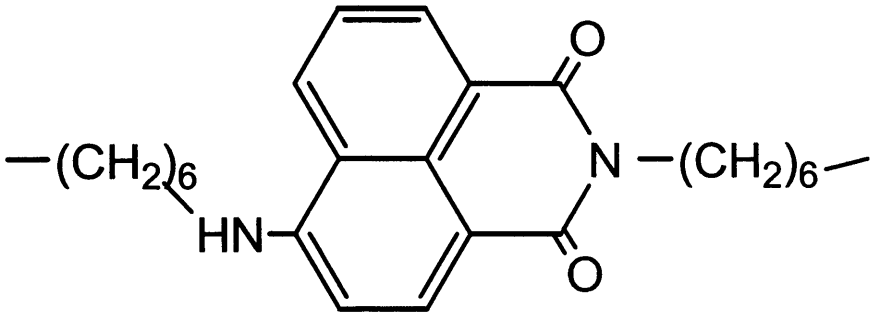
Spot ID	Gel set 1			Gel set 2			Gel set 3		
	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	FC
5704	3864	1345	-2.87	3782	246	-15.39	1787	854	-2.09
5705	8572	3976	-2.16	4787	329	-14.55	921	1526	1.66
5706		16599			54859			44326	
5707		42254			961			1528	
5803	2485			2441			175		
5804	963			1066			540		
5805	3521			3466			1946		
5902	2353			7531			17405		
5903	1177	1111	-1.06		520		406	916	2.25
6102	6286	5919	-1.06	7908	4081	-1.94	7064	4087	-1.73
6103		**9400**			**2071**			**5523**	
6202	15601	8754	-1.78	14838	4019	-3.69	11290	5235	-2.16
6203	1805	4236	2.35	1341	4731	3.53	3352	4597	1.37
6204				6572			3499		
6302	9866	7175	-1.38	10763	6660	-1.62	7208	5309	-1.36
6303				811			2093		
6304		2128			877			912	
6402	2745	7405	2.70	3257	3924	1.20	2599	3716	1.43
6403	343			794			2778		
6502	**3618**	**30820**	**8.52**	**5125**	**11700**	**2.28**	**4768**	**11804**	**2.48**
6503	426			1353			254		
6504		2438						6819	
6601	**148690**	**55831**	**-2.66**	**97792**	**5500**	**-17.78**	**62091**	**5480**	**-11.33**
6604	**11380**	**1443**	**-7.89**	**13491**			**81235**	**376**	**-216.21**
6605	3465	1300	-2.67	1469	1761	1.20	505	1833	3.63
6606		2029			2675			1865	
6702	4030	2102	-1.92	3110	1200	-2.59	7053	1663	-4.24
6802	6996	1866	-3.75	7641	1001	-7.63	3262	1687	-1.93
6901	1844						856		
6902	783								
7101		**5867**			**6863**			**5388**	
7102		5546			3115			3903	
7103		3567			3258			3405	
7302	18061	2031	-8.89	6511	3178	-2.05	4182	2879	-1.45
7303	1885	7354	3.90	838	4644	5.54	2481	3605	1.45
7401	15139	53413	3.53	10952	50592	4.62	16132	37342	2.31
7402	6262	8446	1.35	4060	5101	1.26	7112	3180	-2.24
7403		3126		701	1719	2.45	2410	1554	-1.55
7404		2511		1704	243	-7.02	2381	760	-3.13
7405	2036	4128	2.03	1925	3004	1.56	1915	3413	1.78
7406	377	4018	10.65		4848		1791	3875	2.16
7407		2180			2393			1028	
7408		5668			5266			5691	
7409		4165			1455			1723	
7502	2239	4528	2.02	2274	4547	2.00	631	4415	7.00
7503	572	16949	29.65	1612	1023	-1.58	2258	1371	-1.65
7504		1613			665			703	
7505		3575			2868			2932	
7506		1229			2734			923	
7507					967			1262	
7601	**24304**	**12016**	**-2.02**	**27107**	**5276**	**-5.14**	**29088**	**5964**	**-4.88**

Table C.1 Continued

Spot ID	Gel set 1			Gel set 2			Gel set 3		
	Un-exposed	PHMB exposed	FC	Un-exposed	PHMB exposed	IR	Un-exposed	PHMB exposed	FC
7602	1056	756	-1.40	3797	1051	-3.61	12257	1613	-7.60
7604	2026	1400	-1.45	2218	780	-2.84	709	1299	1.83
7605	877	5935	6.77	3086	2234	-1.38	4247	2161	-1.97
7606					761			853	
7701		13514		1801	8248	4.58	6940	4631	-1.50
7702		3598			2898			1459	
7703		2984			11672			980	
7704		540			1124			1004	
7705		6274			2722			2135	
7801	1567	1969	1.26	12935	1165	-11.11	6020	815	-7.39
7802	332			633			595		
7803	449			980			6571		
7804		2866			3526			22124	
9103					5753			4958	
9302	10427	14116	1.35	16182	56115	3.47	11931	12004	1.01
9304	2598	3702	1.42	4079	2506	-1.63	4265	2815	-1.52
9305					3695			3190	
9401	1688			6050	3267	-1.85	2951	2236	-1.32
9402	1491	1596	1.07	4080	4498	1.10	2103	2766	1.32
9403				1645			2957		
9404					9378			12372	
9405					3903			4776	
9501				2405			729		
9502				2307			5995		
9601				1102			1796		
9602		6219			3819			2641	
9603		3029			8076			7393	
9703	3411	1383	-2.47	5589	2845	-1.96	6087	3374	-1.80
9704	2950			8292			18774		
9801	1269			1728			3704		

Appendix D

Figure D.1 Structure of The fluor, 6-amino-benzo[de]isoquinoline-1,3-dione, attached to flanking hexamethylene chains via the ring N(2) and amino substituent at position 6, replaces 1 in 100 guanidines in the fluorescent PHMB.



Appendix E

Table E.1 Input parameters and computed quantities for the SAM performed on data from microarray experiment (comparison 1 repeated). Results from this analysis can be seen in Table B.2.

Input Parameters	
Imputation Engine	10-Nearest Neighbour Imputer
Data Type	Two Class, unpaired data
Data in log scale?	FALSE
Number of Permutations	100
Blocked Permutation?	FALSE
RNG Seed	1234567
Delta	3.72791
(Upper Cutoff, Lower Cutoff)	(6.24753, -7.92656)
Computed Quantities	
Computed Exchangeability Factor S0	1.14515E-06
S0 percentile	0.02
False Significant Number (Median, 90 percentile)	(0.69717, 6.27452)
False Discovery Rate (Median, 90 percentile)	(0.91733, 8.25595)
Pi0Hat	0.69717

Table E.2 Results of SAM performed on data from microarray experiment (comparison 1 repeated).

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
insA_3	b0275	25.12	4.16E-04	1.66E-05	8.55704	0.92
b1171	b1171	19.99	1.14E-03	5.68E-05	12.85879	0.92
mioC	b3742	19.97	3.43E-04	1.72E-05	4.05252	0.92
yaeD	b0200	19.37	5.76E-04	2.97E-05	8.75764	0.92
mbhA	b0230	18.06	2.25E-04	1.24E-05	2.55689	0.92
b0833	b0833	16.78	4.95E-04	2.95E-05	16.77521	0.92
b1459	b1459	16.11	2.11E-04	1.31E-05	18.18824	0.92
dnaG	b3066	15.02	2.14E-04	1.42E-05	4.85145	0.92
b1825	b1825	14.95	8.84E-05	5.91E-06	52.48427	0.92
yebG	b1848	13.82	9.32E-04	6.75E-05	7.12375	0.92
b1568	b1568	11.54	4.23E-04	3.66E-05	3.68768	0.92
sprT	b2944	11.16	9.93E-05	8.90E-06	7.78541	0.92
yraH	b3142	11.13	2.86E-04	2.57E-05	18.75655	0.92
yhiR	b3499	11.07	2.80E-04	2.53E-05	3.23741	0.92
htrL	b3618	11.04	3.23E-03	2.93E-04	54.94624	0.92
yjfH	b4180	10.98	1.08E-04	9.86E-06	6.56240	0.92
infA	b0884	10.97	2.44E-04	2.22E-05	27.15143	0.92
yjeB	b4178	10.89	1.69E-04	1.55E-05	10.24695	0.92
yaeC	b0197	10.86	3.53E-04	3.25E-05	3.65289	0.92
b1964	b1964	10.76	2.89E-04	2.68E-05	12.93220	0.92
truB	b3166	10.65	1.88E-04	1.77E-05	2.74648	0.92
yceP	b1060	10.43	1.30E-04	1.25E-05	5.22053	0.92
yibA	b3594	10.40	6.68E-04	6.42E-05	11.11964	0.92
yebF	b1847	10.24	3.82E-04	3.73E-05	20.51745	0.92
dsbD	b4136	10.16	1.09E-04	1.07E-05	10.37094	0.92
ybaD	b0413	10.08	5.74E-04	5.69E-05	4.16351	0.92
b1172	b1172	10.07	3.56E-03	3.54E-04	6.39691	0.92
rfbX	b2032	9.77	1.88E-04	1.92E-05	38.74423	0.92
yabC	b0082	9.48	1.77E-04	1.86E-05	5.09878	0.92
marR	b1530	9.35	3.47E-04	3.71E-05	55.95674	0.92
b1841	b1841	8.35	4.80E-05	5.75E-06	15.35633	0.92
rnt	b1652	8.34	1.32E-04	1.59E-05	11.80895	0.92
prfC	b4375	8.29	6.24E-04	7.53E-05	10.67408	0.92
b1578	b1578	8.24	1.92E-04	2.33E-05	5.65474	0.92
yafQ	b0225	8.22	8.69E-05	1.06E-05	5.45651	0.92
ymcD	b0987	8.19	8.12E-05	9.91E-06	7.08585	0.92
topA	b1274	8.18	2.45E-03	3.00E-04	6.88713	0.92
marA	b1531	8.16	5.12E-04	6.27E-05	14.97128	0.92
rpsQ	b3311	7.90	3.00E-04	3.80E-05	3.10794	0.92
vacB	b4179	7.63	1.23E-04	1.61E-05	2.95663	0.92
yefM	b2017	7.60	5.44E-05	7.16E-06	8.33190	0.92

Table E.2 Continued

Gene Name	Gene ID	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	q-value (%)
msbA	b0914	7.37	5.59E-04	7.59E-05	5.87690	0.92
cspA	b3556	7.23	1.93E-04	2.68E-05	13.91328	0.92
b2529	b2529	7.13	7.63E-05	1.07E-05	19.35359	0.92
yrbC	b3192	7.03	6.59E-05	9.39E-06	3.07498	0.92
yahM	b0327	7.01	1.20E-04	1.71E-05	10.24602	0.92
emrE	b0543	6.99	5.36E-04	7.66E-05	17.66087	0.92
ybiI	b0803	6.97	2.39E-05	3.43E-06	5.53271	0.92
abc	b0199	6.81	2.26E-04	3.31E-05	19.16661	0.92
fur	b0863	6.68	1.34E-04	2.00E-05	2.65987	0.92
b1011	b1011	6.68	8.65E-05	1.29E-05	5.89334	0.92
hyfR	b2491	6.55	8.28E-05	1.26E-05	5.22413	0.92
yceA	b1055	6.53	5.04E-05	7.71E-06	2.23561	0.92
proW	b2678	6.51	1.72E-03	2.64E-04	11.17372	0.92
b2506	b2506	6.51	6.39E-04	9.82E-05	42.42916	0.92
yhhI	b3484	6.51	1.65E-05	2.53E-06	9.29848	0.92
gpmB	b4395	6.44	5.89E-05	9.15E-06	1.99501	0.92
osmB	b1283	6.39	1.70E-04	2.66E-05	3.82339	0.92
rcsF	b0196	6.35	3.50E-04	5.52E-05	3.07321	0.92
fepD	b0590	6.34	5.02E-05	7.92E-06	2.38641	0.92
ycfO	b1107	6.32	2.76E-04	4.37E-05	3.79303	0.92
yi22_3	b1996	6.31	7.93E-04	1.26E-04	17.56307	0.92
wbbI	b2034	6.25	1.78E-04	2.85E-05	23.23394	0.92
rmf	b0953	-23.91	-3.95E-04	1.65E-05	0.19778	0.92
ydhO	b1655	-22.33	-5.50E-04	2.46E-05	0.20317	0.92
aceE	b0114	-21.42	-1.85E-04	8.64E-06	0.24811	0.92
cspD	b0880	-16.36	-6.86E-04	4.19E-05	0.05936	0.92
flgB	b1073	-12.48	-5.48E-04	4.39E-05	0.07211	0.92
rbsD	b3748	-12.41	-2.63E-04	2.12E-05	0.11217	0.92
b1085	b1085	-10.08	-1.23E-04	1.22E-05	0.15605	0.92
pepN	b0932	-8.99	-4.74E-04	5.27E-05	0.18414	0.92
cydB	b0734	-8.46	-2.38E-04	2.81E-05	0.17228	0.92
nuoK	b2279	-8.17	-8.61E-05	1.05E-05	0.32111	0.92
kup	b3747	-7.99	-7.62E-05	9.53E-06	0.42731	0.92
gcvT	b2905	-7.93	-8.05E-05	1.02E-05	0.23065	0.92

Table E.3 Results of standard analysis performed on data from microarray experiment (i.e. comparison 1 repeated using microarrays).

Gene	Blattner number	Gene product description	Fold Change
yi21_1	b0360	IS2 hypothetical protein	+ 17.6
ybcK	b0544	orf, hypothetical protein	+ 328.3
ybcU	b0557	bacteriophage lambda Bor homologue	+ 39.4
ybcY	b0562	orf, hypothetical protein	+ 91.1
ybeK	b0651	putative tRNA synthetase	- 14.2
ybfE	b0685	orf, hypothetical protein	+ 30.7
sdhA	b0723	succinate dehydrogenase	- 10.4
cspD	b0880	cold shock protein	- 17.5
ycaL	b0909	putative heat shock protein	+ 225.1
yceE	b1053	putative transport protein	+ 20.4
flgB	b1073	flagellar biosynthesis	- 13.0
flgG	b1078	flagellar biosynthesis	- 18.8
ycfJ	b1110	orf, hypothetical protein	+ 55.9
ycfR	b1112	orf, hypothetical protein	+ 36.5
ycgW	b1160	orf, hypothetical protein	+ 26.5
ycgX	b1161	orf, hypothetical protein	+ 110.7
ynaF	b1376	putative filament protein	- 25.7
ydfA	b1571	orf, hypothetical protein	+ 71.2
ydfC	b1573	orf, hypothetical protein	+ 20.4
yeaZ	b1807	orf, hypothetical protein	+ 78.8
b1825	b1825	orf, hypothetical protein	+ 72.9
fliD	b1924	flagellar biosynthesis	- 15.3
fliS	b1925	flagellar biosynthesis	- 12.9
yedJ	b1962	orf, hypothetical protein	+ 257.8
b1963	b1963	orf, hypothetical protein	+ 350.6
yi22_3	b1996	IS2 hypothetical protein	+ 18.7
gatA	b2094	galactitol-specific enzyme IIA	- 22.4
gatZ	b2095	putative tagatose 6-phosphate kinase 1	- 38.9
gatY	b2096	tagatose-bisphosphate aldolase 1	- 26.3
mglB	b2150	galactose-binding transport protein	- 16.2
nuoG	b2283	NADH dehydrogenase I chain G	- 13.8
nuoC	b2286	NADH dehydrogenase I chain C, D	- 16.9
yfhL	b2562	orf, hypothetical protein	+ 70.9
yfiM	b2586	orf, hypothetical protein	+ 41.4
yfjW	b2642	orf, hypothetical protein	+ 180.4
yqeH	b2846	orf, hypothetical protein	+ 277.2
yqeJ	b2848	orf, hypothetical protein	+ 38.0
yi22_4	b2860	IS2 hypothetical protein	+ 39.6
ygiR	b3016	orf, hypothetical protein	+ 19.5
ygjL	b3081	putative NADPH dehydrogenase	+ 9.4
ygjM	b3082	orf, hypothetical protein	+ 325.2
ygjN	b3083	orf, hypothetical protein	+ 276.2

Table E.3 Continued

Gene	Blattner number	Gene product description	Fold Change
yraH	b3142	putative fimbrial-like protein	+ 30.3
yhcA	b3215	putative chaperone	+ 173.5
yheE	b3324	putative general secretion pathway	+ 82.8
yhhQ	b3471	orf, hypothetical protein	+ 161.9
yhiL	b3490	orf, hypothetical protein	+ 489.5
yhiM	b3491	orf, hypothetical protein	+ 713.6
yhiW	b3515	putative regulatory protein	+ 40.0
yibJ	b3595	orf, hypothetical protein	+ 1036.3
htrL	b3618	involved in LPS biosynthesis	+ 92.2
tnaL	b3707	tryptophanase leader peptide	- 39.7
tnaA	b3708	tryptophanase	- 54.1
atpC	b3731	membrane-bound ATP synthase	- 13.2
atpG	b3733	membrane-bound ATP synthase	- 13.5
atpH	b3735	membrane-bound ATP synthase	+ 5.9
yifA	b3762	orf, hypothetical protein	- 12.5
b3914	b3914	orf, hypothetical protein	+ 191.7
yiiX	b3937	orf, hypothetical protein	+ 261.6
aceB	b4014	malate synthase A	- 18.4
malF	b4033	part of maltose permease, periplasmic	- 31.0
malE	b4034	periplasmic maltose-binding protein	- 35.4
lamB	b4036	phage lambda receptor protein	- 28.2
malM	b4037	periplasmic protein of mal regulon	- 29.3
aspA	b4139	aspartate ammonia-lyase (aspartase)	- 13.1
treB	b4240	PTS enzyme II, trehalose specific	- 26.0
yi21_6	b4272	IS2 hypothetical protein	+ 111.5

Appendix F

Table F.1 Alignment by Blattner number order of the results from the standard analysis of the macroarray replicate 2 versus standard analysis of the microarray replicate 2.

Microarray Rep. 2			Macroarray Rep. 2			Microarray Rep. 2			Macroarray Rep. 2		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			talB	b0008	0.81	*ycfJ*	b1110	2.03			
			dnaK	b0014	0.88				*b1119*	b1119	0.75
			dnaJ	b0015	0.73	*lit*	b1139	1.81			
nhaA	b0019	1.68							*mcrA*	b1159	0.96
			insA_1	b0022	0.80	*ycgW*	b1160	1.83	*b1160*	b1160	1.45
			rpsT	b0023	0.77	*ycgX*	b1161	2.35	*b1161*	b1161	0.89
			murD	b0088	0.69				*b1162*	b1162	0.72
b0135	b0135	1.65	*yadC*	b0135	1.49	*ycgZ*	b1164	2.09			
			ecpD	b0140	0.70	*ymgA*	b1165	2.36			
			b0174	b0174	0.76	*b1169*	b1169	2.18			
abc	b0199	1.61							*b1171*	b1171	1.12
yafT	b0217	2.43	*yafT*	b0217	1.04				*b1172*	b1172	1.19
			tra8_1	b0256	0.72	*ycgY*	b1196	1.66			
			insB_2	b0264	0.72				*b1202*	b1202	0.76
			eaeH	b0297	0.69				*ychF*	b1203	0.80
			b0299	b0299	0.76				*b1228*	b1228	1.43
ykgH	b0310	1.76				*tpr*	b1229	2.34			
yahA	b0315	2.56	*yahA*	b0315	1.16	*yciL*	b1269	1.79			
yahM	b0327	1.93							*cysB*	b1275	0.90
			b0357	b0357	1.45				*osmB*	b1283	0.79
yi22_1	b0361	2.22				*ycjL*	b1298	1.67			
yaiW	b0378	1.66				*ycjF*	b1322	1.92			
			yajI	b0412	0.67	*ydaL*	b1340	1.95			
ylaD	b0459	1.87				*ydaW*	b1361	1.94			
			hha	b0460	1.07	*hslJ*	b1379	1.88			
			ybaJ	b0461	0.82	*ydbH*	b1381	1.85			
			rhsD	b0497	1.49	*ydbA2*	b1405	2.42			
			b0499	b0499	0.95				*rhsE*	b1456	0.99
			ybbD	b0500	0.95				*ydcD*	b1457	1.02
			b0501	b0501	0.94	*b1459*	b1459	2.14			
ybcK	b0544	2.12				*ydcC*	b1460	2.08			
ybcL	b0545	2.17				*narZ*	b1468	2.24			
			b0546	b0546	1.20				*b1522*	b1522	0.77
			b0558	b0558	0.93				*b1527*	b1527	0.77
ybcY	b0562	1.81				*marR*	b1530	2.42			
ybdO	b0603	1.73	*b0603*	b0603	0.81	*marB*	b1532	2.23			
			criR	b0620	0.69				*ydeH*	b1535	0.89
ybeU	b0648	1.83				*ydeJ*	b1537	1.73			
			rhsC	b0700	0.77	*b1543*	b1543	1.94			
b0703	b0703	2.05				*ydfK*	b1544	1.61			
ybfL	b0705	1.77	*b0705*	b0705	0.96	*ydfO*	b1549	2.56	*b1549*	b1549	1.27
ybfD	b0706	2.01	*ybfD*	b0706	0.73				*b1552*	b1552	0.68
			modB	b0764	0.82	*cspB*	b1557	2.79			
ybiS	b0819	1.82				*relF*	b1562	2.17			
b0833	b0833	2.33				*relE*	b1563	2.07			
			grxA	b0849	0.68	*relB*	b1564	2.05			
			potG	b0855	0.69	*ydfA*	b1571	1.86			
infA	b0884	2.19				*ydfB*	b1572	2.09			
ycaL	b0909	2.02				*rspA*	b1581	1.65			
cspG	b0990	2.20				*tus*	b1610	1.78			
			b1021	b1021	0.71	*ydhA*	b1639	1.99	*ydhA*	b1639	0.80
ycdT	b1025	2.02							*b1648*	b1648	0.74

Table F.1 Continued

Microarray Rep. 2			Macroarray Rep. 2			Microarray Rep. 2			Macroarray Rep. 2		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
rnt	b1652	2.02							*rpsP*	b2609	1.08
ydiC	b1684	1.61				*grpE*	b2614	1.94			
b1706	b1706	1.94				*recN*	b2616	1.88			
			b1720	b1720	1.05				*b2642*	b2642	0.88
			b1721	b1721	0.76	*b2657*	b2657	1.92			
b1726	b1726	1.77				*stpA*	b2669	1.65	*stpA*	b2669	0.98
			b1730	b1730	0.88	*proV*	b2677	2.38			
			b1743	b1743	0.75	*b2681*	b2681	1.82			
			b1770	b1770	0.89	*b2689*	b2689	2.00			
b1825	b1825	1.80				*recA*	b2699	1.98	*recA*	b2699	0.76
			htpX	b1829	0.68	*ygbE*	b2749	1.87			
b1836	b1836	1.68				*yqeH*	b2846	2.23			
ptrB	b1845	2.20							*recC*	b2822	0.95
yebF	b1847	2.17							*kduD*	b2842	0.68
			yebG	b1848	1.11				*b2854*	b2854	1.27
yedN	b1934	1.91				*b2862*	b2862	2.16			
			b1935	b1935	0.72	*b2863*	b2863	2.24	*b2863*	b2863	1.35
b1936	b1936	2.06				*exbD*	b3005	1.93			
			yedJ	b1962	0.76				*b3046*	b3046	0.69
b1963	b1963	2.38	*b1963*	b1963	0.98	*b3051*	b3051	1.65			
			b1983	b1983	0.82				*rpsU*	b3065	0.78
			cbl	b1987	1.13	*ygjM*	b3082	2.85			
yeeF	b2014	1.94				*yqjA*	b3095	1.62			
yefM	b2017	1.61							*tdcR*	b3119	0.76
			yefI	b2032	1.16				*yhaB*	b3120	1.11
wbbJ	b2033	2.00				*agaR*	b3131	1.61			
wbbI	b2034	1.97	*yefG*	b2034	1.12	*yraH*	b3142	2.04			
			rfc	b2035	0.69	*deaD*	b3162	2.79	*deaD*	b3162	0.77
rfbX	b2037	2.13	*rfbX*	b2037	1.03				*rpsO*	b3165	0.74
wcaH	b2051	1.90							*infB*	b3168	0.80
gmd	b2053	1.61				*yheE*	b3324	2.29			
			b2081	b2081	0.80	*yhfL*	b3369	1.82			
			yehC	b2110	0.89				*glpD*	b3426	0.87
rsuA	b2183	1.97				*yhhZ*	b3442	2.04	*yhhZ*	b3442	1.40
			rplY	b2185	1.07	*yrhA*	b3443	1.93			
			b2269	b2269	0.80				*yrhB*	b3446	0.78
yfbN	b2273	1.71				*yhhQ*	b3471	2.35			
b2295	b2295	1.87							*rhsB*	b3482	0.92
truA	b2318	2.30	*truA*	b2318	0.68				*yhhH*	b3483	1.01
			vacJ	b2346	0.89				*yhiJ*	b3488	1.37
			evgA	b2369	0.86				*yhiK*	b3489	0.85
			evgS	b2370	0.81	*yhiL*	b3490	2.98	*yhiL*	b3490	1.13
b2506	b2506	1.84				*yhiM*	b3491	2.80			
			xseA	b2509	0.83				*b3515*	b3515	0.72
b2529	b2529	1.78							*yhiX*	b3516	0.85
yfiM	b2586	1.63				*cspA*	b3556	1.63			
			kgtP	b2587	0.80	*yibJ*	b3595	2.87	*yibJ*	b3595	1.18
sfhB	b2594	1.64				*htrL*	b3618	1.81	*b3618*	b3618	0.91
aroF	b2601	1.90							*rfaL*	b3622	0.71
			rplS	b2606	0.86	*rfaZ*	b3624	1.85			
			trmD	b2607	1.03				*yicF*	b3647	0.70
			yfjA	b2608	0.79	*asnC*	b3743	1.70			

Table F.1 Continued

Microarray Rep. 2			Macroarray Rep. 2			Microarray Rep. 2			Macroarray Rep. 2		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
yifK	b3795	1.72							*flgJ*	b1081	-1.17
yigG	b3818	1.65							*flgL*	b1083	-0.75
			yiiG	b3896	0.72				*rne*	b1084	-0.68
			b3913	b3913	1.14				*potD*	b1123	-0.98
b3914	b3914	2.12	*b3914*	b3914	1.25				*b1152*	b1152	-0.99
yiiR	b3921	1.71							*umuC*	b1184	-0.90
yiiX	b3937	2.41							*narJ*	b1226	-0.93
metJ	b3938	1.93							*ycjC*	b1299	-0.81
trmA	b3965	1.87							*pspE*	b1308	-0.92
			rplJ	b3985	0.72	*ynaF*	b1376	-1.49			
			rplL	b3986	0.74				*b1396*	b1396	-0.74
			yjbM	b4048	0.88	*aldA*	b1415	-1.16			
			yjcF	b4066	0.90	*trg*	b1421	-1.02	*trg*	b1421	-0.67
			miaA	b4171	0.78				*ydcH*	b1426	-1.14
			intB	b4271	0.73				*fdnH*	b1475	-0.95
yi21_6	b4272	1.80							*ydfA*	b1571	-1.22
uxuR	b4324	1.93				*pykF*	b1676	-1.00			
									b1752	b1752	-0.89
									ansA	b1767	-0.91
									b1777	b1777	-0.73
			yabI	b0065	-0.86				*b1778*	b1778	-0.73
			nadC	b0109	-0.92				*b1780*	b1780	-0.69
			speD	b0120	-1.03				*b1788*	b1788	-0.79
			fhuB	b0153	-0.72				*manZ*	b1819	-0.99
			b0247	b0247	-1.04				*fliC*	b1923	-0.73
			b0280	b0280	-0.82	*fliD*	b1924	-1.18	*fliD*	b1924	-1.03
			secF	b0409	-0.85	*fliS*	b1925	-1.04			
			bolA	b0435	-0.92				*fliE*	b1937	-0.82
			b0484	b0484	-0.81				*fliF*	b1938	-0.94
			b0598	b0598	-1.14				*b1976*	b1976	-1.10
			rna	b0611	-0.77				*b2000*	b2000	-0.68
			ybeG	b0622	-0.73				*b2007*	b2007	-1.04
			cspE	b0623	-0.89				*yefJ*	b2031	-1.04
			leuS	b0642	-0.85				*b2045*	b2045	-0.72
ybeK	b0651	-1.08							*gatR2*	b2090	-0.71
			gltA	b0720	-0.88	*gatD*	b2091	-1.09			
sdhA	b0723	-1.09				*gatC*	b2092	-1.29			
b0725	b0725	-1.07							*gatB*	b2093	-1.33
			sucA	b0726	-0.72	*gatA*	b2094	-1.19	*gatA*	b2094	-1.33
glnH	b0811	-1.03				*gatZ*	b2095	-1.53			
			ompX	b0814	-0.71	*gatY*	b2096	-1.54			
cspD	b0880	-1.41							*mglC*	b2148	-0.68
			pflB	b0903	-1.51	*mglA*	b2149	-1.27			
			b0937	b0937	-0.81	*mglB*	b2150	-0.99			
			rmf	b0953	-1.57				*fruB*	b2169	-1.24
			cspG	b0990	-0.81				*b2176*	b2176	-1.02
			b1017	b1017	-0.77				*yejA*	b2177	-0.82
yceH	b1067	-1.01							*nuoN*	b2276	-0.84
flgB	b1073	-1.21							*nuoJ*	b2280	-0.91
flgE	b1076	-1.04	*flgE*	b1076	-1.20	*nuoG*	b2283	-1.20			
flgF	b1077	-0.97				*nuoC*	b2286	-1.16			
flgG	b1078	-1.14				*nuoB*	b2287	-0.99			

Table F.1 Continued

Microarray Rep. 2			Macroarray Rep. 2			Microarray Rep. 2			Macroarray Rep. 2		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
pta	b2297	-1.00				*glnA*	b3870	-1.14			
			b2450	b2450	-0.98				*yihN*	b3874	-0.83
			acrD	b2470	-1.07				*fdoG*	b3894	-0.99
			b2484	b2484	-0.86				*ppc*	b3956	-0.69
			b2493	b2493	-0.77				*hupA*	b4000	-0.90
			upp	b2498	-1.03	*aceB*	b4014	-1.32	*aceB*	b4014	-0.74
csiE	b2535	-1.03				*malF*	b4033	-1.49			
			rseC	b2570	-1.01	*malE*	b4034	-1.83			
yfiQ	b2584	-1.08				*malK*	b4035	-1.03			
			ygaG	b2687	-1.19	*lamB*	b4036	-1.70	*lamB*	b4036	-0.83
			srlA_1	b2702	-0.92	*malM*	b4037	-1.48			
			hycC	b2723	-0.92				*ubiC*	b4039	-0.90
			b2767	b2767	-0.88				*yjbR*	b4057	-0.79
			eno	b2779	-0.98				*soxS*	b4062	-0.77
			fucR	b2805	-1.22				*rpiR*	b4089	-0.99
			b2817	b2817	-0.90	*melR*	b4118	-1.09			
			b2869	b2869	-0.80	*melA*	b4119	-0.96			
			fba	b2925	-0.78	*aspA*	b4139	-1.13	*aspA*	b4139	-1.65
			hybB	b2995	-1.40				*yjfN*	b4188	-0.70
			hybA	b2996	-1.59				*yjfO*	b4189	-0.88
			b3012	b3012	-0.71				*cysQ*	b4214	-0.94
			uxaC	b3092	-0.76				*ytfK*	b4217	-0.74
			secG	b3175	-0.81	*treC*	b4239	-1.14			
			sspB	b3228	-1.01	*treB*	b4240	-1.27	*treB*	b4240	-0.90
			yrdB	b3280	-1.17				*treR*	b4241	-0.93
			rpsM	b3298	-0.76				*yjgF*	b4243	-1.01
			yhfM	b3370	-0.78				*yjgF*	b4243	-0.73
			glgB	b3432	-0.85				*hsdM*	b4349	-0.93
			ugpA	b3452	-0.84	*hsdR*	b4350	-0.99			
uspA	b3495	-0.94	*uspA*	b3495	-1.99						
			yhiP	b3496	-1.29						
			hdeB	b3509	-1.33						
hdeA	b3510	-1.11									
dctA	b3528	-1.38									
			yiaC	b3550	-1.23						
			yiaE	b3553	-0.68						
			cspA	b3556	-1.01						
			xylF	b3566	-0.89						
lldP	b3603	-1.13	*lldP*	b3603	-1.06						
			radC	b3638	-0.76						
yidG	b3675	-1.02									
			yidK	b3679	-0.78						
			gyrB	b3699	-0.68						
tnaL	b3707	-1.67	*tnaL*	b3707	-1.26						
tnaA	b3708	-1.66	*tnaA*	b3708	-1.48						
atpC	b3731	-1.29									
atpG	b3733	-0.98									
atpH	b3735	-0.96									
			rbsD	b3748	-0.71						
rbsB	b3751	-0.97									
yifA	b3762	-1.03									
			fadB	b3846	-1.04						

Table F.2 Alignment by Blattner number order of the results from the standard analysis of the macroarray replicate 3 versus standard analysis of the microarray replicate 3.

Microarray Rep. 3			Macroarray Rep. 3			Microarray Rep. 3			Macroarray Rep. 3		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			talB	b0008	0.86	*yccV*	b0966	1.24			
			dnaK	b0014	0.84				*yccE*	b1001	0.73
			nhaA	b0019	0.83	*ycdT*	b1025	1.01			
			ddlB	b0092	0.67	*ycdU*	b1029	1.05			
			yadC	b0135	1.57				*b1030*	b1030	0.69
fhuC	b0151	1.60				*yceE*	b1053	1.30			
			yaeE	b0198	0.85				*dinI*	b1061	0.71
abc	b0199	1.13							*mviN*	b1069	0.88
yafT	b0217	1.57				*ycfJ*	b1110	1.11			
			b0220	b0220	0.72	*ycfR*	b1112	1.21			
dinP	b0231	1.05							*b1138*	b1138	1.03
			phoE	b0241	0.68				*mcrA*	b1159	1.48
			insB_2	b0264	0.75	*ycgW*	b1160	1.29	*b1160*	b1160	1.66
			b0280	b0280	0.89	*ycgX*	b1161	1.61			
			b0299	b0299	0.85				*b1164*	b1164	0.85
yahA	b0315	1.41	*yahA*	b0315	1.02	*b1169*	b1169	1.39			
			b0316	b0316	1.24				*b1170*	b1170	0.76
yahG	b0321	1.05				*b1171*	b1171	1.16	*b1171*	b1171	1.15
yahI	b0323	1.03				*b1172*	b1172	1.04	*b1172*	b1172	1.43
			b0357	b0357	1.86				*b1173*	b1173	0.67
yi21_1	b0360	1.01							*b1202*	b1202	0.67
			b0370	b0370	1.05				*ychF*	b1203	0.80
b0373	b0373	1.11							*b1228*	b1228	1.69
			araJ	b0396	0.68	*tpr*	b1229	1.57			
glnK	b0450	1.20							*ychG*	b1239	1.01
			ybaA	b0456	0.91	*cysB*	b1275	1.06	*cysB*	b1275	0.81
			hha	b0460	1.20				*osmB*	b1283	0.81
			ybaJ	b0461	0.91	*ycjJ*	b1296	1.02			
			apt	b0469	0.85				*b1297*	b1297	0.73
			dnaX	b0470	0.81	*ycjL*	b1298	1.52			
			rhsD	b0497	1.38	*ydaC*	b1347	1.11			
			ybbC	b0498	0.70				*b1368*	b1368	0.71
			b0499	b0499	0.98	*b1371*	b1371	1.12			
			ybbD	b0500	1.19				*ydbA*	b1405	0.85
			b0501	b0501	0.70				*ydbD*	b1407	0.80
			b0532	b0532	1.01	*b1433*	b1433	1.19			
			tra5_2	b0541	0.80	*ydcP*	b1435	1.03			
ybcK	b0544	2.56							*b1453*	b1453	0.81
			b0546	b0546	1.22				*rhsE*	b1456	1.23
			b0551	b0551	0.69				*ydcD*	b1457	1.12
ybcU	b0557	1.20							*b1459*	b1459	0.90
			b0558	b0558	0.79				*b1527*	b1527	0.76
ybcY	b0562	1.03	*b0562*	b0562	0.73	*marR*	b1530	1.44			
			b0648	b0648	0.86				*marA*	b1531	1.02
ybfE	b0685	1.19	*b0685*	b0685	1.37	*b1545*	b1545	1.35			
			rhsC	b0700	0.76	*cspB*	b1557	1.45			
			ybfD	b0706	0.81	*cspF*	b1558	1.50			
			b0805	b0805	0.78	*relF*	b1562	1.03			
			infA	b0884	0.89				*b1566*	b1566	0.92
									b1568	b1568	0.68
ycaL	b0909	1.83									
ycbO	b0936	1.02				*dicA*	b1570	1.03			
fabA	b0954	1.03				*ydfA*	b1571	1.58			

Table F.2 Continued

Microarray Rep. 3			Macroarray Rep. 3			Microarray Rep. 3			Macroarray Rep. 3		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
ydfC	b1573	1.21							*yeiR*	b2173	0.79
dicB	b1575	1.15							*b2174*	b2174	0.86
			ydhA	b1639	1.11	*rsuA*	b2183	1.31			
			b1643	b1643	0.73				*rplY*	b2185	0.82
			lhr	b1653	0.78				*yfaE*	b2236	0.72
			ydhC	b1660	0.69	*ubiX*	b2311	1.20			
ydiC	b1684	1.12				*cvpA*	b2313	1.03			
			b1720	b1720	0.79				*b2326*	b2326	0.68
			b1721	b1721	1.07				*b2345*	b2345	0.68
b1730	b1730	1.19	*b1730*	b1730	0.84	*vacJ*	b2346	1.16	*vacJ*	b2346	1.35
spy	b1743	1.21							*evgS*	b2370	0.86
			b1770	b1770	0.98	*b2380*	b2380	1.56			
b1773	b1773	1.08				*eutI*	b2458	1.47			
			b1785	b1785	0.79				*b2504*	b2504	0.67
			b1786	b1786	0.98	*b2506*	b2506	1.42			
yeaZ	b1807	1.35							*xseA*	b2509	0.69
yebH	b1822	1.35				*b2529*	b2529	1.71			
b1825	b1825	1.73				*b2531*	b2531	1.23			
			b1826	b1826	0.79	*yfhL*	b2562	1.94			
			htpX	b1829	0.75				*yfiE*	b2577	0.69
ptrB	b1845	1.15							*yfiF*	b2581	0.67
			yebE	b1846	0.68	*yfiM*	b2586	1.03			
			yebG	b1848	1.11				*rpsP*	b2609	0.67
			insB_5	b1893	1.09	*grpE*	b2614	1.55			
			b1933	b1933	0.72	*smpA*	b2617	1.34			
			b1935	b1935	0.76	*yfjW*	b2642	1.71	*b2642*	b2642	1.34
b1936	b1936	1.37				*b2657*	b2657	1.19			
yedJ	b1962	1.23							*stpA*	b2669	0.85
b1963	b1963	1.92	*b1963*	b1963	1.17				*ygaC*	b2671	0.70
			b1969	b1969	0.92	*nrdE*	b2675	1.24			
			b1979	b1979	0.86	*proV*	b2677	1.43			
			b1983	b1983	1.08	*proW*	b2678	1.04			
			cbl	b1987	1.07	*b2680*	b2680	1.05			
yi22_3	b1996	1.26							*oraA*	b2698	0.77
			b2027	b2027	0.79				*recA*	b2699	0.77
			yefI	b2032	1.55				*ygcB*	b2761	0.68
			yefH	b2033	0.70				*fucU*	b2804	0.82
			yefG	b2034	1.47	*yqeH*	b2846	1.52			
wbbH	b2035	1.20	*rfc*	b2035	0.69				*b2847*	b2847	0.95
rfbX	b2037	1.26	*rfbX*	b2037	1.45	*yqeJ*	b2848	1.67			
			rfbA	b2039	0.77				*b2849*	b2849	0.93
			rfbD	b2040	0.68				*b2850*	b2850	0.86
wcaH	b2051	1.03				*ygeH*	b2852	1.24			
wza	b2062	1.08							*b2854*	b2854	1.30
b2070	b2070	1.42							*b2857*	b2857	0.70
b2071	b2071	1.71				*yi22_4*	b2860	1.91			
baeR	b2079	1.57							*b2863*	b2863	1.77
			gatR_2	b2090	0.81				*yggE*	b2922	0.76
			yehC	b2110	0.85				*yqgD*	b2941	0.95
			b2145	b2145	0.75				*b2974*	b2974	0.80
yeiL	b2163	1.27				*b2998*	b2998	1.31			
			yeiQ	b2172	0.67	*ygiR*	b3016	1.23			

Table F.2 Continued

Microarray Rep. 3			Macroarray Rep. 3			Microarray Rep. 3			Macroarray Rep. 3		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			b3022	b3022	0.88	*yiiX*	b3937	1.10			
yi22_5	b3045	1.33				*metJ*	b3938	1.49			
			b3046	b3046	0.72	*argE*	b3957	1.18			
ygjM	b3082	1.27				*birA*	b3973	1.57			
ygjN	b3083	1.77	*ygjN*	b3083	0.78				*dinF*	b4044	0.87
			tdcR	b3119	0.82				*yjbM*	b4048	0.99
			yhaB	b3120	1.77				*yjcF*	b4066	1.11
yraH	b3142	1.13	*yraH*	b3142	1.07				*adiY*	b4116	0.87
deaD	b3162	1.17				*dsbD*	b4136	1.04			
			yhbC	b3170	0.96				*yjeQ*	b4161	0.86
			yhbX	b3173	0.81				*yjeF*	b4167	0.72
			dacB	b3182	0.92				*miaA*	b4171	0.87
			yrbI	b3198	0.68				*yjeB*	b4178	0.98
			b3207	b3207	0.71				*yjfI*	b4181	0.81
yhcA	b3215	1.01							*yjfZ*	b4204	0.78
yhcD	b3216	1.22							*b4215*	b4215	0.87
envR	b3264	1.43							*argI*	b4254	0.67
			rpsJ	b3321	0.71				*b4257*	b4257	0.73
yheE	b3324	1.18							*intB*	b4271	0.75
			yheL	b3343	0.67	*yi21_6*	b4272	1.34			
yrfE	b3397	1.15							*yjhB*	b4279	0.69
			glpD	b3426	1.03				*yjhC*	b4280	0.68
insB_6	b3445	1.01	*insB_6*	b3445	1.17				*yjiR*	b4340	0.70
			yrhB	b3446	0.84	*fhuF*	b4367	2.25			
yhhQ	b3471	1.19									
			rhsB	b3482	1.30						
			yhhH	b3483	1.43						
			yhiJ	b3488	1.99				*yabO*	b0058	-0.67
			yhiK	b3489	1.21				*fruR*	b0080	-0.90
yhiL	b3490	1.15	*yhiL*	b3490	1.49				*ftsI*	b0084	-1.07
yhiM	b3491	1.50	*yhiM*	b3491	0.84				*yacF*	b0102	-0.97
			yhiV	b3514	0.97				*lpdA*	b0116	-0.68
yhiW	b3515	1.96	*b3515*	b3515	0.77				*acnB*	b0118	-1.10
			yhiX	b3516	1.34				*yadG*	b0127	-0.72
yibJ	b3595	2.40	*yibJ*	b3595	1.35	*b0139*	b0139	-1.02			
yibG	b3596	1.42				*mesJ*	b0188	-0.90			
htrL	b3618	1.66	*b3618*	b3618	0.97				*yi52_1*	b0259	-0.94
			rfaL	b3622	1.30	*yahO*	b0329	-0.91			
			rfaZ	b3624	0.75				*b0402*	b0402	-1.08
			rfaI	b3627	0.67	*ybeK*	b0651	-1.16			
			rfaS	b3629	0.78	*abrB*	b0715	-1.33			
			rfaP	b3630	0.70	*b0718*	b0718	-1.30			
dnaA	b3702	1.03							*sdhD*	b0722	-0.77
			bglF	b3722	0.72	*sdhA*	b0723	-0.98			
wzzE	b3785	1.02				*ybiH*	b0796	-0.93			
wecB	b3786	1.02							*glnH*	b0811	-0.72
yifK	b3795	1.19				*cspD*	b0880	-1.13			
			yigF	b3817	0.76				*pflB*	b0903	-1.13
			b3875	b3875	0.74				*focA*	b0904	-1.03
			yiiG	b3896	0.96				*b1010*	b1010	-0.96
			b3913	b3913	1.41	*flgB*	b1073	-1.14			
b3914	b3914	2.47	*b3914*	b3914	1.41	*flgC*	b1074	-1.15			

Table F.2 Continued

Microarray Rep. 3			Macroarray Rep. 3			Microarray Rep. 3			Macroarray Rep. 3		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
flgE	b1076	-1.27	*flgE*	b1076	-0.73	*nuoB*	b2287	-1.15			
flgF	b1077	-1.28				*nuoA*	b2288	-1.07			
flgG	b1078	-1.22				*pta*	b2297	-0.96	*pta*	b2297	-0.71
			flgJ	b1081	-0.98	*dsdA*	b2366	-0.89			
b1085	b1085	-1.23							*b2389*	b2389	-0.83
yceC	b1086	-1.06							*b2412*	b2412	-0.74
			b1140	b1140	-0.74				*cysK*	b2414	-0.71
			sapA	b1294	-0.97				*cchA*	b2457	-0.67
ordL	b1301	-1.23							*b2562*	b2562	-0.81
ynaF	b1376	-1.22							*b2629*	b2629	-0.85
			tynA	b1386	-0.68	*emrB*	b2686	-1.40			
trg	b1421	-0.96							*ygaD*	b2700	-0.75
			b1446	b1446	-0.95	*eno*	b2779	-0.91			
b1448	b1448	-0.87				*b2790*	b2790	-0.89			
yddA	b1496	-1.64							*ppdC*	b2823	-0.72
			b1511	b1511	-0.91	*recJ*	b2892	-0.89			
			b1512	b1512	-1.09	*gcvP*	b2903	-0.89			
			b1641	b1641	-0.74	*gcvH*	b2904	-0.87			
sodB	b1656	-1.12							*gcvT*	b2905	-0.72
pykF	b1676	-1.29	*pykF*	b1676	-0.70	*pgk*	b2926	-1.10			
			b1681	b1681	-0.88	*tktA*	b2935	-1.08			
			b1777	b1777	-0.94	*speA*	b2938	-1.17			
			b1778	b1778	-0.87				*b2989*	b2989	-0.97
			b1810	b1810	-0.78				*b2998*	b2998	-0.85
manY	b1818	-1.09				*yqiH*	b3047	-1.19			
manZ	b1819	-0.97				*ygjL*	b3081	-0.96			
			yebJ	b1831	-0.69				*yraL*	b3146	-1.01
pykA	b1854	-1.07							*yhbP*	b3154	-0.80
			b1877	b1877	-0.68				*secG*	b3175	-0.71
			cheZ	b1881	-0.68	*yhbJ*	b3205	-1.18			
b1903	b1903	-1.09				*hopD*	b3335	-0.88			
fliD	b1924	-1.23							*yrfE*	b3397	-0.98
fliS	b1925	-1.10							*uspA*	b3495	-0.72
fliT	b1926	-0.95							*hdeB*	b3509	-0.92
yedF	b1930	-1.01							*hdeA*	b3510	-0.90
			b1967	b1967	-1.85	*dctA*	b3528	-0.87			
			wcaB	b2058	-0.71	*dppB*	b3543	-0.93			
			b2077	b2077	-0.70	*waaI*	b3627	-0.94			
			b2084	b2084	-1.04				*yicJ*	b3657	-0.70
gatC	b2092	-0.91							*dnaN*	b3701	-0.75
			gatB	b2093	-0.82	*tnaL*	b3707	-1.65	*tnaL*	b3707	-1.54
gatA	b2094	-1.15				*tnaA*	b3708	-1.68	*tnaA*	b3708	-1.39
gatZ	b2095	-1.56				*yieL*	b3719	-0.96			
gatY	b2096	-1.27							*yieC*	b3720	-0.86
yehR	b2123	-0.95				*atpC*	b3731	-1.10			
mglB	b2150	-1.08				*atpD*	b3732	-0.88			
glpQ	b2239	-1.03				*atpG*	b3733	-1.32			
yfbK	b2270	-1.30				*atpH*	b3735	-1.08	*atpH*	b3735	-0.79
nuoI	b2281	-0.88				*atpF*	b3736	-1.12	*atpF*	b3736	-0.86
nuoG	b2283	-1.24				*rbsD*	b3748	-1.11	*rbsD*	b3748	-1.38
nuoE	b2285	-0.97				*yifA*	b3762	-1.11			
nuoC	b2286	-1.46							*yifE*	b3764	-0.90

Table F.2 Continued

Microarray Rep. 3			Macroarray Rep. 3		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			yifJ	b3792	-0.91
			yigE	b3815	-1.43
			b3856	b3856	-0.78
fdoH	b3893	-0.98			
fdoG	b3894	-0.97			
			yijI	b3948	-0.69
aceB	b4014	-0.95			
malF	b4033	-1.18			
malE	b4034	-1.45			
			malK	b4035	-1.22
lamB	b4036	-1.18	*lamB*	b4036	-1.08
malM	b4037	-1.30			
			yjbP	b4055	-0.68
			nrfC	b4072	-0.90
melR	b4118	-0.90			
aspA	b4139	-1.38	*aspA*	b4139	-1.06
			b4144	b4144	-0.72
			yjfR	b4192	-1.04
			b4198	b4198	-0.69
ytfG	b4211	-0.89			
			ytfJ	b4216	-0.71
			yjgA	b4234	-0.85
treC	b4239	-0.90			
treB	b4240	-1.82			
yjgF	b4243	-0.88			
pyrI	b4246	-0.88			
idnR	b4264	-0.93			
sgcX	b4305	-1.10			
			uxuR	b4324	-0.96
			mcrB	b4346	-1.15

Table F.3 Alignment by Blattner number order of the results from the standard analysis of the macroarray in triplicate versus standard analysis of the microarray in triplicate. Log IR indicates the log10 of the average fold induction.

Microarray Triplicate			Macroarray Triplicate			Microarray Triplicate			Macroarray Triplicate		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
			dnaK	b0014	0.87				*rfc*	b2035	0.75
			yadC	b0135	1.50				*rfbX*	b2037	1.23
			yahA	b0315	1.10				*yehC*	b2110	0.85
			yaiN	b0357	1.61				*rplY*	b2185	0.99
yi21_1	b0360	1.10							*vacJ*	b2346	1.18
			hha	b0460	1.20				*evgS*	b2370	0.84
			ybaJ	b0461	0.88				*xseA*	b2509	0.83
			rhsD	b0497	1.60	*yfhL*	b2562	1.71			
			b0499	b0499	0.95	*yfiM*	b2586	1.52			
			ybbD	b0500	1.14				*rpsP*	b2609	1.15
			b0501	b0501	0.82	*yfjW*	b2642	1.97	*yfjW*	b2642	1.14
ybcK	b0544	2.44							*stpA*	b2669	0.96
			ybcM	b0546	1.23				*recA*	b2699	0.82
ybcU	b0557	1.30				*yqeH*	b2846	2.21			
ybcY	b0562	1.81				*yqeJ*	b2848	1.62			
ybfE	b0685	1.29							*b2854*	b2854	1.30
			ybfD	b0706	0.97	*yi22_4*	b2860	1.59			
ycaL	b0909	2.07							*b2863*	b2863	1.57
yceE	b1053	1.27				*ygiR*	b3016	1.11			
ycfJ	b1110	1.90							*yqiG*	b3046	0.74
ycfR	b1112	1.46				*ygjM*	b3082	2.40			
			mcrA	b1159	1.34	*ygjN*	b3083	2.28			
ycgW	b1160	1.63	*ycgW*	b1160	1.60				*tdcR*	b3119	0.85
ycgX	b1161	2.05							*yhaB*	b3120	1.69
			b1171	b1171	1.18	*yraH*	b3142	1.66			
			b1172	b1172	1.35	*yhcA*	b3215	2.01			
			b1202	b1202	0.72	*yheE*	b3324	1.91			
			ychF	b1203	0.92				*glpD*	b3426	0.98
			b1228	b1228	1.65				*yrhB*	b3446	0.83
			cysB	b1275	0.91	*yhhQ*	b3471	1.93			
			osmB	b1283	0.87				*rhsB*	b3482	1.15
			rhsE	b1456	1.21				*yhhH*	b3483	1.30
			ydcD	b1457	1.15				*yhiJ*	b3488	1.84
			b1527	b1527	0.78				*yhiK*	b3489	1.06
ydfA	b1571	1.75				*yhiL*	b3490	2.53	*yhiL*	b3490	1.44
ydfC	b1573	1.24				*yhiM*	b3491	2.55			
			ydhA	b1639	1.00	*yhiW*	b3515	1.70	*yhiW*	b3515	0.82
			b1721	b1721	0.92				*yhiX*	b3516	1.28
			b1730	b1730	0.82	*yibJ*	b3595	2.83	*yibJ*	b3595	1.33
			b1770	b1770	0.91	*htrL*	b3618	1.76	*htrL*	b3618	0.93
yeaZ	b1807	1.74							*rfaL*	b3622	1.06
b1825	b1825	1.73							*yiiG*	b3896	1.00
			yebG	b1848	1.14				*b3913*	b3913	1.24
			yedM	b1935	0.74	*b3914*	b3914	2.22	*b3914*	b3914	1.42
yedJ	b1962	2.19				*yiiX*	b3937	2.07			
b1963	b1963	2.50	*b1963*	b1963	1.12				*yjbM*	b4048	0.93
			b1983	b1983	0.99				*yjcF*	b4066	1.01
			cbl	b1987	1.08				*miaA*	b4171	0.89
yi22_3	b1996	1.26							*intB*	b4271	0.74
			yefI	b2032	1.41	*yi21_6*	b4272	1.78			
			yefG	b2034	1.36						

Table F.3 Continued

Microarray Triplicate			Macroarray Triplicate		
Gene name	Blattner No.	Log IR	Gene name	Blattner No.	Log IR
ybeK	b0651	-1.11			
sdhA	b0723	-1.11			
cspD	b0880	-1.26			
			pflB	b0903	-1.27
flgB	b1073	-1.15			
			flgE	b1076	-0.99
flgG	b1078	-1.25			
			flgJ	b1081	-1.05
ynaF	b1376	-1.41			
			b1777	b1777	-0.80
fliD	b1924	-1.24			
fliS	b1925	-1.11			
			gatB	b2093	-1.28
gatA	b2094	-1.26			
gatZ	b2095	-1.57			
gatY	b2096	-1.42			
mglB	b2150	-1.13			
nuoG	b2283	-1.19			
nuoC	b2286	-1.27			
ygjL	b3081	-0.92			
			uspA	b3495	-1.57
			hdeB	b3509	-1.17
tnaL	b3707	-1.69	*tnaL*	b3707	-1.53
tnaA	b3708	-1.73	*tnaA*	b3708	-1.47
atpC	b3731	-1.20			
atpG	b3733	-1.15			
atpH	b3735	-1.00			
yifA	b3762	-1.10	*rbsD*	b3748	-1.08
aceB	b4014	-1.36			
malF	b4033	-1.47			
malE	b4034	-1.62			
lamB	b4036	-1.51	*lamB*	b4036	-1.00
malM	b4037	-1.53			
aspA	b4139	-1.22	*aspA*	b4139	-1.34
treB	b4240	-1.54			

Table F.4 Alignment by Blattner number order of the results of SAM analysis from microarray and macroarray experiments. Parameters for SAM analysis can be seen in Table A.1 (macroarrays) and Table B.1 (microarrays).

Microarray SAM			Macroarray SAM			Microarray SAM			Macroarray SAM		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
			htgA	b0012	3.0				*ybcI*	b0527	1.5
			dnaK	b0014	7.3	*emrE*	b0543	17.7	*b0543*	b0543	2.6
			dnaJ	b0015	4.9				*b0544*	b0544	1.7
			yi82_1	b0017	1.9				*b0545*	b0545	2.3
			rpsT	b0023	3.6				*b0546*	b0546	16.9
			leuO	b0076	2.1				*b0556*	b0556	3.3
yabC	b0082	5.1							*b0557*	b0557	4.0
			ftsL	b0083	2.4				*b0558*	b0558	5.7
			yadC	b0135	31.6				*appY*	b0564	1.9
			htrA	b0161	3.9				*b0570*	b0570	1.6
			frr	b0172	2.4	*fepD*	b0590	2.4			
			yaeL	b0176	3.6				*b0600*	b0600	1.8
			ldcC	b0186	1.8				*b0603*	b0603	5.1
			yaeF	b0193	1.3				*b0619*	b0619	2.7
rcsF	b0196	3.1							*criR*	b0620	3.3
yaeC	b0197	3.7							*nagB*	b0678	1.4
abc	b0199	19.2				*fur*	b0683	2.7			
yaeD	b0200	8.8							*rhsC*	b0700	5.2
yafQ	b0225	5.5							*ybfB*	b0702	1.9
mbhA	b0230	2.6							*ybfC*	b0704	2.7
dinP	b0231	9.3	*dinP*	b0231	1.4	*ybfD*	b0706	7.5			
			yafN	b0232	3.2				*ybgA*	b0707	1.6
			yafO	b0233	1.6				*b0725*	b0725	1.4
			yafP	b0234	3.6				*b0753*	b0753	1.5
			b0235	b0235	1.4				*ybhD*	b0768	2.3
			b0245	b0245	1.9				*b0770*	b0770	1.8
			insB_2	b0264	5.0	*ybiI*	b0803	5.5			
			insB_3	b0274	3.8				*b0807*	b0807	2.5
insA_3	b0275	8.6				*ybiS*	b0819	11.3	*b0819*	b0819	1.5
			ykgB	b0301	1.7	*b0833*	b0833	16.8	*b0833*	b0833	3.6
			b0302	b0302	2.3				*b0847*	b0847	1.4
			yahA	b0315	12.4				*b0859*	b0859	2.0
			b0326	b0326	1.8	*infA*	b0884	27.2			
yahM	b0327	10.2				*msbA*	b0914	5.9			
			adhC	b0356	2.6				*pncB*	b0931	3.2
			b0357	b0357	30.0				*fabA*	b0954	2.7
			b0362	b0362	1.6				*b0955*	b0955	3.2
			b0365	b0365	2.9				*b0986*	b0986	2.1
ybaD	b0413	4.2				*ymcD*	b0987	7.1			
			b0453	b0453	2.0				*b0989*	b0989	3.0
			ybaA	b0456	5.3	*b1011*	b1011	5.9			
			b0458	b0458	2.3				*b1025*	b1025	2.3
			hha	b0460	15.7				*b1031*	b1031	3.1
			ybaJ	b0461	7.5				*b1047*	b1047	2.1
			dnaX	b0470	4.9	*yceA*	b1055	2.2			
			ybaB	b0471	1.9	*yceP*	b1060	5.2	*b1060*	b1060	2.6
			b0482	b0482	1.6				*dinI*	b1061	4.0
			ybbC	b0498	4.1				*yceB*	b1063	1.8
			b0499	b0499	8.8				*yceD*	b1088	2.0
			ybbD	b0500	12.7				*fabF*	b1095	3.1
			b0501	b0501	6.2				*pabC*	b1096	3.2
			b0517	b0517	1.3				*ycfH*	b1100	1.5

Table F.4 Continued

Microarray SAM			Macroarray SAM			Microarray SAM			Macroarray SAM		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
ycfO	b1107	3.8							*b1668*	b1668	2.8
			ycfJ	b1110	2.5				*b1678*	b1678	2.1
			b1111	b1111	1.8				*b1679*	b1679	1.7
			b1122	b1122	2.1				*ydiB*	b1692	2.1
			ycfC	b1132	1.7				*b1720*	b1720	6.3
			lit	b1139	3.0				*b1722*	b1722	2.0
			b1160	b1160	38.9				*b1728*	b1728	3.0
			b1162	b1162	4.6				*b1730*	b1730	6.5
			b1165	b1165	2.8	*spy*	b1743	20.8	*b1743*	b1743	4.8
			b1169	b1169	3.1				*b1762*	b1762	3.3
			b1170	b1170	4.3				*b1770*	b1770	8.0
b1171	b1171	12.9	*b1171*	b1171	14.8				*rnd*	b1804	1.7
b1172	b1172	6.4	*b1172*	b1172	22.1				*b1809*	b1809	3.3
			ychF	b1203	7.4				*b1811*	b1811	2.3
			hemM	b1209	1.6				*pabB*	b1812	1.8
			ychN	b1219	2.0				*b1821*	b1821	3.9
			b1228	b1228	39.5				*yebH*	b1822	3.4
topA	b1274	6.9				*b1825*	b1825	52.5			
			cysB	b1275	7.8				*b1826*	b1826	6.2
			yciH	b1282	2.1				*htpX*	b1829	4.7
osmB	b1283	3.8	*osmB*	b1283	7.0	*b1841*	b1841	15.4			
ydaY	b1366	6.6							*b1844*	b1844	2.0
			b1374	b1374	2.0	*yebE*	b1846	8.4			
			ydbA_2	b1405	5.8	*yebF*	b1847	20.5			
			rhsE	b1456	14.3	*yebG*	b1848	7.1	*yebG*	b1848	13.5
			ydcD	b1457	13.1				*ruvA*	b1861	3.0
b1459	b1459	18.2							*b1933*	b1933	4.1
			b1504	b1504	2.2				*b1935*	b1935	5.5
			b1524	b1524	3.0				*fliR*	b1950	1.8
			b1527	b1527	5.9				*rcsA*	b1951	3.2
marR	b1530	56.0							*yedA*	b1959	2.8
marA	b1531	15.0							*dcm*	b1961	3.4
			b1546	b1546	1.8				*b1963*	b1963	12.7
			b1560	b1560	2.6	*b1964*	b1964	12.9			
			relF	b1562	2.1				*b1968*	b1968	3.2
relE	b1563	10.0							*b1970*	b1970	3.5
			b1566	b1566	3.5				*b1980*	b1980	2.1
			b1567	b1567	2.9				*b1983*	b1983	9.0
b1568	b1568	3.7							*cbl*	b1987	11.9
			ydfC	b1573	4.6	*yi22_3*	b1996	17.6			
			dicB	b1575	3.1				*yeeC*	b2010	1.6
b1578	b1578	5.7							*yeeF*	b2014	2.6
			b1586	b1586	2.7	*yefM*	b2017	8.3			
			b1596	b1596	3.5				*b2027*	b2027	3.9
			b1604	b1604	3.5	*wbbK*	b2032	5.5	*yefI*	b2032	23.6
			gusC	b1615	2.2	*wbbI*	b2034	23.2	*yefG*	b2034	21.3
			ydhA	b1639	8.7				*rfc*	b2035	5.3
			b1642	b1642	2.5	*rfbX*	b2037	38.7			
			b1648	b1648	5.2				*galF*	b2042	2.1
rnt	b1652	11.8							*b2081*	b2081	4.9
			lhr	b1653	4.4				*b2100*	b2100	1.8
			cfa	b1661	2.0				*yehA*	b2108	3.4
			b1667	b1667	3.0				*yehC*	b2110	7.0

Table F.4 Continued

Microarray SAM			Macroarray SAM			Microarray SAM			Macroarray SAM		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
			metG	b2114	3.6				*lysA*	b2838	1.3
			b2174	b2174	6.1				*b2850*	b2850	5.0
rsuA	b2183	35.6	*yejD*	b2183	2.6				*b2851*	b2851	4.7
			rplY	b2185	9.3				*b2854*	b2854	19.3
yejK	b2186	4.2							*b2856*	b2856	3.4
			yejL	b2187	1.9				*b2862*	b2862	3.6
			eco	b2209	2.7				*pepP*	b2908	3.1
			yfaE	b2236	5.0				*ygfE*	b2910	2.5
			b2255	b2255	1.5				*ygfD*	b2918	2.7
			b2271	b2271	1.5	*sprT*	b2944	7.8			
			b2295	b2295	2.5				*yggN*	b2958	2.5
			truA	b2318	4.2				*b3046*	b3046	5.3
			b2339	b2339	3.2				*b3048*	b3048	2.9
			b2340	b2340	3.2				*cca*	b3056	2.1
			b2345	b2345	3.9				*rpsU*	b3065	5.1
			emrY	b2367	1.2	*dnaG*	b3066	4.9			
			evgA	b2369	5.2				*ygjN*	b3083	5.3
			b2466	b2466	1.7				*ygjO*	b3084	2.5
hyfR	b2491	5.2							*tdcB*	b3117	1.8
			b2496	b2496	2.2				*tdcR*	b3119	6.7
b2506	b2506	42.4				*yraH*	b3142	18.8			
			guaB	b2508	2.5				*yhbU*	b3158	2.0
			xseA	b2509	6.1				*rpsO*	b3165	5.8
b2529	b2529	19.4				*truB*	b3166	2.7			
			yfhD	b2558	2.5				*nusA*	b3169	3.0
			yfhC	b2559	1.3				*yhbC*	b3170	6.1
			pdxJ	b2564	1.5				*greA*	b3181	2.8
			recO	b2565	1.6				*yhbZ*	b3183	3.7
			b2584	b2584	1.6				*rpmA*	b3185	2.0
			clpB	b2592	2.6				*rplU*	b3186	2.9
			sfhB	b2594	4.1	*yrbC*	b3192	3.1			
			tyrA	b2600	1.4				*gltF*	b3214	3.2
			aroF	b2601	1.8				*rpsI*	b3230	2.2
			rplS	b2606	4.8				*rplM*	b3231	2.4
			trmD	b2607	7.9				*argR*	b3237	1.7
			yfjA	b2608	5.5				*mreC*	b3250	1.2
			rpsP	b2609	8.7				*def*	b3287	1.6
			yfjB	b2615	1.4				*rplQ*	b3294	2.2
			recN	b2616	2.4				*rpsE*	b3303	1.8
			b2623	b2623	2.4				*rplR*	b3304	2.1
			b2628	b2628	1.9				*rpsN*	b3307	2.8
			b2639	b2639	1.7	*rpsQ*	b3311	3.1			
			b2665	b2665	1.7				*rpmC*	b3312	1.5
			b2667	b2667	2.0				*rplV*	b3315	2.7
			b2668	b2668	2.1				*rplB*	b3317	2.0
			stpA	b2669	8.8				*rplD*	b3319	2.9
proW	b2678	11.2	*proW*	b2678	3.6				*rplC*	b3320	2.5
			oraA	b2698	5.2				*rpsG*	b3341	3.5
			recA	b2699	6.5				*yheL*	b3343	4.0
			b2760	b2760	1.9				*aroB*	b3389	1.2
			ygcA	b2785	2.2				*glpR*	b3423	2.7
			ptr	b2821	2.1				*glpG*	b3424	3.6
ppdC	b2823	4.4							*glpD*	b3426	9.3

Table F.4 Continued

Microarray SAM			Macroarray SAM			Microarray SAM			Macroarray SAM		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
			rpoH	b3461	2.9				*yjfZ*	b4204	4.2
			b3472	b3472	2.5				*b4257*	b4257	5.6
			yhhS	b3473	2.4				*intB*	b4271	5.4
			yhhT	b3474	2.3				*yi41*	b4278	3.5
			rhsB	b3482	12.7				*yjhB*	b4279	3.7
			yhhH	b3483	17.0				*yjhC*	b4280	3.6
yhhI	b3484	9.3							*yjhO*	b4305	2.8
			yhiJ	b3488	46.3				*fimB*	b4312	2.6
			yhiK	b3489	10.3				*rimI*	b4373	1.9
			yhiM	b3491	4.2				*yjjG*	b4374	2.0
yhiR	b3499	3.2				*prfC*	b4375	10.7			
			gor	b3500	2.1				*yjjI*	b4380	1.7
			arsB	b3502	3.0	*gpmB*	b4395	2.0			
			slp	b3506	2.2						
			yhiF	b3507	2.6						
			gadA	b3517	2.3	*aceE*	b0114	0.2			
			b3524	b3524	2.9				*b0518*	b0518	0.4
cspA	b3556	13.9							*ybeK*	b0651	0.3
			avtA	b3572	1.3				*b0701*	b0701	0.3
yibA	b3594	11.1				*cydB*	b0734	0.2			
			yibJ	b3595	19.9	*cspD*	b0880	0.1			
htrL	b3618	54.9	*b3618*	b3618	8.4				*b0881*	b0881	0.3
			rfaL	b3622	8.6				*trxB*	b0888	0.3
			yidR	b3689	2.5				*pflB*	b0903	0.1
			yidS	b3690	1.2	*pepN*	b0932	0.2			
			rnpA	b3704	2.0	*rmf*	b0953	0.2			
mioC	b3742	4.1							*rimJ*	b1066	0.5
			asnC	b3743	2.7				*flgA*	b1072	0.3
			rep	b3778	3.7	*flgB*	b1073	0.1			
			yigG	b3818	4.0				*flgF*	b1077	0.3
			yiiG	b3896	8.7				*flgG*	b1078	0.3
			yiiL	b3901	1.8				*flgH*	b1079	0.4
			yiiM	b3910	2.1				*flgJ*	b1081	0.1
			b3913	b3913	15.6	*b1085*	b1085	0.2			
trmA	b3965	7.0							*minE*	b1174	0.4
			rplA	b3984	3.3				*ychE*	b1242	0.4
			yjaA	b4011	3.4				*pspB*	b1305	0.6
			xylE	b4031	1.2				*b1312*	b1312	0.5
			yjbM	b4048	8.2				*fnr*	b1334	0.8
			yjcF	b4066	9.7				*b1420*	b1420	0.3
			proP	b4111	3.3				*b1483*	b1483	0.4
			basS	b4112	2.4	*ydhO*	b1655	0.2			
dsbD	b4136	10.4							*flhA*	b1879	0.5
			frdA	b4154	2.0				*gatZ*	b2095	0.3
			yjeQ	b4161	5.2				*nuoM*	b2277	0.5
			miaA	b4171	7.3	*nuoK*	b2279	0.3	*nuoK*	b2279	0.4
			hflX	b4173	3.2				*nuoH*	b2282	0.6
			hflC	b4175	2.3				*pta*	b2297	0.2
			purA	b4177	2.7				*b2483*	b2483	0.4
yjeB	b4178	10.2							*yfiD*	b2579	0.3
vacB	b4179	3.0							*hycB*	b2724	0.4
yjfH	b4180	6.6							*eno*	b2779	0.2
			yjfI	b4181	4.3						

Table F.4 Continued

Microarray SAM			Macroarray SAM		
Gene name	Blattner No.	FC	Gene name	Blattner No.	FC
gcvT	b2905	0.2			
			pgk	b2926	0.4
			hybG	b2990	0.5
			hybB	b2995	0.2
			sufI	b3017	0.7
			yhfQ	b3374	0.5
			pckA	b3403	0.4
			uspA	b3495	0.1
			yhjY	b3548	0.7
			radC	b3638	0.3
			yicE	b3654	0.5
			tnaL	b3707	0.0
			tnaA	b3708	0.0
kup	b3747	0.4			
rbsD	b3748	0.1			
			murB	b3972	0.7
hupA	b4000	0.2			
			malK	b4035	0.1
			soxR	b4063	0.9
			ytfA	b4205	0.4
			nadR	b4390	0.5

Publications

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 318 (2004) 397–404
BBRC
www.elsevier.com/locate/ybbrc

Cooperativity in the binding of the cationic biocide polyhexamethylene biguanide to nucleic acids

Michael J. Allen,[1] Andrew P. Morby, and Graham F. White*

Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3US, UK

Received 6 April 2004
Available online 22 April 2004

Abstract

The interaction between the broad-spectrum antimicrobial agent, polyhexamethylene biguanide (PHMB), and various nucleic acids was investigated. Titration of either single- or double-stranded 100-bp DNA, or mixed-molecular weight marker DNA, or tRNA with PHMB caused precipitation of a complex between nucleic acid and PHMB in which the nucleotide/biguanide ratio was always close to unity. Binding of PHMB was highly cooperative, with apparent Hill coefficients 10.3–14.6. When a fluorescent derivative of PHMB was titrated with increasing amounts of nucleic acid, all four forms of nucleic acid caused strong polarisation of fluorescence, demonstrating the association with PHMB. The intensity and broad-spectrum binding of PHMB to all forms of nucleic acid has significant implications for the mechanism of action of this biocide.

Keywords: Biocide; Polyhexamethylene biguanide; Cationic surfactant; DNA binding; Cooperativity

Polyhexamethylene biguanide (PHMB) is a broad-spectrum antibacterial agent that also possesses biocidal action against some fungi and protozoa. Preparations of PHMB are mixtures of polymeric biguanides (Fig. 1) with a molecular weight range of 400–8000 representing polymers of $n = 2$ to $n = 40$, respectively, with an average size of $n = 11$. It has been widely used for many years as an antiseptic in medicine and its current applications also include swimming pool sanitisation, the treatment of cooling systems to prevent infection by *Legionella* [1], solid surface cleaners in the food industry, as a treatment against fungi [2] and *Acanthamoeba* [3–5] in infective keratitis, as a disinfecting contact lens solution [6], as an antibacterial mouth rinse [7,8], as a durable anti-odour finish in textiles [9], impregnation of gauze wound-dressing to prevent *Pseudomonas* infection [10], and the treatment of hatching eggs to prevent *Salmonella* infection [11,12].

Khunkitti et al. [13] observed clusters of densely stained precipitates in *Acanthamoeba castellanii* treated with high concentrations of PHMB, and also noted that PHMB treatment produced increased amounts of phosphorus inside the cells compared with untreated controls, and these accumulations were often confined to cell walls and nuclei [14]. Reduced membrane permeability causing retention of phosphorus, coagulation of proteins, and aggregation of phospholipids have been considered as possible causes of elevated phosphorus but the possibility of association between PHMB and nucleic acids has not been not considered hitherto, despite rapid growth in the literature in the last 20 years [15] on the interaction of DNA with the natural polyamines, putrescine, spermine, and spermidine (for a review, see Cohen [16]), and with a variety of other synthetic polycationic compounds that are currently being developed as vehicles for non-viral transfection of DNA into cells for therapeutic purposes (for a review, see [17]).

Here we show that PHMB interacts strongly and cooperatively with nucleic acids in vitro. In addition, we show that the structural status of the nucleic acid affects the nature of its initial interactions with PHMB.

* Corresponding author. Fax: +44-29-2087-4116.
E-mail address: whitegf1@cardiff.ac.uk (G.F. White).
[1] Present address: Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, UK.

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.04.043

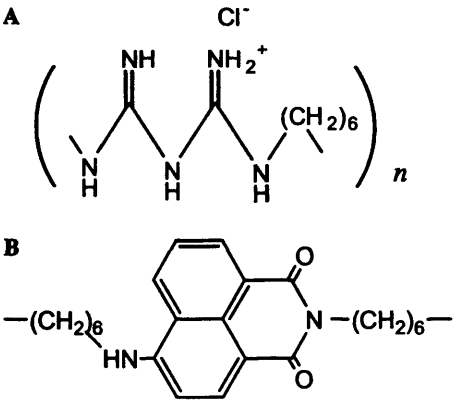


Fig. 1. Structure of PHMB. (A) Repeating unit of hexamethylene biguanide with $n = 2$–40, average 11; end-groups may be amine, cyanoguanide or guanide. (B) The fluor, 6-amino-benzo[de]isoquinoline-1,3-dione, attached to flanking hexamethylene chains via the ring N(2) and amino substituent at position 6, replaces 1 in 100 biguanides in the fluorescent PHMB.

Materials and methods

Materials. A 100-bp sequence of DNA was randomly selected from the *Escherichia coli* K12 MG1655 genome. The sequence was part of the *yfiA* gene, from bases 3001 to 3100 of *E. coli* MG1655, section 236 of the 400 in the complete genome [18]. Single-stranded (ss) and double-stranded (ds) versions were purchased from Sigma (Poole, Dorset, UK). A *Hin*dIII digest of Lambda DNA marketed as a molecular weight marker (125–23,130 bp, Promega UK, Southampton, UK) was used as a mixed-molecular weight sample of dsDNA. Yeast tRNA, from Sigma, was a mixture of approximately 250 different tRNAs ranging in size from 71 to 106 bases, mainly 72–74 bases long (70%), with the mode at 72 bases.

PHMB was kindly provided by Avecia (Blackley, Manchester) as a mixture of homologues with n (Fig. 1) ranging from 2 to 40 and with a mean value of 11. Fluorescent-PHMB, also from Avecia, was of the same composition except that 1% of all biguanide units was replaced by the fluor, 6-amino-benzo[de]isoquinoline-1,3-dione, attached to flanking hexamethylene chains via the ring N(2) and amino substituent at position 6 (see Fig. 1). The synthetic route ensured that each homologue ($n = 2$–40) was labelled and this was confirmed by coincidence of UV and fluorescent elution profiles of effluents from GPC column chromatography. The fluorescence was not quenched by added PHMB. The parent fluor, 1,8-naphthalimide, was obtained from Lancaster Chemicals (Lancaster, UK).

Precipitation experiments. Aliquots (100–150 μl) of stock aqueous solution of a nucleic acid (either dsDNA, ssDNA, *Hin*dIII-digested Lambda DNA or yeast tRNA, containing the same concentration of nucleic acid expressed as nucleotides, typically 0.3 mM) were mixed with appropriate volumes (0–50 μl) of stock PHMB solutions to produce 150 μl final volumes containing 0.25 mM nucleotide and incremental concentrations of PHMB up to 0.3 mM biguanide units. After mixing for 10 s at 20 °C, precipitated material was sedimented by centrifugation at 20,000*g* for 25 min. Supernatants were diluted with 0.35 ml of water and UV absorbances at 236 and 260 nm were determined in 1 cm quartz cuvettes using a Hewlett Packard diode array spectrophotometer. Using known extinction coefficients of nucleic acid and PHMB at these wavelengths, the concentrations of nucleic acid and PHMB remaining in solution were estimated using simultaneous equations.

Fluorescence polarisation. Aliquots (3 μl) of either DNA or tRNA stock solution (1.67 mM nucleotides) were added to 2 ml of fluorescent-PHMB (25 μM in biguanide units) in 3-ml quartz fluorimetry cuvettes (Hellman) with constant stirring at 30 °C and after each addition, fluorescence measurements were made in a Cary Eclipse fluorimeter (Varian). Excitation was for 0.5 s at 455 nm through a 20 nm slit in the vertical plane. Emissions were measured at 535 nm in the vertical and horizontal planes. For the two controls, the procedure was repeated except that either (a) aliquots of the stock nucleic acid solutions were replaced by the same aliquots of water or (b) fluorescent-PHMB solution was replaced with a solution of the parent fluor, 1,8-naphthalimide (excitation and emission wavelengths, 340 and 395 nm, respectively), at 0.25 μM (equivalent to the total fluor concentration in the fluorescent-PHMB solution).

Polarisation index was calculated as $p = (F_v - F_h)/(F_v + F_h)$, where F_v and F_h were the fluorescence intensities in the vertical (parallel to the excitation plane) and horizontal planes, respectively. The polarisation index of free fluorescent-PHMB, p_f, and that of fully bound fluorescent-PHMB, p_b, were determined for fluorescent-PHMB in the absence of nucleic acid and in the presence of excess nucleic acid, respectively. The respective molar fluorescence constants Q_f and Q_b (where $Q = (F_v + F_h)$/[molar concentration]) of free and bound forms were determined in a similar way. These constants and the observed value of p for a given mixture enabled the calculation [19] of the ratio, R, of bound to free fluorescent PHMB as:

$$R = \frac{Q_f}{Q_b}\left(\frac{p - p_f}{p_b - p}\right).$$

Finally, since $R = L_b/L_f$, where L_b and L_f are the concentrations of bound and free PHMB, respectively, and $L_b + L_f = [\text{PHMB}]_{tot}$ (the total concentration of fluorescent-PHMB present), L_b and L_f were calculated using Eqs. (1) and (2), respectively.

$$L_b = [\text{PHMB}]_{tot}\left(\frac{R}{1+R}\right), \quad (1)$$

$$L_f = [\text{PHMB}]_{tot}\left(\frac{1}{1+R}\right). \quad (2)$$

Results

Precipitation experiments

Titrating dsDNA (100 bp, 2.5 μM strands, i.e., 0.25 mM nucleotide phosphate residues) with increasing concentrations of PHMB caused the concentrations of free PHMB (measured as dissolved monomer) and of dissolved DNA to change as shown in Fig. 2A. At the equivalence point (i.e., addition of 0.25 mM PHMB biguanide units to 0.25 mM nucleotide phosphate residues), residual concentrations of DNA and PHMB were about 5% and 10% of totals added, respectively, indicating extensive precipitation of each from solution. The straight line of unit slope shows the expected concentration of dissolved PHMB, assuming no loss by precipitation. The extent to which measured values fell below this line during the titration shows the loss of PHMB by precipitation, which increased modestly as PHMB was added. At the same low concentrations of added PHMB, the concentration of dissolved DNA was similarly reduced by precipitation. However, as concentrations of added PHMB approached 0.2 mM further equal increments in PHMB produced marked reductions in both dissolved DNA and dissolved PHMB.

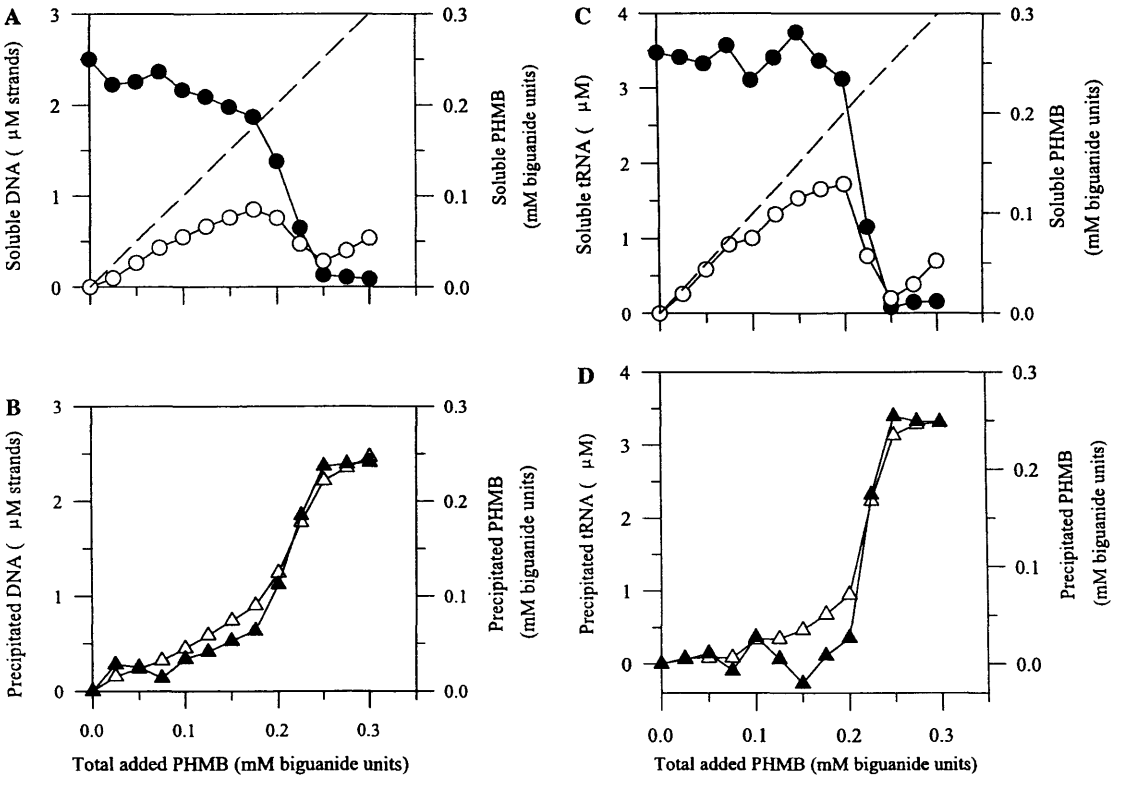


Fig. 2. Titration of either DNA (A,B) or tRNA (C,D) with increasing concentrations of PHMB to form insoluble complexes. PHMB concentration is expressed as molarity of biguanide units. Initial total concentration of dsDNA (100 bp) was 2.5 μM strands (i.e., 0.25 mM nucleotide phosphate residues). (A,C) Concentrations in solution: filled circles, free (soluble) nucleic acid; open circles, free (soluble) PHMB. Broken line is the line of unit slope, i.e., expected soluble PHMB in the absence of any precipitation. (B,D) Concentrations lost by precipitation: filled triangles, complexed nucleic acid, calculated as [total initial nucleic acid minus measured soluble nucleic acid]; open triangles, complexed PHMB calculated as [added PHMB minus measured soluble PHMB].

These results indicated a strong complex formation between DNA and PHMB that appears virtually complete when the total PHMB concentration (as monomer) was 0.25 mM. Fig. 2B shows concentrations of the DNA–PHMB complex measured as either precipitated-DNA (initial concentration minus measured dissolved concentration) or precipitated-PHMB (total PHMB added minus measured dissolved concentration), each as a function of added PHMB. Close correspondence between the curves in Fig. 2B supports the notion of complex formation between DNA and PHMB. The plot of precipitated-PHMB concentration against precipitated-DNA concentration (not shown) was linear with a slope of 106, indicating 106 PHMB biguanide units associated per 100 bp, i.e., equivalence between nucleotide units and biguanide units in the complex. Moreover, the sigmoidal shape of the curves in Fig. 2B indicated a cooperativity in the binding of PHMB to DNA, i.e., initial binding of PHMB to DNA promoted stronger binding of further PHMB.

Similar curves were obtained when double-stranded DNA was replaced by single-stranded DNA of the same length and at the same concentration of strands (not shown). As with dsDNA, there was a close correspondence between precipitated-DNA and precipitated-PHMB, with 98 biguanides bound per 100-base strand, i.e., 0.98 biguanides per nucleotide residue.

To determine whether DNA fragment-size was critical in this phenomenon, a mixture of fragments of dsDNA derived by *Hin*dIII digestion of Lambda DNA was used. The sample (a DNA molecular size marker for calibrating electrophoresis gels) contained an equimolar mixture of eight fragments ranging in size from 125 to 23,130 bp (see Materials). For the titration with PHMB, the concentration of digested DNA was adjusted so that the total nucleotide concentration (0.2 mM, corresponding to 2.1 nM of each double-stranded fragment) was similar to that used in the previous experiments (0.25 mM). Titration curves (not shown) were very similar to those shown in Fig 2A, and with the addition of PHMB equivalent to a total of 0.25 mM monomer, the residual soluble concentrations of PHMB and DNA were 10% and <5%, respectively. The corresponding complex-formation curves (not shown) paralleled the data in Fig. 2B with 1.18 biquanides bound per nucleotide residue, and a sigmoidal form characteristic of

cooperativity. These results indicated that complex formation via a cooperative binding of PHMB to DNA occurred with a range of sizes of DNA.

Addition of PHMB concentrations in the range equivalent to 0–0.075 mM biguanides to tRNA produced no reduction in soluble tRNA, and no disappearance of any added PHMB (Fig. 2C), indicating no precipitation under these conditions. In the added PHMB concentration range 0.1–0.2 mM, there was a slight reduction in dissolved concentration, but at higher concentrations (around 0.2 mM PHMB monomer, 80% of the total nucleotide concentration) there was a very marked precipitation of all tRNA over a narrow range of PHMB concentrations (0.2–0.25 mM). Complex-formation curves (Fig. 2D) measured as either precipitated-tRNA or precipitated-PHMB were very similar (with 0.95 biguanide units bound per nucleotide residue, assuming the length of tRNA as the modal value of 72, see Materials and methods), and showed very strong cooperativity.

Fluorescence polarisation

In order to investigate interactions between PHMB and DNA at lower concentrations at which precipitation did not occur, polarisation of the fluorescence from fluorescent-PHMB during its titration with nucleic acids was measured (Fig. 3A). Addition of increasing concentrations of either DNA or tRNA produced increasing polarisation, indicating progressive restriction of rotation of fluorescent-PHMB molecules, consistent with its binding to the much larger nucleic acid molecules. Control experiments using the same procedure but in which (a) water replaced the stock solutions of nucleic acid and (b) free fluor (1,8-naphthalimide) replaced fluorescent-PHMB, produced no increase in the baseline polarisation (Fig. 3A).

Fig. 3B shows values of concentration of bound PHMB (calculated from values of p as described in Materials and methods and expressed as biguanide unit equivalents) plotted against total nucleic acid added (expressed as nucleotide equivalents added). For DNA, the titration curve closely followed the line of unit slope up to 10 μM DNA nucleotides, i.e., one biguanide unit was bound for every nucleotide unit added. At higher concentrations of added DNA, the number of bound biguanide residues exceeded the number of nucleotide residues added, so that almost all 25 M biguanide units present were in complex form when only about 17 μM nucleotides had been added.

For tRNA, at concentrations of added tRNA nucleotides up to 17.5 μM, the amount of complexed PHMB fell below the line of unit slope, showing that only about 60–70% of nucleotides were complexed with biguanide units (Fig. 3B). However, between 17.5 and 22.5 μM nucleotides, the binding of biguanides strongly

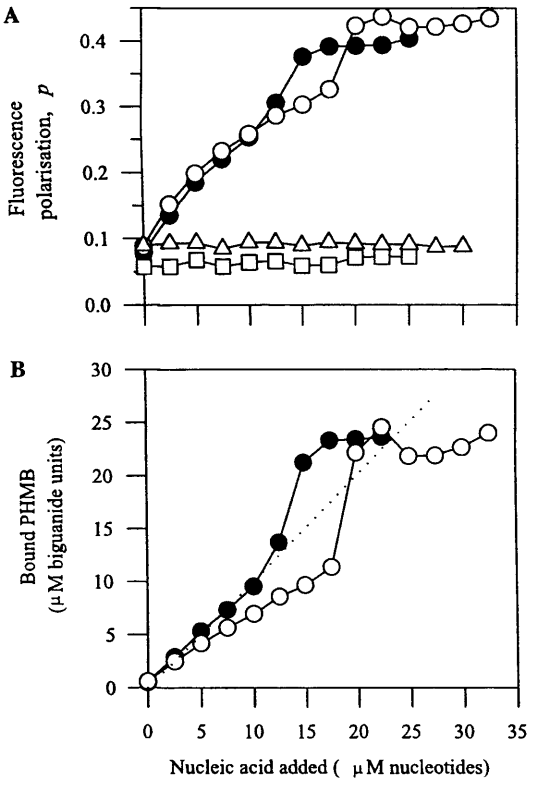


Fig. 3. Fluorescence polarisation of fluorescent-PHMB during titration with DNA (filled circles) or tRNA (open circles). (A) Fluorescence polarisation was calculated as $p = (F_v - F_h)/(F_v + F_h)$, where F_v and F_h were the fluorescence intensities in the vertical (parallel to the excitation plane) and horizontal planes, respectively. Aliquots (3 μl) of either DNA or tRNA stock solution were added to fluorescent-PHMB, and fluorescence measurements made at 30 °C between additions (see text for details). In the control experiments, either the stock nucleic acid solutions were replaced by the same aliquots of water (open squares), or the fluorescent-PHMB was replaced with 1,8-naphthalimide at the equivalent concentration (open triangles). See text for details. (B) Bound PHMB was calculated from p as described in the text. The dotted line is the line of unit slope.

recovered to the same point as with DNA, i.e., all PHMB biguanide units bound to less than an equivalent amount of nucleotides.

Discussion

At micromolar concentrations of polymeric PHMB and nucleic acid molecules, single-stranded DNA, double-stranded DNA of different lengths, and tRNA all interacted with PHMB to remove a proportion of both species (nucleic acid and PHMB) from solution as a precipitate. Given the polyanionic nature of the (deoxy)ribose phosphate backbones, and the polycationic nature of PHMB (Fig. 1), electrostatic interaction is likely to be the dominant factor in complex formation. This is consistent with the parallel precipitation of

bi-guanides and nucleotide phosphates, invariably on a 1:1 basis throughout the titrations (Figs. 2–4).

Recent work has shown that the optimum ratio of polyethyleneimine/DNA needed to maximise gene translocation into cells [20] corresponded to a N/P ratio of 4.5. The spacing between positive charges on polyethyleneimine is 3 bonds, making for a spacing between phosphate-bound nitrogens of $4.5 \times 3 = 13.5$ bonds. Thus, for this polymeric cation, 13.5-bond loops between phosphate-bound cationic groups optimise translocation. Examination of the structure of PHMB (Fig. 1) shows that there are in fact 13 C–N or C–C bonds between successive positively charged imino groups. Thus, it may be that the "intra-cation spacing" in PHMB is already optimised for DNA binding.

The sigmoidal form of the curves in Figs. 2B and D shows that initial binding of PHMB promotes the stronger binding of further PHMB. The occurrence of this cooperativity with ssDNA (not shown) demonstrates it does not arise simply from increased PHMB binding following separation of DNA strands. In enzyme kinetics, cooperativity is assessed using the Hill equation, one form of which is given by

$$\log(Y/(1-Y)) = h\log[\mathrm{L}]_{\mathrm{T}} + h\log K, \quad (3)$$

where Y is the fraction of occupied sites, $[\mathrm{L}]_{\mathrm{T}}$ is the total concentration of substrate (ligand) added, K is a constant, and h is the Hill coefficient, a measure of cooperativity. Apparent Hill coefficients are shown in Table 1. These values are particularly high; for enzymes, values are always well below 10 and typically in the range 1–3. Thus, the precipitation arising from binding of PHMB to DNA is strongly cooperative.

Table 1
Measures of cooperativity for binding of PHMB to nucleic acids to form a precipitated complex

Nucleic acid		Apparent Hill coefficient[a]
Type	Size, characteristics	
DNA	100 bp duplex	14.8
DNA	100 bases, single strand	10.3
DNA	Mixed sizes, 125–23,130 bp	14.6
tRNA	Mixture from yeast, modal length is 72 bases	14.5

[a] Value for non-cooperative binding is 1; values increase as cooperativity increases.

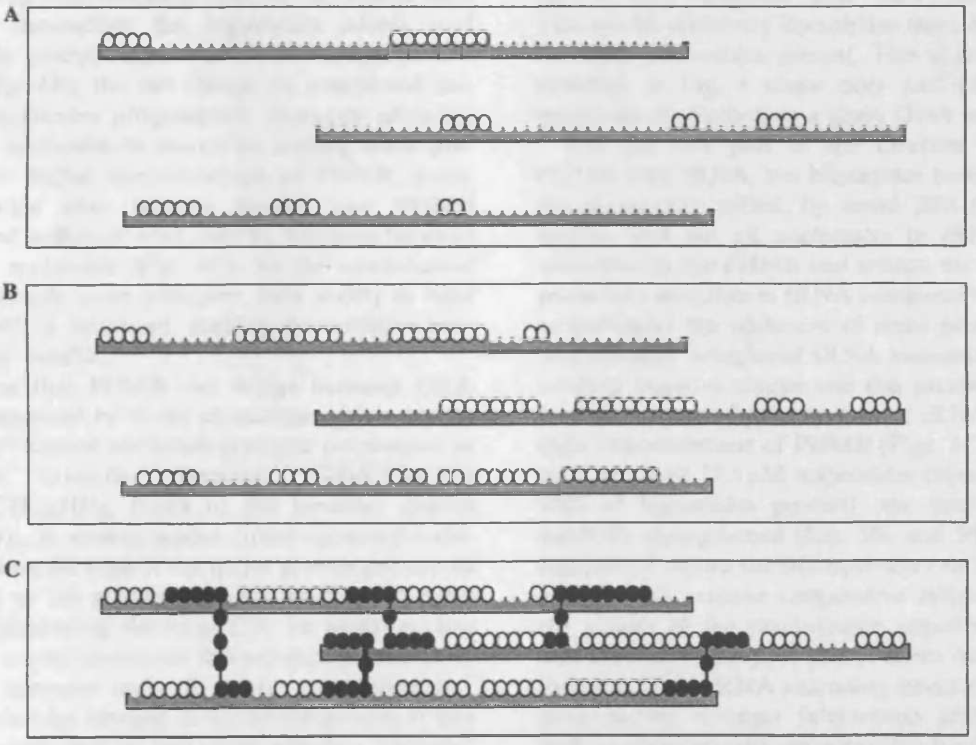


Fig. 4. Proposed mechanism for precipitation of nucleic acid by addition of PHMB. Long grey bars are nucleic acid molecules, each cup-shape representing a nucleotide phosphate residue capable of binding positively charged biguanide groups (circles) that occur in PHMB oligomers. Filled circles represent PHMB molecules that bridge between nucleic acid molecules. (A) PHMB at low concentration occupies a small fraction of the available sites on nucleic acid. At intermediate concentrations (B), PHMB causes significant decrease in net charge and promotes association of nucleic acid molecules. At higher concentration of PHMB (C), PHMB molecules find sufficient binding sites only by bridging across nucleic acids. See text for details.

Although the Hill equation is based loosely on deviations from a hyperbolic saturation curve, it is essentially empirical and constants derived using it should not be interpreted in terms of any specific mechanism of cooperativity [21]. Nevertheless, values of the Hill coefficient may be taken as an upper limit on the number of binding sites. This interpretation for the defined length DNA (100 bp) predicts a maximum of 10.3–14.8 PHMB-binding sites per 100 bp molecule (Table 1). To achieve a fully complexed 100-bp DNA by binding PHMB containing 11 biguanides (the average length), there need to be $100/11 = 9.1$ PHMB molecules bound per strand, i.e., on average there are 9.1 binding sites per 100 bp. This is close to and entirely consistent with the maximum predicted from the Hill analysis and strengthens the hypothesis that PHMB-nucleic acid binding is based on a 1:1 association between biguanide and phosphate residues. On the other hand, the 72 bases in tRNA should provide, on average, $72/11 = 6.5$ PHMB-binding sites whereas the Hill coefficient for PHMB binding to tRNA is much higher (14.5, Table 1). However, this exceptionally high cooperativity probably arises from additional factors (see below).

To account for cooperativity, we propose the mechanism illustrated in Fig. 4. At low concentration of PHMB (Fig. 4A), available sites on the nucleic acid significantly outnumber the biguanides added, and there is little precipitation. As PHMB concentration increases (Fig. 4B), the net charge on complexed nucleic acids molecules progressively decreases allowing nucleic acid molecules to associate, causing some precipitation. At higher concentrations of PHMB, available nucleotide sites become limited and PHMB molecules find sufficient sites only by bridging between nucleic acid molecules (Fig. 4C). As the cross-linked complexes become more extensive, their ability to bind further PHMB is enhanced, making the complex progressively less soluble.

The notion that PHMB can bridge between DNA strands is supported by X-ray crystallographic evidence for binding of natural and other synthetic polyamines to DNA. For example, spermine, $NH_2(CH_2)_3NH(CH_2)_4NH(CH_2)_3NH_2$, binds to the hexamer duplex DNA, $d(CG)_3$, in several modes. Some spermine molecules bind along the edge of the major groove and across the entrance to the groove [22], and mediate contacts between neighbouring duplexes [23]. In addition, low temperature crystal structures showed that besides these "interhelix" spermine molecules, there are "intrahelix" spermine molecules binding in the minor groove of two duplexes stacked end to end, thus creating "infinite" duplexes [24]. More recent data show that several other polyamines also mediate association of DNA molecules in the crystal. Thus, spermidine, $NH_2(CH_2)_3NH(CH_2)_4NH_2$, although shorter than spermine, still binds two flanking DNA duplexes [25], and thermospermine $(NH_2(CH_2)_3NH(CH_2)_3NH(CH_2)_5NH_2$, an isomer of spermine) and the synthetic analogue N^1-[2-(2-aminoethylamino)-ethyl]-ethane-1,2-diamine $(NH_2(CH_2)_2NH(CH_2)_2NH(CH_2)_2NH_2)$ each link three duplexes to each other [26,27]. If these comparatively short polyamines can bridge between duplexes then there is every possibility that the long PHMB molecules, with on average more cationic groups per molecule, will do so equally, if not more, effectively.

Fluorescence polarisation experiments in which fluorescent-PHMB was titrated with DNA showed that, up to about the mid-titration point, the amount of biguanide converted to the complex form was equal to the total amount of nucleotides added (Fig. 3B). This implies that every nucleotide site on the added DNA molecules was associated with a biguanide residue, and is consistent with a strong association between DNA and PHMB and with the 1:1 binding ratio deduced from the precipitation experiments. Remarkably, further increments of DNA were able to bind more than an equivalent amount of biguanides. We propose that in this later stage, all nucleotides sites are fully occupied and some molecules of PHMB are bound to DNA through some but not all of the biguanides present in the molecule, i.e., PHMB molecules are bound only along a part of their structure, with "tails" into the medium. This would effectively immobilise more biguanides than there are nucleotides present. This is analogous to the situation in Fig. 4 where only part of some PHMB molecules is attached to a given DNA molecule.

For the first part of the titration of fluorescent-PHMB with tRNA, the biguanides bound fell short of the nucleotides added, by about 30% (Fig. 3A). This implies that not all nucleotides in tRNA are readily accessible to the PHMB and reflects the more extensive secondary structure in tRNA compared with DNA, and in particular the occlusion of some phosphate groups. The partially complexed tRNA molecules must carry a residual negative charge and this presumably accounts for the absence of precipitation of tRNA by even quite high concentrations of PHMB (Figs. 2C and D). However, at about 17.5 μM nucleotides (equivalent to about 70% of biguanides present), the binding of PHMB suddenly strengthened (Fig. 3B) and PHMB was fully complexed before the full equivalent of nucleotides was added. This extreme cooperative behaviour paralleled the results of the precipitation experiments (Figs. 2C and D) and we propose that it arises from sudden degradation of the tRNA secondary structure caused by the progressively stronger interactions with PHMB, thus making more phosphates accessible for PHMB binding.

Hitherto, mechanisms for the biocidal action of PHMB have focussed primarily on the readily observable changes occurring at the cell envelope, including loss of LPS, changes in membrane integrity, and loss of function of membrane proteins [27–31]. A number of

cytoplasmic changes have also been observed [13,14], but these have not been incorporated into a mechanistic theory. Nevertheless, they clearly imply ingress of PHMB to the cytoplasm, and this, together with the results presented here, that PHMB can bind tightly and cooperatively to DNA and RNA, opens a clear possibility that the binding of PHMB to nucleic acids inside the cell may be an important contributor to the bacteriostatic and biocidal action of this compound. This prospect is lent credence by the well-known interaction of DNA with natural polyamines (spermine, spermidine) that achieves significant compaction of DNA and as a result has important implications in all aspects of cell biology, including cell growth and division, replication and repair of DNA, synthesis and processing of RNA, and protein synthesis (see [15] for a review). PHMB-nucleic acid complex formation might also account for the observed aggregation of phosphorus near the cell wall and nuclei in PHMB-treated Acanthamoeba [14].

Given that transcriptional regulation is paramount in enabling prokaryotic adaptation, growth, and survival in the face of ever-changing and potentially hostile environments, the possibilities for PHMB to interfere in prokaryotic cell-function by binding to both DNA and RNA are legion and hard to ignore.

Acknowledgments

We thank Dr. A. Collins and Dr. L.P. O'Malley (Avecia, Manchester, UK) for gifts of PHMB and fluorescent PHMB.

References

[1] J.M. Kusnetsov, A.I. Tulkki, H.E. Ahonen, P.J. Martikainen, Efficacy of three prevention strategies against legionella in cooling water systems, J. Appl. Microbiol. 82 (1997) 763–768.

[2] C.R. Messick, S.L. Pendland, M. Moshirfar, R.G. Fiscella, K.J. Losnedahl, C.A. Schriever, P.C. Schreckenberger, In vitro activity of polyhexamethylene biguanide (PHMB) against fungal isolates associated with infective keratitis, J. Antimicrob. Chemother. 44 (1999) 297–298.

[3] R. Donoso, J.J. Mura, M. Lopez, *Acanthamoeba* keratitis treated with propamidine and polyhexamethyl biguanide, Rev. Med. Chil. 130 (2002) 396–401.

[4] T.B. Gray, K.A. Gross, R.T.M. Cursons, J.F. Shewan, *Acanthamoeba*-Keratitis—a sobering case and a promising new treatment, Aust. N. Z. J. Ophthalmol. 22 (1994) 73–76.

[5] S. Narasimhan, H.N. Madhavan, L.K. Therese, Development and application of an in vitro susceptibility test for *Acanthamoeba* species isolated from keratitis to polyhexamethylene biguanide and chlorhexidine, Cornea 21 (2002) 203–205.

[6] K. Hiti, J. Walochnik, E.M. Haller-Schober, C. Faschinger, H. Aspock, Viability of *Acanthamoeba* after exposure to a multipurpose disinfecting contact lens solution and two hydrogen peroxide systems, Br. J. Ophthalmol. 86 (2002) 144–146.

[7] M. Rosin, A. Welk, O. Bernhardt, M. Ruhnau, F.A. Pitten, T. Kocher, A. Kramer, Effect of a polyhexamethylene biguanide mouthrinse on bacterial counts and plaque, J. Clin. Periodontol. 28 (2001) 1121–1126.

[8] M. Rosin, A. Welk, T. Kocher, A. Majic-Todt, A. Kramer, F.A. Pitten, The effect of a polyhexamethylene biguanide mouthrinse compared to an essential oil rinse and a chlorhexidine rinse on bacterial counts and 4-day plaque regrowth, J. Clin. Periodontol. 29 (2002) 392–399.

[9] J.D. Payne, D.W. Kudner, A durable antiodor finish for cotton textiles, Text. Chem. Color. 28 (1996) 28–30.

[10] A. Cazzaniga, V. Serralta, S. Davis, R. Orr, W. Eaglstein, P.M. Mertz, The effect of an antimicrobial gauze dressing impregnated with 0.2-percent polyhexamethylene biguanide as a barrier to prevent *Pseudomonas aeruginosa* wound invasion, Wounds 14 (2002) 169–176.

[11] N.A. Cox, J.S. Bailey, M.E. Berrang, Bactericidal treatment of hatching eggs I. Chemical immersion treatments and *Salmonella*, J. Appl. Poult. Res. 7 (1998) 347–350.

[12] N.A. Cox, M.E. Berrang, R.J. Buhr, J.S. Bailey, Bactericidal treatment of hatching eggs II. Use of chemical disinfectants with vacuum to reduce *Salmonella*, J. Appl. Poult. Res. 8 (1999) 321–326.

[13] W. Khunkitti, A.C. Hann, D. Lloyd, J.R. Furr, A.D. Russell, Biguanide-induced changes in *Acanthamoeba castellanii*: an electron microscopic study, J. Appl. Microbiol. 84 (1998) 53–62.

[14] W. Khunkitti, A.C. Hann, D. Lloyd, J.R. Furr, A.D. Russell, X-ray microanalysis of chlorine and phosphorus content in biguanide-treated *Acanthamoeba castellanii*, J. Appl. Microbiol. 86 (1999) 453–459.

[15] H.M. Wallace, Polyamines and their role in human disease—an introduction, Biochem. Soc. Trans. 31 (2003) 354–355.

[16] S.S. Cohen, A Guide to Polyamines, Oxford University Press, Oxford, 1998, pp. 185–230.

[17] V. Vijayanathan, T. Thomas, T.J. Thomas, DNA nanoparticles and development of DNA delivery vehicles for gene therapy, Biochemistry 41 (2002) 14085–14094.

[18] F.R. Blattner, G. Plunkett 3rd, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, N.W. Davis, H.A. Kirkpatrick, M.A. Goeden, D.J. Rose, B. Mau, Y. Shao, The complete genome sequence of *Escherichia coli* K-12, Science 277 (1997) 1453–1474.

[19] W.B. Dandliker, H. Mao-Lin, J. Levin, B.R. Rao, Equilibrium and kinetic inhibition assays based upon fluorescence polarization, in: J.L. Langone, H. Van Vunakis (Eds.), Methods in Enzymology: Immunochemical Techniques Part C, Academic Press, New York, 1981, pp. 3–28.

[20] C.H. Lee, Y.H. Ni, C.C. Chen, C.K. Chou, F.H. Chang, Synergistic effect of polyethylenimine and cationic liposomes in nucleic acid delivery to human cancer cells, Biochim. Biophys. Acta 1611 (2003) 55–62.

[21] A. Cornish-Bowden, Fundamentals of Enzyme Kinetics, second ed., Portland Press, London, 1995.

[22] R.V. Gessner, C.A. Frederick, G.J. Quigley, A. Rich, A.H.-J. Wang, The molecular structure of the left-handed Z-DNA double helix at 1.0 Å resolution, J. Biol. Chem. 264 (1989) 7921–7935.

[23] M. Egli, L.D. Williams, Q. Gao, A. Rich, Structure of the pure-spermine form of Z-DNA (magnesium free) at 1-Å resolution, Biochemistry 30 (1991) 11388–11402.

[24] D. Bancroft, L.D. Williams, A. Rich, M. Egli, The low temperature crystal structure of the pure-spermine form of Z-DNA reveals binding of a spermine molecule in a minor groove, Biochemistry 33 (1994) 1073–1086.

[25] H. Ohishi, I. Nakanishi, K. Inubushi, G. van der Marel, J.H. van Boom, A. Rich, A.H.J. Wang, T. Hakoshima, K. Tomita, Interaction between the left-handed Z-DNA and polyamine. 2. The crystal structure of the d(CG)(3) and spermidine complex, FEBS Lett. 391 (1996) 153–156.

[26] H. Ohishi, K. Suzuki, M. Ohtsuchi, T. Hakoshima, A. Rich, The crystal structure of *N*-1-[2-(2-amino-ethylamino)-ethyl]-ethane-1,2-diamine (polyamines) binding to the minor groove of $d(CGCGCG)_2$, hexamer at room temperature, FEBS Lett. 523 (2002) 29–34.

[27] H. Ohishi, N. Terasoma, I. Nakanishi, G. van der Marel, J.H. van Boom, A. Rich, A.H.J. Wang, T. Hakoshima, K. Tomita, Interaction between left-handed Z-DNA and polyamine. 3. The crystal structure of the $d(CG)_{(3)}$ and thermospermine complex, FEBS Lett. 398 (1996) 291–296.

[28] P. Broxton, P.M. Woodcock, P. Gilbert, A study of the antibacterial activity of some polyhexamethylene biguanides towards *Escherichia coli* ATCC 8739, J. Appl. Bacteriol. 54 (1983) 345–353.

[29] P. Broxton, P.M. Woodcock, P. Gilbert, Interaction of some polyhexamethylene biguanides and membrane phospholipids in *Escherichia coli*, J. Appl. Bacteriol. 57 (1984) 115–124.

[30] T. Ikeda, S. Tazuke, C.H. Bamford, A. Ledwith, Interaction of polymeric biguanide biocide with phospholipid membranes, Biochim. Biophys. Acta 54 (1984) 796–799.

[31] P. Gilbert, D. Pemberton, D.E. Wilkinson, Synergism within polyhexamethylene biguanide biocide formulations, J. Appl. Bacteriol. 69 (1990) 593–598.