GENOME SEQUENCING AND FUNCTIONAL GENOMICS IN *Bacillus subtilis*: THE EXPRESSION AND ROLES OF THE SUPEROXIDE DISMUTASES AND NADH DEHYDROGENASES

A Thesis Submitted to the University of Dublin for the Degree of Doctor of Philosophy

By

Elizabeth Scanlan

Department of Genetics

University of Dublin

Trinity College

Dublin 2

May 2002
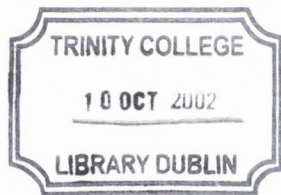
DECLARATION

This thesis has not been previously submitted to this or any other university for examination for a higher degree. The work presented here is entirely my own except where otherwise acknowledged. This thesis may be made available for consultation within the university library. It may be copied or lent to other libraries for purposes of consultation.

Elizabeth Scanlan

May, 2002

Acknowledgements

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

SUMMARY

The work in this thesis spans genome sequencing of *Bacillus subtilis*, functional analysis of genes of unknown function and more detailed characterisation of the function and regulation of specific genes. Approximately 0.7% of the *B. subtilis* genome was sequenced as part of the work for this thesis. The completed genome sequence was published in Nature in 1997 (Kunst *et al.*, 1997). 40% of the predicted open reading frames encoded genes of unknown function and the *B. subtilis* functional analysis (BFA) project therefore began to gather expression and phenotypic data on 1,275 strains that contained pMUTIN4 insertions into such genes. The growth and expression profiles of 13 genes of unknown function in the region sequenced were determined during the BFA project as part of the work for this thesis. In our laboratory 1,148 strains from the BFA collection were tested for their responses to the oxidising agents H_2O_2 and paraquat. Further characterisation was carried out on 2 of these strains that had increased resistance to the superoxide radical generating agent paraquat: *yjlC* and *yjlD* mutants. The *yjlD* paralogues *yutJ* and *yumB* were also investigated. In addition, genes encoding proteins involved in detoxification of superoxide radicals: *sodA, sodF* and *yojM* were studied. The genes chosen for investigation are all involved in the oxidative stress status of the *B. subtilis* cell. The *yjlD* gene has two paralogues, *yutJ* and *yumB* and they encode putative type 2 NADH dehydrogenases (NDH-2). These enzymes have been shown to generate intracellular superoxide radicals and H_2O_2 during aerobic growth in *Escherichia coli*. Null mutants in both *yjlC* and *yjlD* are more resistant to paraquat than the wild type strain on solid medium. These mutants also grow slower than the wild type. It is suggested that this phenotype reflects the ease with which the flavin moiety of the *yjlD*- encoded NDH-2 donates electrons to oxygen via the redox cycling compound, paraquat. The *yjlD* gene was found to comprise an operon with *yjlC*. Expression profiles showed the operon was highly expressed during exponential growth. Expression initiated from a σ^A-type promoter and the transcript included a 257 base leader sequence before the start codon of *yjlC*. In *yjlC* and *yjlD* null backgrounds, expression from the *yjlCD* promoter is higher than for the wild type background, implying that the *yjlCD* gene products directly or indirectly regulate their own expression. This negative autoregulation effect occurs at the transcriptional level. On NADH dehydrogenase activity gels a *yjlCD* specific band of activity is present in wild type protein extract but disappears in a *yjlCD* null background. This indicates that the product of this operon has NADH dehydrogenase activity. The function of *yjlC* is unknown and it has no significant similarities to any genes in the databases. Growth

experiments in which expression of *yjlC* and *yjlD* could be switched on or off by inducible promoters showed that the expression of both genes, but not either gene alone, could partially restore the growth defect of a *yjlCD* mutant strain. This implies that expression of *yjlC* is crucial for the function of YjlD. Expression profiles of *yutJ* and *yumB* showed *yutJ* is expressed early in the growth curve and *yumB* is expressed during sporulation. The *yutJ* transcript initiates from a σ^A-type promoter and the *yumB* transcript initiates from both σ^{F-} and σ^E-type promoters, showing that it is expressed in both the mother cell and the forespore.

The genome encodes three superoxide dismutase genes: *sodA*, *sodF* and *yojM*. The *sodA* gene is monocistronic. It is highly expressed during exponential growth from a σ^A-type promoter and transcription is induced from this promoter after paraquat addition. SodA is also part of the general stress response and is induced by ethanol from a σ^B promoter. Expression from the *sodA* promoter is affected by null mutations in the upstream gene *yqgC* and the iron homeostasis regulator *fur*. In contrast, expression of *sodF* and *yojM* is very low under all conditions tested and even in the absence of *sodA*. Activity of SodA, but not SodF or YojM is visible on non denaturing PAGE gels stained for superoxide dismutase activity. Results obtained in this thesis imply that the *yjlCD* operon encodes the principal respiratory NADH dehydrogenase and that this enzyme is a source of oxidative stress during aerobic growth. SodA is the principal superoxide dismutase that is involved in both superoxide and general stress responses.

PUBLICATIONS ARISING FROM WORK DURING THIS THESIS

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Danchin, A., and et al. (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249-256.

Scanlan, E., Hecker, M., and Devine, K.M. (2001) The oxidative stress response. In *Functional Analysis of Bacterial Genes; A Practical Manual*. Schumann, W., Ehrlich, S.D. and Ogasawara, N. (eds). Chichester, UK: Wiley, pp. 215-226.

Scanlan, E., O'Reilly, M., Andersen, K.K., Noone, D., and Devine, K.M. (2001) Screening for increased sensitivity or resistance and level of *lacZ* reporter gene expression in response to the oxidising agents hydrogen peroxide and methyl viologen. In *Functional Analysis of Bacterial Genes; A Practical Manual*. Schumann, W., Ehrlich, S.D. and Ogasawara, N. (eds). Chichester, UK: Wiley, pp. 227-230.

Scanlan, E., Karamata, D., and Devine, K.M. The type-2 NADH dehydrogenases of *Bacillus subtilis*. Manuscript in preparation.

Scanlan, E., and Devine, K.M. The *sodA* encoded superoxide dimutase of *B. subtilis* is a member of the general stress and superoxide stress responses. Manuscript in preparation.

ABBREVIATIONS

aa	Amino acids
Acn	Aconitase
ATP	Adenosine triphosphate
BFA	*Bacillus subtilis* Functional Analysis Project
BLAST	Basic local alignment search tool
BLM	Basic limitation medium
CHP	Cumene hydroperoxide
CRE	Catabolite- responsive element
Cu,ZnSOD	Copper, zinc superoxide dismutase
ddNTP	Dideoxynucleoside triphosphate
DEPC	Diethyl pyrocarbonate
dNTP	Deoxynucleoside triphosphate
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
FAD	Flavin adenine dinucleotide
FeSOD	Iron superoxide dismutase
GCG	Genetics Computer Group
GRAS	Generally regarded as safe
GSNO	S-nitrosoglutathione
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	Kilobase pairs
LacZ	β-galactosidase
LB	Luria Bertani
MCP	Methyl accepting chemotaxis proteins
MCS	Multiple cloning site
mJ	Millijoules
MnSOD	Manganese superoxide dismutase
MOPS	3-Morpholinopropanesulfonic acid
NAD	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NADP	Nicotinamide adenine dinucleotide phosphate (oxidised)
NADPH	Nicotinamide adinucleotide phosphate (reduced)
NDH	NADH dehydrogenase

ABBREVIATIONS

OD	Optical density
ONPG	o-Nitrophenyl β-D-galactopyranoside
ORF	Open reading frame
PCR	polymerase chain reaction
PQ	Paraquat
RBS	Ribosome binding site
ROS	Reactive oxygen species
Rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SIN-1	3-Morpholinosydnonimine
SiRase	Sulfite reductase
SM	Schaeffer's sporulation medium
SNAP	5-Nitroso-N-acetylpenicillamine
SOD	Superoxide dismutase
tBOOH	*tert*-butyl hydroperoxide
Temed	N,N,N',N'-tetramethylethylenediamine
XBAP	X11 version, Sequence assembly program (Staden package)
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
XNIP	X11 version, Nucleotide interpretation program (Staden package)

CHAPTER 1

INTRODUCTION

1.1 GENOME SEQUENCING AND FUNCTIONAL ANALYSIS

The *Bacillus subtilis* genome is 4,214,810 bases in length and contains 4,100 genes (Kunst *et al.*, 1997). It was the seventh eubacterial genome sequence to be published. Today it is one of 60 eubacterial species whose genome sequences have been completely determined and are available in GenBank. Genetic and physical maps of the *B. subtilis* genome were available when the genome sequencing project was initiated in what became a joint effort between European and Japanese research groups. The sequencing strategy employed involved assigning regions between loci of known sequence to each of the groups. Each group cloned and sequenced the DNA within their region. In contrast, some other sequencing projects including that of *Haemophilus influenzae* used a shotgun approach, successfully assembling the genome sequence from sequences of randomly generated genomic subclones (Fleischmann *et al.*, 1995). This approach was also used in the biggest sequencing project of all, that of the human genome. Celera Genomics used a whole genome shotgun approach and the International Human Genome Sequencing Consortium used a hierarchical shotgun strategy where large DNA fragments of known position were subjected to shotgun sequencing. Celera Genomics and the public international consortium jointly announced the completion of a 'working draft' of the human genome sequence on June 26th 2000.

The completion of the *B. subtilis* genome sequence revealed that approximately 40% of the predicted genes were of unknown function. The *B. subtilis* Functional Analysis Project involving 18 European and 12 Japanese laboratories undertook mutant construction and phenotypic analysis of mutant strains to provide clues about the role of these genes in the cell. The European consortium generated 1,275 strains containing pMUTIN4 insertions in genes of unknown function. As part of the primary characterisation of these genes, LacZ assays and phenotypic analysis in response to a broad range of stimuli were carried out. The work in this thesis spans genome sequencing and subsequent efforts to characterise genes of unknown function. Over 29 kilobases of the *B. subtilis* genome was sequenced. Genes of unknown function involved in bioenergetics (*yjlCD, yumB, yutJ*) and detoxification (*yojM, sodF*), as well as a gene whose function was already known (*sodA*) were investigated.

1.2 OXIDATIVE STRESS: FORMATION AND TOXICITY OF REACTIVE OXYGEN SPECIES (ROS)

Aerobic growth is energetically more advantageous than anaerobic growth but these energetic benefits do not come without a price. Incomplete reduction of molecular oxygen leads to the generation of highly toxic reactive oxygen species which can rapidly disrupt the integrity of the cell, causing damage to membranes, proteins and DNA. Cells capable of aerobic respiration have evolved numerous mechanisms designed to protect against ROS. These protective mechanisms are designed to cope with the basal level of ROS generated during normal aerobic metabolism and with surges of ROS generated by conditions such as radiation, the presence of redox cycling drugs and macrophage activation. Cells therefore have a constitutive ability to deal with ROS and an inducible capability to deal with surges of ROS. This latter phenomenon is called the oxidative stress response.

1.2.1 The superoxide radical

During respiration, four electrons are passed through the intermediates of the respiratory chain and are finally transferred to molecular oxygen to yield water (Reaction 1).

$$4H^+ + 4e^- + O_2 \longrightarrow 2H_2O \qquad \text{(Rxn 1)}$$

However in some instances molecular oxygen can be incompletely reduced resulting in the formation of potentially damaging intermediates. The superoxide anion is formed by the single electron reduction of molecular oxygen. This reaction is catalysed by transition metals with unpaired electrons such as iron (Reaction 2).

$$Fe^{2+} + O_2 \longrightarrow Fe^{3+} + O_2^{\cdot-} \qquad \text{(Rxn 2)}$$
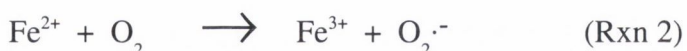
Superoxide radical formation is also propagated by redox cycling compounds such as paraquat (PQ) and quinones. When taken up into the cell, these compounds can become reduced by a single electron. The reduced molecule then transfers its electron to oxygen to form the superoxide radical (Reaction 3).

$$PQ \longrightarrow PQ^-$$

$$PQ^- + O_2 \longrightarrow PQ + O_2^{\cdot-} \qquad \text{(Rxn 3)}$$
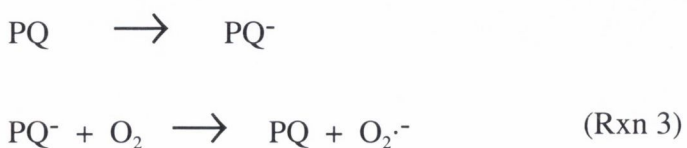
In *Escherichia coli* generation of increased intracellular levels of superoxide radicals, either through hyperbaric oxygen, mutation of superoxide dismutases or treatment with quinones, leads to auxotrophy for branched chain, aromatic and sulfur amino acids and

hypermutability of DNA (Farr *et al.*, 1986). Initially these phenotypes were something of a mystery since superoxide radicals do not react in aqueous solution with bases, amino acids or sugars. It has now been established that the toxicity of superoxide radicals is closely linked with iron metabolism.

$$[4Fe-4S]^{2+} + O_2\cdot^- + 2H^+ \longrightarrow [4Fe-4S]^{3+} + H_2O_2 \quad (\text{Rxn 4})$$

$$[4Fe-4S]^{3+} \longrightarrow [3Fe-4S]^+ + Fe^{2+} \quad (\text{Rxn 5})$$

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH^- + OH\cdot \quad (\text{Rxn 6})$$
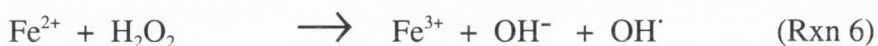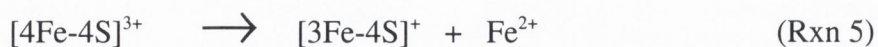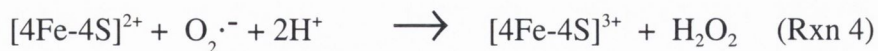
The superoxide radical oxidises the $[4Fe-4S]^{2+}$ clusters of enzymes such as dihydroxyacid dehydratases and aconitases (Flint *et al.*, 1993; Gardner and Fridovich, 1991) (Reaction 4). The resultant oxidised Fe-S cluster is unstable and releases a ferrous ion into the cytosol of the cell (Reaction 5). This leads to inactivation of enzymatic activity and accounts for the branched chain amino acid auxotrophy. The released ferrous ion can generate toxic hydroxyl radicals through the Fenton reaction (Reaction 6) and thereby oxidise DNA, accounting for the hypermutability phenotype (Keyer and Imlay, 1996). The cellular defence against superoxide radicals is mediated by superoxide dismutases (SODs) that detoxify the radicals as shown in Reaction 7 via reduction and oxidation of their active site metals (Figure 1.1).

$$O_2\cdot^- + O_2\cdot^- + 2H^+ \longrightarrow H_2O_2 + O_2 \quad (\text{Rxn 7})$$
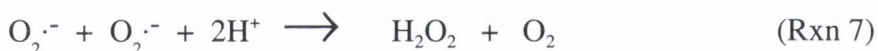
1.2.2 Hydrogen peroxide

Hydrogen peroxide is generated by the two electron reduction of oxygen. It can also be formed by the dismutation reaction of the superoxide anion by superoxide dismutase (Reaction 7). Hydrogen peroxide diffuses easily across the bacterial cell membrane in contrast to the superoxide anion. Hydrogen peroxide can then react with ferrous iron in the cell leading to the formation of hydroxyl radicals via the Fenton reaction (Reaction 6). The extent of this reaction depends on the availability of the free transition metals that catalyse the reaction. Hydrogen peroxide is detoxified by catalases which dismute two molecules of H_2O_2 using one as a reductant and one as an oxidant (Reaction 8), or peroxidases which use H_2O_2 as an oxidant and another substrate as a reductant (AH_2, Reaction 9).

$$2H_2O_2 \longrightarrow O_2 + 2H_2O \quad (\text{Rxn 8})$$

4
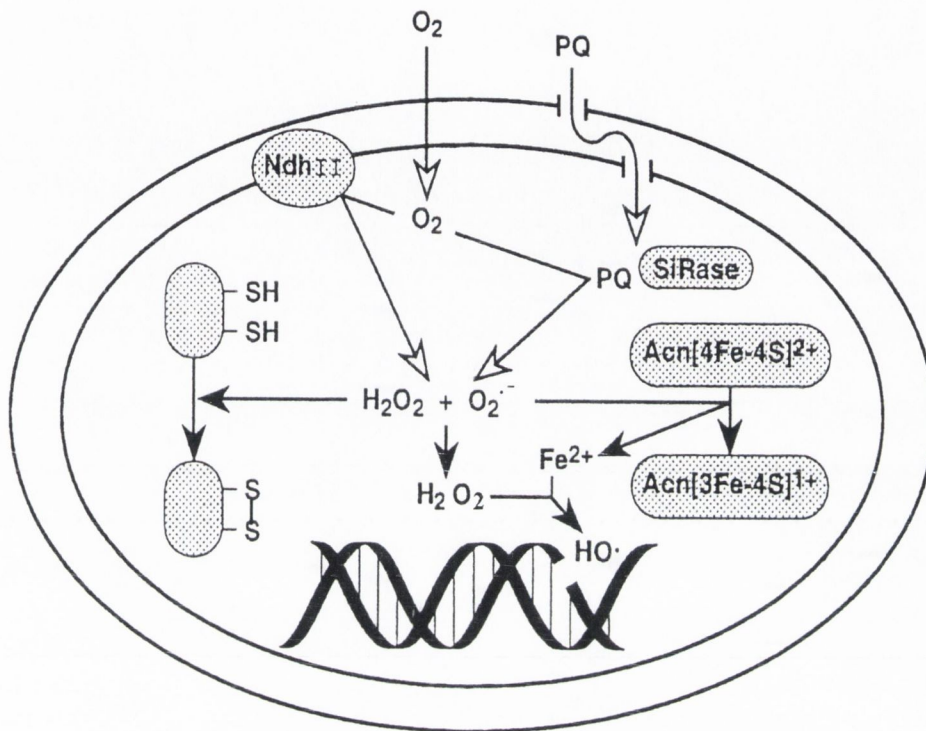
Figure 1.1 Illustration of mechanisms of cell damage by oxidative stress in *E. coli*. Under aerobic conditions oxygen diffuses into cells. The flavoproteins type 2 NADH dehydrogenase (NdhII) and sulfite reductase (SiRase) are responsible for incomplete reduction of oxygen forming superoxide and H_2O_2. Superoxide formation can be accelerated by redox cycling drugs such as paraquat (PQ). Superoxide can inactivate proteins containing vulnerable Fe-S clusters such as aconitase (Acn) releasing iron. Reduced iron can react with H_2O_2 to form the hydroxyl radical that can damage DNA. H_2O_2 also damages proteins by oxidising cysteine residues and thereby causing disulfide bridge formation. This diagram is from Storz and Imlay, 1999.

$$H_2O_2 + AH_2 \longrightarrow 2H_2O + A \qquad \text{(Rxn 9)}$$

Experiments in *E. coli* have shown that autooxidation of components of the respiratory chain, particularly NADH dehydrogenase type 2, produces superoxide and H_2O_2 in the cell. Sulfite reductase is another enzyme which autooxidises in *E. coli*. Autooxidation of exposed reduced flavin cofactors of these enzymes results in the production of superoxide and H_2O_2 (Messner and Imlay, 1999) (Figure 1.1).

1.2.3 The hydroxyl radical

The hydroxyl radical is a highly reactive and highly toxic oxidant that can react instantaneously with a wide variety of biological compounds. It is formed by the Fenton reaction (Reaction 6) (Figure 1.1). It is a very short-lived species and so there is no specific biological detoxification mechanism for this radical. Instead the cellular defence mechanisms are aimed at preventing accumulation of the superoxide radicals and peroxides that are required for generation of the hydroxyl radical. Hydroxyl radicals cause damage to DNA, proteins, lipids and carbohydrates. These radicals can attack DNA at both the sugar and base residues. Sugar residue attack leads to base loss and strand breakage as a result of sugar fragmentation. Base damage resulting from radical attack includes the formation of thymine glycol and 8-hydroxyguanine and opening of the imidazole ring of adenine and guanine (Demple and Linn, 1982; Jovanovic and Simic, 1989). Thymine glycol stops DNA replication whereas 8-hydroxyguanine results in mispairing, leading to a lesion in the daughter strand DNA (Clark and Beardsley, 1987; Shibutani *et al.*, 1991). In addition, peroxidation chain reactions initiated by hydroxyl radicals cause lipid damage. Peroxidation of membrane phospholipids alters membrane function and leads to the formation of secondary radicals and hydroperoxides (for review Riley, 1994). Hydroxyl radicals also induce protein damage resulting in fragmentation or aggregation of proteins. Protein denaturation by hydroxyl radicals results in greater susceptibility to proteolytic systems in the cell (Davies, 1987). With this repertoire of reactions that cause cellular damage, it is evident that the cell must devote considerable resources towards preventing hydroxyl radical formation.

1.2.4 Nitric oxide

Nitric oxide (NO·) and other reactive nitrogen species are cytotoxic and can act as signalling molecules and induce gene expression. NO· is produced by denitrifying bacteria during the conversion of nitrite to dinitrogen. Though *B. subtilis* is not known to be a

denitrifier it is likely to be exposed to NO˙ produced by denitrifiers in the soil and NO˙ has recently been shown to be a biologically relevant inducer of gene expression in *B. subtilis* (Nakano, 2002). In addition, recent work has shown evidence for the production of nitric oxide by a nitric oxide synthase protein of *B. subtilis*, YflM, in a manner similar to mammalian nitric oxide synthases (Adak *et al.*, 2002). Reactive nitrogen species and reactive oxygen species are also produced by activated macrophages. Reaction of $O_2\cdot^-$ with NO˙ results in the formation of peroxynitrite and consequently peroxynitrous acid (OONO˙/HOONO). These agents can readily cross membranes and may exacerbate damage caused by macrophage generated extracellular $O_2\cdot^-$ which cannot cross membranes due to its charge. Peroxynitrite oxidises iron-sulfur clusters and thiols leading to protein damage (Reaction 10).

$$[4Fe\text{-}4S]^{2+} + OONO^- + H^+ \longrightarrow [3Fe\text{-}4S]^{1+} + HO^- + NO_2\cdot + Fe^{2+} \qquad (Rxn\ 10)$$
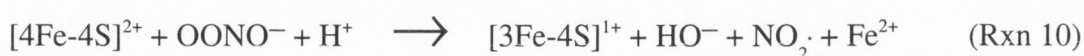
In *E. coli* the citric acid cycle enzyme aconitase has an NO˙ sensitive iron sulfur cluster (Gardner *et al.*, 1997). Oxidation of iron-sulfur clusters of sensitive dehydratases including aconitase and fumarase A also causes the release of iron into the cytoplasm and this released iron could localise to the DNA of the cell causing DNA damage (Keyer and Imlay, 1996). NO˙ can also directly nitrosylate iron-sulfur clusters causing their disruption. This type of NO˙ mediated damage is used to activate the SoxR protein of *E. coli* (Ding and Demple, 2000). Also, under aerobic conditions, NO˙ causes deamination of DNA bases and consequent DNA mutation (Wink *et al.*, 1991).

1.3 BACTERIAL DEFENCES AGAINST OXIDATIVE STRESS

Exposure of bacteria to oxidative stress induces expression of a wide variety of genes encoding proteins with very diverse functions. The coordinate regulation of this diverse group of genes is designed to cope with the many facets of cell exposure to oxidising species. These include removal and neutralisation of the ROS, protection of DNA and metal-containing enzymes, repair and/or removal of damaged DNA, proteins and lipids. The constituent genes of the oxidative stress response that encode these functions can be classified into three categories based on the type of stressor involved: (i) the superoxide stress response; (ii) the peroxide stress response and (iii) the general stress response. These responses overlap with cellular systems for protection against disulfide stress, heat shock and nitric oxide stress. These responses are described with particular emphasis on the systems of *B. subtilis* and *E. coli*.

1.3.1 The superoxide stress response
1.3.1.1 The SoxRS-MarRAB-Rob regulon of *E. coli*

The response of *E. coli* to superoxide radicals is mediated by the SoxRS regulon. SoxR and SoxS are encoded by two genes that are divergently transcribed from overlapping promoters. Both proteins bind to DNA. SoxR is a MerR-like DNA-binding protein while SoxS is an AraC-like DNA-binding protein. SoxR is a homodimer with two bound iron-sulfur clusters which can exist in both oxidised and reduced states (Ding *et al.*, 1996; Ding and Demple, 1997). SoxR can therefore function as a sensor of superoxide stress. This is illustrated in Figure 1.2. In the reduced form, SoxR binds to the intergenic DNA repressing its own synthesis. However this reduced form of the protein cannot activate transcription of its only known target gene, *soxS*. The SoxR protein is oxidised via its [2Fe-2S] clusters in response to fluxes of superoxide radicals. SoxR can also be activated by nitrosylation of its [2Fe-2S] clusters (Ding and Demple, 2000) and by any mechanism which interferes with its maintenance in the reduced state (Touati, 2000). The oxidised/activated form of SoxR activates transcription of the *soxS* gene (Ding *et al.*, 1996; Hidalgo *et al.*, 1998). SoxS in turn activates transcription of the target genes of the regulon that cope with the superoxide flux. SoxS binds to the target gene promoters and activates expression (Li and Demple, 1994). SoxS represses expression of its own promoter, a capability that probably functions to attenuate the extent of the superoxide response (Nunoshiba *et al.*, 1993). A second attenuation mechanism may function through cellular thiols such as thioredoxin and glutaredoxin that can reduce SoxR returning it to the non-active form (Ding and Demple, 1998).

The constituent genes of the SoxRS regulon are involved in a wide variety of processes that reflect the varied cellular effects of the superoxide radical. The theme of their activities is to inactivate superoxide radicals and to replenish, repair or degrade damaged molecules. There are three superoxide dismutases (SODs) in *E. coli*, each with a different metal at the active site. Only the manganese form of the enzyme, MnSod, is known to be part of the SoxRS regulon. MnSod, encoded by *sodA*, is found interacting with DNA (Steinman *et al.*, 1994) and in the cytoplasm. The *zwf* gene which encodes glucose-6-phosphate dehydrogenase is also a member of the regulon. Since this enzyme generates reducing equivalents in the form of NADPH, increased levels would function to counteract the oxidising environment and restore redox homeostasis in the cytoplasm (Salvemini *et al.*, 1999). The ferridoxin:NADPH oxidoreductase enzyme, encoded by the *fpr* gene, may function to maintain proteins containing Fe-S clusters in their reduced state. The *fumC*
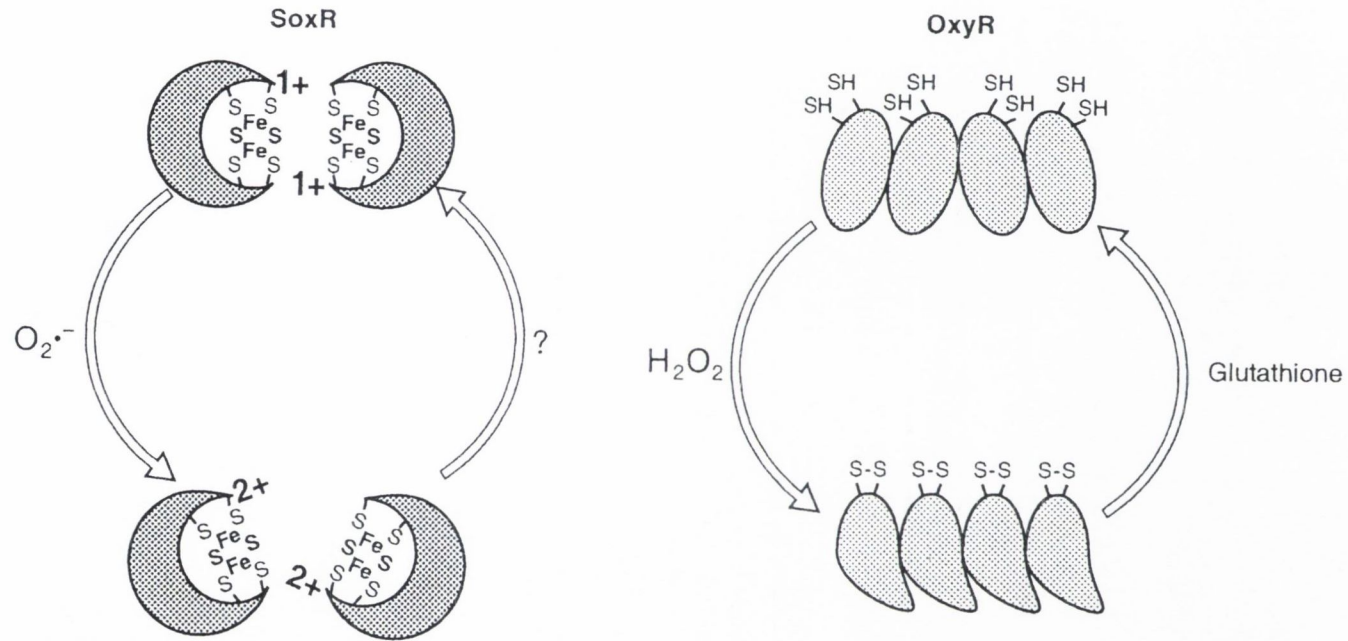
Figure 1.2 Illustration of the activation of the SoxR and OxyR regulators of *E. coli*.

SoxR becomes activated by oxidation of its Fe-S clusters by superoxide. The deactivation mechanism is not known OxyR becomes activated by oxidation of its cysteines by H_2O_2 which causes disulfide bond formation. It can be rereduced by glutathione which it a member of the OxyR regulon. This diagram is from Storz and Imlay, 1999.

gene encodes an oxidation resistant form of fumarase that may replace the existing form of the enzyme. The *acnA* gene encodes aconitase, an Fe-S containing enzyme that is susceptible to damage by superoxide. Its inclusion in the SoxRS regulon may function to replenish the normal levels of this enzyme. The *nfo* gene encodes the DNA repair enzyme endonuclease IV, that may function to repair DNA damaged by superoxide radicals. Other genes of the regulon include *micF* (antisense RNA repressor of OmpF), *fldA* and *fldB* (flavodoxins A and B), *inaA* (unknown function), *marRAB* (antibiotic resistance operon), *ribA* (GTP cyclohydrolase), *nfsA* (nitrofuran reductase), *tolC* (outer membrane porin), *acrAB* (drug efflux pump), *rimK* (modifier of 30S ribosomal subunit protein S6, lies immediately downstream of *nfsA*) and *fur* (iron uptake repressor). Transcriptional profiling, using arrays, of the *E. coli* response to superoxide stress mediated by paraquat identified 112 genes whose expression was altered in response to paraquat. This included some genes whose expression was activated by paraquat and expression of *soxS* that had not been previously identified as members of the *soxRS* regulon. These included *lpxC* (an essential cell envelope gene), *map* (methionine aminopeptidase), *nfnB* (nitrofuran reductase), *ptsG* (glucose phosphotransferase enzyme) and *ybjC* (unknown function) (Pomposiello *et al.*, 2001). In view of the large and multifunctional nature of the SoxRS regulon, it is not surprising that expression of many of the constituent genes are also under the control of alternative regulators. The *sodA* gene for example is also regulated by MarA, Fur, ArcA, Fnr and IHF (Compan *et al.*, 1993). The *fumC* gene is another example, its expression is also regulated by MarA, ArcAB and σ^s (Storz *et al.*, 1999). The regulatory network is further complicated by the fact that the regulators *fur* and *marRAB* are activated by SoxS. The transcription of *fur* and its upstream gene *fldA* (a flavodoxin) is induced by paraquat in a *soxRS* dependent manner (Zheng *et al.*, 1999). Fur is a transcriptional regulator which usually represses its target genes by binding to 'Fur boxes' in an Fe(II) dependent manner. Fur regulates genes involved in the uptake of iron into the cell, such as siderophores. As many of the toxic effects of superoxide stress are caused by iron it is logical that superoxide stressed cells induce the Fur regulator thereby repressing expression of iron uptake genes. Over expression of *soxS* was observed to cause increased *marRAB* transcription, though no reverse effect of *marA* on *soxS* transcription has been seen (Miller *et al.*, 1994). Despite this relationship, genes can be regulated by both MarA and SoxS independently of each other (Rosner and Slonczewski, 1994; Greenberg *et al.*, 1991). There is considerable overlap between the genes regulated by MarA and SoxS, and these two regulatory proteins are structurally similar. The *marA* gene encodes a transcriptional activator and is under the negative control of the repressor MarR. Induction

of the regulon occurs after inactivation of MarR upon exposure to agents such as salicylate. Genes regulated by both MarA and SoxS include *zwf, fpr, fumC, micF, nfo, inaA, sodA, acrAB, ribA, rimK* and *tolC* (Ariza *et al.*, 1994; Greenberg *et al.*, 1991; Jair *et al.*, 1995; Rosner and Slonczewski, 1994; Barbosa and Levy, 2000). A third transcriptional activator, Rob, is also known to bind to promoters of genes of the *mar* regulon. Rob is a homologue of MarA and SoxS and its over expression leads to multiple antibiotic resistance (Ariza *et al.*, 1995). It has been suggested that differential regulation by the three activators may be due to their ability to form complexes with RNA polymerase in the absence of DNA. The resulting binary complexes may scan the DNA and then bind promoters of the regulon. This mechanism of activation does not exclude an RNA polymerase capture mechanism, whereby the activator bound to the DNA recruits the RNA polymerase to initiate transcription at the promoter (Martin *et al.*, 2002).

1.3.1.2 The SODs of *E. coli*

In addition to the *sodA* encoded MnSOD, *E. coli* has two other SODs : the *sodB* encoded iron binding SOD (FeSOD) and the *sodC* encoded copper and zinc binding SOD (Cu,Zn SOD). MnSODs and FeSODs are structurally similar and are thus thought to have evolved from a common ancestral SOD, whereas the Cu,ZnSODs are distinct and postulated to have evolved later when the atmosphere was replete with oxygen and thus soluble copper(II) was available for incorporation into the metalloprotein. Zinc plays a structural role in the Cu,ZnSODs (Bannister *et al.*, 1991). The *sodA* gene is repressed by metallated Fur but, in contrast, *sodB* is unusually activated by metallated Fur. This may occur partly at the level of mRNA stability and partly by transcriptional activation (Dubrac and Touati, 2000), though these may be indirect effects (Dubrac and Touati, 2002). The histone like factors IHF and H-NS have a negative effect on *sodB* expression. The IHF effect is seen in the presence and absence of Fur, but the H-NS effect is only evident in the absence of Fur (Dubrac and Touati, 2000). The MnSOD and FeSOD are differently located within the cell, probably reflecting distinct functions. The MnSOD associates with the DNA and the FeSOD is found at highest concentrations close to the inner membrane (Steinman *et al.*, 1994). Cu,ZnSODs have so far been isolated in eukaryotes and localised to the periplasms of Gram negative bacteria. The *E. coli* Cu,ZnSOD is periplasmically located and presumably functions to protect unknown sensitive targets of this compartment from exogenous superoxide possibly originating from activated macrophages, some other external source, or from superoxide generated within the periplasm (Benov and Fridovich,

1994; Imlay and Imlay, 1996). The *sodC* gene is anaerobically repressed by Fnr and induced in stationary phase by RpoS (Gort *et al.*, 1999).

1.3.1.3 The superoxide response of *B. subtilis*

In contrast to *E. coli* the cellular response to superoxide radicals is less well characterised in *B. subtilis*. A single SOD, designated SodA, has been detected in vegetative cells and spores. SodA is presumed to be a manganese binding enzyme, based on amino acid similarities, conservation of metal binding sites and the observation of increased activity with manganese supplementation of the medium (Inaoka *et al.*, 1998). SodA has been shown by 2D protein gel analysis to be paraquat inducible (Antelmann *et al.*, 1997b). A *sodA* mutant is sensitive to paraquat during normal growth, but this mutation does not affect the resistance of spores to oxidants (Casillas-Martinez and Setlow, 1997). Interestingly, SodA appears to be involved in spore coat formation: in *sodA* mutants, the inner layer of the spore coat is incomplete and the outer layer lacks its usual striated appearance. However spores of *sodA* mutants have normal heat and lysozyme resistance (Henriques *et al.*, 1998). There are two further SOD homologues in *B. subtilis*: SodF is similar to Fe/MnSODs and YojM is similar to CuZnSODs. 2D protein gel analysis has also shown that IlvC, LeuC and CysK are superoxide inducible (Antelmann *et al.*, 1997b). CysK is an enzyme of the cysteine biosynthetic pathway and IlvC and LeuC are involved in branched chain amino acid synthesis. LeuC has a 4Fe-4S cluster and its induction may relate to inactivation of this cluster by superoxide anions. Interestingly, it has been observed that manganese plays an important role in protection of growing *B. subtilis* cells against superoxide radicals. Supplementation of the medium with Mn(II) reverses some of the deleterious effects of a *sodA* mutation. It is not clear whether this effect is mediated by Mn(II) alone or a complex incorporating Mn(II) (Inaoka *et al.*, 1999). The bacterium *Lactobacillus plantarum*, which lacks SOD, accumulates intracellular Mn(II) to a high level and uses the Mn(II) ion itself to catalyze the dismutation of the superoxide radical albeit less efficiently than MnSOD (Archibald and Fridovich, 1981). An analogous Mn(II) catalysed non-enzymic superoxide dismutation mechanism has been suggested for *Neisseria gonorrhoeae* (Tseng *et al.*, 2001) indicating that Mn(II) provides a biologically significant defence against superoxide radicals. If this is the case for *B. subtilis* the superoxide response may be linked to MntR, the central regulator of manganese homeostasis. MntR is sensitive to manganese levels in the cell and can act as an activator or repressor of manganese transport genes (Que and Helmann, 2000).

10

1.3.2 The peroxide stress response

There are regulons in both *B. subtilis* and *E. coli* that respond specifically to increased peroxide fluxes within the cell. In *E. coli*, the peroxide inducible regulon is controlled by OxyR. There is an analogous system in *B. subtilis* that is controlled by PerR. However the redox sensing mechanism and the mode of regulation effected by OxyR and PerR are quite distinct. These regulons are described in the following sections.

1.3.2.1 The OxyR regulon of *E. coli*

OxyR is a LysR type regulator. Characteristically for this family of regulators, *oxyR* is divergently transcribed from a gene whose expression it regulates, namely *oxyS*. Upon redox activation, the OxyR protein stimulates transcription of *oxyS* together with the other genes of the regulon. Since neither transcription nor translation of *oxyR* is significantly induced by increased levels of H_2O_2, a post-translational mode of activation was inferred (Storz *et al.*, 1990). OxyR has the potential to exist in both reduced and oxidised forms. In the normal reducing environment of the cell, OxyR exists in a reduced state that is not transcriptionally active. Perturbation of the redox homeostasis by oxidative stress leads to the formation of an intramolecular disulfide bond that converts OxyR into a form capable of transcriptional activation (Zheng *et al.*, 1998). This is illustrated in Figure 1.2. The activated form of OxyR stimulates expression of many genes including those encoding glutaredoxin 1 (*grxA*) and glutathione reductase (*gorA*). This helps restore the normal reducing environment of the cell thereby providing a form of negative feedback regulation (Zheng *et al.*, 1998). When normal redox homeostasis is established, OxyR returns to the fully reduced and transcriptionally inactive form.

1.3.2.2 The PerR regulon of *B. subtilis*

In contrast to *E. coli,* the regulator of the peroxide stress response of *B. subtilis*, PerR, is a repressor. The possibility of the existence of a peroxide stress regulon was first recognised when a derepressed mutant was isolated by growth of *B. subtilis* in the presence of increasing levels of H_2O_2 (Hartford and Dowds, 1994). Expression of the genes encoding catalase, alkylhydroperoxide reductase and MrgA were found to be inducible by H_2O_2 and in stationary phase by divalent metal ion limitation (Chen *et al.*, 1995). The *perR* gene was identified as a locus that is required for the H_2O_2 and metal ion inducibility (Bsat *et al.*, 1998). PerR is homologous to Fur, a metalloregulatory protein whose structure consists of a DNA binding domain and a metal binding domain. PerR contains zinc and has a second metal binding site which preferably contains either iron or manganese forming PerR:Zn,Fe

or PerR:Zn,Mn (Herbig and Helmann, 2001). When the second metal binding site contains bound iron or manganese PerR is activated, binds DNA (at 'Per boxes') and thus represses genes of the PerR regulon. The iron-containing form of PerR is very sensitive to oxidation by H_2O_2 and the manganese-containing form is quite insensitive to H_2O_2. This fact explains the prevention of post exponential induction of the Per regulon that was observed in the presence of Mn(II) (Chen *et al.*, 1995). In contrast, sensitivity of PerR:Zn,Fe to H_2O_2 leads to inactivation of the protein, relieves repression via DNA binding and results in induction of the PerR regulon genes (Herbig and Helmann, 2001). The mechanism of PerR:Zn,Fe sensitivity to H_2O_2 remains to be elucidated. Possible mechanisms include the oxidation of bound FeII to FeIII that may cause its release from PerR and thus derepression. Alternatively, hydroxyl radical formation resulting from bound FeII oxidation by H_2O_2 may cause oxidation of the PerR protein and the damage caused to PerR may prevent DNA binding and repression at Per boxes. A third possible mechanism is suggested by the observation that DTT can restore PerR DNA binding under some conditions. This points to a possible role of oxidation at cysteine residues. Reaction of H_2O_2 with a cysteine coordinated to Fe/Mn could cause release of the metal and disulfide bond formation (Herbig and Helmann, 2001). A similar mechanism has been described for Hsp33 in *E. coli* (Ruddock and Klappa, 1999; Jakob *et al.*, 2000).

1.3.2.3 Binding sites of PerR and OxyR

The operator sites for PerR and OxyR have been identified. The PerR box is a 14 base pair inverted repeat which is the site of PerR binding to DNA (Chen *et al.*, 1995; Herbig and Helmann, 2001). All strongly induced Per regulon genes are preceded by one or more Per boxes: the *ahpCF* operon has three; *mrgA* has one; the *hemAXCDBL* operon has two and *perR* itself has two (Herbig and Helmann, 2002) and PerR negatively regulates its own expression (Herbig and Helmann, 2001). The sequences recognised by the oxidised and reduced forms of OxyR were established by footprinting. Oxidised OxyR recognises four tetramers, ATAG, one in each of four adjacent major grooves. These contacts differ to those observed when reduced OxyR binds; in this case the protein binds to these sequences in two pairs of major grooves and the pairs of grooves are separated by one helical turn (Toledano *et al.*, 1994).

1.3.2.4 Constituent genes of the PerR and OxyR regulons

The constituent genes of the OxyR and PerR regulons identified in these two bacteria encode very similar functions. Members of the OxyR regulon include *katG* (catalase HPI),

ahpCF (alkyl hydroperoxide reductase), *dps* (DNA protection protein) and *gorA* (glutathione reductase). Genes of the PerR regulon include the equivalents *katA* (catalase), *ahpCF* (alkyl hydroperoxide reductase) and *mrgA* (a *dps* homologue). Microarray analysis of the hydrogen peroxide stimulon in *E. coli* has revealed new OxyR regulon genes including *hemH* (heme biosynthesis), *sufABCDSE* (possibly involved in Fe-S cluster repair), *yaaA, yaiA, ybjM* and *yljA* (genes of unknown function) (Zheng *et al.*, 2001). Both OxyR and PerR regulate *fur*, once again underlining the fundamental importance of regulating levels of iron in the cell. The *fur* gene of *B. subtilis* is preceded by a Per box and transcription is repressed by the manganese containing form of Per but not the iron containing form, therefore repression is not alleviated by H_2O_2 (Herbig and Helmann, 2001). This observation may indicate that the physiological role of PerR:Zn,Mn is to provide a mechanism of regulation by PerR which does not involve H_2O_2. In contrast, in *E. coli, fur* is induced by H_2O_2 in an OxyR dependent manner (Zheng *et al.*, 1999). This is easier to rationalise than the situation in *B. subtilis,* as repression of genes encoding iron uptake proteins by elevated levels of Fur under peroxide stress conditions reduces the levels of iron available for participation in the Fenton reaction with H_2O_2. In *Staphylococcus aureus* which has genes encoding Fur and PerR homologues their interaction mirrors that of *B. subtilis* as PerR represses *fur* transcription in an Mn-dependent manner and *fur* transcription is not upregulated by H_2O_2 (Horsburgh *et al.*, 2001). Clearly regulation of *fur* in *E. coli, S. aureus* and *B. subtilis* is complex and intimately linked with oxidative stress responses.

1.3.2.5 The *oxyS* gene of *E. coli*

The *oxyS* gene of *E. coli* encodes a small untranslated regulatory RNA which has been shown to control expression of eight genes (and possibly many more) of diverse function (Altuvia *et al.*, 1997). Expression of *oxyS* is induced by hydrogen peroxide in an OxyR dependent manner. One of the genes whose translation is inhibited by OxyS is *fhlA*, a transcriptional activator of formate metabolism. Antisense sequences in two loops of OxyS bind to two sites in the *fhlA* mRNA, the first overlapping the ribosome binding site and the second within the coding sequence, to form a stable kissing complex which inhibits translation of the message (Argaman and Altuvia, 2000). A subset of the regulatory effects of increased *oxyS* levels is mediated through RpoS, the alternative sigma factor that is involved in general stress and stationary phase phenomena in *E. coli*. The *oxyS* RNA inhibits translation of the *rpoS* message (Zhang *et al.*, 1998). Translation of *rpoS* mRNA depends on the protein host factor I (HF-I) which is thought to bind to the mRNA thereby

facilitating translational initiation through a conformational change (Muffler *et al.*, 1996). The *oxyS* RNA also binds to the HF-1 protein and *rpoS* repression by *oxyS* is HF-1 dependent. This suggests that *oxyS* represses RpoS synthesis by obstructing HF-1 mediated translation initiation (Zhang *et al.*, 1998). Several genes have been shown to be controlled both by OxyR and RpoS including *katG* (Ivanova *et al.*, 1994), *dps* (Altuvia *et al.*, 1994) and *gorA* (Becker-Hapak and Eisenstark, 1995). Therefore increased hydrogen peroxide levels will lead to induction of these genes through OxyR and repression of their synthesis through RpoS. The rationale for this regulatory integration of the OxyR and RpoS regulons may be, as suggested by Altuvia *et al.* (Altuvia *et al.*, 1997), that induction of oxidative stress genes by OxyR renders their induction by RpoS redundant. This is an example of integration and coordinate regulation of the constituent genes of two regulons.

1.3.2.6 The catalases and peroxidases of *E. coli* and *B. subtilis*

E. coli and *B. subtilis* have more than one catalase. Both have a single catalase whose expression is induced by H_2O_2, *katG* in *E. coli* and *katA* in *B. subtilis*. In addition, *E. coli* expresses the *katE* encoded catalase HPII in stationary phase under the control of RpoS (Tanaka *et al.*, 1997). *B. subtilis* has two additional catalases: *katB* encodes a σ^B-dependent catalase (Engelmann *et al.*, 1995) and the *katX* encoded catalase is under the control of both σ^F and σ^B. KatX is present in dormant spores and functions to protect the germinating spore against H_2O_2 (Petersohn *et al.*, 1999). Under non-stress conditions in *E. coli*, alkyl hydroperoxide reductase is the primary scavenger of endogenous H_2O_2. It was found that a *katE katG E. coli* mutant could detoxify H_2O_2 at the same efficiency as wild type, but an *ahpCF katE katG* triple mutant secreted considerable amounts of H_2O_2 and also had a growth defect (Costa Seaver *et al.*, 2001). These experiments led to the conclusion that high concentrations of H_2O_2 are more efficiently detoxified by catalases as the cell cannot provide AhpCF with its reductant, NADH, at sufficient amounts to achieve the rates of H_2O_2 detoxification of the catalases. However at low levels of H_2O_2 AhpCF is kinetically more efficient. It is suggested that H_2O_2 may be the primary substrate of *E. coli* AhpCF *in vivo* (Costa Seaver *et al.*, 2001).

The endogenous substrates of alkyl hydroperoxide reductases *in vivo* are unknown but are suggested to be nucleic acid hydroperoxides or lipid hydroperoxides produced during lipid peroixidation of unsaturated fatty acids. In *B. subtilis* organic hydroperoxides are clearly hazardous *in vivo* as in addition to the *ahpCF* genes the cell also contains two genes for organic hydroperoxide resistance *ohrA* and *ohrB*. These belong to the conserved

OsmC/Ohr family of proteins and are similar to the *E. coli* OsmC, though the OsmC protein is functionally distinct from Ohr (Atichartpongkul *et al.*, 2001). In *B. subtilis* the *ohrA* gene and the downstream gene *ohrR* are convergently transcribed. OhrR is a MarR type transcriptional repressor of *ohrA* that senses organic hydroperoxides. The *ohrA* gene is strongly induced by t-butyl hydroperoxide and cumene hydroperoxide. The *ohrB* gene is divergently transcribed from *ohrR* and is part of the σ^B regulon. It is strongly induced by ethanol and salt and only weakly induced by t-butyl hydroperoxide and cumene hydroperoxide (Volker *et al.*, 1998; Fuangthong *et al.*, 2001). A mutant in *ohrA* and not *ohrB* or *ahpC* is sensitive to cumene hydroperoxide. The double *ohrAB* mutant is more sensitive to cumene hydroperoxide than *ohrA* or *ohrB* alone and is also H_2O_2 and paraquat sensitive (Fuangthong *et al.*, 2001). The sensing mechanism of OhrR and the detoxification mechanism of OhrA and OhrB remain to be elucidated. The role of Ohr-type proteins in organic hydroperoxide detoxification was previously seen in *Xanthomonas campestris* (Mongkolsuk *et al.*, 1998) and has also been described in *Pseudomonas aeruginosa* (Ochsner *et al.*, 2001).

1.3.3 The general stress response

The general stress response in *B. subtilis* is mediated by the alternative sigma factor sigma B (σ^B) (Price, 2002). This response is so-called because it can be induced by exposure of cells to a variety of stressors including heat, ethanol, acids, salt and starvation. The *sigB* operon encodes genes involved in activation of σ^B. These proteins comprise a signal transduction cascade with two branches, one of which responds to physical stress and the other to nutritional stress. The ultimate effect of these modules is to control the level of interaction between σ^B and its cognate anti-sigma factor RsbW. The constituent genes of the σ^B regulon encode a wide diversity of proteins (called general stress proteins, GSPs) of both known and unknown function (for review see Hecker and Volker, 2001). A characteristic of these genes is that they are often expressed from a promoter additional to the σ^B promoter. Therefore they can be expressed under a variety of conditions with the σ^B promoter conferring stress-sensitive expression. There are a number of genes of the σ^B regulon that encode proteins with the capability to cope with reactive oxygen species. These genes can be induced by general stress conditions thereby providing protection against oxidative stress, though the σ^B regulon is not induced by oxidative stress. The general oxidative stress response has particularly been observed for glucose starved cells growing aerobically (Dowds, 1994). Glucose-starved cells were found to be significantly more resistant than exponentially growing cells to challenge by a lethal level of hydrogen

peroxide. Such resistance was not observed in response to amino acid or phosphate starvation. Since *sigB* mutants do not show this effect, the non-specific oxidative stress can be attributed to constituent genes of the σ^B regulon (Engelmann and Hecker, 1996). Some characterised members of the σ^B regulon are candidates for contributing to this general oxidative stress resistance for example the *katE* gene of *B. subtilis* that encodes a catalase. However the precise contribution of KatE to stationary phase-induced oxidative resistance in *B. subtilis* remains to be established since this response in glucose-starved *katE* mutants is the same as wild-type cells (Engelmann and Hecker, 1996). The Dps group of proteins bind to DNA and protect it from oxidative damage. The MrgA protein is a member of this family and its expression is responsive to hydrogen peroxide through the PerR regulon. A second member of this family was identified and called general stress protein Dps. This *dps* gene is a member of the σ^B regulon and its expression is responsive to general stress conditions but not to hydrogen peroxide (Antelmann *et al.*, 1997a). This protein does contribute to the resistance of stationary phase cells to oxidative stress since *dps* mutants fail to develop this resistance (Antelmann *et al.*, 1996).

E. coli cells also exhibit stationary phase-associated resistance to stresses including hydrogen peroxide, salt and heat. This response is mediated by RpoS, a sigma factor involved in adaptation to starvation, stress and stationary phase (Jenkins *et al.*, 1988; Jenkins *et al.*, 1990). However it has been established that RpoS is not only associated with stationary phase phenomena. It has been described how expression of *rpoS* is downregulated by *oxyS* when the peroxide regulon is activated during exponential growth. RpoS also mediates osmotic control of gene expression during exponential growth (Hengge-Aronis *et al.*, 1993). Constituent genes of the RpoS regulon relevant to oxidative stress resistance include *dps*, *katE* and *xthA* (encoding exonuclease III). In some cases however the homologous genes are regulated differently in the two bacteria. For example in *B. subtilis, mrgA* is a member of the PerR regulon and its homologue *dps* is a member of the SigB regulon. In *E. coli* the *dps* gene is under the control of both OxyR and SigS (Altuvia *et al.*, 1994). However, in common with *B. subtilis*, *dps* mutants of *E. coli* also fail to develop starvation induced resistance to hydrogen peroxide (Almiron *et al.*, 1992). This indicates that Dps plays an essential role in expression of the general oxidative stress response upon entry of cells into stationary phase, despite the apparent lack of sequence specificity with which Dps proteins bind to DNA.

1.3.4 Defences against ROS in spores of *B. subtilis*

The KatX catalase of *B. subtilis* is the only catalase present in spores (Casillas-Martinez and Setlow, 1997). However this catalase is not responsible for hydrogen peroxide resistance of the spore but rather plays an important role in hydrogen peroxide detoxification during germination (Casillas-Martinez and Setlow, 1997; Bagyan *et al.*, 1998). SodA, MrgA and alkyl hydroperoxide reductase are also not involved in spore resistance to oxidative stress (Casillas-Martinez and Setlow, 1997). Even spores which are mutated in *perR* and thus overexpressing all the genes which provide protection against oxidative stress in growing cells do not significantly differ in their oxidative stress resistance to wild type spores (Casillas-Martinez *et al.*, 2000). Instead, spore DNA is substantially protected from peroxide stress by the binding of alpha- and beta-type small acid soluble proteins and their synthesis marks one of two significant increases in hydrogen peroxide resistance during sporulation. The second increase in hydrogen peroxide resistance occurs at a later stage of sporulation coinciding with the accumulation of dipicolinic acid (DPA) (Setlow *et al.*, 1993). Other aspects of sporulation including the impermeability and dehydration of the spore and the lack of NADH/NAD(P)H are also likely to contribute to the oxidative stress resistance of the spore. Consistent with this idea, spores which lack DPA and have a higher core water content are less resistant to H_2O_2 than wild type spores (Paidhungat *et al.*, 2000).

1.3.5 The relationship between oxidative stress, disulfide stress and heat shock

During normal growth the reducing environment of the cytosol maintains cysteines in reduced (thiol) form. The thioredoxin super-family of proteins maintain this thiol–disulfide homeostasis within the cell (Aslund and Beckwith, 1999). Two systems are involved in *E. coli*: (i) thioredoxin (*trxA*) thioredoxin 2 (*trxC*) and their cognate thioredoxin reductase (*trxB*) and (ii) glutaredoxins (*grxA, grxB, grxC*), glutathiones (*gshA, gshB*) and glutathione reductase (*gorA*). The thioredoxin and glutaredoxin proteins readily donate reducing equivalents from their two conserved reduced cysteines thereby maintaining the cysteines of cytosolic proteins in a reduced form. The resultant oxidised thioredoxin is regenerated by the cognate reductases (Rietsch and Beckwith, 1998). There are nine genes (*tlp, trxA, ydbP, ydfQ, ykvV, yolI, yosR, ytpP* and *yusE*) encoding thioredoxin-like proteins and four (*trxB, ycgT, ypdA, yumC*) encoding thioredoxin reductase -like proteins in the *B. subtilis* genome (Kunst *et al.*, 1997). Interestingly *trxA*, encoding thioredoxin, is an essential gene. Expression of *trxA* is driven by σ^A and σ^B promoters, making it a member of the general stress regulon (Scharf *et al.*, 1998).

However there are instances when disulfide bond formation is necessary: for example in the cytosol, a disulfide bond is formed during the catalytic cycle of the enzyme ribonucleotide reductase. In addition many secreted proteins require disulfide bond formation for proper folding and activity. Formation of these bonds in *E. coli* is effected by a number of membrane and periplasmically located thiol:disulfide oxidoreductases (*eg.* DsbA, DsbB, DsbC and DsbD) (Rietsch and Beckwith, 1998). *B. subtilis* also contains genes for thiol:disulfide oxidoreductases: *bdbA, bdbB, bdbC* and *bdbD* and mutation in either *bdbC* or *bdbD* result in loss of competence for DNA uptake indicating that formation of a disulfide bond by these enzymes is essential for competence development (Bolhuis *et al.*, 1999; Meima *et al.*, 2002).

The redox homeostasis of the cell can be perturbed by cellular exposure to oxidising agents. Evidence exists that these conditions are conducive to cytosolic disulfide formation. Alkaline phosphatase, an enzyme that requires disulfide bond formation for activity, is normally located in the periplasm. Removal of the signal sequence results in a cytoplasmically located inactive enzyme which becomes activated under oxidative stress conditions (Prinz *et al.*, 1997). This suggests that aberrant disulfide bond formation may occur in other cytoplasmically located proteins leading to misfolding and/or loss of activity. However, in *E. coli,* the increased intracellular oxidising environment is sensed by two proteins that are normally inactive but become activated upon disulfide bond formation. The activities of these proteins, OxyR and Hsp33, assist in counteracting the effects of the oxidising environment and function to restore normal redox homeostasis. As already outlined, OxyR is inactive in the fully reduced form but becomes active as a transcriptional activator upon disulfide bond formation as a result of peroxide stress. OxyR stimulates expression of many genes including *grxA* (glutaredoxin 1), *gorA* (glutathione reductase) and *trxC* (thioredoxin 2) (Ritz *et al.*, 2000). These gene products help to restore the normal redox homeostasis of the cell and their induction shows that disulfide stress and oxidative stress are linked. They also function to attenuate expression of the OxyR regulon by returning OxyR to the fully reduced (and transcriptionally inactive) form. Hsp33 is a redox regulated chaperone. Its expression is induced by heat shock and its chaperone function is induced by oxidative stress via disulfide bond formation. When normal redox homeostasis is restored Hsp33 becomes reduced and is inactive as a chaperone. The active Hsp33 chaperone binds stably to folding intermediates preventing aggregation. It directly contributes to cellular protection against oxidative stress and protects proteins from oxidative stress *in vitro* (Jakob *et al.*, 1999). Zinc binding

has been shown to be essential for the Hsp33 redox switch. The reduced form of Hsp33 coordinates a zinc atom with its four conserved cysteines. Upon oxidation the cysteines form two intramolecular disulfide bonds and the zinc is released (Jakob *et al.*, 2000).

It has been firmly established in *E. coli* and in *Saccharomyces cerevisiae* that oxidative stress is a major contributor to the lethality caused by heat stress (Benov and Fridovich, 1995; Dukan and Nystrom, 1998; Davidson *et al.*, 1996; Davidson and Schiestl, 2001). For example, yeast cells grown anaerobically are up to 20,000 times more resistant to thermal stress than cells grown aerobically, an effect that is not mediated by increased levels of heat shock proteins (Davidson *et al.*, 1996). In addition, *E. coli* SOD null mutants grown aerobically are more sensitive to thermal stress than the isogenic parent. However the thermotolerance of the SOD null mutant and the isogenic parent are identical when both are grown anaerobically (Benov and Fridovich, 1995). It is evident that the heat and oxidative stresses both affect cytosolic redox homeostasis with consequent implications for disulfide bond formation. Therefore aspects of heat, oxidative and disulphide stresses may be unified by their effects on cellular redox homeostasis.

As already stated there are two systems for maintenance of the reducing environment of the *E. coli* cytoplasm: the thioredoxin system and the glutaredoxin system. A *trxB, gorA* double mutant which lacks both systems grows very slowly under aerobic conditions and accumulates suppressors at a high rate (Bessette *et al.*, 1999). The suppressor mutation was linked to the *ahpC* locus by transposon insertions. The mutation was found to be an expansion of a triplet nucleotide repeat sequence by a single repeat. Insertion of the bases TCT at this locus which in the wild type contains four TCT repeats results in the addition of a phenylalanine to the AhpC protein. This has a dramatic effect on protein activity. It results in the conversion of AhpC from a peroxidase to a disulfide reductase (Ritz *et al.*, 2001). The mutant form of AhpC (denoted AhpC*) accepts electrons from AhpF like the wild type enzyme, and appears to direct electrons to the glutaredoxin pathway thereby changing the source of electrons of this pathway from NADPH to NADH. The triplet repeat also occurs in the *ahpC* genes of *Salmonella typhi*, *Klebsiella pneumoniae* and *Shigella flexneri* underlining the evolutionary advantage it is likely to confer (Ritz *et al.*, 2001).

1.3.6 The nitric oxide response

Nitric oxide has been recognised as an important signalling molecule. Nitric oxide produced by mammalian endothelial cells on the inside walls of blood vessels activates guanylate cyclase to release cyclic GMP. Cyclic GMP phosphorylates myosin light chains causing muscle relaxation and vasodilation (Palmer *et al.*, 1987; Ignarro *et al.*, 1987; Furchgott and Jothianandan, 1991). The identification that NO· mediated this effect came several years after the identification of an endothelium derived relaxation factor (EDRF) (Furchgott and Zawadzki, 1980). It was initially surprising that such a reactive and short lived species could play such an important physiological role. NO· is now recognised as a significant signalling molecule and has also been identified as a neurotransmitter (O'Dell *et al.*, 1991). In addition, the cytotoxic effects of NO· are also harnessed by activated macrophages which release large amounts of reactive species, including NO·, to defend against invading microbes. Another source of NO· is denitrifying bacteria that reduce nitrate to dinitrogen via nitrite, NO· and nitrous oxide and thereby produce NO· during metabolism. *B. subtilis* does not denitrify but respires anaerobically by nitrate ammonification, where nitrate is converted to ammonia via nitrite (Hoffmann *et al.*, 1998). Nevertheless reactive nitrogen intermediates are abundant in the soil, the natural environment of *B. subtilis*. An endogenous source of NO· may also be significant for *B. subtilis*. It has been reported that a *B. subtilis* nitric oxide synthatase, YflM, can generate NO·, though the regulation and role of this enzyme *in vivo* remains to be elucidated (Adak *et al.*, 2002).

Flavohemoglobins play a significant role in the bacterial response to NO·. These proteins are well conserved among distantly related organisms including animals, plants and bacteria suggesting the fundamental importance of NO· detoxification. Flavohemoglobins are two domain proteins, consisting of a hemoglobin domain and a domain for binding NAD(P)H and FAD. Like the globins of higher organisms, flavohemoglobins bind O_2. They are involved in the nitrosative stress response of *E. coli, Salmonella typhimurium, Alcaligenes eutrophus, Mycobacterium tuberculosis* and *B. subtilis* (Gardner *et al.*, 1998; Crawford and Goldberg, 1998; Hu *et al.*, 1999; Nakano, 2002). The function of these proteins in NO· detoxification became clear when a nitric oxide dioxygenase activity was demonstrated for flavohemoglobin in *E. coli*. The flavohemoglobin catalyzes the conversion of NO· to either nitrate or nitrous oxide depending on whether conditions are aerobic or anaerobic (Gardner *et al.*, 1998; Gardner *et al.*, 2000). In *E. coli* the flavohemoglobin is NO· inducible and a *hmp* (gene encoding flavohemoglobin) mutant is

NO˙ sensitive and paraquat sensitive (Gardner *et al.*, 1998; Membrillo-Hernandez *et al.*, 1999). Transcription of *hmp* is induced by NO˙ in *B. subtilis*. This upregulation occurs in both a ResDE-dependent and unknown ResDE-independent manner (Nakano, 2002). The ResDE two component system is involved in aerobic and anaerobic respiration (Nakano and Zuber, 1998). Transcription of *hmp* is also induced by oxygen limitation and this induction is also ResDE-dependent (LaCelle *et al.*, 1996). In *M. tuberculosis* the transcription of *hmp* increases upon oxygen limitation and in stationary phase there is a small increase in transcript after nitrosative stress (Hu *et al.*, 1999). The expression of the *hmp* gene of *S. typhimurium* is increased by NO˙ and expression is repressed by Fur (Crawford and Goldberg, 1998). The FAD group is proposed to transfer electrons from NAD(P)H to the ferric heme group, thus reducing it from heme-Fe^{3+} to heme-Fe^{2+} which then binds O_2 forming heme-Fe^{2+}-O_2. This is probably converted to heme-Fe^{3+}- $O_2^{˙-}$ which reacts with NO˙ to produce $2NO_3^-$. The overall reaction catalysed by the flavohemoglobins is thought to be as in Reaction 11 (Gardner *et al.*, 1998; Poole and Hughes, 2000).

$$2NO˙ + 2O_2 + NAD(P)H \longrightarrow 2NO_3^- + NAD(P)^+ + H^+ \qquad \text{(Rxn 11)}$$

In addition to flavohemoglobins other gene products, many of which overlap with the oxidative stress response, protect the cell against nitrosative stress. For example the periplasmically located Cu,ZnSOD of *S. typhimurium* is important for cellular defences against activated macrophages and a *sodC* mutant has reduced survival in macrophages and attenuated virulence in mice. Inhibitors of the macrophage NADPH oxidase or NO synthase eliminate this effect. The *sodC* mutant is sensitive to extracellular superoxide and this is exacerbated in the presence of NO˙ (De Groote *et al.*, 1997). Further overlap with the oxidative stress response occurs in *E. coli* where it has been shown that NO˙ activates both SoxR and OxyR. NO˙ activates SoxR by nitrosylation of its iron-sulfur centres generating dinitrosyl-iron-dithiol clusters and this form of SoxR activates transcription of *soxS* almost as effectively as when the iron-sulfur clusters are oxidised (Ding and Demple, 2000). The genes of the soxRS regulon may also serve important functions in protection against the toxicity of NO˙. The dismutation of the superoxide radical means it is not available for peroxynitrite formation with NO˙. Another member of the regulon, glucose 6-phosphate dehydrogenase encoded by *zwf* could function to provide NADPH for the flavohemoglobin. Certain S-nitrosothiols also cause transcriptional activation of OxyR by S-nitrosylation (Hausladen *et al.*, 1996). Other genes which have an effect have been discovered in *S. typhimurium*, where a *metL* mutant was found to be susceptible to S-

nitrosothiols and had attenuated virulence in mice (De Groote *et al.*, 1996). The *metL* gene product is involved in homocysteine biosynthesis and the sensitivity phenotype of the mutant was overcome by addition of homocysteine (De Groote *et al.*, 1996). The experimental evidence suggested that S-nitrosothiols interacted with homocysteine. In *E. coli* elevated levels of homocysteine result in decreased expression of *hmp*. Conversely depletion of homocysteine levels by addition of the S-nitrosothiol GSNO is purported to increase *hmp* expression via the binding of MetR, a LysR family DNA binding protein which has homocysteine as its cofactor, to the *hmp* promoter (Membrillo-Hernandez *et al.*, 1999). In another example, and a further overlap with the oxidative stress response, AhpC of *S. typhimurium* was found to protect the cell from reactive nitrogen intermediates independently of AhpF (Chen *et al.*, 1998).

1.4 THE RESPIRATORY NADH DEHYDROGENASES

NADH dehydrogenases are components of the respiratory chain. These enzymes carry out the initial oxidation of their substrate, NADH, and feed electrons into the quinone pool. The reduced quinones deliver electrons to the terminal oxidoredutases and these enzymes transfer the electrons to the terminal electron acceptors. Sources of electrons in *E. coli* include formate, hydrogen, lactate, succinate, pyruvate and glycerol-3-phosphate (Gennis and Stewart, 1996). In *B. subtilis* succinate and glycerol-3-phosphate can donate electrons to the quinone pool. *B. subtilis* has only one type of quinone, menaquinone, in the quinone pool whereas *E. coli* has benzoquinone, menaquinone, demethylmenaquinone and ubiquinone. *B. subtilis* can utilize oxygen, nitrite and nitrate as terminal electron acceptors (Nakano and Zuber, 1998). *E. coli* has a much wider range of potential terminal electron acceptors including oxygen, nitrite, nitrate, fumarate, DMSO (dimethyl sulfoxide) and TMAO (trimethylamine-N-oxide). When *E. coli* is growing aerobically it uses two types of terminal oxidases: cytochrome bo_3 (under aerobic conditions) and cytochrome bd (under microaerophilic conditions) (Gennis and Stewart, 1996). *B. subtilis* growing aerobically uses three (and a possible fourth) cytochromes as terminal oxidases: cytochrome caa_3, cytochrome bd and cytochrome aa_3 and possibly YthAB the gene product of *ythA* and *ythB* though this potential cytochrome remains to be characterised (Winstedt and von Wachenfeldt, 2000). In fact the *B. subtilis* respiratory chain branches after menaquinone reduction. Menaquinol is oxidised by one of three terminal oxidases (cytochrome bd, cytochrome aa_3 or possibly YthAB) or by menaquinol:cytochrome *c* oxidoreductase (the *bc* complex). The *bc* complex is oxidised by *c*-type cytochromes or directly by cytochrome caa_3, the terminal oxidase of this branch of the respiratory chain.

Cytochromes can act as proton pumps thereby generating an energy storing proton gradient across the membrane.

Therefore the role of the NADH dehydrogenases is to oxidise NADH to NAD^+ and feed electrons to the quinones in the first step of the respiratory electron transport chain. Transmembrane NADH dehydrogenases can be coupled to a proton motive force and thus harness the free energy generated during this step of respiration. The energy of the gradient can then be used for various activities of the cell including ATP synthesis, motility and solute uptake. Another important function of NADH dehydrogenases is to regenerate NAD^+.

1.4.1 Two types of respiratory NADH dehydrogenase

NADH is a substrate for two distinct NADH dehydrogenases: NDH-1 and NDH-2. NDH-1 is a 14 subunit enzyme and it is homologous to the mitrochondrial complex 1. In contrast NDH-2 is a single polypeptide. NDH-1 is a transmembrane protein complex that uses the free energy generated by its enzymatic activity to pump protons across the membrane thus creating an electrochemical gradient. In contrast NDH-2 is a membrane associated protein which does not pump protons. NDH-1 has several Fe-S clusters and a FMN cofactor whereas NDH-2 does not have Fe-S clusters and uses FAD as a cofactor. NDH-1 can be inhibited by mitochondrial complex 1 inhibitors such as rotenone whereas NDH-2 is rotenone insensitive (Yagi, 1993). *E. coli* has both NDH-1 and NDH-2 type NADH dehydrogenases, *Paracoccus denitrificans*, *Campylobacter jejuni* and *Helicobacter pylori* appear to contain only NDH-1 (Calhoun *et al.*, 1993; Yano and Yagi, 1999; Smith *et al.*, 2000). In contrast, *B. subtilis* and *Sulfolobus acidocaldarius* have only NDH-2 which in *Sulfolobus* exists as a dimer (Bergsma *et al.*, 1982; Wakao *et al.*, 1987). *Mycobacterium tuberculosis* appears to have NDH-1 as well as NDH-2 (Cole *et al.*, 1998). In this bacterium NDH-2 is of clinical importance as mutations in the NDH-2 gene, *ndh,* confer resistance to the drug isoniazid (INH) a drug used in tuberculosis chemotherapy (Miesel *et al.*, 1998; Lee *et al.*, 2001). INH works by inhibiting the biosynthesis of mycolic acids and thus disrupting the bacterial cell envelope. To carry out its bacteriocidal function INH must first be converted to an active form by the *katG* encoded peroxidase. The active INH reacts with the NADH bound to InhA, an enzyme involved in mycolic acid synthesis. This forms a INH-NAD adduct covalently attached to InhA that causes its inactivation (Rozwarski *et al.*, 1998). Mutations in *katG* and *ahpC* and mutations which cause overproduction of InhA have been found to cause INH resistance in clinical isolates (Heym

et al., 1999; Telenti *et al.*, 1997; Musser *et al.*, 1996). More recently mutations in *ndh* have been discovered that lead to INH resistance (Miesel *et al.*, 1998; Lee *et al.*, 2001). A mechanism is proposed whereby defects in *ndh* result in an increase in NADH concentration since NDH-2 is not generating NAD^+. Increased NADH may disrupt KatG activation of INH or may result in displacement of INH-NAD from the InhA enzyme causing its activation (Miesel *et al.*, 1998). This is surprising as the *M. tuberculosis* genome contains an *nuoA-N* gene cluster which would be expected to encode the components of the 14 subunit NDH-1 type NADH dehydrogenase and therefore compensate for any mutations in *ndh*. However *ndh* mutants of *Mycobacterium smegmatis* have very low NADH dehydrogenase activity and it appears that NDH-2 is responsible for the majority of NADH dehydrogenase activity in this bacterium. NDH-1 is thought to be inactive, at least under the conditions used in experiments (Miesel *et al.*, 1998).

1.4.2 Aerobic versus anaerobic growth

B. subtilis was once thought to be a strict aerobe. Recently it was discovered that it can grow anaerobically either by fermentation or using nitrate or nitrite as a terminal electron acceptor (Nakano and Zuber, 1998). Fermentation is stimulated by the presence of pyruvate and the products of fermentation include ethanol, lactate, acetoin and 2,3-butanediol (Nakano *et al.*, 1997). *B. subtilis* can respire anaerobically using nitrate or nitrite as terminal electron acceptors. Oxygen limiting conditions are sensed by ResE, the sensor kinase of the ResDE, a two component system that is involved in early cellular responses to oxygen limitation. This leads to phosphorylation of ResD, the response regulator. Phosphorylated ResD upregulates expression of *fnr* from the promoter immediately upstream of the *fnr* gene. Fnr is a conserved global regulator of anaerobic gene expression with a redox sensitive Fe-S cluster. FNR induces expression of its own upstream gene *narK* (and thereby induces its own expression from the *narK* promoter). The *narK* gene product is responsible for nitrite extrusion. FNR also upregulates *narGHJI* expression (Nakano *et al.*, 1996). The *narGHJI* genes encode nitrate reductase (Cruz Ramos *et al.*, 1995). Nitrate reductase reduces nitrate to nitrite which is further reduced to ammonia by nitrite reductase (encoded by *nasDE*). The expression of *nasDEF* is induced by ResD under anaerobic conditions (Nakano *et al.*, 1998). These genes are also induced aerobically by TnrA under nitrogen limiting conditions (Wray *et al.*, 1996). Under these circumstances the function of nitrite reduction is to allow nitrogen incorporation into biomolecules. Similarly the alternative nitrate reductase encoded by *nasBC* acts as an assimilatory nitrate reductase under nitrogen limiting conditions when cells are growing

aerobically (Ogawa *et al.*, 1995). The ResDE two component system of *B. subtilis* also functions in phosphate starvation (Sun *et al.*, 1996a). In addition to *fnr* and *nasDEF*, other genes known to be regulated by ResDE include *catA* (Sun *et al.*, 1996b), *ctaBCDEF* (Liu and Taber, 1998), *hmp* (Nakano, 2002), *lctE* (Cruz Ramos *et al.*, 2000), the *sbo-alb* operon (Nakano *et al.*, 2000) and its own operon *resABCDE* (Sun *et al.*, 1996b). ResE is a membrane protein with two transmembrane domains. The signal that it senses is unknown. It acts as both a phosphatase and a kinase in controlling the activity of ResD. ResE probably acts primarily as a phosphatase under aerobic conditions and as a kinase when cells are grown anaerobically (Nakano *et al.*, 1999). The regulatory proteins ResDE and Fnr are central to anaerobiosis in *B. subtilis* though there is evidence that additional regulatory proteins and pathways are involved. The probable two component system YclJ/YclK is anaerobically induced though its role is not known (Ye *et al.*, 2000). YwiD also appears to have a regulatory role since *ywiD* expression is induced by Fnr and is required for induction of *hemN* and *hemZ* (Homuth *et al.*, 1999). The mechanism of activation of Fnr in *B. subtilis* is probably similar to that of *E. coli*. In *E. coli* Fnr exists in a monomeric inactive form during aerobic growth. Under oxygen limiting conditions the Fe-S cluster becomes reduced and Fnr becomes an active dimer (Khoroshilova *et al.*, 1997). Conserved N-terminal cysteines are thought to coordinate the Fe-S cluster (Nakano and Zuber, 2002). In *E. coli* a two component system, the ArcA-ArcB system, controls gene expression in response to changes in oxygen conditions. This is a complex two component system. The sensor kinase ArcB autophosphorylates in response to several signals. ArcB can transfer the phosphoryl group to conserved residues within ArcB and finally to the response regulator ArcA (Ishige *et al.*, 1994). Alternatively phosphorylated ArcB can phosphorylate ArcA directly without transphosphorylation (Tsuzuki *et al.*, 1995). ArcA can repress genes involved in aerobic respiration and induce those involved in anaerobic respiration .

1.4.3 Regulation of NADH dehydrogenases in *E. coli*

Experiments in *E. coli* have shown the regulation of the *nuoA-N* genes encoding NDH-1 to be complex. The locus is transcriptionally regulated in response to ArcA (anaerobic repression), NarL (anaerobic induction in the presence of nitrate), FNR (weak anaerobic repression) and IHF (weak anaerobic repression) (Bongaerts *et al.*, 1995). Transcriptional regulation also occurs in response to electron acceptors and electron donors. Expression is induced during anaerobiosis when fumarate is acting as an electron acceptor and during aerobic growth when C_4 dicarboxylates (*e.g.* succinate) are used as a carbon source

25

(Bongaerts *et al.*, 1995). The *nuoA-N* locus is also regulated in a growth phase dependent manner by the growth phase responsive regulator Fis. Fis binds to the *nuoA-N* promoter region and induces expression particularly during exponential growth. This ensures an abundance of NDH-1, which is coupled to the proton motive force, and thereby ATP generation during exponential growth when high ATP levels are required (Wackwitz *et al.*, 1999).

In *E. coli ndh* is regulated by the Fnr protein which represses transcription under anaerobic or oxygen limiting conditions (Spiro *et al.*, 1989; Green and Guest, 1994). Fnr binds to the *ndh* promoter at two sites and binding at both sites is necessary for complete repression. Fnr represses *ndh* transcription by preventing the α-subunit of the RNA polymerase from contacting the DNA at these two sites (Meng *et al.*, 1997). However *ndh* also has more than one level of regulation. In the absence of Fnr, *ndh* is activated during anaerobic growth in rich medium by the amino acid response regulator Arr which binds to two sites in the promoter overlapping the Fnr binding sites (Green *et al.*, 1997). The integration host factor (IHF) and the histone like protein HU also bind the *ndh* promoter and IHF represses transcription by preventing open complex formation (Green *et al.*, 1997). The growth phase regulator Fis is also involved in *ndh* regulation which it probably coordinates with its role as an activator of *nuoA-N*. When Fis is present at low levels (stationary phase) it activates transcription of *ndh* and when it is present at high levels (exponential phase) it represses *ndh* transcription. In contrast, during exponential phase Fis activates transcription of the *nuoA-N* genes of the coupled NDH-1 thereby ensuring maximal ATP generation during periods of rapid growth (Green *et al.*, 1996). The fact that Fis and IHF are involved in the regulation of both *ndh* and *nuoA-N* indicates that the two NADH dehydrogenases are coordinately regulated.

1.4.4 Regulation of NADH dehydrogenases in *B. subtilis*

B. subtilis does not contain an operon homologous to the *nuoA-N* gene cluster. However it contains three *ndh* genes: *yjlD, yutJ* and *yumB*, whose expression and regulation has been investigated as part of the work for this thesis. An NADH dehydrogenase was previously purified from *B. subtilis*. The enzyme was membrane associated, but based on its hydrophobicity and polarity was judged not to be an intrinsic membrane protein (Bergsma *et al.*, 1982). It was found to be located at the inner surface of the cytoplasmic membrane (Bergsma *et al.*, 1981). The enzyme consists of a single subunit and has FAD as a cofactor (Bergsma *et al.*, 1982). More recently, analysis of the global response of the transcriptome

and the proteome has shown YjlD is exclusively produced under aerobic conditions (Marino *et al.*, 2000). Microarray analysis has shown that *yjlCD* is downregulated during anaerobic nitrate or nitrite respiration or fermentation (Ye *et al.*, 2000). Though microarray analysis showed a slightly elevated level of expression of *yjlCD* in a *resDE* mutant (Ye *et al.*, 2000), protein analysis showed that the anaerobic repression of YjlD did not require *fnr* or *resDE* (Marino *et al.*, 2000). Transcriptome and proteome analysis indicated the *yjlCD* operon is under negative stringent control, that is, the expression of *yjlCD* is repressed during the stringent response and this repression is *relA* dependent (Eymann *et al.*, 2002). The stringent response occurs in the cell upon starvation and amino acid limitation. In *E. coli* the regulator of the stringent response is encoded by *relA*. RelA is a guanosine tetra and pentaphosphate ((p)ppGpp) synthetase that is bound to the ribosome. When an uncharged tRNA arrives at the ribosome RelA becomes activated and catalyses GTP phosphorylation (Chatterji *et al.*, 2001). The (p)ppGpp mediates the stringent response and has been shown to induce *rpoS* (Gentry *et al.*, 1993). Like *E. coli* the *relA* gene product in *B. subtilis* can activate and repress gene expression (Wendrich and Marahiel, 1997). The mechanism of regulation is not defined. However it has been suggested that negative control in *E. coli* is due to (p)ppGpp binding to β and β' subunits of RNA polymerase and positive control is mediated by redistribution of RNA polymerase as a result of down regulation of many genes normally expressed during growth, although this is probably not the only mechanism (Toulokhonov *et al.*, 2001; Barker *et al.*, 2001a; Barker *et al.*, 2001b). It is not surprising that the *yjlCD* operon is under negative stringent control as NADH dehydrogenase functions in growing and respiring cells and these are the functions which are repressed in cells undergoing the stringent response. However it is likely that this is not the only regulator of the *yjlCD* genes. The regulator of anaerobic repression of *yjlCD*, for example, remains to be elucidated.

1.4.5 NADH dehydrogenases and oxidative stress

In *E. coli* the NDH-2 enzyme is linked to the oxidative stress status of the cell because this enzyme is one of the primary sources of superoxide and H_2O_2 production during aerobic respiration (Messner and Imlay, 1999). NDH-2 of *E. coli*, but not NDH-1, was identified as the source of superoxide and H_2O_2 production via autooxidation of its reduced FAD cofactor. It was suggested that this does not occur in other flavin containing proteins due to their hydrophobicity or because the flavin is not exposed and is protected from oxidation by the protein structure (Messner and Imlay, 1999). Also, the acceleration of ROS production from the respiratory chain at high temperatures as well as high oxygen levels

may partly explain the apparent overlap between oxidative stress and heat shock responses. As NDH-2 enzymes all have a similar structure this enzyme may also be a source of oxidative stress in *B. subtilis*, especially as the genome contains three *ndh*-like genes. Results presented in this thesis point to a paraquat diaphorase activity of YjlD. It may be that YjlD generates superoxide and H_2O_2 like NDH-2 in *E. coli* and paraquat simply accelerates the transfer of electrons to oxygen. This implies that the toxicity of paraquat is due not only to superoxide generation but also the diversion of electrons from the respiratory chain. The paraquat diaphorase activity of NAD(P)H oxidoreductases has been observed in *E. coli* (Liochev and Fridovich, 1994; Liochev *et al.*, 1994). It has been suggested that the degree of oxidative stress may depend on the amount of autooxidizable enzymes present in the cell in addition to environmental conditions (Messner and Imlay, 1999). This observation points to a close relationship between the NDH-2 enzyme of the respiratory chain and the redox status of the cell.

1.5 AIMS OF THIS THESIS

Genome sequencing and functional analysis carried out as part of the work for this thesis is described in Chapter 3. The aim of the work described in Chapters 4 and 5 was the characterisation of two families of genes in *B. subtilis*: the NADH dehydrogenases and the SODs. Both SODs and NADH dehydrogenases affect the oxidative stress status of the cell. Chapter 5 describes investigation into the expression, regulation and role of the type 2 NADH dehydrogenases encoded by *yjlCD*, *yutJ* and *yumB*. The products of these genes are potential sources of oxidative stress during respiration. Chapter 4 describes work carried out on the SODs of *B. subtilis* encoded by *sodA*, *sodF* and *yojM*. The function of SODs is in detoxification of superoxide radicals. The expression and regulation of these genes was investigated

CHAPTER 2

MATERIALS AND METHODS

2.1 GROWTH AND MAINTENANCE OF BACTERIAL STRAINS

Escherichia coli and *Bacillus subtilis* strains were routinely maintained on Luria Bertani (LB) (Miller, 1972) agar plates for short term storage and as glycerol stocks at −70°C for long term storage. To make glycerol stocks strains were grown in LB liquid medium and glycerol was added to a final concentration of 25%. *E. coli* and *B. subtilis* were grown in liquid medium at 220 rpm and 37°C in a Gallencamp orbital shaker. 5-Bromo-4-chloro-3-indolyl-D-galactoside (X-gal) was added when necessary to a concentration of 40μg/ml. Isopropyl -β-D-thiogalactopyranoside (IPTG) was included in the medium as indicated in the text to a concentration of 1mM. Xylose was added to the medium when indicated in the text to a final concentration of 0.5%.

2.1.1 Antibiotic concentrations

Antibiotics were used during growth and for selection of strains where appropriate. For *E. coli* ampicillin was used at 100μg/ml. For *B. subtilis* the antibiotics used and their final concentrations were as follows: Erythromycin 1μg/ml and 150 μg/ml; chloramphenicol 3μg/ml; spectinomycin 100 μg/ml and kanamycin 5 μg/ml.

2.1.2 Growth media

The growth media used and their composition were as follows:

Luria Bertani Medium (LB):

(Miller, 1972)

For 1 litre:

10g tryptone

5g yeast extract

5g NaCl

Schaeffers Spoulation Medium (SM):

(Schaeffer *et al.*, 1965)

For 1 litre:

Nutrient broth 8g

KCl 1g

$MgSO_4$ 0.25g

After autoclaving the following were added from sterile stocks:

$CaCl_2$ to 1mM

MnCl2 to 0.01mM

$FeSO_4$ to 0.001mM

Bacillus Functional Analysis Minimal Medium:

(Yoshida *et al.*, 2000)

For 1 litre:

200ml 5x minimal salts

20ml   20% glucose

5ml   1% L-tryptophan

50ml   4% L-glutamine

2ml   2mg/ml $FeCl_3$

2ml   0.1mg/ml $MnSO_4$

10ml   100x trace elements

5x minimal salts:

57mM K_2SO_4

310mM K_2HPO_4

220mM KH_2PO_4

17mM NaCitrate

4mM $MgSO_4$

100x trace elements:

5mM $CaCl_2$

1.25mM $ZnCl_2$

0.25mM $CuCl_2$

0.25mM $CoCl_2$

0.25mM Na_2MoO_4

The final concentration of glucose in this medium was 0.4%. For glucose starvation conditions this medium was used with glucose at a final concentration of 0.05%.

Basic Limitation Medium (BLM):

As in Stulke *et al.*, 1993 except glucose was increased to 0.4% final concentration and $CaCl_2$ was excluded.

50mM Tris

15mM $(NH_4)_2SO_4$

8mM $MgSO_4$

27mM KCl

7mM NaCitrate

was adjusted to pH7.5

After autoclaving the following were added from sterile stocks:

$FeSO_4$ to 1μM

$MnSO_4$ to 10μM

Potassium glutamate to 4.5mM

Glucose to 0.4%

KH_2PO_4 to 0.6mM

Tryptophan to 160μg/ml

Solid medium was made by adding agar to a final concentration of 1.5% w/v.

2.2 DNA SEQUENCING

2.2.1 DNA sequencing reactions

DNA sequencing reactions were carried out using either fluorescently labelled forward and reverse M13 primers or fluorescently labelled ddNTPs. Labelled M13 primers were used with either the Genpak PyroSeq kit (Genpak; New Milton, UK) or the Amersham Thermosequenase kit (Amersham; Bucks., UK) following the manufacturers instructions. Custom primers were used with labelled terminators in the Prism DyeTerminator kit (Perkin Elmer; Warrington, UK) according to the manufacturers instructions. Completed sequencing reactions were purified by running through a Sephadex G-50 column and concentrated in a SpeedVac. Reactions were run on a 6% denaturing polyacrylamide gel (Seqagel; National Diagnostics, Atlanta, GA, USA) on an ABI373A automated sequencer (Applied Biosystems; Foster City, CA, USA).

2.2.2 DNA sequencing of large genomic fragments

A 29.324 kb segment of the *B. subtilis* chromosome between *ykwC* and *rok* (formerly *ykuW*) was sequenced. The region between *ykuA* and *ykuO* was cloned by chromosome walking. The plasmid pCheV3 was made by Dr. Susanne Krogh Devine. This plasmid contains a fragment of *cheV,* of which the sequence was known, cloned into the integrating plasmid pDIA5304 (Glaser *et al.*, 1993). The plasmid pCheV3 was integrated into the *B. subtilis* chromosome by single crossover and chloramphenicol resistant transformants were selected. This strain was called EScheV. Restriction digests were performed on transformant chromosomal DNA using enzymes with a single site within the polylinker of

pDIA5304. The size of the chromosomal fragment between the restriction site in the polylinker and the closest site in the adjacent *B. subtilis* chromosomal DNA was determined by Southern analysis. The fragment of known *cheV* DNA had been cloned into the *Hind*III site. Therefore restriction sites downstream of *Hind*III could be used to clone chromosomal DNA upstream of *cheV* and restriction sites upstream could be used to clone DNA located downstream of *cheV* on the chromosome. Southern blotting analysis indicated that *Sal*I was an appropriate enzyme to clone the region upstream of *cheV* which was 6.020kb in size. *Eco*RI was suitable to clone the downstream region of chromosomal DNA which was 13.618kb in size. Chromosomal DNA from strain EScheV was digested with *Sal*I and *Eco*RI and subsequently ligated to generate pCheVSalI and pCheVEcoRI respectively. The recircularized plasmids were cloned in *E. coli* strain TP611 (Glaser *et al.*, 1993). In this strain ColE1 plasmids are maintained at a low copy number to prevent high levels of expression of *B. subtilis* genes in *E. coli*. Random fragments of plasmid pCheVSalI and plasmid pCheVEcoRI were subcloned into pUC19 by Dr. Mary O' Reilly. DNA sequencing of inserts in these libraries was carried out using M13 forward and reverse primers. The DNA sequence data generated was assembled using the XBAP programme from the Staden package.

To sequence the remainder of this region long range PCR products were generated by Dr. Mary O' Reilly. Two long range products were generated. The first was a 7.802 kb fragment between *ptsI* and *ykuA*, synthesised with the primers LR36 (5'-GCTGAGTGGGCGAAGCTTGTCAATGAACCG-3') and Lib1 (5'-GCTTGTAAGCCTGGTCTCCTTTCAGATTGC-3'). The second was a 8.235kb fragment between *ykuO* and *mobA*, synthesised with the primers C10L (5'-CCCCTGTTATCCGGTCGTTATGCCTTCACG-3') and LR46 (5'-TATTGCCCCGTCAGTCCCGCTTGAGGG-3'). These PCR fragments were digested with single enzymes. The LR36/Lib1 fragment was digested with *Xba*I, *Pst*I and *Hind*III. The C10L/LR46 fragment was digested with *Xba*I, *Eco*RI, *Acc*I, *Sac*I, *Kpn*I and *Pst*I. The resulting fragments were cloned into pUC19 and the ends of these fragments sequenced using universal M13 primers. The resulting sequence data was used to design oligonucleotides that primed DNA sequencing reactions using the long range PCR products as the template.

2.2.3 DNA sequence assembly

The DNA sequence data generated was assembled using the XBAP programme from the Staden package. The DNA sequence was translated in all six open reading frames using the XNIP programme of the Staden package. The amino acid sequence of ORFs was annotated by comparing ORFs greater than 30 amino acids with sequence from the GenBank database using the BLASTP programme (Altschul *et al.*, 1990). The entire piece of genomic DNA sequenced was 29.324kb in length and was submitted to the EMBL nucleotide sequence database under the accession number AJ222587.

2.2.4 Sequence analysis

Routine DNA sequence analysis was carried out using GCG programmes (Genetics Computer Group, Inc., Wisconsin Package, Madison, WI, USA). Multiple sequence alignments were carried out using the ClustalW programme (Thompson *et al.*, 1994) from the www service at the European Bioinformatics Institute (http://www2.ebi.ac.uk/clustalw). Pretty printing and shading of multiple alignments was carried out using Boxshade 3.21 at http://www.ch.embnet.org/software/BOX_form.html.

2.3 DNA MANIPULATIONS

Routine DNA modifications and cloning procedures such as restriction digests, ligations, phenol/chloroform extractions, ethanol precipitations, agarose gel electrophoresis, gel purifications, Southern blotting and CsCl density gradient centrifugation were carried out according to standard methods (Sambrook *et al.*, 1989).

2.3.1 Polymerase chain reaction (PCR)

PCR reactions were usually carried out in 25-50 µl volumes containing PCR buffer, 10 picomoles of each primer, 200µM dNTPs, approximately 5ng template DNA and Taq polymerase (Invitrogen; Paisley, UK) or Pfu polymerase (Stratagene; Amsterdam, The Netherlands).

2.3.2 Plasmid DNA preparation

Small scale plasmid DNA preparation was carried out by the boiling lysis method (Holmes and Quigley, 1981). Large scale plasmid DNA preparation was carried out according to Sambrook (Sambrook *et al.*, 1989). Plasmid DNA which was used to generate libraries for genome sequencing was purified by CsCl gradient density centrifugation.

2.3.3 *B. subtilis* chromosomal DNA preparation

To isolate *B. subtilis* chromosomal DNA cells, 0.5ml overnight culture was pelleted by spinning for 1 minute at 12,000rpm. The supernatant was removed and the pellet resuspended in 300µl of solution I (0.05M Tris pH8, 0.05M EDTA, 25% sucrose). 100µl lysozyme solution (20mg/ml in solution I) was added and the mixture was incubated at 37°C for 15 minutes. 400µl solution II was added (0.01M Tris pH8, 0.005M EDTA, 1%SDS). 20µl RNase solution (10 mg/ml) was added and the mixture was incubated at 37°C for a further 30 minutes. A phenol extraction was carried out followed by a phenol/ chloroform extraction. The DNA was then precipitated by addition of an equal volume of isopropanol and centrifugation for 10 minutes at 12,000rpm. The DNA pellet was dissolved in 300µl 0.1M potassium acetate and re-precipitated by adding 700µl cold ethanol followed by 10 minutes centrifugation. The DNA pellet was washed in 70% ethanol, air dried and dissolved in 50µl sterile water.

2.3.4 Bacterial transformations

E. coli was transformed by the calcium chloride method (Sambrook *et al.*, 1989). *B. subtilis* was transformed by a modified version of the method of Anagnostopoulos and Spizizen (Anagnostopolous and Spizizen, 1961). 250µl of an overnight culture was added to 2.5ml Broth I (Spizizen's minimal salts containing 0.5% glucose, 5mM $MgSO_4$, 0.02% yeast extract, 0.02% casamino acids and 0.05 mg/ml tryptophan). This culture was grown for four and a half hours. 1ml of this culture was added to 4ml Broth II (Spizizen's minimal salts containing 0.5% glucose, 5mM $MgSO_4$, 0.01% yeast extract, 0.01% casamino acids and 0.005 mg/ml tryptophan) and mixed. Plasmid (0.5-1.0µg) or chromosomal (10-20ng) DNA was added to 0.5ml of the Broth II mix and the cells/DNA mix was grown for a further 90 minutes. Transformants were selected on LB agar plates containing the appropriate antibiotic.

2.4 STRAIN CONSTRUCTION

2.4.1 Inactivation of genes by pMUTIN4 insertion

The integrating plasmid pMUTIN4 was used to generate insertional inactivations of genes and transcriptional *lacZ* fusions (Vagner *et al.*, 1998). The integration of pMUTIN4 into a target gene is illustrated in Figure 2.1. The pMUTIN4 vector has an *E. coli* origin of replication, so it can replicate in *E. coli* but not in *B. subtilis*. The ampicillin resistance gene allows selection in *E. coli* and the erythromycin resistance gene allows selection in *B. subtilis*. There is a multiple cloning site immediately upstream of the promoterless *lacZ*
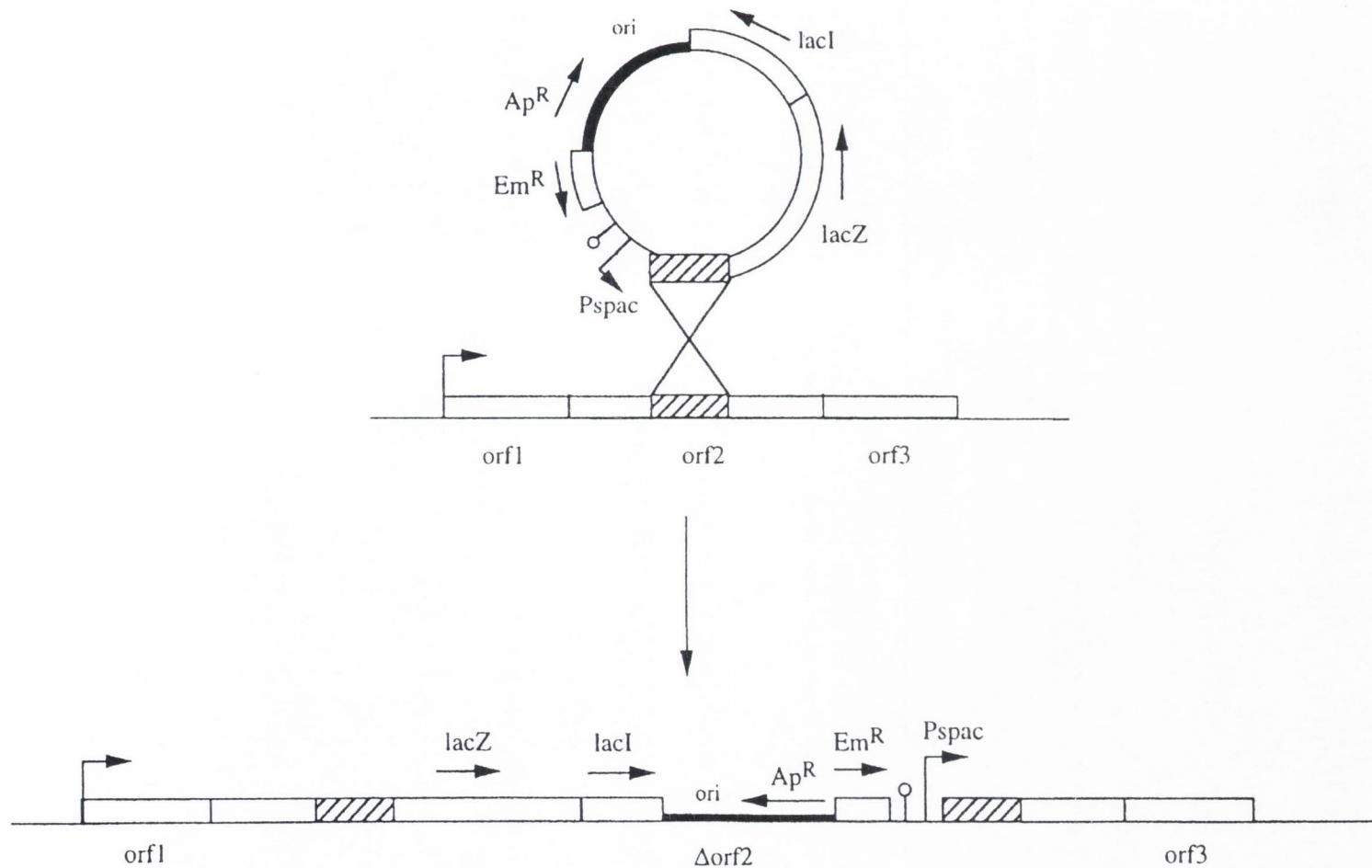
Figure 2.1 Diagram of pMUTIN4 integration into a target gene.
Genes of orf1, orf2 and orf3 are depicted as white boxes. Hatched boxes indicate the internal fragment of the target gene which is cloned into pMUTIN4 and directs recombination into the *B. subtilis* chromosome. The pMUTIN4 vector containing the cloned fragment of orf2 is integrated into the chromosome by a single crossover event. Arrows show direction of transcription. Right angled arrows show promoters. The lollipop represents a terminator. This diagram is from Vagner *et al.* (1998)

gene that allows generation of transcriptional fusions when chromosomal fragments are cloned in these sites. The *lacZ* gene has a *B. subtilis spoVG* ribosome binding site (RBS) and there are three stop codons in each of the three frames upstream of the RBS to prevent the generation of translational fusions with upstream genes. The P_{SPAC} promoter has SPO1 phage RNA polymerase recognition sequences. The P_{SPAC} promoter also contains the O1 *lac* operator (Yansura and Henner, 1984). The LacI protein is constitutively expressed in pMUTIN4 and it represses the P_{SPAC} promoter. This repression is prevented by IPTG. In order to create an insertional mutant using pMUTIN4 an internal fragment of the gene of interest is cloned into the MCS. After propagation in *E. coli* the recombinant plasmid is transformed into the *B. subtilis* chromosome by a single crossover as illustrated in figure 2.1. Transformants are selected on erythromycin. Transformants are obtained if the plasmid integrates into the chromosome at the homologous site generating a null mutant and a transcriptional *lacZ* fusion to the promoter of the gene. Genes downstream in the same operon can be induced by IPTG to avoid polar effects on expression. The pMUTIN4 vector can also be used to put a gene of interest under the control of the P_{SPAC} promoter thus creating a conditional mutant. In this case a PCR product overlapping the 5' end of the gene, including its ribosome binding site, is cloned into pMUTIN4. The subsequent single crossover places the gene under the inducible control of the P_{SPAC} promoter.

Strains BFS1831- BFS1841, BFS1862 and BFS1863

As part of the *B. subtilis* functional analysis project insertional mutants were constructed using pMUTIN4 in thirteen of the open reading frames in the region sequenced between *ykwC* and *rok* (formerly *ykuW*). Mutants were made in the following genes: *ykuB*, *ykuC*, *ykuD*, *ykuE*, *ykuG*, *ykuH*, *ykuI*, *ykuJ*, *ykuK*, *ykuL*, *ykuM* (now called *ccpC*), *ykuU* and *ykuV*. The same strategy was used for the creation of all these strains with the exception of the *ykuJ* mutant strain BFS1838. Internal fragments of the target genes were generated by PCR with primers which included *Eco*RI and *Bam*HI restriction sites to facilitate cloning. The PCR fragments were cloned into pMUTIN4 and the resultant plasmids were propagated in *E. coli*. These plasmids were transformed into *B. subtilis*, where the recombinant plasmid integrated into the chromosome by a Campbell-type event at the region of homology. The chromosomal DNA arrangement at these loci was subsequently verified by PCR with combinations of oligonucleotides internal and external to the integrated plasmid DNA. In the case of *ykuJ*, as the predicted open reading frame is small (237 base pairs) it was decided to place the gene under P_{SPAC} control rather than generating an insertional mutant. The cloning strategy used was the same as outlined above except

rather than cloning an internal fragment of the gene, the 5' end of the ORF including the ribosome binding site was cloned into pMUTIN4. The primers used to generate these constructs are listed below beside the strain name. The name of the target gene in each case is in parenthesis.

BFS1831 (*ykuB*):   Lib36 5'-CCGGAATTCAAATCCAAGGTTCAGCC-3'
                      Lib37 5'-CGCGGATCCGCTTTAGGATTTCGATAC-3'

BFS1832 (*ykuC*):   Lib38 5'-CGCGGATCCATGAGCGTGCCGATTAC-3'
                      Lib39 5'-CCGGAATTCCGGATTCCAAAAGAAGC-3'

BFS1833 (*ykuD*):   Lib40 5'-CCGGAATTCTCAATCGGAGCGAAGAC-3'
                      Lib41 5'-CGCGGATCCGTGGGCGGTAAAAAAGG-3'

BFS1834 (*ykuE*):   Lib42 5'-CGCGGATCCAAGTCATCAAGCGAAGC-3'
                      Lib43 5'-CCGGAATTCCTAAACCTGATCTCATTG-3'

BFS1835 (*ykuG*):   Lib44 5'-CGCGGATCCGTTGAACCTAAAAAACCG-3'
                      Lib45 5'-CCGGAATTCGAAAGATGCGTTTGCTG-3'

BFS1836 (*ykuH*):   Lib46 5'-CCGGAATTCAGAGTATTTGGCTGGGC-3'
                      Lib47 5'-CGCGGATCCCCATACATTGAACCGAC-3'

BFS1837 (*ykuI*):   Lib48 5'-CCGGAATTCCTGAAAGTGTCACAGCC-3'
                      Lib49 5'-CGCGGATCCAGCCCTGAAAATAGCGG-3'

BFS1838 (*ykuJ*):   Lib50 5'-CCGGAATTCGCTTTTTGAAAGGAGTC-3'
                      Lib51 5'-CGCGGATCCTCTCACCATTCACTTCG-3'

BFS1839 (*ykuK*):   Lib52 5'-CGCGGATCCCCATCAAGAATATGAGCGG-3'
                      Lib53 5'-CCGGAATTCTGCATCGTACAGGCAAG-3'

BFS1840 (*ykuL*):   Lib54 5'-CCGGAATTCCCTACCGTTTACATGGC-3'
                      Lib55 5'-CGCGGATCCGTCATTGATATGAAGGCGC-3'

BFS1841 (*ykuM*):   Lib56 5'-CGCGGATCCCAATATCTTCGATGCAGG-3'
                      Lib57 5'-CCGGAATTCGATCTCGCTCATAACCG-3'

BFS1862 (*ykuU*):   LibU1 5'-CCGGAATTCTGATGCAGAAGTCATCG-3'
                      LibU2 5'-CGCGGATCCAAAACGCCATATTCCCG-3'

BFS1863 (*ykuV*):   LibV1 5'-CCGGAATTCGATGCCGCAAGTGAATG-3'
                      LibV2 5'-CGCGGATCCCAAAAATCGGCTGAGTG-3'

Strain ES998

Strain ES998 contains a pMUTIN4 insertion into *yutJ*. A PCR fragment was amplified using the primers utj1 (5'-CCGGAATTCCCATCGATCAGTCCCGC-3') and utj2 (5'-CGCGGATCCGGCGTCTGCAGAAATCG-3') and cloned into the *Eco*R1 and *Bam*H1

sites in pMUTIN4 to generate the plasmid pES998. This plasmid was transformed into *B. subtilis*. Transformants contained a single copy of the pES998 plasmid integrated into the chromosome at the *yutJ* locus.

Strain ES999

ES999 contains a pMUTIN4 insertion into *yumB*. The internal fragment cloned into the *Eco*RI and *Bam*HI sites of pMUTIN4 was synthesized by PCR using the primers umb1 (5'-CCGGAATTCTTGAGCTGCAGTTTGCG-3') and umb2 (5'-CGCGGATCCCCCATACAACGGTTTGC-3'). This generated the plasmid pES999 which was integrated into the *B. subtilis* chromosome generating strain ES999.

Strain BFS841

BFS841 was generated in D. Karamata's laboratory as part of the *B. subtilis* functional analysis project. The strain contains a pMUTIN4 insertion into *yjlD*. The plasmid pjlD.1 is pMUTIN4 with an internal fragment of *yjlD* (genome coordinates 1,298,764-1,299,025) cloned into the *Bam*HI and *Eco*RI sites. This plasmid was used to generate BFS841 via a single crossover insertion into the chromosome.

Strain BFS842

Strain BFS842 was generated in the laboratory of D. Karamata. The plasmid pjlC.1.1 was constructed by cloning an internal fragment of *yjlC* (genome coordinates 1,298,126-1,298,324) into the *Bam*HI and *Eco*RI sites of pMUTIN4. Single crossover insertion of pjlC.1.1 into *B. subtilis* generated BFS842.

Strain ES001

This strain contains a pMUTIN4 insertion into *sodF*. It was constructed by transformation of the plasmid pES001 into *B. subtilis*. A PCR fragment was synthesized using sodF1 (5'-CCGGAATTCCGATGGCCTGAACAAAG-3') and sodF2 (5'-CGCGGATCCTTTTTTAGAGGCCTGGG-3') and cloned into the *Eco*RI and *Bam*HI sites of pMUTIN4 generating pES001. This plasmid was integrated into the *B. subtilis* chromosome generating ES001.

Strain ES002

Strain ES002 contains a pMUTIN4 insertion into *sodA*. An internal *sodA* fragment was synthesized by PCR using the primers

soda1 (5'-CCGGAATTCGCGGACACGCGAACCAC-3') and
soda2 (5'-CGCGGATCCCACAACAAGCCATGCCC-3'). This fragment was cloned into
the *Eco*RI and *Bam*HI sites of pMUTIN4 to create pES002. This plasmid was used to
generate ES002 by a single crossover.

Strain ES003

This strain contains a pMUTIN4 insertion into *ilvC*. The plasmid used to generate the
strain is pES003. It contains a cloned PCR fragment at *Eco*RI-*Bam*HI which was
synthesized by PCR using the primers
ilvc1 (5'-CCGGAATTCACATATGAGCAAGGAGC-3') and
ilvc2 (5'-CGCGGATCCTCAGGCTGATAACCTGC-3'). This plasmid was integrated into
the chromosome at the *ilvC* locus generating ES003.

Strain ES004

ES004 contains a pMUTIN4 insertion into *cysK*. The plasmid pES004 is pMUTIN4
containing a cloned internal fragment of *cysK* at *Eco*RI-*Bam*HI. The PCR fragment was
synthesized with the primers
cysk1 (5'-CCGGAATTCGTGCTGACTCCTGGTGC-3') and
cysk2 (5'-CGCGGATCCGGTATGCTTCCTTCAGC-3'). The strain ES004 was
generated by a single crossover insertion of pES004 into the *B. subtilis* chromosome.

Strain CM991

CM991 was made by Colin Murtagh under my supervision. It contains a pMUTIN4
insertion into *dhbB*. The plasmid used to generate the strain was pCM991. It is pMUTIN4
with a cloned internal fragment of *dhbB* synthesized with the primers
dhbb1 (5'-CCGGAATTCAATGGAGATACAGCGCG-3') and
dhbb2 (5'-CGCGGATCCGCACAGCGTCCAGCCGC-3'). Integration of pCM991 into
the *B. subtilis* chromosome generated CM991.

Strain CM992

CM992 was made by Colin Murtagh under my supervision. This strain contained a
pMUTIN4 insertion into *leuC*. The PCR fragment used to drive the insertion was made
with the primers leuc1 (5'-CCGGAATTCTCCTTGCTGTCATCGGC-3') and leuc2 (5'-
CGCGGATCCGCCGTCAAGGACGATAG-3'). This fragment was cloned into

pMUTIN4 to make pCM992. This plasmid was transformed into *B. subtilis* by single crossover to generate CM992.

Strain ES31

In this strain *yojM* was placed under the control of the P_{SPAC} promoter. A PCR fragment which overlapped the 5' end of *yojM* and the ribosome binding site of this gene was synthesized using pspacyojm1 (5'-CCGGAATTCCAAGGGGGAGACAGTATG-3') and pspacyojm2 (5'-CGCGGATCCCCGGCCTGACACAAGAACC-3'). This was cloned into pMUTIN4 to make pES31. When pES31 was transformed into ES0007 (section 2.4.5) it resulted in ES31 which contains the *yojM* gene under P_{SPAC} control.

Strain ES32

ES32 contains *sodF* under the control of P_{SPAC}. A PCR fragment which overlapped the 5' end of the *sodF* gene and contained the ribosome binding site of this gene was synthesized using pspacsodf1 (5'-CCGGAATTCCTATTGAAAAGGAGACAGC-3') and pspacsodf2 (5'-CGCGGATCCCGCGTCTTTCACTTGTATC-3') and was cloned into pMUTIN4 to generate pES32. This plasmid was transformed into ES007 (section 2.4.5) to generate ES32, a *sodA* mutant in which *sodF* is under P_{SPAC} control.

2.4.2 Inactivation of genes by pMUTIN4XZ insertion

The vector pMUTIN4XZ was made by Brian Jester and it is a derivative of pMUTIN4 where the *lacZ* gene has been excised. The pMUTIN4 plasmid was digested with *Sal*I and *Bsi*WI. This results in complete excision of the *lacZ* gene. The plasmid was then blunt ended using Klennow and religated to make pMUTIN4XZ (B. Jester pers comm). This vector allows inactivation of genes by insertion of pMUTIN4XZ via a single crossover, but does not result in the generation of a *lacZ* fusion. This is useful for the creation of mutant strains where a *lacZ* fusion is already present in the chromosome.

Strain ES18

This strain is a *yqgB* mutant generated by insertion of pMUTIN4XZ. An internal fragment of *yqgB* was made with the primers
pM4yqgB1 (5'-CCGGAATTCATGAGAAAAAAGCCACGG-3') and
pM4yqgB2 (5'-CGCGGATCCCCCGCTGCTGTTACGAG-3') and cloned into pMUTIN4XZ to generate pES18. When pES18 was transformed into *B. subtilis* it resulted in ES18 which contains an insertional inactivation of *yqgB*.

Strain ES19

ES19 contains a pMUTIN4XZ insertion into *yqgC*. The plasmid used to generate this strain was pES19. It was made by cloning a PCR product synthesized with the primers pM4yqgC1 (5'-CCGGAATTCGCCGCAGTATTATTCGC-3') and pM4yqgC2 (5'-CGCGGATCCCTGCGCAAACAGCCCCG-3') into pMUTIN4XZ. This plasmid was transformed into *B. subtilis* to generate ES19.

2.4.3 Generation of transcriptional *lacZ* fusions using pDG268

The pDG268 plasmid is an integrating vector that allows integration of transcriptional *lacZ* fusions into the *B. subtilis* chromosome at the *amyE* locus by a double crossover event between the front and back halves of *amyE* (Antoniewski *et al.*, 1990). PCR products were cloned at sites in front of the *lacZ* gene. The plasmid contains an *E. coli* origin of replication and an ampicillin resistance gene and can therefore be amplified in *E. coli*. Two regions of homology to *amyE* flank the *lacZ* gene and a chloramphenicol resistance gene for selection in *B. subtilis*. When the plasmid is linearised these regions drive the double crossover integration at the amylase locus. Strains were verified by examining the amylase phenotype on LB plates which contained 0.2% starch. Where integration at the *amyE* locus has occurred strains have an amylase deficiency phenotype that can be observed when plates are flooded with iodine (0.1% iodine, 0.1%KI$_2$). Strains were also verified by PCR using the primers amy1 (5'-GCTCGGGCTGTATGACTGG-3') and amy2 (5'-CGTATCATGCGACTCTACCC-3'). These primers amplify a 368 base pair fragment from the wild type *amyE* gene but do not produce a product in strains which contain a double crossover integration at the *amyE* locus.

Strain ES991

This strain was generated using pDG268 and it contains the upstream untranslated region of *katA* transcriptionally fused to *lacZ* at the amylase locus. The PCR fragment containing the *katA* promoter region was synthesized with the primers
Kam1 (5'- CGCGGATCCCATGTTATCACCTCTTGG-3') and
Kam2 (5'- CCGGAATTCGAGCCTTCGGCGGTCCG-3') and cloned into the *Eco*RI and *Bam*HI sites of pDG268, using restriction sites included in the primer sequences, to make pES991. This plasmid was linearised with *Sca*I and integrated into *B. subtilis* by double crossover generating ES991.

Strain ES993

This strain was generated using pDG268 and it contains the upstream untranslated region of *ahpC* transcriptionally fused to *lacZ* at the amylase locus. The PCR fragment containing the *ahpC* promoter region was synthesized with the primers

Aam1 (5'- CCGGAATTCTTCTATCCCAAACAACAG-3') and

Aam2 (5'- CGCGGATCCCCTAAAAATGTATTAGAAAGC-3') and cloned into the *Eco*RI and *Bam*HI sites of pDG268, using restriction sites included in the primer sequences, to make pES993. This plasmid was linearised with *Sca*I and integrated into *B. subtilis* by double crossover generating ES993.

Strain ES9910

This strain was generated using pDG268 and it contains the upstream untranslated region of *yumB* transcriptionally fused to *lacZ* at the amylase locus. The PCR fragment containing the putative *yumB* promoter region was synthesized with the primers

umb3 (5'-CCGGAATTCGCCTCCTAAATGAAAATTG-3') and

umb4 (5'-CGCGGATCCAGACCCATAATAAAATC-3') and cloned into the *Eco*RI and *Bam*HI sites of pDG268, using restriction sites included in the primer sequences, to make pES9910. This plasmid was linearised with *Xho*I and integrated into *B. subtilis* by double crossover generating ES9910.

Strain ES9911

This strain, generated with pDG268, contains the upstream untranslated region of *yutJ* fused to *lacZ* at the amylase locus. The PCR product containing this region was made with the primers utj3 (5'-CCGGAATTCTCACCTCACATATGTCC-3') and utj4 (5'-CGCGGATCCATCCGAAGAGTATGATAG-3') and was cloned into the *Eco*RI and *Bam*HI sites of pDG268 generating pES9911. This plasmid was linearised with *Sca*I and integrated by double crossover into the amylase locus generating ES9911.

Strain ES9912

ES9912 contains the upstream untranslated region of the *yjlCD* operon fused to *lacZ* at the amylase locus. This promoter region of *yjlCD* was synthesized using the primers jlc1 (5'-GCCGAAGCTTAGGGGAGAAATTGCTGG-3') and

jlc2 (5'-CGCGGATCCCTCCTTTACTATAAGCAG-3') and was cloned into the *Hind*III and *Bam*HI sites of pDG268 to produce pES9912. This plasmid was digested with *Xho*I and integrated by double crossover into the amylase locus generating ES9912.

Strain ES14

In ES14 the untranslated region upstream of *sodA* which contains the promoters of *sodA* is fused to *lacZ* at the amylase locus. The primers used to synthesize the PCR product which was used to make this strain were

psoda1 (5'-CCGGAATTCGAAATGCTGGCGGCAGG-3')

and psoda2 (5'-CGCGGATCCCCGTAAGCCATGATAATTCC-3'). The PCR product was cloned into the *Eco*RI and *Bam*HI sites of pDG268 to make pES14. The plasmid was linearised using *Pst*I and transformed into *B. subtilis* where it integrated by double crossover into the amylase locus generating ES14.

Strain ES15

ES15 contains the untranslated region immediately upstream of *yojM* fused to *lacZ* at the amylase locus. It was not apparent whether the *yojM* promoter was contained in this region or upstream of *yojL* so both fusions were constructed (see below). The PCR fragment used to make this strain was synthesized using the primers

yojm3 (5'-CCGGAATTCAAATGAGTATGACAAAAAGC-3') and

yojm4 (5'-CGCGGATCCTGTCCGCATCCGGCAAC-3') and cloned into the *Eco*R1 and *Bam*HI sites of pDG268 to make pES15. This plasmid was linearised using *Pst*I and transformed into *B. subtilis* generating ES15.

Strain ES38

ES38 contains the untranslated region upstream of *yojL* fused to *lacZ* at the amylase locus. The *yojL* gene may be in an operon with *yojM* and this region may therefore contain the promoter sequences for this operon. The PCR fragment used to make this strain was made with the primers PyojL1 (5'-CGCGGATCCCATTCATTCAACCTCCTAAAC-3') and PyojL2 (5'-CCGGAATTCGCCATTTCATAAGAATGC-3') and cloned into pDG268 to produce pES38. This plasmid was digested with *Pst*I and transformed into *B. subtilis* generating ES38.

2.4.4 Generation of translational *lacZ* fusions using pAC5

The pAC5 vector allows the construction of translational fusions to *lacZ* and single copy integration at the *amyE* locus (Martin-Verstraete *et al.*, 1992). PCR products cloned at sites in front of the *lacZ* gene result in fusion to the ninth codon of LacZ. The vector can be amplified in *E. coli* and contains an ampicillin resistance gene. It also contains a

chloramphenicol resistance gene for selection in *B. subtilis*. Two regions of homology drive double crossover integration into the *amyE* locus which is verified as in section 2.4.3.

Strain ES22

This strain contains a translational fusion of the *sodA* promoter region to *lacZ* and was generated using pAC5. The same PCR product that was used for creation of ES14 was cloned into pAC5 to produce pES22. This plasmid was digested using *Pst*I and integrated by double crossover into the *amyE* locus to generate ES22.

2.4.5 Gene inactivation by insertion of antibiotic cassettes

Genes encoding resistance to kanamycin and spectinomycin were cloned into the *E. coli* plasmid vectors BluescriptKS⁺ and pMTL23 respectively to generate pDG780 and pDG1727 by Guerout-Fleury *et al.* (Guerout-Fleury *et al.*, 1995). These cassettes were used to generate and select for mutations in *B. subtilis*.

Strain ES007

ES007 contains a spectinomycin cassette inserted into *sodA*. Two PCR fragments were synthesized to make the plasmid pES007. The first overlapped the start of *sodA* and was made using the primers soddco1 (5'-CGCGGATCCTTGTTACATTGAATGAAG-3') and soddco2 (5'-AAAACTGCAGGCAAGAGCAGTGTTTCC-3') which contained restriction sites to facilitate cloning. This PCR fragment was digested with *Bam*HI and *Pst*I and cloned into pDG1727 (Guerout-Fleury *et al.*, 1995) to produce the plasmid pES007a. The second PCR fragment overlapped the 3' end of the *sodA* gene and was made using soddco3 (5'-CCGCTCGAGAGATTCTCCGCTTTCAG-3') and soddco4 (5'-CCGCTCGAGCCGAGATATGTAAAAAG-3') which contain *Xho*I sites. The fragment was digested with *Xho*I and cloned into pES007a to make pES007. The plasmid was screened by PCR to ensure that the cloned fragments were in the correct orientation. The pES007 plasmid was digested with *Stu*I and transformed into *B. subtilis* generating ES007. Double crossover integration into the *sodA* locus conferred spectinomycin resistance and resulted in the deletion of 437 bases of the *sodA* gene.

Strain ES17

ES17 is a *fur* mutant strain which was generated using pDG780 (Guerout-Fleury *et al.*, 1995). A PCR fragment from the middle of the *fur* gene was amplified using the primers fur1 (5'-CGCGGATCCGAGATCGGTCTCGCTAC-3') and fur2

(5'-CCGGAATTCCCATGCACACCAAGTGG-3'). The product was digested with *Bam*HI and *Eco*RI and ligated into pDG780. This plasmid was called pES17. This plasmid was transformed into *B. subtilis*. A single crossover into the *fur* locus generated a kanamycin resistant Fur mutant. This strain was called ES17.

Strain ES13

Chromosomal DNA from ES001 was transformed into ES007 to create ES13. This strain contains *sodF* and *sodA* null mutations and a transcriptional *lacZ* fusion to the 5' end of *sodF*.

Strains ES16 and ES39

Chromosomal DNA prepared from ES007 was transformed into ES15 and ES38 to generate ES16 and ES39 respectively. These strains contain *sodA* null mutations and the upstream intergenic region of either *yojM* or *yojL* is transcriptionally fused to *lacZ* at the amylase locus.

Strains ES20 and ES21

Chromosomal DNA from ES18 and ES19 was transformed into ES14 to create ES20 and ES21 respectively. These strains carry mutations in either *yqgB* or *yqgC* and both have a transcriptional fusion of the *sodA* promoter region to *lacZ* at the amylase locus.

Strains ES23, ES24, ES25 and ES40

Chromosomal DNA from ES17 was transformed into ES001, ES15, ES14 and ES38 to generate ES23, ES24, ES25 and ES40 respectively. These strains are all *fur* mutants. ES23 is a *sodF* mutant and transcriptional *lacZ* fusion. ES24 contains the upstream untranslated region of *yojM* fused to *lacZ* at the amylase locus. ES25 contains the *sodA* promoter region fused to *lacZ* at the amylase locus. ES40 contains the upstream untranslated region of *yojL* fused to *lacZ* at the amylase locus.

Strain ES37

Chromosomal DNA from PB344 (Boylan *et al.*, 1993) was transformed into ES14 to generate ES37. PB344 contains a spectinomycin cassette inserted into *sigB* (Boylan *et al.*, 1993). ES37 contains a transcriptional *lacZ* fusion to the *sodA* promoter region at the amylase locus in a *sigB* mutant background.

2.4.6 Placing genes under the control of P_{xyl} using pX

The pX plasmid is a xylose inducible expression vector (Kim *et al.*, 1996). It contains an *E. coli* origin of replication and an ampicillin resistance gene which allows amplification in *E. coli*. The chloramphenicol resistance cassette allows selection after integration in *B. subtilis*. Two regions of homology to the *amyE* gene allow integration by double crossover into the amylase locus. The vector contains the *xylR* repressor gene and the promoter of the xylose utilization operon of *Bacillus megaterium*. This allows induction of expression of any gene which has been cloned downstream of the xylose promoter upon addition of xylose to the medium.

Strain ES27

The *yjlC* open reading frame was amplified by PCR with the primers pXyjlC1 (5'-CGCGGATCCATGCCAGAAACAATCGATC-3') and pXyjlC2 (5'-CGCGGATCCATATCAATTTGAATCTGAG-3'). The PCR fragment was digested with *Bam*HI and cloned into pX. Primers pXF (5'-AAGGGGGAAATGACAAATGGTC-3') and pXR (5'-TCCGTCGCTATTGTAACCAGTTC-3') were used to screen for the correct orientation of the insert. This plasmid, called pES27, was digested with *Sca*I and transformed into *B. subtilis*. A double crossover integration at the amylase locus generated ES27. In this strain *yjlC* is under P_{xyl} inducible control.

Strain ES28

The *yjlD* open reading frame was amplified by PCR using the primers pXyjlD1 (5'-CGCGGATCCATGTCAAAACATATTGTC-3') and pXyjlD2 (5'-CGCGGATCCTAAAAGGATTAGTAAGCC-3') and subsequently digested with *Bam*HI and cloned into the *Bam*HI site of pX and PCR screened for correct orientation, creating pES28. This plasmid was linearised using *Hpa*I and transformed into *B. subtilis* generating ES28. In this strain the linearised plasmid integrated into the chromosome by a double crossover at the amylase locus. The *yjlD* gene is under P_{xyl} control at this locus.

Strain ES29

Chromosomal DNA prepared from BFS841 was transformed into ES28 to create ES29. ES29 contains a pMUTIN4 insertional mutant of *yjlD*. The *yjlD* gene is under P_{xyl} control at the amylase locus. The *yjlC* gene is intact and under the control of its natural promoter.

Strain ES30

Chromosomal DNA from BFS842 was transformed into ES27 to generate ES30. This strain therefore contains a pMUTIN4 insertional mutant of *yjlC*. The pMUTIN4 insertion also has a polar effect on *yjlD* and this gene is under P_{SPAC} control. The *yjlC* gene is under P_{xyl} control at the amylase locus.

2.5 TRANSCRIPTIONAL ANALYSIS

2.5.1 Preparation of RNA

To prepare total RNA *B. subtilis* cells were harvested at designated time points as indicated in the relevant results sections. 10ml cells were aliquoted, centrifuged for 1 minute at 4°C, snap frozen in a dry ice ethanol bath and either stored at −80°C or processed immediately. The cell pellet was resuspended in 200μl DEPC treated water and the suspension was transferred to a tube containing 1ml TRI reagent (Sigma-Aldrich; Dorset, UK) and 0.5g glass beads (Biospec; Bartlesville, OK, USA). The tubes were shaken for 3x 45 second pulses in a Fastprep bead beater (Bio101, Biospec; Bartlesville, OK, USA). The tubes were then cooled on ice and 200μl chloroform was added. The contents were mixed and the tubes spun for 5 minutes. The upper layer was removed and added to an equal volume of chloroform. The mixture was mixed, centrifuged and the upper layer was removed and added to an equal volume of acid 5:1 phenol/chloroform (Sigma-Aldrich; Dorset, UK). This was again mixed, centrifuged and the aqueous layer was removed to a new tube. One tenth the volume of 3M sodium acetate and an equal volume of isopropanol was then added to precipitate the RNA. The precipitation mix was centrifuged for 15 minutes and the pellet was washed in 80% ethanol and resuspended in DEPC treated water. RNA concentration was determined by measurement of the OD_{260}. RNA was then divided into 25μg aliquots for subsequent procedures.

2.5.2 Northern blot analysis

25μg aliquots of total RNA were run on 1.2% formaldehyde gels for Northern blot analysis according to standard procedures (Sambrook *et al.*, 1989). RNA was transferred onto positively charged nylon membrane (Pall Gelman, Ann Arbor, MI, USA). Membranes were crosslinked (150mJ) and baked for 30 minutes at 80°C. Membranes were hybridised with digoxigenin labelled probes generated using a PCR Dig labelling mix (Roche, East Sussex, UK) according to the manufacturers instructions. Transcripts were detected using the digoxigenin detection kit (Roche, East Sussex, UK). The gene whose sequence was

used as a probe and the sequence of the primers used to synthesise the probe for Northern hybridisation are as follows:

sodA: sodA1 5'-CCGGAATTCGCGGACACGCGAACCAC-3'

sodA2 5'-CGCGGATCCCACAACAAGCCATGCCC-3'

yojM: yojM1 5'-CCGGAATTCTCCGCAAACAGCCTGCG-3'

yojM2 5'-CGCGGATCCGGCGCATTCATTATGAC-3'

hmp: hmp1 5'-GAAGAGCAAGCCGGCGGC-3'

hmp2 5'-GAATGAACAGAATTTGCCG-3'

yjlB: jlb1 5'-AGGATTCCGAATCACCC-3'

jlb2 5'-CATATCGTACTGAACGCC-3'

yjlC: jlc3 5'-AAATGCGTCTGTCAGCC-3'

jlc4 5'-GCATGCGAAGAAGGCCG-3'

yjlD: jlc5 5'-ATTGCGGAAGTAAGCTC-3'

jlc6 5'-GATCCGTCTTTCAGATCG-3'

yutJ: utj1 5'-CCGGAATTCCCATCGATCAGTCCCGC-3'

utj2 5'-CGCGGATCCGGCGTCTGCAGAAATCG-3'

yumB: umb1 5'-CCGGAATTCTTGAGCTGCAGTTTGCG-3'

umb2 5'-CGCGGATCCCCCATACAACGGTTTGC-3'

2.5.3 Primer extension analysis

Primer extension analysis was carried out with 25μg samples of total RNA. The RNA was annealed to a radioactively labelled primer for 30 minutes at 55°C in 15μl 1x Superscript II buffer. The temperature was adjusted to 48°C and 5μl primer extension mix containing 1x Superscript II buffer, 20mM dithiothreitol, 2mM dNTPs and 200U of Superscript II (Invitrogen, Paisley, UK) was added to each sample. The samples were incubated at 48°C for 30 minutes. The samples were precipitated overnight at −20°C. The pellet was resuspended in 8μl 50% 10mM Tris-HCl-1mM EDTA (pH8)-50% stop solution (U.S. Biochemicals, Cleveland, OH, USA). Samples were denatured immediately prior to running on a 6% denaturing polyacrylamide gel (Seqagel; National Diagnostics, Atlanta, GA, USA). Primer extension samples were run alongside sequencing reactions generated using the same labelled primer and an appropriate PCR product as the template. Sequencing reactions were carried out using the fmol DNA cycle sequencing system (Promega; Southampton, UK). The gene whose promoter region was examined and the primers used were as follows

yjlCD: jlc8 5'-GGAAAGGGTAGTGCCTCTCC-3'

  jlc9 5'-AAGATCTTGCTGGCTTTGGC-3'

yumB: umb4 5'-CGCGGATCCAGACCCATAATAAAATC-3'

  umb6 5'-CCTGCACCTAAAATCACG-3'

  umb7 5'-ACCGTCATCAATCCGCC-3'

yutJ: utj4 5'- CGCGGATCCATCCGAAGAGTATGATAG-3'

  utj5 5'-CCCGTAACCTCCGCCGATC-3'

sodA: sodPE1 5'-TCCTTGTCGATATGCGG-3'

  sodPE2 5'-ATGTACTGAAATGCCGC-3'

2.6 MEASUREMENT OF β– GALACTOSIDASE ACTIVITY

1ml aliquots of cells were collected for measurement of β-galactosidase activity. Cells were spun at 12,000rpm for 2 minutes. The supernatant was discarded and the cell pellet stored at −20°C until the assay was carried out. The cell pellet was lysed for 25 minutes in either 250µl or 500µl of Z buffer (0.06M Na_2HPO_4, 0.04M NaH_2PO_4, 0.01M KCl, 1mM $MgSO_4$) which contained 10µg/ml lysozyme and 1µg/ml DNase. The cells were then spun at 12,000rpm for 2 minutes and the supernatant was transferred to new tubes. Appropriate aliquots were added to reaction mixtures containing o–nitrophenyl-β-D-galactopyranoside (ONPG) as a substrate. The final concentration of ONPG in the reaction mix was 0.8mg/ml and the final volume of the reaction mix was 150µl. The reaction was carried out for a fixed length of time at 28°C and stopped by the addition of 120µl of the reaction mix to 60µl of 1.2M Na_2CO_3. The OD_{420} of the Na_2CO_3/reaction mixtures was then read. The protein concentration of the same cell lysates used for the β-galactosidase assay was measured using BioRad microassays (BioRad, Hercules, CA, USA) following the manufacturers instructions. One unit of activity was defined as 1 nanomole of ONPG hydrolysed per minute per mg of protein. Units were calculated using the formula:

$$\text{Activity in nmole/min/mg} = \frac{OD_{420} \times 2.2 \times 1.5}{\text{mg/ml protein} \times \text{volume sampled in ml} \times \text{mins} \times 0.00486}$$

2.2 is the correction factor for reading samples in a microtitre plate rather than a cuvette. It was determined from calibration curves of samples measured in microtitre plates compared to those measured in cuvettes. This accounts for the difference in pathlength between OD measurements in microtitre plates and OD measurements in cuvettes.

1.5 is the correction factor for the dilution with Na_2CO_3. The molar extinction coefficient of ONPG is 4860.

This formula was used for calculation of LacZ activities for all the assays except those which are described in Chapter three. For these assays activity was measured in Miller units where β-galactosidase activity is measured as a function of OD_{600} (nmole/min/OD_{600}) (Ferrari et al., 1986).

$$\text{Activity in Miller units} = \frac{OD_{420} \times 1.5 \times 222.2}{\text{mg protein} \times \text{volume sampled in ml} \times \text{mins}}$$

where mg protein was calculated from the OD_{600} as follows:

μg protein= OD_{600} x 83

mg protein=$\dfrac{\mu\text{g protein}}{\text{ml}}$ x $\dfrac{\text{ml used in assay}}{1000}$

2.7 PHENOTYPIC ANALYSIS

2.7.1 *B. subtilis* Functional Analysis Project phenotypic tests

As part of the *B. subtilis* Functional Analysis project 1,148 mutant strains were screened for their response to the oxidising agents hydrogen peroxide and paraquat in Kevin Devine's laboratory. Screening was carried out in collaboration with Dr. Mary O' Reilly, David Noone and Kasper Krogh Andersen. The phenotype of these strains was examined on solid medium and compared to the wild type phenotype. Gradient plates were used for these tests which are described in detail in Functional Analysis of Bacterial Genes, A Practical Manual (Scanlan et al., 2001).

2.7.2 Disc assays

Disc assays were carried out by a Bauer-Kirby disc diffusion method (Bauer et al., 1966) to determine the zone of inhibition measurements for various strains in response to toxic chemicals. 25ml LB agar (1.5%) plates were poured, allowed to solidify and prewarmed at 37°C. Cultures of strains to be tested were grown to exponential phase (OD_{600}=0.2). 100μl culture was then added to 2ml molten top agar (LB+ 0.7% agar) and gently mixed. The top agar/culture mix was poured onto the 25ml LB plates. 10μl of the chemical being tested was pipetted onto an antimicrobial susceptibility test disc (Oxoid, Hampshire, UK). The concentration of the chemicals used was as follows: hydrogen peroxide: 0.88M; t-butyl hydroperoxide: 0.2M; cumene hydroperoxide: 0.2M; paraquat: 0.5M. After the top

agar had solidified forceps were used to place the disc onto the centre of the plate. Plates were incubated at 37°C overnight. The zone of inhibition was the diameter of the circle of growth inhibition measured in millimetres.

2.7.3 Stress responses in liquid media

2.7.3.1 Stress induction with paraquat, hydrogen peroxide, tBOOH and external superoxide

An overnight culture was diluted 100-fold in fresh LB broth and grown to OD_{600}=0.2-0.3. The culture was then split and the inducing agent was added to one half of the culture, but not to the control uninduced half. Cells were harvested at the time of the split (T_0) and from both cultures 5, 10, 20, 30, 45, 60, 90 and 120 minutes after the split, for measurement of OD_{600} and for β-galactosidase assays. For external superoxide stress (using xanthine/xanthine oxidase) and for induction with tBOOH, cells were harvested up to 60 minutes after T_0. Samples were harvested for RNA preparation and analysis at time points indicated in the results. The final concentration of paraquat, hydrogen peroxide and tBOOH in the induced cultures was 100μM. To generate extracellular superoxide the enzyme xanthine oxidase (Roche; East Sussex, UK) and its substrate hypoxanthine (Sigma-Aldrich; Dorset, UK) were used. Hypoxanthine (dissolved in 1N NaOH) was added to both the induced and non induced cultures to a final concentration of 1mM. Xanthine oxidase was added to the induced culture only, to a final concentration of 50mU/ml. These amounts of substrate and enzyme have been shown to generate 9.8μM superoxide/minute (Nunoshiba *et al.*, 1993).

2.7.3.2 Stress induction with ethanol and nitric oxide

BFA minimal medium (section 2.1.2) was used for stress inductions with ethanol and nitric oxide. Fresh medium was inoculated with an overnight culture to an OD_{600} of 0.05 and grown to exponential phase (OD_{600}=0.5). The culture was then split and the inducing agent was added to one half of the culture. Cells were harvested at the time of the split (T_0) and from both induced and non induced cultures 5, 10, 20, 30, 45, 60, 90 and 120 minutes after the split, for measurement of OD_{600} and for β-galactosidase assays. Samples were harvested for RNA preparation at time points indicated in the results. The final concentrations of the inducing agents were: ethanol: 4% (v/v); 3-morpholinosydnonimine (SIN-1): 100μM; S-nitroso-N-acetylpenicillamine (SNAP): 200μM and S-nitrosoglutathione (GSNO): 50μM. SIN-1 and SNAP were from Sigma (Dorset, UK). 500mM stocks of GSNO were prepared by dissolving 1.537g glutathione (Sigma; Dorset,

UK) in 7.5ml 1N HCl and then adding 0.346g $NaNO_2$ (Sigma; Dorset, UK). The solution was then neutralised with 712.5µl 10N NaOH to pH7.5. H_2O was added to bring the final volume to 10ml. Aliquots were stored at –80°C for up to three weeks (De Groote et al., 1995).

2.7.3.3 Paraquat and tBOOH sensitivity

To measure paraquat sensitivity in liquid medium LB cultures were grown to an OD_{600} of 0.2-0.3. The culture was divided and paraquat, to a final concentration of 75mM, was added to one half of the culture. Aliquots were removed at the time the cultures were divided (T_0) and from both induced and non induced cultures 10, 20, 30, 40 and 50 minutes after the split and plated out on LB plates at appropriate dilutions. Percentage survival was calculated as: (colony count for the induced culture/colony count of the non-induced culture) x100 at each time point.

To examine viability after t-butyl hydroperoxide (tBOOH) challenge cells in early exponential phase (OD_{600}~0.2) were exposed to 100mM tBOOH. At T_0 and 10, 20 and 30 minutes after exposure, cells were plated out on LB agar plates. Percentage survival was calculated as: (Colony count at time point after tBOOH addition/Colony count at T_0) x 100.

2.7.4 Paraquat resistance on solid medium

Cells were grown to exponential phase to examine paraquat resistance on solid medium. Then cells were plated out at appropriate dilutions onto LB plates which contained no paraquat, 175µM, 350µM, 525µM and 700µM paraquat. Plates were incubated overnight and the colonies were counted.

2.8 NON DENATURING PROTEIN GEL ELECTROPHORESIS

2.8.1 Protein preparation

Cells were harvested for protein preparation in 10ml aliquots, spun for 5 minutes at 5000rpm and the pellet was resuspended in 980µl of sonication buffer (10mM Tris, 150mM NaCl pH8). 10µl cocktail III protease inhibitor (Calbiochem, San Diego, CA, USA) was added along with 10µl lysozyme (20mg/ml in sonication buffer). The mixture was incubated for 30 minutes on ice. The mixture was then sonicated for 15 second periods until it became clear. The mix was then centrifuged for 20 minutes at 8,000rpm. The supernatant (soluble fraction) was removed and the protein concentration was

measured using BioRad microassays. The insoluble pellet was resuspended in 1% Triton (in sonication buffer) and incubated on ice for 20 minutes followed by centrifugation for 20 minutes at 8,000 rpm to remove any protein which had not been solubilised.

2.8.2 Protein gel electrophoresis

5% stacking gels and 10% or 15% separating gels were used for non denaturing protein gel electrophoresis. Gels were made and run according to standard procedures (Bollag and Edelstein, 1991).

2.8.3 Staining for NADH dehydrogenase activity

To detect NADH dehydrogenase activity the native gels were washed for 30 minutes in MOPS buffer (100mM MOPS-NaOH pH8). Then the gels were incubated in MOPS buffer with nitroblue tetrazolium (final concentration 0.5mg/ml) and NADH (1mM) until the blue bands had fully developed (~30minutes). Enzymes which can oxidise NADH under these conditions will form a band on the gel. The blue colour is formazan resulting from the enzymatic reduction of the tetrazolium salt (Howitt *et al.*, 1999).

2.8.4 Staining for superoxide dismutase activity

A two step washing procedure was used to detect superoxide dismutase activity on native protein gels (Beauchamp and Fridovich, 1971). Washes were carried out in the dark. Firstly, the gels were washed in 0.025% nitroblue tetrazolium (Sigma, Dorset, UK); 0.01% riboflavin (Sigma, Dorset, UK) in H_2O for 20 minutes. The second wash was in 1% Temed for a further 20 minutes. Then gels were exposed to the light. Nitroblue tetrazolium is reduced by superoxide radicals to a blue formazan. Superoxide radicals are generated by riboflavin in a light catalyzed reaction with temed as the electron donor. Achromatic bands in the gel indicate the presence of superoxide dismutase activity which has prevented superoxide radicals from reducing nitroblue tetrazolium.

Table 2.1
Bacterial strains and plasmids used in this work

Strain or Plasmid	Relevant characteristics	Source/Reference[a]
E. coli strains		
TG-1	*supE hsdΔ5 thi Δ(lac-proAB)* F'[*traD36 proAB⁺ lacI^q lacZΔM15*]	Laboratory stock
TP611	*recBC hsdR⁻M⁻ cyab10 pcn*	(Glaser *et al.*, 1993)
B. subtilis strains		
168	*trpC2*	Laboratory stock
EScheV	*trpC2 cheV*::pCheV3 Cm^r	pCheV3→168
BFS1831	*trpC2 ykuB*::pDU1831 (*ykuB'-lacZ*) Erm^r	pDU1831→168
BFS1832	*trpC2 ykuC*::pDU1832 (*ykuC'-lacZ*) Erm^r	pDU1832→168
BFS1833	*trpC2 ykuD*::pDU1833 (*ykuD'-lacZ*) Erm^r	pDU1833→168
BFS1834	*trpC2 ykuE*::pDU1834 (*ykuE'-lacZ*) Erm^r	pDU1834→168
BFS1835	*trpC2 ykuG*::pDU1835 (*ykuG'-lacZ*) Erm^r	pDU1835→168
BFS1836	*trpC2 ykuH*::pDU1836 (*ykuH'-lacZ*) Erm^r	pDU1836→168
BFS1837	*trpC2 ykuI*::pDU1837 (*ykuI'-lacZ*) Erm^r	pDU1837→168
BFS1838	*trpC2 ykuJ*::pDU1838 (P_SPAC-*ykuJ*) Erm^r	pDU1838→168
BFS1839	*trpC2 ykuK*::pDU1839 (*ykuK'-lacZ*) Erm^r	pDU1839→168
BFS1840	*trpC2 ykuL*::pDU1840 (*ykuL'-lacZ*) Erm^r	pDU1840→168
BFS1841	*trpC2 ykuM*::pDU1841 (*ykuM'-lacZ*) Erm^r	pDU1841→168

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
BFS1862	*trpC2 ykuU*::pDU1862 (*ykuU'-lacZ*) Erm^r	pDU1862→168
BFS1863	*trpC2 ykuV*::pDU1863 (*ykuV'-lacZ*) Erm^r	pDU1863→168
BFS841	*trpC2 yjlD*::pjlD.1 (*yjlD'-lacZ*) Erm^r	pjlD.1→168
BFS842	*trpC2 yjlC*::pjlC.1.1 (*yjlC'-lacZ*) Erm^r	pjlC.1.1→168
CM991	*trpC2 dhbB*::pCM991 (*dhbB'-lacZ*) Erm^r	pCM991→168
CM992	*trpC2 leuC*::pCM992 (*leuC'-lacZ*) Erm^r	pCM992→168
PB344	*trpC2 sigB*::*spc* Spc^r	(Boylan *et al.*, 1993)
ES991	*trpC2 katA*::pES991 (P$_{katA}$- *lacZ*) Cm^r	pES991→168
ES993	*trpC2 ahpC*::pES993 (P$_{ahpC}$-*lacZ*) Cm^r	pES993→168
ES998	*trpC2 yutJ*::pES998 (*yutJ'-lacZ*) Erm^r	pES998→168
ES999	*trpC2 yumB*::pES999 (*yumB'-lacZ*) Erm^r	pES999→168
ES9910	*trpC2 amyE*::pES9910 (P$_{yumB}$-*lacZ*) Cm^r	pES9910→168
ES9911	*trpC2 amyE*::pES9911 (P$_{yutJ}$-*lacZ*) Cm^r	pES9911→168
ES9912	*trpC2 amyE*::pES9912 (P$_{yjlCD}$-*lacZ*) Cm^r	pES9912→168
ES001	*trpC2 sodF*::pES001 (*sodF'-lacZ*) Erm^r	pES001→168
ES002	*trpC2 sodA*::pES002 (*sodA'-lacZ*) Erm^r	pES002→168
ES003	*trpC2 ilvC*::pES003 (*ilvC'-lacZ*) Erm^r	pES003→168
ES004	*trpC2 cysK*::pES004 (*cysK'-lacZ*) Erm^r	pES004→168

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
ES007	*trpC2 sodA*::*spc* Spc^r	pES007→168
ES13	*trpC2 sodA*::*spc sodF*::pES001 (*sodF'-lacZ*) Spc^r Erm^r	ES001→ES007
ES14	*trpC2 amyE*::pES14 (P$_{sodA}$-*lacZ*) Cm^r	pES14→168
ES15	*trpC2 amyE*::pES15 Cm^r	pES15→168
ES16	*trpC2 amyE*::pES15 *sodA*::*spc* Cm^r Spc^r	ES007→ES15
ES17	*trpC2 fur*::pES17 Km^r	pES17→168
ES18	*trpC2 yqgB*::pES18 Erm^r	pES18→168
ES19	*trpC2 yqgC*::pES19 Erm^r	pES19→168
ES20	*trpC2 amyE*::pES14 (P$_{sodA}$-*lacZ*) *yqgB*::pES18 Cm^r Erm^r	ES18→ES14
ES21	*trpC2 amyE*::pES14 (P$_{sodA}$-*lacZ*) *yqgC*::pES19 Cm^r Erm^r	ES19→ES14
ES22	*trpC2 amyE*::pES22 (P$_{sodA}$-*lacZ*) Cm^r	pES22→168
ES23	*trpC2 sodF*::pES001 (*sodF'*-lacZ) fur::pES17 Erm^r Km^r	ES17→ES001
ES24	*trpC2 amyE*::pES15 *fur*::pES17 Cm^r Km^r	ES17→ES15
ES25	*trpC2 amyE*::pES14 (P$_{sodA}$-*lacZ*) *fur*::pES17 Cm^r Km^r	ES17→ES14
ES27	*trpC2 amyE*::pES27 (P$_{xyl}$-*yjlC*) Cm^r	pES27→168
ES28	*trpC2 amyE*::pES28 (P$_{xyl}$-*yjlD*) Cm^r	pES28→168
ES29	*trpC2 amyE*:: pES28 (P$_{xyl}$-*yjlD*) *yjlD*::pjlD.1 Cm^r Erm^r	BFS841→ES28
ES30	*trpC2 amyE*:: pES27 (P$_{xyl}$-*yjlC*) *yjlC*::pjlC.1.1 (P$_{SPAC}$-*yjlD*) Cmr Erm^r	BFS842→ES27

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
ES31	*trpC*2 *yojM*::pES31 (P$_{SPAC}$-*yojM*) Erm^r	pES31→168
ES32	*trpC*2 *sodF*::pES32 (P$_{SPAC}$-*sodF*) Erm^r	pES32→168
ES37	*trpC*2 *amyE*::pES14 (P$_{sodA}$-*lacZ*) *sigB*::*spc* Cm^r Spc^r	PB344→ES14
ES38	*trpC*2 *amyE*::pES38 Cm^r	pES38→168
ES39	*trpC*2 *amyE*::pES38 *sodA*::*spc* Cm^r Spc^r	ES007→ES38
ES40	*trpC*2 *amyE*::pES38 *fur*::pES17 Cm^r Km^r	ES17→ES38
Plasmids		
pDIA5304	pBluescript SK$^+$ with a chloramphenicol resistance cassette cloned into the *Nae*I site (Ap^r Cm^r)	(Glaser *et al.*, 1993)
pMUTIN4	Integration vector containing the promoterless *E. coli lacZ* gene and the inducible Pspac promoter (Ap^r Erm^r)	(Vagner *et al.*, 1998)
pMUTIN4XZ	pMUTIN4 derivative where the *lacZ* gene has been excised (Ap^r Erm^r)	B. Jester, unpublished data
pDG268	Integration vector for the introduction of transcriptional fusions to *lacZ* by double crossover at the *amyE* locus (Ap^r Cm^r)	(Antoniewski *et al.*, 1990)
pAC5	Integration vector for the introduction of translational fusions to the ninth codon of the *lacZ* gene by double crossover at the *amyE* locus (Ap^r Cm^r)	(Martin-Verstraete *et al.*, 1992)
pDG780	pBluescript KS$^+$ with a kanamycin resistance cassette cloned into the *Cla*I site (Ap^r Km^r)	(Guerout-Fleury *et al.*, 1995)

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pDG1727	pMTL23 with a spectinomycin resistance cassette cloned into the *Hind*III/*Bgl*II sites (Ap^r Spc^r)	(Guerout-Fleury *et al.*, 1995)
pX	Integration vector which allows genes to be placed under the control of a xylose inducible promoter at the amylase locus (Ap^r Cm^r)	(Kim *et al.*, 1996)
pCheV3	pDIA5304 containing a PCR amplified fragment of *cheV* cloned into *Hind*III	S. Krogh Devine, unpublished data
pCheVSalI	Rescued plasmid obtained by restriction digestion of EScheV chromosomal DNA with *Sal*I and ligation of the digested DNA	This work
pCheVEcoRI	Rescued plasmid obtained by restriction digestion of EScheV chromosomal DNA with *Eco*RI and ligation of the digested DNA	This work
pDU1831	pMUTIN4 containing a PCR amplified internal fragment of *ykuB* cloned into *Eco*RI-*Bam*HI	This work
pDU1832	pMUTIN4 containing a PCR amplified internal fragment of *ykuC* cloned into *Eco*RI-*Bam*HI	This work
pDU1833	pMUTIN4 containing a PCR amplified internal fragment of *ykuD* cloned into *Eco*RI-*Bam*HI	This work
pDU1834	pMUTIN4 containing a PCR amplified internal fragment of *ykuE* cloned into *Eco*RI-*Bam*HI	This work

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pDU1835	pMUTIN4 containing a PCR amplified internal fragment of *ykuG* cloned into *Eco*RI-*Bam*HI	This work
pDU1836	pMUTIN4 containing a PCR amplified internal fragment of *ykuH* cloned into *Eco*RI-*Bam*HI	This work
pDU1837	pMUTIN4 containing a PCR amplified internal fragment of *ykuI* cloned into *Eco*RI-*Bam*HI	This work
pDU1838	pMUTIN4 containing a PCR amplified fragment of the 5' end of *ykuJ* cloned into *Eco*RI-*Bam*HI	This work
pDU1839	pMUTIN4 containing a PCR amplified internal fragment of *ykuK* cloned into *Eco*RI-*Bam*HI	This work
pDU1840	pMUTIN4 containing a PCR amplified internal fragment of *ykuL* cloned into *Eco*RI-*Bam*HI	This work
pDU1841	pMUTIN4 containing a PCR amplified internal fragment of *ykuM* cloned into *Eco*RI-*Bam*HI	This work
pDU1862	pMUTIN4 containing a PCR amplified internal fragment of *ykuU* cloned into *Eco*RI-*Bam*HI	This work
pDU1863	pMUTIN4 containing a PCR amplified internal fragment of *ykuV* cloned into *Eco*RI-*Bam*HI	This work

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pjlD.1	pMUTIN4 containing a PCR amplified internal fragment of *yjlD* cloned into *Eco*RI-*Bam*HI	D. Karamata, unpublished data
pjlC.1.1	pMUTIN4 containing a PCR amplified internal fragment of *yjlC* cloned into *Eco*RI-*Bam*HI	D. Karamata, unpublished data
pCM991	pMUTIN4 containing a PCR amplified internal fragment of *dhbB* cloned into *Eco*RI-*Bam*HI	C. Murtagh, unpublished data
pCM992	pMUTIN4 containing a PCR amplified internal fragment of *leuC* cloned into *Eco*RI-*Bam*HI	C. Murtagh, unpublished data
pES991	pDG268 containing a PCR amplified fragment spanning the upstream untranslated region of *katA* cloned into *Eco*RI-*Bam*HI	This work
pES993	pDG268 containing a PCR amplified fragment spanning the upstream untranslated region of *ahpC* cloned into *Eco*RI-*Bam*HI	This work
pES998	pMUTIN4 containing a PCR amplified internal fragment of *yutJ* cloned into *Eco*RI-*Bam*HI	This work
pES999	pMUTIN4 containing a PCR amplified internal fragment of *yumB* cloned into *Eco*RI-*Bam*HI	This work
pES9910	pDG268 containing a PCR amplified fragment spanning the upstream untranslated region of *yumB* cloned into *Eco*RI-*Bam*HI	This work

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pES9911	pDG268 containing a PCR amplified fragment spanning the upstream untranslated region of *yutJ* cloned into *Eco*RI-*Bam*HI	This work
pES9912	pDG268 containing a PCR amplified fragment spanning the upstream untranslated region of *yjlCD* cloned into *Hind*III-*Bam*HI	This work
pES001	pMUTIN4 with a PCR amplified internal fragment of *sodF* cloned into *Eco*RI-*Bam*HI	This work
pES002	pMUTIN4 containing a PCR amplified internal fragment of *sodA* cloned into *Eco*RI-*Bam*HI	This work
pES003	pMUTIN4 containing a PCR amplified internal fragment of *ilvC* cloned into *Eco*RI-*Bam*HI	This work
pES004	pMUTIN4 containing a PCR amplified internal fragment of *cysK* cloned into *Eco*RI-*Bam*HI	This work
pES007	pDG1727 containing a PCR-amplified fragment spanning the 5' end of *sodA* cloned into *Bam*HI/*Pst*I and a second PCR-amplified fragment spanning the 3' end of *sodA* cloned into *Xho*I	This work
pES14	pDG268 containing a PCR-amplified fragment containing the *sodA* control region cloned into *Eco*RI and *Bam*HI	This work

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pES15	pDG268 with a PCR-amplified fragment containing the *yojM* 5' untranslated region cloned into *Eco*RI and *Bam*HI	This work
pES17	pDG780 containing a PCR amplified internal fragment of *fur* cloned into *Eco*RI and *Bam*HI	This work
pES18	pMUTIN4XZ containing a PCR amplified internal fragment of *yqgB* cloned into *Eco*RI-*Bam*HI	This work
pES19	pMUTIN4XZ containing a PCR amplified internal fragment of *yqgC* cloned into *Eco*RI-*Bam*HI	This work
pES22	pAC5 containing a PCR-amplified fragment containing the *sodA* control region cloned into *Eco*RI and *Bam*HI	This work
pES27	pX containing a PCR amplified fragment of the *yjlC* open reading frame cloned into *Bam*HI	This work
pES28	pX containing a PCR amplified fragment of the *yjlD* open reading frame cloned into *Bam*HI	This work
pES31	pMUTIN4 containing a PCR amplified fragment of the 5' end of *yojM* including the ribosome binding site cloned into *Eco*RI-*Bam*HI	This work
pES32	pMUTIN4 containing a PCR amplified fragment of the 5' end of *sodF,* including the ribosome binding site, cloned into *Eco*RI-*Bam*HI	This work

Table 2.1 -Continued

Strain or Plasmid	Relevant characteristics	Source/Reference
pES38	pDG268 with a PCR-amplified fragment containing the *yojL* 5' untranslated region cloned into *Eco*RI and *Bam*HI	This work

[a] → indicates strain construction by transformation of the strain on the right of the arrow with the plasmid or chromosomal DNA on the left of the arrow.

CHAPTER 3

THE *ykwC-rok* REGION OF THE *B. subtilis* CHROMOSOME: SEQUENCING, ANNOTATION AND FUNCTIONAL ANALYSIS

3.1 INTRODUCTION

Bacillus subtilis is one of the best characterised bacteria, following a long history of investigative research. The traditional interest in *B. subtilis* derived from three features: (i) it secretes large quantities of enzymes that are used in the food and brewing industries, (ii) it sporulates and therefore was a simple model for development and (iii) it can become competent naturally, and a laboratory procedure was established for transformation thereby facilitating genetic analysis. These factors, as well as the availability of comprehensive genetic and physical maps, led to the decision to sequence the *B. subtilis* genome. The sequencing project was initiated by 5 European laboratories. By the time the completion of the genome sequence was announced in 1997, 34 laboratories from Europe, Japan and Korea and two biotechnology companies had participated in the project. The complete genome comprised 4,214,810 base pairs encoding 4,100 putative open reading frames approximately 40% of which encode genes of unknown function (Kunst *et al.*, 1997).

The knowledge of the genome sequence has allowed the functional analysis of genes of unknown function by phenotypic and expression analysis of their mutant strains. These experiments were carried out by European and Japanese consortia during the course of the *B. subtilis* Functional Analysis Project. Furthermore, the genome sequence has allowed the global analysis of the transcriptome using micro and macroarrays. It has also facilitated proteomics, as the spots on 2D protein gels can now be identified by comparison of N-terminal sequences with the genome sequence. Whole genome analysis has been carried out for heat shock, salt and ethanol stress, the stringent response, catabolite repression, early and middle sporulation and for many two component system regulons and the ComK regulon (Helmann *et al.*, 2001; Petersohn *et al.*, 2001; Price *et al.*, 2001; Eymann *et al.*, 2002; Moreno *et al.*, 2001; Fawcett *et al.*, 2000; Kobayashi *et al.*, 2001; Ogura *et al.*, 2001; Ogura *et al.*, 2002). This wealth of data contributes to knowledge of how the cell works and brings closer the elucidation of the complete formula of life of a bacterium.

3.2 GENOME SEQUENCING: RESULTS AND DISCUSSION

The *B. subtilis* chromosomal region between *xre* and *fruA* was sequenced in Kevin Devine's laboratory. Within this region I sequenced a 29.324 kb section between *ykwC* and *rok* (formerly *ykuW*). The cloning and sequencing strategies used included chromosome walking, long range PCR and oligonucleotide walking (Chapter 2, section 2.2.2). Open reading frames (ORFs) were annotated as described in Chapter 2 (section 2.2.3 and 2.2.4). The DNA sequence of the fragment was submitted to the EMBL nucleotide sequence database under the accession number AJ222587. Intergenic DNA regions were searched for promoters using the consensus sequences described by Haldenwang (1995). Intergenic regions were searched for promoters with a maximum of two mismatches from the consensus, except for the σ^K consensus where only one mismatch was allowed. The spacer region between the –35 and –10 was allowed to vary by one base (plus or minus) from that described by Haldenwang (1995). The results are shown in Table 3.1. The presence of promoter consensus sequences does not, however, conclusively signify that the relevant gene is under the control of that promoter, as other regulatory elements may be operative at that locus. In addition, the expression profiles of some of the genes examined show differing levels of expression of genes that appear to be in an operon together (*eg*. The *ykuH-ykuI* operon, section 3.2.12). This indicates the existence of as yet unidentified regulators of gene expression. Therefore the promoters identified here based on consensus searches have limitations in their predictive value. Identification and location of rho-independent terminators has been found to be the most reliable method to deduce operon structure bioinformatically. The terminators in the *ykwC-rok* region are shown in Figure 3.1. A bioinformatic analysis of each operon is presented. Genes are grouped in putative operons. If the genome coordinates and protein sizes are given for more than one gene the information is listed starting with the first gene of the operon. For promoters, the consensus sequence used in the search and the potential promoter sequence(s) found are detailed in Table 3.1. Putative ribosome binding sites (RBS) are presented with the Shine-Dalgarno sequence complementary to the 3' end of the 16S rRNA of *B. subtilis* in capitals and the start codon underlined in upper case bold text. The sequence of the 3' end of the 16S rRNA is UCUUUCCUCCACUAG (McLaughlin *et al.*, 1981). The sequences of rho- independent terminators are shown. The stem of the stem loop of the terminator is in uppercase bold letters and is underscored >>>> <<<<. The loop is in lower case letters. The run of T residues is underscored ^^^^. Paralogues are defined as homologues in the *B. subtilis* genome where the level of similarity as defined by
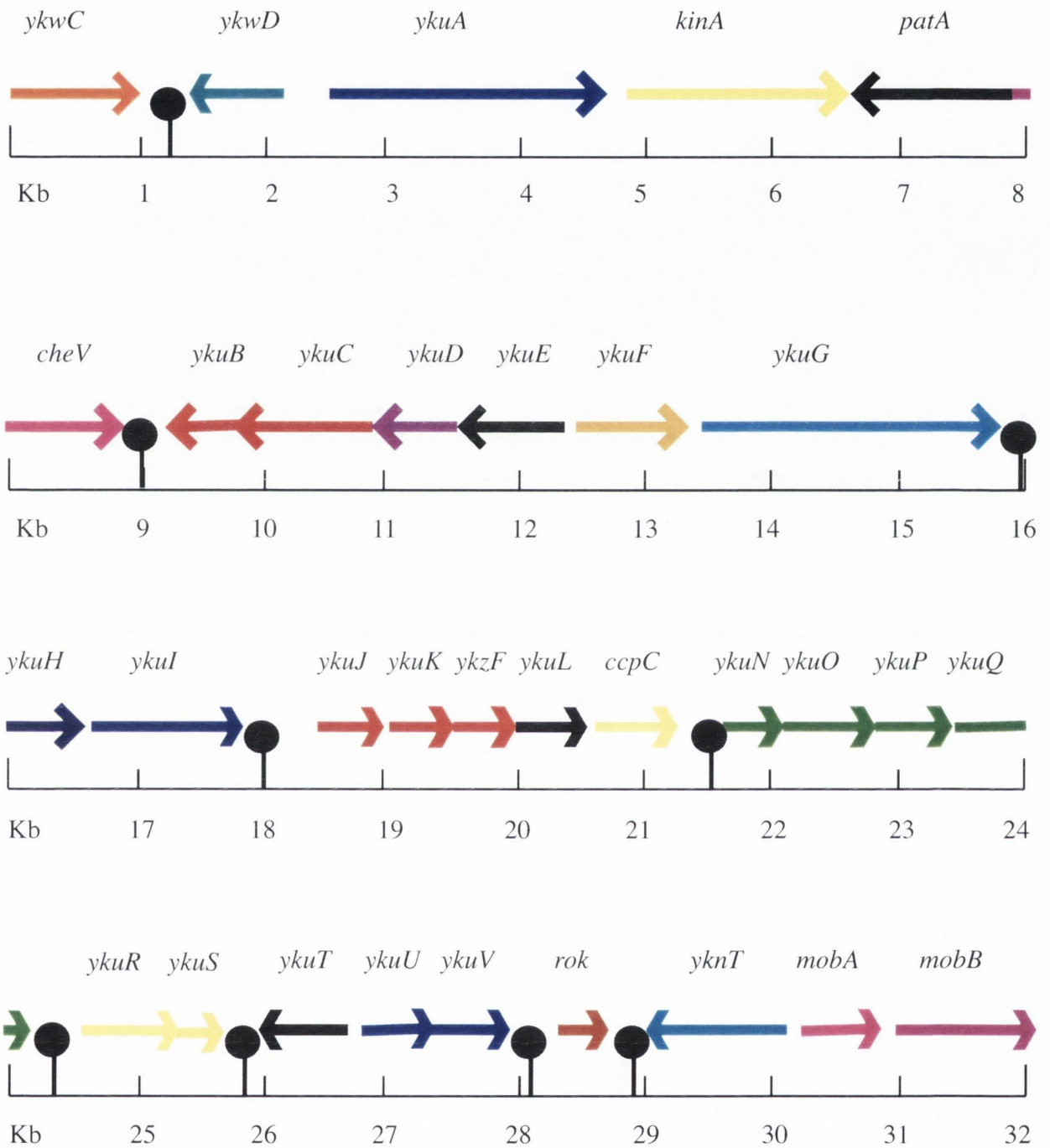
Figure 3.1 Diagrammatic representation of the open reading frame organisation in the region of the *B. subtilis* genome between *ykwC* and *mobB*. Arrows represent open reading frames. Lollipops represent rho- independent transcriptional terminators.

the e (expect) value is e^{-06} or less. This value defines the likelihood of the match occuring by chance in a database of a particular size.

3.2.1 The *ykwC* operon

Putative operon: monocistronic

Genome coordinates (direction): 1465037-1465900 (+)

Size (amino acids): 288

Potential promoters: None detected

RBS: AtctgGGAGGaatgacc**TTG**

Terminator: (1465902-1465940) bidirectional

<div align="center">

5'-AAAAAAA**GATTGCCTGT**acgaa**ACAGGCAATC**TTTTTTT-3'

^^^^^^^^^^^>>>>>>>>>>>   <<<<<<<<<<<^^^^^^^^^^

</div>

Paralogues: *yfjR, gntZ, yqjI, yqeC*

Predicted function: *ykwC* is a gene of unknown function that is similar to 3-hydroxybutyrate dehydrogenase (*e.g.* 3-hydroxyisobutyrate dehydrogenase, *Clostridium acetobutylicum*, level of similarity $2e^{-86}$). This enzyme is involved in valine catabolism. The *ykwC* gene is predicted to be monocistronic as the open reading frame is immediately followed by a putative rho-independent transcriptional terminator. This gene is regulated by the catabolite repression in a CcpA-dependent manner. This repression occurs both in the presence or absence of glucose, however the genes are more highly repressed in the absence of glucose and this gives the overall effect of glucose activation (Moreno *et al.*, 2001).

3.2.2 The *ykwD* operon

Putative operon: monocistronic

Genome coordinates (direction): 1465945-1466715 (-)

Size (amino acids): 257

Potential promoters: σ^H

RBS: AGGAGGTtAcgaactaa**ATG**

Terminator: The *ykwC* terminator could act as a terminator for *ykwD*

Paralogues: *ylbC*

Predicted function: The *ykwD* gene is divergently transcribed from the upstream gene *ykuA* (Figure 3.1). Its function is unknown and BLAST searches reveal similarities to other proteins of unknown function in *B. subtilis* and other organisms. The upstream

region contains a potential σ^H promoter sequence (Table 3.1). Genes under σ^H control are expressed postexponentially and are often involved in competence and sporulation.

3.2.3 The *ykuA* operon

Putative operon: monocistronic

Genome coordinates (direction): 1467109-1469163 (+)

Size (amino acids): 685

Potential promoters: σ^H

RBS: AgaGgGGTaacCagaa**GTG**

Terminator: None detected

Paralogues: *pbpA, spoVD, pbpB, yrrR*

Predicted function: The *ykuA* gene is homologous to penicillin binding proteins (*e.g.* PBP2b, *Staphylococcus aureus*, e-117). This group of enzymes are involved in peptidoglycan biosynthesis. β-Lactam antibiotics disrupt peptidoglycan biosynthesis by binding to these enzymes. The upstream untranslated region contains a σ^H promoter sequence with one mismatch from the consensus (Table 3.1). Recent array experiments give a further clue to the regulation of *ykuA*. Transcription of *ykuA* increased in cells expressing a constitutively active form of SpoOA during growth (Fawcett *et al.*, 2000). It remains to be established whether *ykuA* is controlled directly or indirectly by SpoOA.

3.2.4 the *kinA* operon

Putative operon: monocistronic

Genome coordinates (direction): 1469330-1471147 (+)

Size (amino acids): 606

Potential promoters: σ^G, σ^H

RBS: AAAGGAGGgatTCt**GTG**

Terminator: None detected

Paralogues: *kinE, kinC, kinB* and other sensor kinases

Predicted function: Sporulation in *B. subtilis* is controlled by a two component signal transduction phosphorelay. KinA is one of the sensor histidine kinases that participates in phosphorylation of SpoOA in the phosphorelay (Grimshaw *et al.*, 1998). The exact nature of the regulation of *kinA* and the signals to which it responds have not been elucidated. Recent studies have pointed to the involvement of an ATP-binding domain of *kinA* called PAS-A (one of three PAS domains of *kinA*) in signal sensing. Mutation of this region of

the protein, leading to increased affinity for ATP, results in stimulation of sporulation (Stephenson and Hoch, 2001).

3.2.5 The *patA* operon

Putative operon: monocistronic

Genome coordinates (direction): 1471164-1472339 (-)

Size (amino acids): 392

Potential promoters: σ^A, σ^E, σ^G, σ^L

RBS: AAGGtGaTGAatt**ATG**

Terminator: None detected

Paralogues: *alaT, aspB, ykrV* and others

Predicted function: This monocistronic gene is predicted to be an aminotransferase. Aminotransferases (also known as transaminases) catalyse the transfer of amino groups from α-amino to α-keto acids. There are several potential sigma factor promoter sequences upstream of the *patA* start codon including those for σ^A, σ^G, σ^E and σ^L (Table 3.1). Genes under σ^A control are usually housekeeping genes. Genes controlled by σ^G are often expressed late in the forespore. Genes controlled by σ^E are expressed early in the mother cell. Genes controlled by σ^L are often degradative enzymes. Recent evidence suggests that *patA* is under σ^E control. Transcriptional analysis of early to middle sporulation shows that *patA* expression is higher in the wild type than in both a *SpoOA* mutant and in a *sigF* mutant two hours after the end of exponential growth (Fawcett *et al.*, 2000). These results are consistent as σ^E expression requires the prior activity of σ^F.

3.2.6 The *cheV* operon

Putative operon: monocistronic

Genome coordinates (direction): 1472906-1473814 (+)

Size (amino acids): 303

Potential promoters: σ^D

RBS: AatAgcGAGGgatTgaa**TTG**

Terminator: (1473820-1465940) bidirectional

```
        5'-AAAAACAGCCGTTGCCagaaagaGGCACGGCTGTTTTT-3'
           ^^^^^^^>>>>>>>>>>>>>>       <<<<<<<<<<<<<^^^^^
```

Paralogues: Other response regulators including *yycF, yvrH, phoP, cssR* and *yrkP*

Predicted function: *cheV* is σ^D dependant and it is involved in chemotaxis. The σ^D promoter sequence has two mismatches from the consensus and has been verified

experimentally (Rosario *et al.*, 1994). The chemotaxis system is well characterised. Transmembrane methyl accepting chemotaxis proteins (MCPs), which are the receptors of the system, are coupled via the adaptor protein CheW to the central two component histidine kinase CheA. Phosphorylated CheA subsequently transfers the phosphoryl group to the three response regulators of the network. The response regulator CheB demethylates the MCPs when phosphorylated. The main response regulator responsible for motility, CheY, when phosphorylated, binds to the flagellar switch resulting in counter clockwise movement which results in smooth swimming. CheV is the third response regulator. Phosphorylated CheV is important for the response to the chemoattractant asparagine. CheV is a two domain protein. The N-terminal part of the protein contains a CheW-like domain and the C-terminal contains a two component receiver domain. The CheW-like domain of the CheV protein is important for the stability of the phosphorylated form of the protein and for the pace of the phospho-transfer reaction from CheA. In addition, CheV is also partially functionally redundant with the adaptor protein CheW. While mutants of either protein alone have reduced chemotaxis, a double mutant is non-chemotactic (Karatan *et al.*, 2001). Therefore CheV can partially compensate for CheW. CheV is a very interesting protein as it functions to couple the MCPs to CheA and this coupling is not dependent on phosphorylation of the receiver domain. In addition CheV is required in its phosphorylated form to function as a response regulator for adaptation to asparagine (Karatan *et al.*, 2001). The stop codon of *cheV* is followed by a probable rho-independent terminator sequence.

3.2.7 The *ykuC-ykuB* operon

Putative operon: dicistronic

Genome coordinates (direction): 1474454-1475743; 1473864-1474325 (-)

Size (amino acids): 430; 154

Potential promoters: *ykuC*: None detected

                       ykuB: σ^A

RBS: *ykuC*: AAGGAGcgttTtt**ATG**

       ykuB: AAAGGAGagaAata**ATG**

Terminator: The *cheV* terminator could act as a terminator for *ykuB*

Paralogues: *ykuC*: *yfiS, yjbB, yvqJ*

             ykuB: None

Predicted function: The *ykuC* open reading frame is homologous to various macrolide efflux determinants (*e.g.* Macrolide efflux determinant, *Streptococcus pneumoniae*, level

of similarity 9e^{-15}). These proteins mediate resistance to macrolide antibiotics via an efflux pump. The *ykuB* ORF has no significant homologies to any other proteins in the databases. It is preceded by a potential σ^A promoter sequence which has two mismatches from the consensus (Table 3.1). It may share the *cheV* terminator sequence. Interestingly, *ykuB* has been described as being stress inducible (Petersohn *et al.*, 2001).

Expression: The expression profiles obtained for *ykuC* and *ykuB* are shown in Figure 3.2. For *ykuC,* the LacZ activity reached a maximum of 8 units during exponential growth. This is in contrast to the *ykuB* profile. The *ykuB-lacZ* fusion is maximally expressed during exponential growth when it attains 61 units of activity, consistent with the presence of a σ^A promoter. This suggests either that the genes are not in the same operon or that additional regulatory signals are involved in *ykuB* expression.

3.2.8 The *ykuD* operon

Putative operon: monocistronic

Genome coordinates (direction): 1475822-1476313 (-)

Size (amino acids): 164

Potential promoters: σ^K

RBS: AaAAAGGAGGatgaggaaa**CTG**

Terminator: None detected

Paralogues: *yqjB*

Predicted function: The YkuD protein is a spore protein and it contains a cell wall binding motif that is speculated to function in forespore protein assembly. Northern analysis has revealed a 0.9 kb transcript which is sufficient to encode the predicted size of this protein (164 aa) but too small to also include either *ykuE* or *ykuC* mRNA (Kodama *et al.*, 2000). Thus *ykuD* is predicted to be monocistronic. YkuD is expressed from T_4 of sporulation. Northern analysis with sigma factor mutants has shown the transcript disappears in *sigF*, *sigE*, *sigG* and *sigK* mutants. However the transcript was present in a *gerE* mutant. This implied that *ykuD* is regulated by σ^K. GerE is a DNA-binding protein. It is involved in regulation of some genes that are regulated by σ^K. In the case of *ykuD* it appears that GerE has a negative effect, as *ykuD* transcript is increased in a *gerE* mutant (Kodama *et al.*, 2000). The upstream region of *ykuD* contains a σ^K consensus sequence which has three mismatches to the consensus. There is also a possible GerE binding site which has a single base difference from the consensus (Table 3.1) (Takamatsu *et al.*, 1998). In addition, microarray analysis of two component systems has revealed a further level of *ykuD* regulation. This gene is upregulated by the YdfH-YdfI two component

(A)
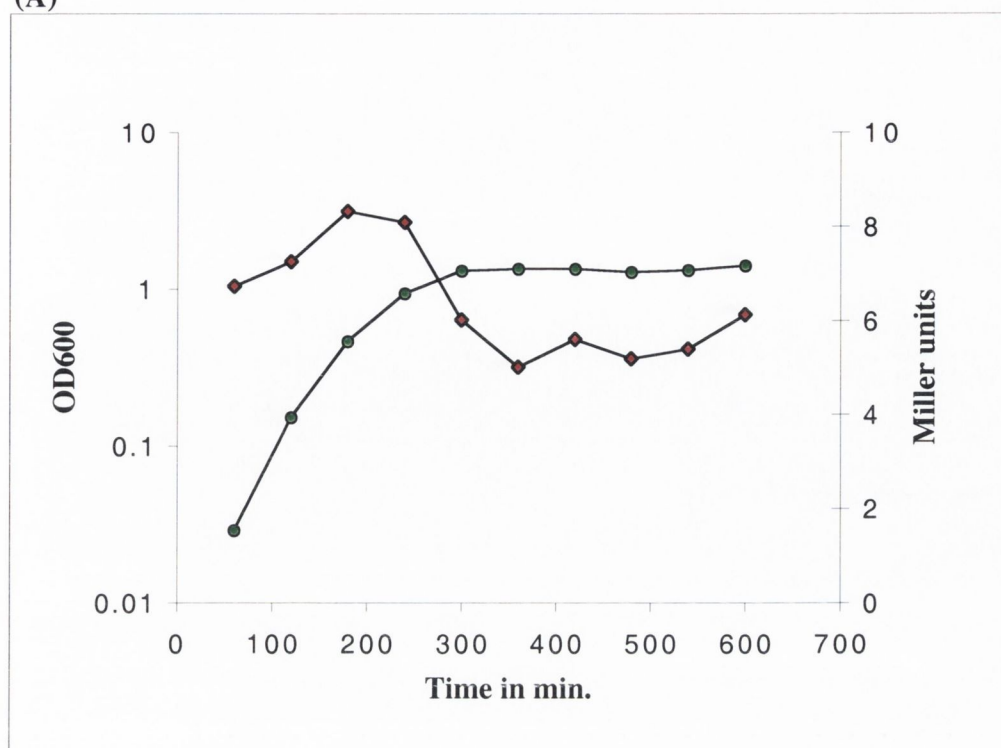
(B)

Figure 3.2 Growth curves and LacZ expression profiles of strains (A) BFS1832 (*ykuC-lacZ*) and (B) BFS1831 (*ykuB-lacZ*).
Cell growth was monitored by measuring OD_{600} and is shown by green circles. LacZ activity was measured in Miller units and is represented by red diamonds. Time is presented in minutes after inoculation. The scale of the right y-axis is different in the two graphs.

system. In these experiments gene expression was examined in cells where the response regulator was overproduced and the cognate sensor kinase was deleted (Kobayashi *et al.*, 2001).

Expression: The *ykuD-lacZ* fusion was lowly expressed during exponential growth but reached 8 activity units in stationary phase (Figure 3.3). This profile is consistent with published results on *ykuD* that show that this is a sporulation gene expressed from T_4 of sporulation onwards (Kodama *et al.*, 2000).

3.2.9 The *ykuE* operon

Putative operon: monocistronic

Genome coordinates (direction): 1476373-1477233 (-)

Size (amino acids): 287

Potential promoters: None detected

RBS: AAAGGttGgtttgaa**ATG**

Terminator: None detected

Paralogues: *ykoQ, ypbG*

Predicted function: The *ykuE* ORF is similar to other proteins of unknown function in *B. subtilis* (*ykoQ* and *ypbG*) as well as in other organisms. It also contains a region similar to the Pfam02549 protein motif which is shared by a family of proteins involved in DNA repair (**http://www.sanger.ac.uk/Pfam**).

Expression: The *ykuE-lacZ* fusion is very lowly expressed reaching a maximum of only 5 units of activity during stationary phase (Figure 3.3).

3.2.10 The *ykuF* operon

Putative operon: monocistronic

Genome coordinates (direction): 1477376-1478137 (+)

Size (amino acids): 254

Potential promoters: σ^A, σ^E, σ^G, σ^K

RBS: AttGGAGGgataaac**ATG**

Terminator: None detected

Paralogues: *gdh, ycdF, yxbG* and other dehydrogenases

Predicted function: It is unclear whether *ykuF* is monocistronic or transcribed as part of an operon with *ykuG*. There is an untranslated intergenic region of 300 base pairs between the two genes and it is possible it contains regulatory elements for expression of *ykuG*. YkuF is homologous to glucose dehydrogenases (e.g. Glucose-1-dehydrogenase, *Bacillus*

(A)

(B)

(C)
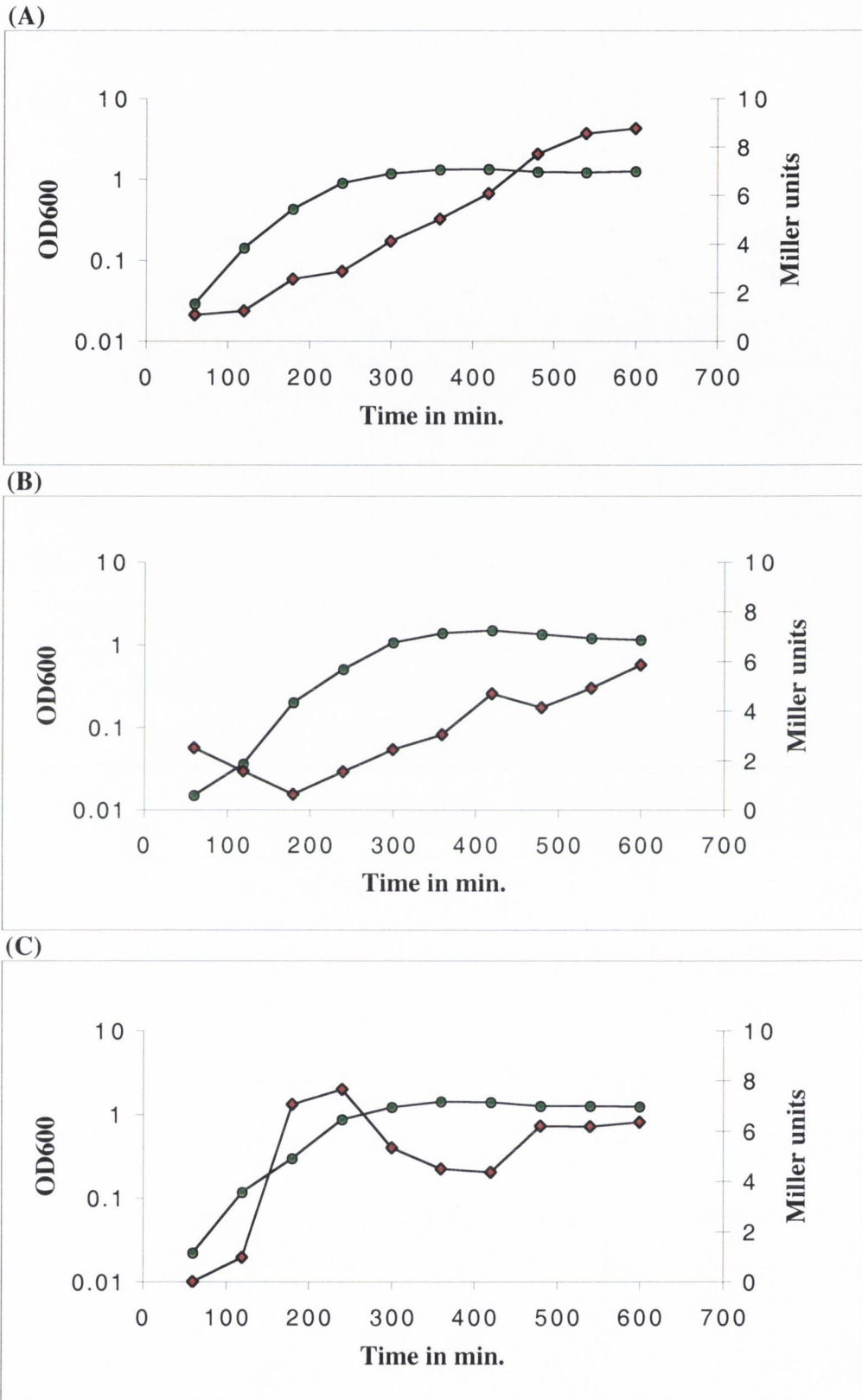
Figure 3.3 Growth curves and LacZ expression profiles of strains (A) BFS1833 (*ykuD-lacZ*) ,
(B) BFS1834 (*ykuE-lacZ)* and (C) BFS1835 (*ykuG-lacZ).*
Cell growth was monitored by measuring OD600 and is shown by green circles. LacZ activity
was measured in Miller units and is represented by red diamonds. Time is presented in
minutes after inoculation.

megaterium, level of similarity 8e^{-35}). Its sequence is similar to the Pfam00106 family of short chain dehydrogenases. Potential promoter sequences motifs include those for σ^E, σ^K, σ^G and σ^A (Table 3.1).

3.2.11 The *ykuG* operon

Putative operon: monocistronic

Genome coordinates (direction): 1478441-1480720 (+)

Size (amino acids): 760

Potential promoters: σ^G

RBS: GGAGGaaggtaatata**ATG**

Terminator: (1480767-1480801) bidirectional

> 5'-**CAAGCGAACG**ttaaaactggt**CGTTCGCTTG**TTTT-3'
> >>>>>>>>>>>>>   <<<<<<<<<<<<<^^^^^

Paralogues: *ybfG*

Predicted function: *ykuG* is a gene of unknown function with similarities to other genes of unknown function. Similarities to sequences in the Pfam database (Pfam 01471) suggest that *ykuG* may contain peptidoglycan binding domains. The intergenic region contains a possible σ^G promoter motif (Table 3.1). The gene is followed by a rho-independent terminator sequence.

Expression: The *ykuG-lacZ* fusion shows very low expression reaching a maximum level of 6 activity units during exponential growth (Figure 3.3).

3.2.12 The *ykuH-ykuI* operon

Putative operon: dicistronic

Genome coordinates (direction): 1480853-1481398; 1481554-1482774 (+)

Size (amino acids): 182; 407

Potential promoters: *ykuH*: σ^G

  ykuI: None detected

RBS: *ykuH*: AttAAGGAGtgatTCa**TTG**

  ykuI: AcAAAGataGaGAagcaggtgataac**ATG**

Terminator: (1482778-1482799)

> 5'-**CGGC**tgaaag**GCCG**TTTTTTTT-3'
> >>>>>>   <<<<<^^^^^^^^^^

Paralogues: *ykuH*: None

  ykuI: *ykoW*

Predicted function: The two genes, *ykuH* and *ykuI* possibly constitute a dicistronic operon. The putative operon is preceded by a potential σ^G motif. The function of *ykuH* is unknown and it does not have any similarities to other proteins. The function of *ykuI* is also unknown. YkuI contains an EAL domain (Pfam 00563) which is a conserved domain of unknown function in bacteria. After the *ykuI* ORF there is a possible terminator sequence.

Expression: The pattern of expression for both *ykuH* and *ykuI* is similar although the level of expression is not (Figure 3.4). For both genes the highest level of expression is observed during exponential growth and expression decreases 4-5 fold in stationary phase. For *ykuH* the highest level is 23 activity units, whereas for *ykuI* it is 55 units. It is interesting that the higher expression level is observed for the promoter distal gene.

3.2.13 The *ykuJ-ykuK-ykzF* operon

Putative operon: tricistronic

Genome coordinates (direction): 1483423-1483659; 1483772-1484287; 1484424-1484618 (+)

Size (amino acids): 79; 172; 65

Potential promoters: *ykuJ*: σ^A, σ^F, σ^G, σ^H

　　　　　　　　ykuK: None detected

　　　　　　　　ykzF: σ^K

RBS: *ykuJ*: GAAAGGAGtcttTtttttaga**ATG**

　　　ykuK: AAAGGAGtgtcgtg**ATG**

　　　ykzF: AAAcagGGgGcctga**TTG**

Terminator: None detected

Paralogues: *ykuJ*: None

　　　　　　ykuK: None

　　　　　　ykzF: None

Predicted function: Three small ORFs constitute this putative operon. Both *ykuJ* and *ykuK* are similar to genes of unknown function in other organisms. However YkzF has no similarity to any proteins in current databases. The *ykuJ* promoter region contains several promoter motifs, these are detailed in Table 3.1.

Expression: For both *ykuJ* and *ykuK* the highest expression is observed in exponential growth, although *ykuK* reaches a higher level of activity (65 units) than *ykuJ* (49 units) and *ykuJ* is more highly expressed in the early in the growth curve (Figure 3.5).

(A)
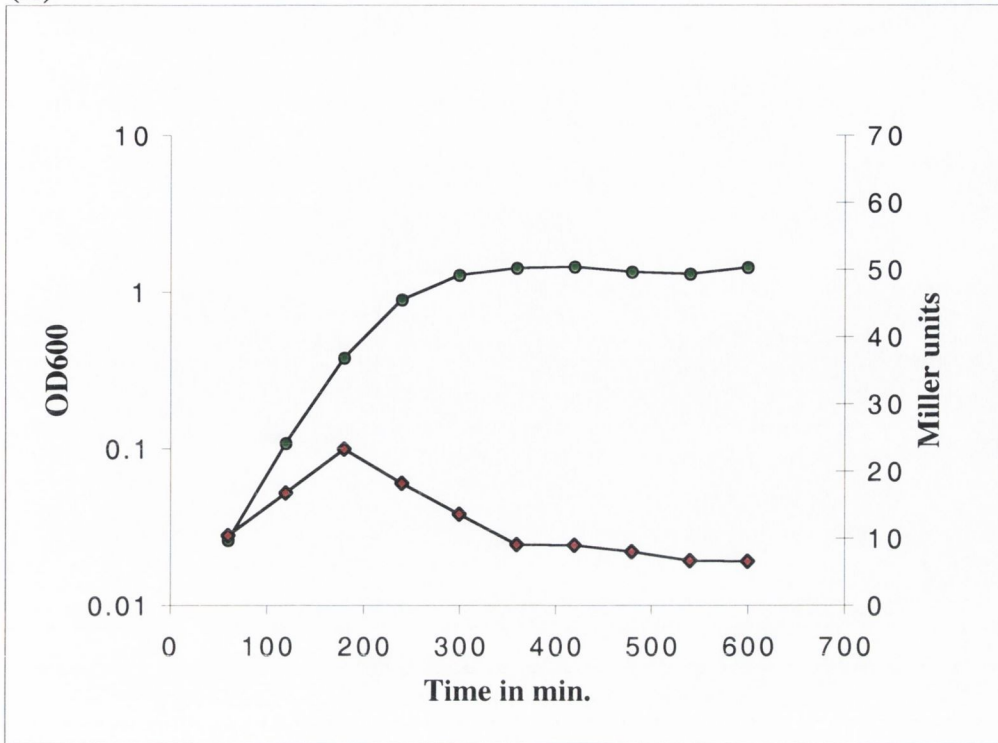
(B)
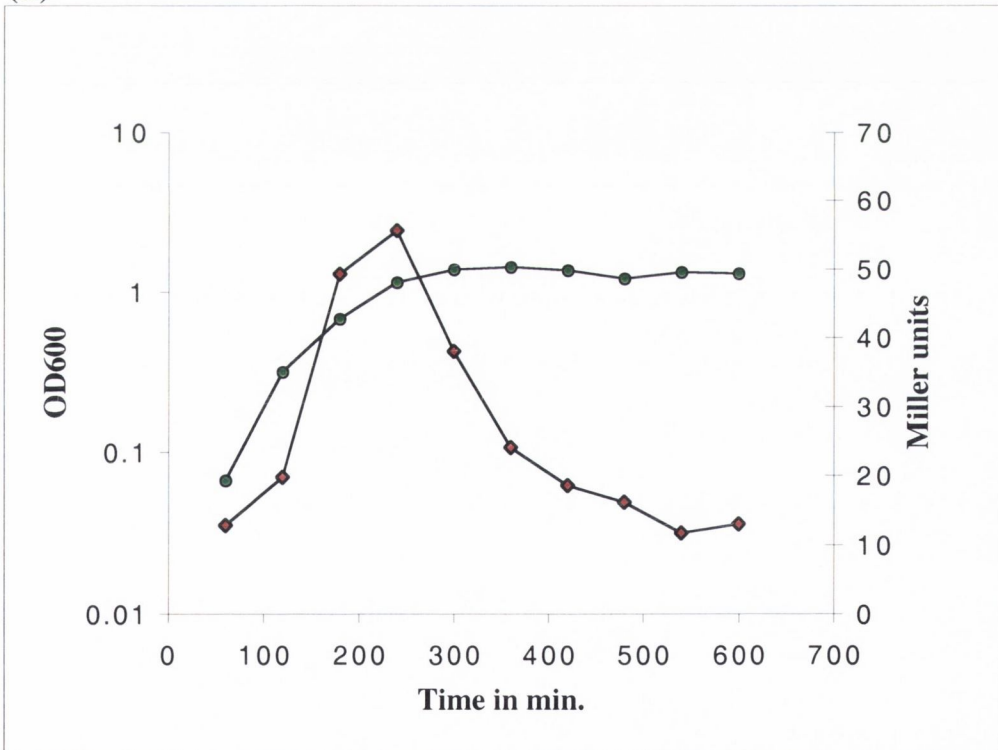
Figure 3.4 Growth curves and LacZ expression profiles of strains (A) BFS1836
 (*ykuH-lacZ*) and (B) BFS1837 (*ykuI-lacZ*).
Cell growth was monitored by measuring OD600 and is shown by green circles. LacZ activity
was measured in Miller units and is represented by red diamonds. Time is presented in
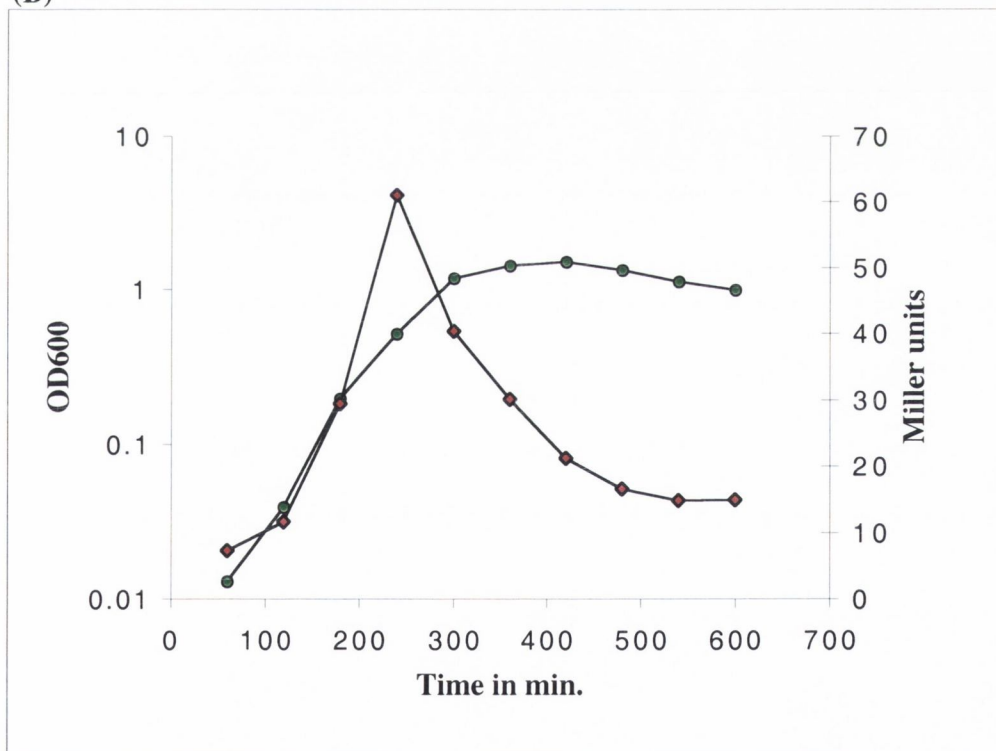minutes after inoculation.

Figure 3.5 Growth curves and LacZ expression profiles of strains (A) BFS1838 (*ykuJ-lacZ*) and (B) BFS1839 (*ykuK-lacZ*).

Cell growth was monitored by measuring OD600 and is shown by green circles. LacZ activity was measured in Miller units and is represented by red diamonds. Time is presented in minutes after inoculation.

3.2.14 The *ykuL* operon

Putative operon: monocistronic

Genome coordinates (direction): 1484759-1485199 (+)

Size (amino acids): 147

Potential promoters: σ^A, σ^K

RBS: AAAGGgGtcGAaaac**ATG**

Terminator: None detected

Paralogues: None

Predicted function: This is a gene of unknown function which is similar to other hypothetical genes of unknown function from other organisms. It contains a CBS domain (Pfam00571). No function has yet been ascribed to CBS domains These domains are found in a wide variety of proteins including inosine monophosphate dehydrogenases. The *ykuL* ORF is preceded by a σ^A motif and a σ^K motif (Table 3.1).

Expression: The *ykuL* gene is expressed during exponential growth and reaches a peak of accumulation (20 units of activity) at the transition point. Expression decreases in stationary phase (Figure 3.6).

3.2.15 The *ccpC* (*ykuM*) operon

Putative operon: monocistronic

Genome coordinates (direction): 1485351-1486229 (+)

Size (amino acids): 293

Potential promoters: None detected

RBS: AAAGGAGGacgaC**ATG**

Terminator: (1486233-1486258)

<div align="center">

5'-**AGACAGC**gaggt**GCTGTCT**TTTTTTT-3'

\>>>>>>>>   <<<<<<<<^^^^^^^^^

</div>

Paralogues: *yraN, ywfK, yrdQ* and other LysR type regulators

Predicted function: CcpC is a LysR-type transcriptional regulator. It participates in carbon catabolite repression and anaerobic repression of *citB* (aconitase) and *citZ* (citrate synthase) (Jourlin-Castelli *et al.*, 2000). The *ccpC* coding sequence is followed by a rho-independent terminator.

Expression: A high level of expression was observed for the *ccpC- lacZ* fusion which was highly expressed throughout exponential growth with a maximal accumulation of 152 units in late exponential phase (Figure 3.6). As described, CcpC is a LysR type transcriptional regulator which represses *citB* and *citZ* during carbon catabolite conditions and

Figure 3.6 Growth curves and LacZ expression profiles of strains (A) BFS1840 (*ykuL-lacZ*) and (B) BFS1841 (*ykuM-lacZ*).
Cell growth was monitored by measuring OD600 and is shown by green circles. LacZ activity was measured in Miller units and is represented by red diamonds. Time is presented in minutes after inoculation. The scale of the right y-axis is different in the two graphs.

anaerobiosis. This level of LacZ activity appears high for a regulator protein. A possible explanation for such high activities may be that many LysR-type regulators often negatively autoregulate their own expression. In this *lacZ* fusion strain the *ccpC* gene is inactivated by the insertion of pMUTIN4. Therefore the expression profile observed may be the unrepressed level, and normal, wild type levels may be lower. Sonenshein *et al.* confirmed that CcpC autoregulates its own expression (Jourlin-Castelli *et al.*, 2000).

3.2.16 The *ykuN-ykuO-ykuP-ykuQ* operon

Putative operon: four cistrons

Genome coordinates (direction): 1486344-1486817; 1486810-1487700; 1487719-1488252; 1488280-1488987 (+)

Size (amino acids): 158; 297; 178; 236

Potential promoters: *ykuN*: σ^A, σ^K

   ykuO: None detected

   ykuP: None detected

   ykuQ: None detected

RBS: *ykuN*: AAtGGgGtgataac**ATG**

   ykuO: AacAAGGAaaaaATCc**ATG**

   ykuP: AgAGGAGGaacaaggaa**TTG**

   ykuQ: AGGAGGactatacat**ATG**

Terminator: (1489010-1489037)

   5'-**AAGGCGTGGA**taatata**TCCACGCCTT**aTT-3'
   >>>>>>>>>>>>   <<<<<<<<<<<< ^^

Paralogues: *ykuN*: *ykuP*

   ykuO: None

   ykuP: *ykuN*, *yvgR*

   ykuQ: *cysE*, *yvfD*, *maa*, *yvoF*, *gcaD*

Predicted function: It appears likely that these four genes are all in the same operon. There is a possible σ^A promoter motif in the control region of the operon. There is also a possible σ^K- type promoter. The *ykuN* gene is probably a flavodoxin and is homologous to flavodoxins in other organisms (*e.g.* Flavodoxin, *S. pneumoniae,* level of similarity 1e^{-21}). The *ykuN* and *ykuO* ORFs overlap by three amino acids. The *ykuO* gene is of unknown function and displays only low homologies to other unknown proteins. There is only a short 15 base pair non-coding region between the stop codon of *ykuO* and the methionine of *ykuP*. Interestingly *ykuP* is a paralogue of *ykuN*, and so is also a probable flavodoxin.

There is a 24 base pair intergenic region separating *ykuP* and *ykuQ*. The fourth gene of the operon *ykuQ* is similar to tetrahydrodipicolinate succinylases in other organisms (*e.g.* tetrahydrodipicolinate succinylase, *Bacillus halodurans,* level of similarity 2e^{-92}). This enzyme is involved in amino acid and cell wall biosynthesis. The *ykuQ* ORF is followed by a potential rho- independent terminator sequence. Intriguingly, microarray analysis revealed that *ykuO,* but none of the other genes of this putative operon, was upregulated upon overproduction of the response regulator of the two component system YclK-YclJ (Kobayashi *et al.*, 2001). Also, microarray analysis of gene expression under anaerobic conditions showed the *ykuNOP* genes were upregulated during anaerobic growth on nitrate or nitrite (Ye *et al.*, 2000).

3.2.17 The *ykuR-ykuS* operon

Putative operon: dicistronic

Genome coordinates (direction): 1489060-1490181; 1490246-1490488 (+)

Size (amino acids): 374; 81

Potential promoters: *ykuR*: σ^G

                          ykuS: σ^G

RBS: *ykuR*: AGgcAGGAGGaatagaa**ATG**

         ykuS: AAcGGAGGTGttCaaaa**ATG**

Terminator: (1490493-1490525) bidirectional

             5'-AAAAAA**CCGAAG**Caaaaa**GCTTCGGT**TcTTTT-3'
             ^^^^^^^^^^>>>>>>>>>>>   <<<<<<<<<<^ ^^^^^

Paralogues: *ykuR*: *yntL, yxeP, yhaA, amhX*

                  ykuS: None

Predicted function: These two genes are possibly cotranscribed as they are separated by only a 60 base pair intergenic sequence. The *ykuR* gene is preceded by a possible σ^G motif. The protein has a conserved sequence homologous to the Pfam 01546 family of peptidases. It is similar to hippurate hydrolases and amino acid hydrolases from other organisms (*e.g.* N-acyl-L-amino acid amidohydrolase, *Listeria innocua*, level of similarity e^{-124}). The second gene of this putative operon, *ykuS*, is a gene of unknown function. It has no significant homologies to any proteins in the current databases. Though the intergenic region upstream of *ykuS* is small it does contain a possible σ^G motif (Table 3.1). In addition, analysis of transcription during mid sporulation has shown that the expression of *ykuS* is higher in the wild type than in both *spoOA* and *sigF* mutants (Fawcett *et al.*, 2000).

Therefore *ykuS* may be directly or indirectly under σ^F control and *ykuR* and *ykuS* may be regulated independently of each other.

3.2.18 The *ykuT* operon

Putative operon: monocistronic

Genome coordinates (direction): 1490531-1491331 (-)

Size (amino acids): 267

Potential promoters: σ^B, σ^G

RBS: AAAGGAGcTcAaC**ATG**

Terminator: The *ykuS* terminator could act as a terminator for *ykuT*

Paralogues: *yfkC, yhdY*

Predicted function: The *ykuT* gene is monocistronic and displays similarities to genes of unknown function in many organisms. It has conserved sequence (Pfam 00924) similar to proteins that form mechanosensitive ion channels. This type of channel has been implicated in osmoregulation in bacteria (Martinac, 2001). Consistently, *ykuT* has been described as a σ^B-dependent general stress gene. This gene is upregulated by ethanol and heat stress (the induction ratios for the wild type/*sigB* mutant were 1.7 and 3.2 respectively). However its involvement in osmotic stress is confirmed by a greater induction ratio after salt stress (induction ratio wt/*sigB* of 9.9) (Petersohn *et al.*, 2001). The upstream region contains a σ^B consensus sequence. There is also a σ^G motif (Table 3.1).

3.2.19 The *ykuU-ykuV* operon

Putative operon: dicistronic

Genome coordinates (direction): 1491568-1492107; 1492182-1492640 (+)

Size (amino acids): 180; 153

Potential promoters: *ykuU*: σ^A, σ^K

  ykuV: σ^G

RBS: *ykuU*: AaAttGGAGGgataCat**ATG**

  ykuV: AaAtAGGAGGgaAaaagca**ATG**

Terminator: (1492652-1492681)

  5'-**AAAAACCTGA**cctca**TCAGGTTTTTT**aTTT-3'
  >>>>>>>>>>>>   <<<<<<<<<<<<^ ^^^

Paralogues: *ykuU*: *ahpC, ygaF*

  ykuV: *yneN, ykvV, ybdE*

Predicted function: These two genes may be coordinately regulated as they appear to have similar oxidoreductase functions. However *ykuV* but not *ykuU,* is induced by ethanol and heat stress in a σ^B independent manner (Petersohn *et al.*, 2001). This implies *ykuV* is regulated independently to *ykuU* under these conditions. The *ykuU* gene is a conserved protein of the AhpC-TSA family (Pfam 00578). This family contains proteins related to alkyl hydroperoxide reductases and thiol-specific antioxidants. Alkyl hydroperoxide reductases, such as *B. subtilis* AhpCF, convert alkyl hydroperoxides to the corresponding alcohol and water. This reaction is carried out via thiol-disulphide exchange on two conserved cysteine residues. BLAST searching reveals YkuU is similar to 2-Cys peroxiredoxins from a variety of organisms. The *ykuV* gene is similar to thioredoxins. Thioredoxins are a family of proteins that efficiently reduce disulphide bonds in a wide variety of proteins via two active site cysteines. They often serve to maintain the thiol-disulphide balance in the cytoplasm. The promoter region of *ykuU* contains potential σ^A and σ^K motifs (Table 3.1). One potential sigma factor motif was detected upstream of *ykuV*, that of σ^G. The *ykuV* gene is followed by a rho-independent terminator.

Expression: The *ykuU* gene is expressed throughout the growth curve with a peak of 30 units at the mid exponential point of the growth cycle. This pattern is very similar to that of *ykuV* though the level of expression is lower for *ykuV* (8 units) (Figure 3.7).

3.2.20 The *rok* (*ykuW*) operon

Putative operon: monocistronic

Genome coordinates (direction): 1493094-1493666 (+)

Size (amino acids): 191

Potential promoters: σ^G, σ^H, σ^K

RBS: GAtAGGgGGTacattgagga**ATG**

Terminator: (1493677-1493713)

<pre>
       5'-AAAAACTGCTTGgcatcgctCAAGCAGTTTTTTTaTT-3'
          ^^^^^^^^^>>>>>>>>>   <<<<<<<<<^^^^^^^^^ ^^
</pre>

Paralogues: None

Predicted function: This gene functions as a repressor of *comK*, the competence transcription factor (Hoa *et al.*, 2002). It is preceded by possible σ^G, σ^H and σ^K promoter motifs. The gene is followed by a terminator.

Figure 3.7 Growth curves and LacZ expression profiles of strains (A) BFS1862 (*ykuU-lacZ*) and (B) BFS1863 (*ykuV-lacZ*).

Cell growth was monitored by measuring OD_{600} and is shown by green circles. LacZ activity was measured in Miller units and is represented by red diamonds. Time is presented in minutes after inoculation. The scale of the right y-axis is different in the two graphs.

3.3 PHENOTYPIC ANALYSIS: RESULTS AND DISCUSSION

The European consortium, supported by the European Union Biotechnology programme carried out a functional analysis of over 1,100 genes of unknown function. Firstly, each gene was inactivated by insertion mutagenesis using the pMUTIN4 plasmid (Chapter 2, Section 2.4.1) (Vagner *et al.*, 1998). This plasmid carries a *lacZ* reporter gene so that the pattern of expression of the gene can be examined. In the region between *ykwC* and *rok* insertional mutants were constructed in thirteen of the open reading frames. Mutants were made in the following genes: *ykuB, ykuC, ykuD, ykuE, ykuG, ykuH, ykuI, ykuJ, ykuK, ykuL, ccpC (ykuM), ykuU* and *ykuV*. Growth curves were carried out in Schaeffer's sporulation medium for each mutant and expression profiles were determined in this null genetic background. Secondly, the collection of mutants was subjected to phenotypic tests. These were designed so that the involvement of the gene in a wide variety of cellular processes could be examined. These included nucleotide and nitrogen metabolism, sporulation, germination, glucose regulation, catabolite repression to name a few. In our laboratory tests were carried out to screen the mutant collection for increased resistance or sensitivity to the oxidising agents paraquat and hydrogen peroxide. We also examined *lacZ* reporter gene expression in response to these oxidising agents (Chapter 2, section 2.7.1) (Scanlan *et al.*, 2001). We tested 1,148 mutant strains for their responses to the oxidising agents paraquat and hydrogen peroxide. Of these, 35 strains displayed a phenotype different to the wild type. Detailed results can be seen on the Micado homepage at **http://locus.jouy.inra.fr/cgi-bin/genomic/madbase_home.pl**. Of the strains tested, 7 were more sensitive than wild type to paraquat and 8 were more resistant. 8 strains appeared to have increased expression of the *lacZ* reporter gene when grown on paraquat. 12 strains were more sensitive to hydrogen peroxide than wild type. Two of the strains which exhibited paraquat resistance, BFS841 and BFS842 mutants in *yjlD* and *yjlC* respectively (constructed in D. Karamata's laboratory) were studied further. The results of this work is described in Chapter five. Of the mutant strains I constructed and submitted to Micado, 6 showed a low level of *lacZ* induction in response to osmotic stress (NaCl) in tests carried out in the laboratory of G. Rapaport. The mutated genes in these strains were: *ykuB, ykuC, ykuI, ykuJ, ykuK* and *ykuL*.

3.4 CONCLUSIONS

This chapter describes work carried out during the *B. subtilis* genome sequencing and functional analysis projects. Over 29 kb of the *B. subtilis* genome was sequenced between *ykwC* and *rok*. There was no concentration of genes involved in any one particular cellular function in this region. In the region sequenced there are 24 genes of unknown function. The function of five genes in the region is known. The *patA* gene function has not been shown experimentally but it is probably an aminotransferase based on sequence similarities. The *cheV* and *kinA* genes have been characterised, they are involved in chemotaxis and sporulation respectively. Since the completion of the genome sequencing two genes of unknown function in this region have been assigned function: *ykuM* has been renamed *ccpC* and *ykuW* has been renamed *rok*. The *ccpC* gene is a regulator of catabolite repression and *rok* encodes a repressor of *comK*. The availability of the genome sequence has transformed *B. subtilis* research and has facilitated global transcriptome and proteome analysis of gene expression.

The *B. subtilis* Functional Analysis Project was the first coordinated attempt to exploit the genome sequence data and to assign function to the large number of genes of unknown function. The project succeeded in the generation of a mutant collection of 1,275 strains containing null mutants in genes of unknown function. These strains also contain transcriptional *lacZ* fusions in the disrupted gene. The collection is a valuable resource and its creation is a considerable achievement. LacZ assays and phenotypic tests were carried out on the strains of the collection during the project. Phenotypic analysis has in many cases pointed towards the function of a gene as in the case of *yjlC* and *yjlD* as described in Chapter 5, although the lack of a phenotype does not mean that the gene under study is not involved in a particular process, as the existence of paralogues for many genes indicates that gene functions overlap in many cases. In addition, work carried out S.D. Ehrlich's laboratory established the identity of many essential genes which is very valuable and interesting information. The expression data generated during the project is complemented by recent micro- and macroarray analysis. Future work in establishing gene function will be carried out at this global level, or by specific analysis of particular genes or groups of genes in a piece meal manner

Table 3.1
Potential sigma factor motifs identified in the region between *ykwC* and *rok*.

Gene	Promoter Motifs (position on the genome)	Consensus Promoter Sequence[1,2] Potential Promoter Sequence[3]
ykwC	None	
ykwD	SigH (1466840)	SigH -rvaggawwt→14nt←mgaat- -AAcGGATTT→14nt←AcAAT-
ykuA	SigH (1466827)	SigH -rvaggawwt→14nt←mgaat- -AAAGtAAAT→14nt←AGAAT-
kinA	SigG (1469245)	SigG -gmatr→18nt←catwmta- -GAATA→17nt←CATACTA-
	SigH (1469239)	SigH -rvaggawwt→14nt←mgaat- -GAAGGAgAa→14nt←CGAAT-
patA	SigA (1472403)	SigA -ttgaca→17nt←tataat- -gTGACA→16nt←TAaAAT-
	SigE (1472452)	SigE -kmataww→14nt←catacamt- -GtATgTT→13nt←CATACAAT-
	SigG (1472389)	SigG -gmatr→18nt←catwmta- -GCtaG→19nt←CATTATA-
	SigL (1472576)	SigL -tggcac→5nt←ttgcannn- -TGGCAC→6nt←cTtCATTT-
cheV	SigD (1472840)	SigD -taaa→15nt←gccgatat- -TtcA→15nt←GCCGATAT-
ykuB	SigA (1474394)	SigA -ttgaca→17nt←tataat- -TTtgCA→17nt←TATAAT-
ykuC	None	
ykuD	SigK (1476338)	SigK -ac→17nt←catannnta- -Ag→17nt←CAatACTTA-
	GerE binding site(1476350)	GerE -rwwtrggynnyy- -GgATGGGCTCCT-
ykuE	None	
ykuF	SigA (1477306)	SigA -ttgaca→17nt←tataat- -cTGACA→18nt←TAgAAT-
	SigE (1477332)	SigE -kmataww→14nt←catacamt- -GAATATT→14nt←CATtCACa-
	SigG (1477299)	SigG -gmatr→18nt←catwmta- -GCATG→18nt←tATAATg-
	SigK (1477309)	SigK -ac→17nt←catannnta- -AC→17nt←gATAGAATA-
ykuG	SigG (1478190)	SigG -gmatr→18nt←catwmta- -GAATG→18nt←CcTTCTt-
ykuH	SigG (1480752)	SigG -gmatr→18nt←catwmta- -GAATA→18nt←CgTTAaA-
ykuI	None	

Table 3.1- Continued

Gene	Promoter Motifs (position on the genome)	Consensus Promoter Sequence[1,2] Potential Promoter Sequence[3]
ykuJ	SigA (1482834)	SigA `-ttgaca→17nt←tataat-` `-TTGAtA→16nt←TtTAAT-`
	SigF (1483239)	SigF `-gcatr→15nt←ggmrarmtw-` `-GCATG→15nt←ctCAAACTA-`
	SigG (1483130)	SigG `-gmatr→18nt←catwmta-` `-GAtTA→19nt←CATAAaA-`
	SigH (1483190)	SigH `-rvaggawwt→14nt←mgaat-` `-AAAGGcATT→14nt←AGAtT-`
ykuK	None	
ykzF	SigK (1484333)	SigK `-ac→17nt←catannnta-` `-AC→16nt←CtTATTCTA-`
ykuL	SigA (1484681)	SigA `-ttgaca→17nt←tataat-` `-TcGACA→17nt←TATtAT-`
	SigK (1484628)	SigK `-ac→17nt←catannnta-` `-AC→18nt←CAaAGAGTA-`
ccpC	None	
ykuN	SigA (1486268)	SigA `-ttgaca→17nt←tataat-` `-TTGACA→16nt←TtTAAa-`
	SigK (1486283)	SigK `-ac→17nt←catannnta-` `-At→17nt←CATATGATA-`
ykuO	None	
ykuP	None	
ykuQ	None	
ykuR	SigG (1488992)	SigG `-gmatr→18nt←catwmta-` `-GAAaA→19nt←gATAATA-`
ykuS	SigG (1490197)	SigG `-gmatr→18nt←catwmta-` `-GAATA→17nt←CAcTCTA-`
ykuT	SigB (1491401)	SigB `-rggwttra→14nt←gggtat-` `-AcGTTTAA→14nt←GGGTAa-`
	SigG (1483130)	SigG `-gmatr→18nt←catwmta-` `-GAtTA→19nt←CATAAaA-`
ykuU	SigA (1491449)	SigA `-ttgaca→17nt←tataat-` `-TTGACt→17nt←TATAAT-`
	SigK (1491546)	SigK `-ac→17nt←catannnta-` `-AC→17nt←CATATGGcA-`
ykuV	SigG (1492138)	SigG `-gmatr→18nt←catwmta-` `-GCATA→18nt←tAaAATA-`

Table 3.1- Continued

Gene	Promoter Motifs (position on the genome)	Consensus Promoter Sequence[1,2] Potential Promoter Sequence[3]
rok	SigG (1492869)	SigG -gmatr→18nt←catwmta- -GtAaG→18nt←CATTATA-
	SigH (1492973)	SigH -rvaggawwt→14nt←mgaat- -tAAGGAATT→15nt←AtAAT-
	SigK (1492967)	SigK -ac→17nt←catannnta- -Ag→16nt←CATAGAGTA-

[1]The consensus sequence used in the search is written in lower case next to the name of the sigma factor it represents.

[2]M: A or C;

N: A, G, C or T;

R: A or G;

V: A, G or C;

W: A or T;

Y: C or T;

Z: T or G

[3]The potential promoter sequence(s) found are written under the consensus. The bases which agree with the consensus are in upper case letters, those which differ from the consensus are in lower case letters.

CHAPTER 4

THE SUPEROXIDE DISMUTASES OF *B. subtilis*

4.1 INTRODUCTION

The *sodA* gene encodes the highly expressed vegetative superoxide dismutase of *B. subtilis*. Some characterisation of this enzyme has been carried out (Casillas-Martinez and Setlow, 1997; Inaoka *et al.*, 1998, 1999). The completion of the *B. subtilis* genome sequence revealed two additional ORFS with homology to superoxide dismutases: *sodF* and *yojM*.

SodA is the principal superoxide dismutase of *B. subtilis*. The *sodA* gene is 606 base pairs in length and the coding sequence is followed by a rho-independent terminator sequence suggesting that the operon is monocistronic (Figure 4.1). It is highly expressed throughout exponential growth and mutation of *sodA* leads to slow growth in minimal medium and sensitivity to paraquat (Inaoka *et al.*, 1999). The *B. subtilis* SodA is predicted to use manganese as its catalytic metal based on homology to other bacterial manganese SODs (Inaoka *et al.*, 1998). SODs which bind either manganese or iron are similar in primary sequence to each other and distinct from SODs which bind copper and zinc. It is likely that SODs evolved according to the availability of oxygen in the environment. It has been suggested that iron was initially the metal used at the active site of the enzyme as the reducing atmosphere of the earth meant this metal was available in the soluble, reduced (FeII) form. As the earths atmosphere became oxygenated, the amount of available FeII would have become less and resulted in evolution of the enzyme so manganese (MnIII) could be incorporated as the catalytic metal. Copper in its reduced form (CuI) is less soluble than oxidised copper (CuII) and consequently was probably unavailable to organisms until the environment became aerobic (Bannister *et al.*, 1991). Iron/manganese SODs are often specific for one metal. However in some bacteria the SOD is cambialistic and binds either metal depending on availability (e.g. *Propionibacterium shermanii* (Meier *et al.*, 1982), *Bacteroides fragilis* (Gregory and Dapper, 1983), *Streptococcus mutans* (Martin *et al.*, 1986)).

The *sodF* gene encodes a probable superoxide dismutase. The amino acid sequence is highly homologous to other iron/manganese superoxide dismutases. The gene upstream of *sodF* is *sqhC* which encodes squalene-hopene cyclase (Figure 4.2). This is an enzyme involved in the synthesis of hopanoids, a class of eubacterial lipids. The stop codon of *sqhC* overlaps the fourth amino acid codon of the *sodF* ORF, therefore *sqhC* is probably co-expressed with *sodF*. The putative control region upstream of *sqhC* is 100 bases in
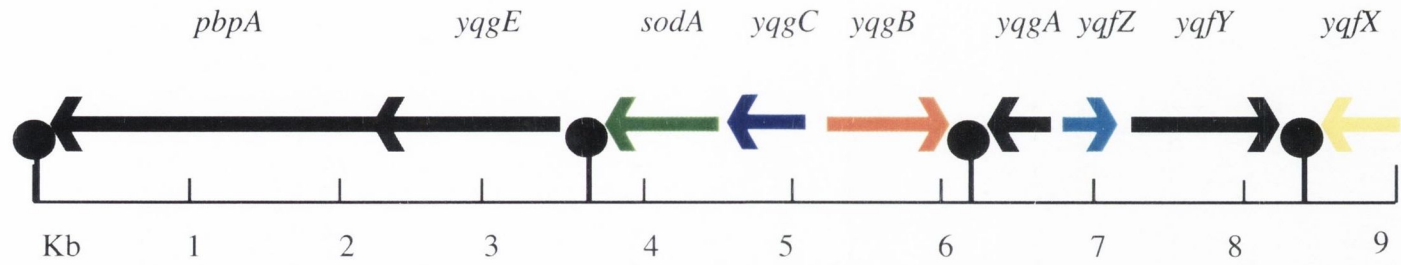
Figure 4.1 Organisation of the ORFs in the vicinity of the *sodA* gene of *B. subtilis*.
Arrows represent genes and are proportional to gene length. Lollipops represent rho-independent terminators.
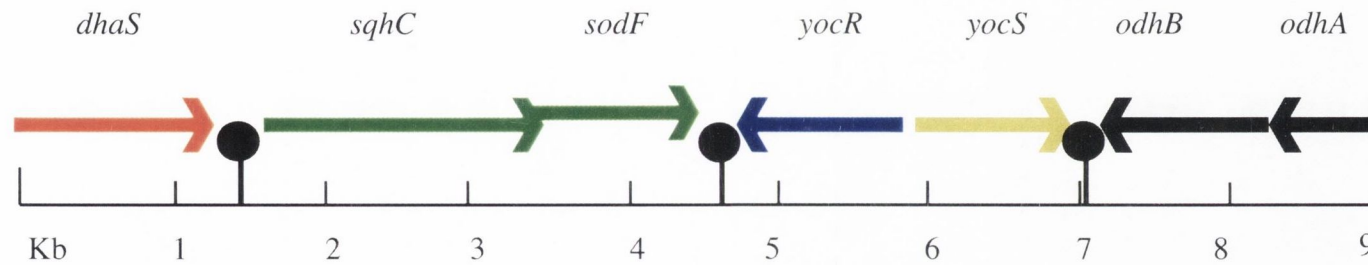
Figure 4.2 Organisation of the ORFs in the vicinity of the *sodF* gene of *B. subtilis*.
Arrows represent genes and are proportional to gene length. Lollipops represent rho-independent terminators.
The *sodF* gene lies almost 10 kilobases away from *yojM*. It is probably in an operon with *sqhC*.

length but does not contain any recognizable promoter consensus sequence. The *sodF* ORF is followed by a rho-independent terminator sequence.

The *yojM* gene, categorised as being a gene of unknown function, shares similarity with copper-zinc type superoxide dismutases (Cu,ZnSODs) in other organisms. The operon structure is shown in Figure 4.3. The upstream gene, *yojL*, is similar to cell wall binding proteins. It is followed by a stem loop structure comprising a stem region of only nine bases that is not followed by a run of T nucleotides. This suggests that it is a regulatory stem loop feature rather than a rho-independent terminator. Therefore *yojL* may be in the same operon as *yojM* and the region upstream of *yojL* may contain the promoter of this operon. There are 63 bases between the stop codon of *yojM* and the start of the next gene *yojN* which is similar to nitric oxide reductase. It is tempting to speculate on the juxtaposition of *yojN* and *yojM* since nitric oxide is a potentially dangerous species that can react with the superoxide radical to from peroxynitrite resulting in damage to DNA and proteins. The next gene, *yojO*, overlaps the end of *yojN* by 19 amino acids. This is a gene of unknown function. It is followed by two genes, *odhA* and *odhB*, which code for oxoglutarate dehydrogenase. These genes are probably in a separate dicistronic operon. They have been reported to be repressed by glucose and both are preceded by potential CRE sites (Moreno *et al.*, 2001). CRE sites are specific for binding of the CcpA protein which mediates catabolite repression in *B. subtilis*. The region between the terminator of *yojK* (the gene upstream of *yojL*) and the initiation codon of *yojL* comprises 343 bases. Cu,ZnSODs are quite distinct from the Mn/FeSODs which are very similar to each other. Copper, rather than iron or manganese is the redox active metal in these enzymes. The role of the zinc ligand is structural. The Cu,ZnSOD is cytosolic in eukaryotes and in eubacteria is found in the periplasm of Gram negative bacteria, where it presumably reacts with superoxide generated in this compartment. The role of this type of superoxide dismutase in Gram positive bacteria is unknown.

The superoxide stress response is not well elucidated in *B. subtilis*. It is clearly regulated differently from the superoxide stress response in *E. coli,* therefore it was decided to investigate the three superoxide dismutases of *B. subtilis*. Only SodA has been characterised in any detail and the manner in which its expression is regulated has not been defined. The two other SODs encoded in the genome, *sodF* and *yojM,* have not been investigated. In this chapter investigation into the regulation and expression of the three SODs and their possible roles in cellular physiology is presented.
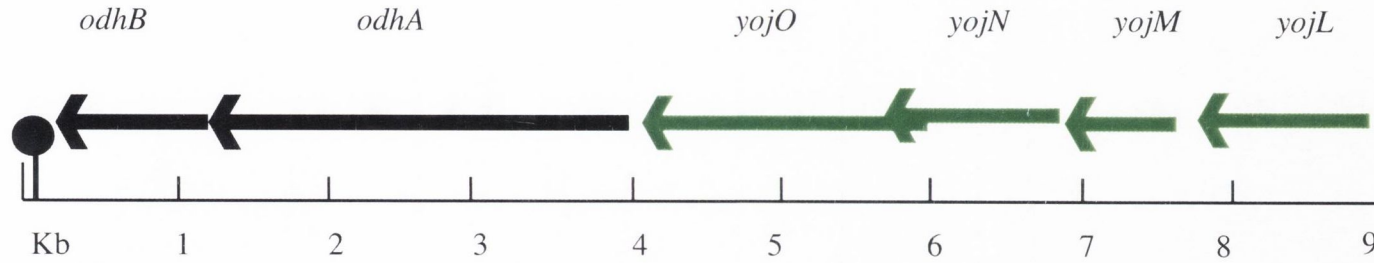
Figure 4.3 Organisation of the ORFs in the vicinity of the *yojM* gene of *B. subtilis*.

The *yojM* gene is possibly in an operon with *yojLNO*.

The downstream genes, *odhAB*, are in a seperate operon and are subject to catabolite repression (Moreno, M.S. *et al*. 2001)

4.2 RESULTS

4.2.1 Multiple alignments of SodA, SodF and YojM.

A comparison of the *B. subtilis* SodA and SodF proteins with iron/manganese SODs shows a high degree of homology suggesting that they belong to this family of SODs. Iron/manganese SODs are structurally similar. The functional enzyme is dimeric or tetrameric. Each subunit consists of two domains: an N terminal α-helical domain and a C-terminal α and β fold. Figure 4.4 shows an alignment of SodA with manganese binding SODs from other organisms. Three histidines and an aspartate (His-28, His-82, Asp-164, His-168 numbering refers to *B. subtilis sodA*) indicated with arrows in Figure 4.4 are responsible for binding the metal ligand. However sequence comparisons have shown that MnSODs can be distinguished from FeSODs at the primary sequence level (Parker and Blake, 1988). Residues specific to manganese binding SODs are Gly-77, Gly-78, Phe-85, Gln-149 and Asp-150 marked by triangles in Figure 4.4 (numbering refers to *B. subtilis sodA*). MnSODs also contain a residue with an aliphatic side chain at the equivalent of position 80 in *B. subtilis sodA*. There is an alanine residue at this position in all the examples in Figure 4.4. These active site residues and are conserved in the *B. subtilis* SodA.

The *B. subtilis* SodF appears to have a longer N-terminal region than most bacterial Mn/FeSODs, a feature it shares with its close relative *Bacillus halodurans* (Figure 4.5). However this region is not predicted to be a signal peptide by the analysis program SignalP V1.1 (http://www.cbs.dtu.dk/services/SignalP/) (Nielsen *et al.*, 1997). The four metal ligand residues characteristic of Mn/FeSODs are present in SodF (the coordinates in *B. subtilis* SodF are His-104, His-152, Asp-236 and His-240). It is not clear by examination of sequence homologies whether the *B. subtilis* SodF binds iron or manganese or both as its catalytic metal. In BLAST searches SodF is homologous to cambialistic SODs (*e.g. Pyrobaculum aerophilum,* level of similarity $7e^{-49}$, *Propionibacterium shermanii,* level of similarity $9e^{-47}$). The amino acids which specify iron or manganese binding are not completely conserved in SodF. In Figure 4.5 the residues of the *E. coli* SodB which specify iron binding are indicated by diamonds. They are Ala-69, Gln-70, Trp-72, Tyr-77, Ala-142 and Gly-143) (Parker and Blake, 1988). At the equivalent positions *B. subtilis* SodF has: Ala (typical of FeSODs); Gly (typical of MnSODs at this position); Tyr (atypical); Phe (typical of MnSODs at this position); His (atypical); Gln (atypical).

Figure 4.4 Multiple alignment of MnSODs from *B. subtilis*, *E. coli* and *Bacillus licheniformis* and superoxide dismutases which probably bind manganese from *Bacillus halodurans* and *Listeria monocytogenes*. Where half or more of the residues in a column are identical/similar they are shaded. White letters in black boxes denote identical residues. Black letters in grey boxes denote similar residues. Arrows indicate the four metal ligand residues (His-28, His-82, Asp-164, His-168 numbering is *B.subtilis* SodA residues). Residues marked with a triangle (∇) distinguish the M binding SOD from the FeSOD.

```
B.subtilisSodF      1  MKRES-YQAEMFNWCEALKDQ---IQKRGQLD--QFEDQIDKMIEALEDDQ------TTE
B.haloduransSod     1  MSIEWGQLHAWVNRVEEAWNQNKALLRSESVDTFLFEETLERLKEQLQNDGEPTEIGELA
P.aerophilumSodF    1  ----------------------------------------------------------
P.shermaniiSod      1  ----------------------------------------------------------
E.coliSodB          1  ----------------------------------------------------------

                                                                          ↓
B.subtilisSodF     49  EDWYKQAAALYRDITESDDTSERRAYVPIGKHVLPKLPYKYSALEPYISRDIMILHHTKH
B.haloduransSod    61  EHLYEEWARLFSEQTEAGGSNGMRRPVPIGGHRLPPLPYPYEALEPYIDREIMRLHHQKH
P.aerophilumSodF    1  -----------------------MVTTKRYTLPPLPYAYNALEPYISAEIMQLHHQKH
P.shermaniiSod      1  ------------------------MVYTLPELPYDYSALEPYISGEIMELHHDKH
E.coliSodB          1  -------------------------MSFELPALPYAKDALAPHISAETIEYHYGKH

                                                          ♦♦ ♦  ↓     ♦
B.subtilisSodF    109  HQSYVDGLNKAESELKKARATKNYDLITHWERELAFHCAGHVLHSIFWFSMHP--NGKRE
B.haloduransSod   121  HQSYVEGLNKAEKEMERARRTNHFQLIKEWEREAAFNCAGHVLHSIFWEVMSP--HGGGE
P.aerophilumSodF   36  HQGYVNGANAALEKLEKFRKGEAQIDIRAVLRDLSFHLNGHILHSIFWPNMAPPGKGGGR
P.shermaniiSod     33  HKAYVDGANTALDKLAEARDKADEGAINKLEKDLAFNLAGHVNHSVFWKNMAPKGSAPER
E.coliSodB         32  HQTYVTNLNNLIKG-TAFEGKSLEEIIRSSEGGVFNNAAQVWNHTFYWNCLAP--NAGGE

                                                                       ♦♦
B.subtilisSodF    167  PTGALFQMIDLSFGSYSAFKEHFTQASKKVEGVGWAILVWAPRSGRLEILTAEKHQLFSQ
B.haloduransSod   179  PKGELRRQIERDFGSFARFKNHFSQAAEKVEGVGWAILVWAPRAHRLEILQAEFHQNLSQ
P.aerophilumSodF   96  PGGKIADLINKFGSFEKFKEEFSQAAKNVEGVGWAILVEPLEEQLLILQIEKHNLMHA
P.shermaniiSod     93  PTDELGAAIDEFFGSFDNMKAQFTAAATGIQGSGWASLVWDPLGKRINTLQFYDHQNNLP
E.coliSodB         89  PTGKVAEAIAASFGSFADFKAQFTDAAIKNFGSGWTWLVKNS-DGKLAIVSTSNAGTPLT

                                  ↓        ↓
B.subtilisSodF    227  WDVIPLLPLDVWEHAYYLQYKNDRASYVDHWWNVVDWREAEKRFEQAK--EVVWKLY
B.haloduransSod   239  QDAIPLLPLDVWEHAYYLQYKNERKDYLENWWNVVNWDAVEKRTNAAK--QLRWTPY
P.aerophilumSodF  156  ADAQVLLALDVWEHAYYLQYKNDRGSYVDNWWNVVNWDDVERRLQKALNGQLALKL-
P.shermaniiSod    153  AGSIPLLQLDVWEHAYYLQYKNVKGDYVKSWWNVVNWDDVALRTSEAR---VA----
E.coliSodB        148  TDATPLLTVDVWEHAYYIDYRNARPGYLEHFWALVNWEFVAKNLAA-----------
```

Figure 4.5 Multiple alignment of probable Mn/FeSODs: SodF, from *Bacillus subtilis*, the FeSOD, SodB, from *Escherichia coli*, the cambialistic S of *Propionibacterium shermanii* and the archaeon *Pyrobaculum aerophilum* and the SOD of *Bacillus halodurans* whose metal binding specificity undetermined. Where half or more of the residues in a column are identical/similar they are shaded. White letters in black boxes denote identical residues. Black letters in grey boxes denote similar residues. Arrows indicate the four metal ligand residues. Residues marked with a diamond (♦) those which specify that the *E. coli* SodB has iron as its metal ligand (Section 4.2.1).

In Figure 4.6 a multiple alignment showing the homology between *yojM* and Cu,ZnSODs in other bacteria and also the bovine Cu,ZnSOD is presented. This and previously reported alignments point towards a common ancestor for the eukaryotic and eubacterial Cu,ZnSOD (Imlay and Imlay, 1996). Most bacterial forms of the Cu,ZnSOD have a leader peptide sequence. It has been suggested that the Cu,ZnSOD arose in the Gram negative bacterial periplasm and was laterally transferred into the eukaryotic cell (Imlay and Imlay, 1996). However this hypothesis clearly does not account for a Gram positive version of the enzyme. In Figure 4.7 an alignment of *yojM* and the *Bos taurus* Cu,ZnSOD is presented showing important residues. Four histidine residues, His-44, His-46, His-61 and His-118 (numbering refers to bovine SOD) have been shown to be involved in binding the copper ion in the bovine Cu,ZnSOD. Also indicated in Figure 4.7 are the zinc ligands: His-61, His69, His-78 and Asp-81 (Bannister *et al.*, 1991). Interestingly two of the conserved His residues are not found in YojM. Instead it has a tyrosine at the His-61 position and a proline at the His-61 position. Two further residues are shown in Figure 4.7. Aspartic acid at position 122 binds His-44 and His-69 an interaction that is important for the structure of the enzyme and the positioning of the metal ligands (numbering refers to bovine Cu,ZnSOD). In addition, arginine at position 141 is involved in positioning the superoxide radical at the active site of the enzyme. These conserved amino acids are found in YojM, again suggesting that it is a member of the Cu,ZnSOD family. The crystallographic structure of bovine Cu,ZnSOD among others (including human and yeast) have shown it is a Greek-key β-barrel which contains eight antiparallel β-strands and three major loops (Tainer *et al.*, 1982; Djinovic *et al.*, 1992a; Parge *et al.*, 1992; Djinovic *et al.*, 1992b). Imlay and Imlay suggest that the same structure is conserved in bacterial Cu,ZnSODs as there is such a good level of conservation of the important amino acids (Imlay and Imlay, 1996). The β-strands are marked in Figure 4.7 by overlining. Also indicated in bold and underlined are two cysteine residues which form a disulfide bond and so joins the first loop and the terminal β-strand. Clearly, with the exception of two histidine residues, YojM has many of the conserved amino acids typical of Cu,ZnSODs from bacteria and eukaryotes. Therefore it would be surprising if this SOD was not functional in detoxification of superoxide in the *B. subtilis* cell.

4.2.2 Paraquat induction of genes of the superoxide stress regulon.

Paraquat is a redox cycling compound. Upon uptake into the cell, paraquat is reduced by a single electron that it subsequently transfers to oxygen to generate the superoxide radical. Two dimensional gel analysis was previously used to characterise the response of *B.*

```
C.crecentusSodC    1  -----------MIRLSAAAALGLAAALAASP-----------------ALAQTSATAVVKAG
E.coliSodC         1  ---------------MKRFSLAILALVVATGAQA----------------ASEKVEMNLVTSQG
R.solanacearumSodC 1  MKQLVIGLAAIGLMAGCASNAGTCASTGAP----------------AAGTSALATLAPKS
BovineSodC         1  -------------------------------------------------ATKAVCVLK--G
B.subtilisYojM     1  ---------MHRLLLMMLTALGVAGCGQKKPPDPPNRVPEKKVVELSAFGHHVQLVNR
A.aeolicusSodC2    1  ---------MKKLSGVLAGSLLLISAS----------------FSQDLKAHAELINT


C.crecentusSodC    35  DGKDAGAVTVTEAPHG-VLLKLELKGLT-PGWHAAHFHEKGDCG-------TPDFKSAG
E.coliSodC         34  VGQSIGSVTITETDKG-LEFSPDLKALP-PGHGFHIHAKGSCQPATKDGKASAAESAGG
R.solanacearumSodC 45  GSNVQGTIKLLQQGDSRVAAAVDIAGLPPNGMFGFHVHEKGDCS-------SPDGMSAGG
BovineSodC         11  DGPVQGTIHFEAKGDT-VVVTGSITGLT-EGDHGFHVHQFGDN-------TQGCTSAGP
B.subtilisYojM     51  LGKAVGFIEIKESDDEGLDIHISANSLRPGASLGFHIYEKGSCVR------PDFESAGG
A.aeolicusSodC2    33  LGEVIGKAELIETN-SGVLIKLNAKGLPPNAELAFHIHERGCKP------PTFKSAKG


C.crecentusSodC    86  HVHTAATTVHGLLNPDANDSGDLPNIFAAADGAATAEIYSP-LVSLKGAGGRPALLDADG
E.coliSodC         92  HLDPQNTGKHEGP-EGAGHLGDLPALVVNNDGKATDAVIAPRLKSLDEIKDKALMVHVGG
R.solanacearumSodC 98  HFNP-TGQPHGDPRSGPBHAGDPMLQSDASGKAVGSIVLNGVTLTPGATSLVGRAVIVH
BovineSodC         61  HFNP-LSKKHGGPKDEERHVGDLGNVTADKNGVAIVDIVDPLISLS-GEYSIIGRTMVVH
B.subtilisYojM     104 PFNP-LNKEHGFNNPMGHHAGDLPNLEVGADGKVDVIMNAPDTSLRKGSKL--NILDEDG
A.aeolicusSodC2    85  HFNP-YGKKHGLLNPEGPHAGDIPNIYTDDKGNVRVQVLNPFVTLKKGEKN--SIFKEGG


C.crecentusSodC    145 SSIVVHANPDDHKTQPIGGAGARVACGVIL------
E.coliSodC         151 DNMSDQP-------KPLGGGGERYACGVIK------
R.solanacearumSodC 157 AGMDDY------KTQPTGNSGGRVACGVI-VAN---
BovineSodC         119 EKPDDLGRGGNEESTKTGNAGSRLACGVIGIAK---
B.subtilisYojM     161 SAFIIHEQADDYLTNPSGNSGARIVCGALLGNNEKQ
A.aeolicusSodC2    142 TALVIHSGPDDYKSDPAGNAGKRLACGVIR------
```

Figure 4.6 Multiple alignment of *B. subtilis* YojM and putative CuZnSODs from *Ralstonia solanacearum* and *Aquifex aeolicus* and the defined CuZnSODs of *Escherichia coli*, *Caulobacter crescentus* and *Bos taurus*. Where half or more of the residues in a column are identical/similar they shaded. White letters in black boxes denote identical residues. Black letters in grey boxes denote similar residues.

```
                                                                  _____        _____
B.subtilis YojM    MHRLLLLMMLTALGVAGCGQKKPPDPPNRVPEKKVVETSAFGHHVQLVNREGKAVGFIEI 60
Bovine SodC        -------------------------------------------ATKAVCVLKGDGPVQGTIHF 20
                                                              :  *  ::: :*  .  *  *.:
                                                              1         10        20


                                            Cu Cu              Cu/Zn         Zn
                                             ↓  ↓                ↓            ↓
     ____    _____        _____   _____           _____     _____
B.subtilis YojM    KESDDEGLDIHISANSLRPGASLGFHIYEKGSCVRPDFESAGGPFNPLNKEHGFNNPMGH 120
Bovine SodC        EAKGDT-VVVTGSITGLTEG-DHGFHVHQFGDNTQG-CTSAGPHFNPLSKKHGGPKDEER 77
                   :  ..*   :  :   *  ..*   *  . ***::: *. .:   ***  ****.*:** :    :
                         30          40          50          60          70


Zn Zn                                                   Cu
↓  ↓                                                     ↓
   _____        _____                   _____  ↓
B.subtilis YojM    HAGDLPNLEVGADGKVDVIMNAPDTSLKKGSKLNILDEDGSAFIIHEQADDY------LT 174
Bovine SodC        HVGDLGNVTADKNGVAIVDIVDPLISLS--GEYSIIG---RTMVVHEKPDDLGRGGNEES 132
                   *.*** *:..:*. *: *  :  **.  .:  .*:.  :::**:.**        :
                   80          90          100         110         120         130


                    ↓
             _____
B.subtilis YojM    NPSGNSGARIVCGALLGNNEKQ 196
Bovine SodC        TKTGNAGSRLACG-VIGIAK-- 151
                   . :**:*:*:.** ::*   :
                       140         150
```

Figure 4.7 Alignment of the putative *Bacillus subtilis* Cu,ZnSOD and the *Bos taurus* Cu,ZnSOD.
* indicates identical residues in both sequences of the alignment : indicates conserved substitutions . indicates semiconserved substitutions. Amino acids important for function are indicated with an arrow. Those involved in binding the copper or zinc ligand are specified Cu or Zn. Conserved cysteines involved in an intramolecular disulfide bond are in bold and underlined. Lines over the sequence denote the position of the β- strands of the bovine CuZnSOD. Coordinates under the sequence refer to the bovine CuZnSOD.

subtilis to paraquat induced superoxide stress (Antelmann *et al.*, 1997; Bernhardt *et al.*, 1999). In these studies it was found that paraquat exposure induced SodA, amino acid biosynthetic proteins IlvC, LeuC and CysK, and DhbB which is involved in siderophore synthesis. To further investigate the superoxide stress inducible regulon of *B. subtilis* it was decided to examine whether *sodA, dhbB, leuC, cysK, ilvC* and *sodF* were inducible by paraquat at the transcriptional level. The *sodF* gene was also chosen as it is a superoxide dismutase homologue.

Mutagenic insertions and transcriptional *lacZ* fusions of *sodA, dhbB, leuC, cysK, ilvC* and *sodF* were generated using the pMUTIN4 integrating plasmid (Vagner *et al.*, 1998). The pMUTIN4 insertion strains which were generated were named ES001 (insertion into *sodF*), ES002 (insertion into *sodA*), ES003 (insertion into *ilvC*), ES004 (insertion into *cysK*), CM991 (insertion into *dhbB*) and CM992 (insertion into *leuC*). The expression of each gene in response to superoxide stress conditions was determined. Cultures were grown to an OD_{600} of 0.2 and then divided, no addition was made to the control half of the culture, paraquat (to a final concentration of 100µM) was added to the second half of the culture. Samples were harvested at fixed time points up to 60 minutes after the split and assayed for LacZ activity (Chapter 2 section 2.7.3). The results of these experiments are shown in Figure 4.8. The *sodF-lacZ* transcriptional fusion is very lowly expressed and little difference in expression is observed between the paraquat-induced and the control cultures (Figure 4.8(A)). In contrast, expression of the *sodA-lacZ* fusion increases after addition of paraquat (Figure 4.8 (B)). A 1.5-2 fold increase in LacZ activity is observed in the paraquat-induced culture compared to the non-induced control. The increase in *lacZ* activity becomes apparent 45 minutes after addition of paraquat. This delay before increased expression is seen may be due to the fact that paraquat must be taken up into the cell and then undergo redox cycling before the superoxide radical is produced. In addition to the difference in *sodA* expression between the cultures it is also apparent that paraquat inhibits growth in the *sodA* mutant strain, a phenotype that had been observed previously (Inaoka *et al.*, 1999). Expression of *ilvC-lacZ* is low and the small difference between induced and control cultures is probably not significant (Figure 4.8 (C)). Expression of the *cysK-lacZ* fusion is high in all samples, but there is no induction of expression in the paraquat stressed culture (Figure 4.8 (D)). The *dhbB* and *leuC lacZ* fusions have low levels of expression with no differences between the induced and control cultures (Figure 4.8 (E) and (F)). These results show that *sodA* is paraquat inducible, while the other genes tested are not induced by paraquat at the transcriptional level under these conditions.
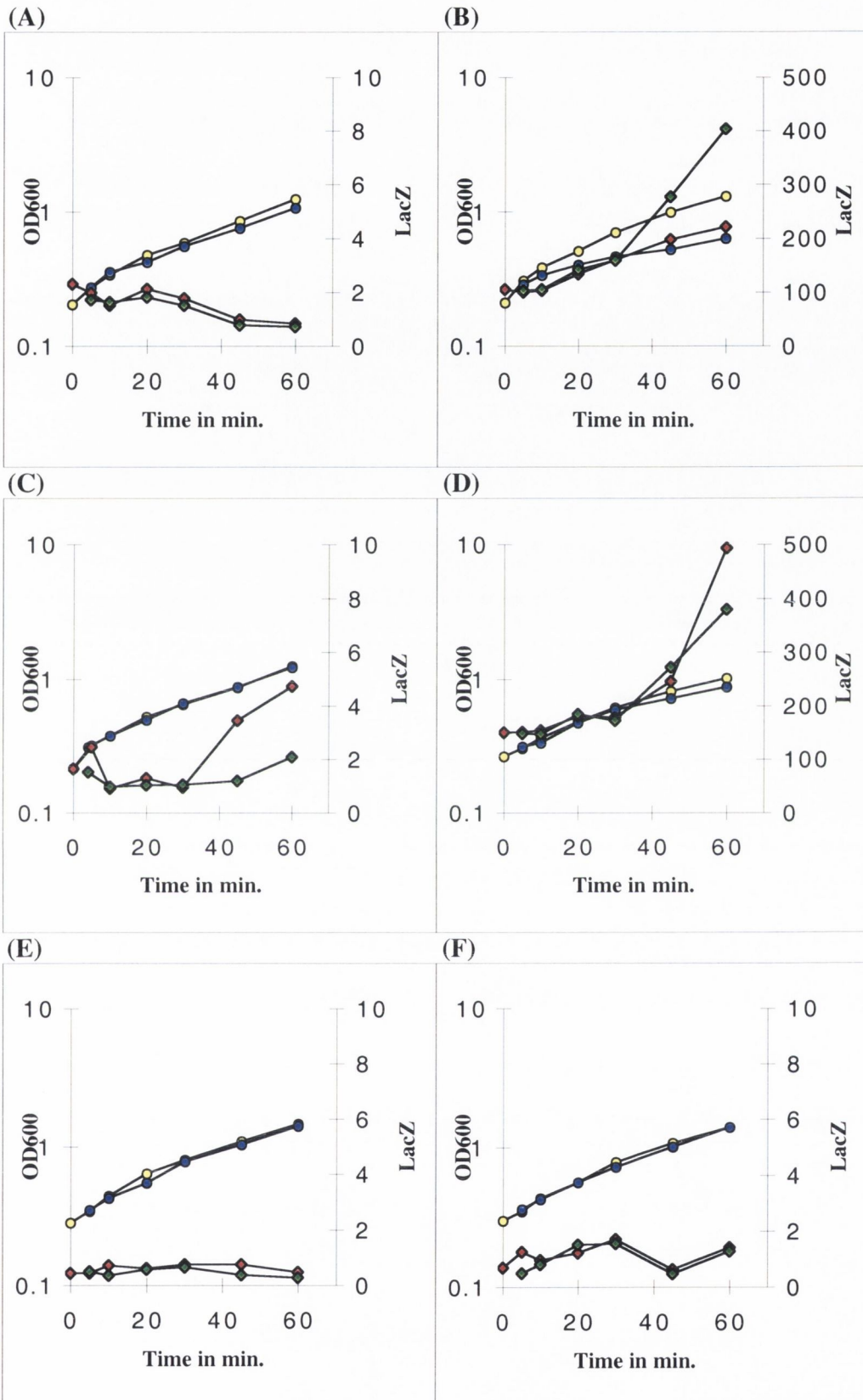
Figure 4.8 Growth and LacZ expression profiles after paraquat induction.
Growth curves and expression profiles of strains (A) ES001 (*sodF'-lacZ*),
(B)ES002 (*sodA'-lacZ*) , (C)ES003 (*ilvC'-lacZ*), (D)ES004 (*cysK'-lacZ*),
(E)CM991(*dhbB'-lacZ*) and (F)CM992 (*leuC'-lacZ*) in LB medium with and
without 100μM paraquat induction. Growth was measured at OD600 and is
shown by yellow (non induced) and blue (induced) circles. LacZ activity was
measured as nmole ONPG/min/mg and is shown by red (non induced) and
green (induced) diamonds. Time is presented in minutes after paraquat addition.

4.2.3 Oxidant sensitivities of strains carrying null mutations in genes of the superoxide stress regulon.

The sensitivities of the mutant strains of *sodA*, *dhbB*, *leuC*, *cysK*, *ilvC* and *sodF* to various oxidants were established using disc assays. The disc assays were carried out as described in chapter 2 (Chapter 2 section 2.7.2). The oxidants tested were: hydrogen peroxide (H_2O_2), paraquat, tert-butyl hydroperoxide (tBOOH) and cumene hydroperoxide (CHP). Wild type *B. subtilis* was also exposed to these agents as a control. The results of three independent experiments (carried out in duplicate each time) are shown in Figure 4.9. The results show that the *cysK* mutant (ES004) is sensitive to H_2O_2, while the remaining strains have zones of inhibition similar to the wild type (Figure 4.9 (A)). The *sodA* mutant (ES002) is sensitive to paraquat. This result is consistent with the growth inhibition effect observed in liquid culture after paraquat addition (Figure 4.8 (B)) and with previously reported results (Inaoka *et al.*, 1999; Casillas-Martinez and Setlow, 1997). However the remainder of the strains show wild type sensitivity to paraquat (Figure 4.9 (B)). Sensitivity to tBOOH varies a little more between strains than for the other oxidants. However it appears that only the *sodA* mutant (ES002) differs significantly in sensitivity from the wild type (Figure 4.9 (C)). In contrast, none of the mutant strains show altered sensitivity to CHP (Figure 4.9 (D)). It is interesting that although CHP and tBOOH are both organic hydroperoxides, the *sodA* mutant is sensitive to tBOOH but not to CHP.

It was decided to determine whether genes of the PerR regulon, especially *ahpCF* which can detoxify tBOOH, could be upregulated by tBOOH as well as H_2O_2. Two genes of the Per regulon were chosen for analysis: *ahpC* and *katA*. Transcriptional *lacZ* fusions of the promoters of *ahpC* and *katA* were placed at the amylase locus, (strains ES991 and ES993). The strains were assayed for LacZ activity after addition of tBOOH and expression levels were compared to those of the non-induced culture. The results are shown in Figure 4.10. At the point of splitting the culture the expression level of *ahpC-lacZ* was 26 LacZ units. In the control culture expression increases to 56 units during the course of the experiment. However addition of tBOOH induces higher expression with levels reaching 78 units in this culture. Within 20 minutes of tBOOH addition, the LacZ activity level has almost doubled in the induced culture in comparison to the control (41 units in the non-induced culture and 78 units in the induced culture). Transcription from the *katA* promoter also increases after tBOOH addition in comparison to the control culture, but the level of LacZ activity in both cultures is low, 5 units in the non-induced culture and 10 units in the induced culture within 20 minutes of tBOOH addition. This is a much lower induction
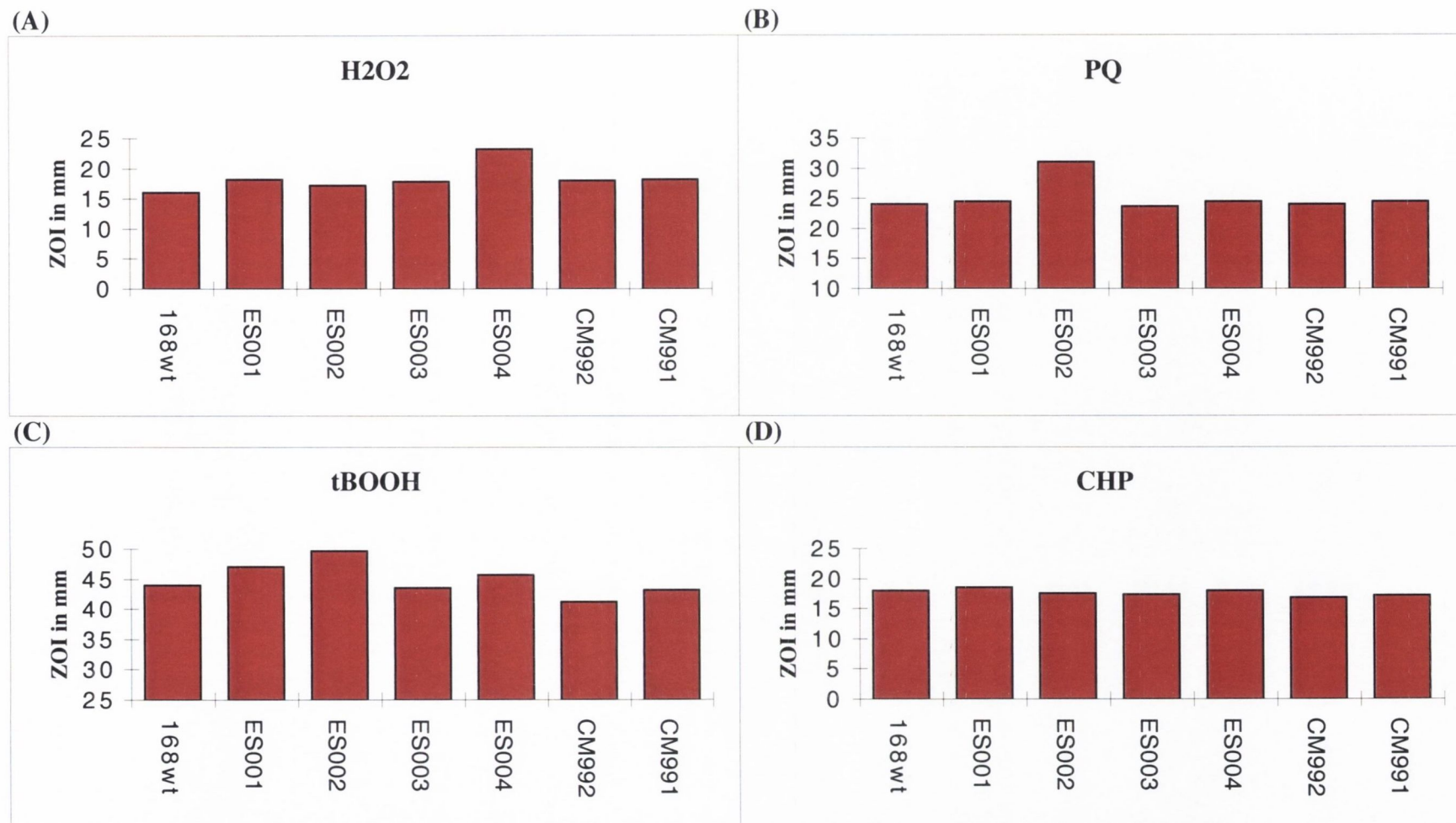
Figure 4.9 Disc assays for oxidant sensitivity of strains carrying null mutants in superoxide regulon genes.
Disc assays were carried out on LB plates as described in Chapter 2 (2.7.2) on the 168 wild type strain and the null mutant strains:
ES001 *(sodF* ::pMUTIN4), ES002 *(sodA* ::pMUTIN4), ES003 *(ilvC* ::pMUTIN4), ES004 *(cysK* ::pMUTIN4),
 CM992 *(leuC* ::pMUTIN4) and CM991 *(dhbB* ::pMUTIN4) with the oxidants: (A) Hydrogen peroxide (H2O2) (B) Paraquat (PQ)
(C) t-Butyl hydroperoxide (tBOOH) and (D) Cumene hydroperoxide (CHP). The y-axis shows the zone of inhibition measured
in millimetres. The strain names are listed on the x-axis.

(A)

(B)
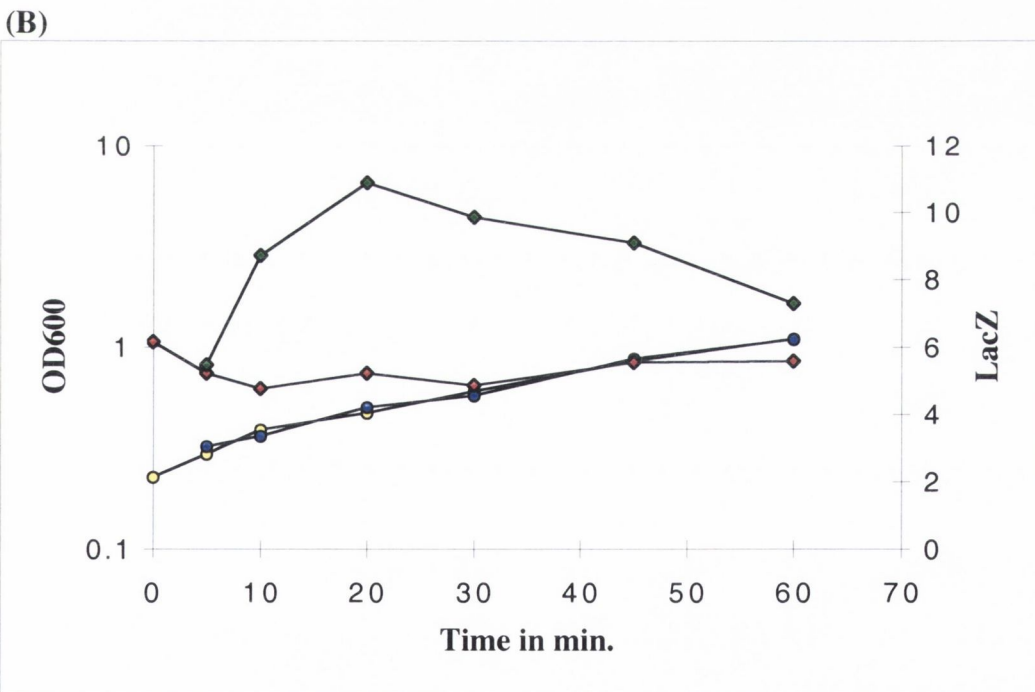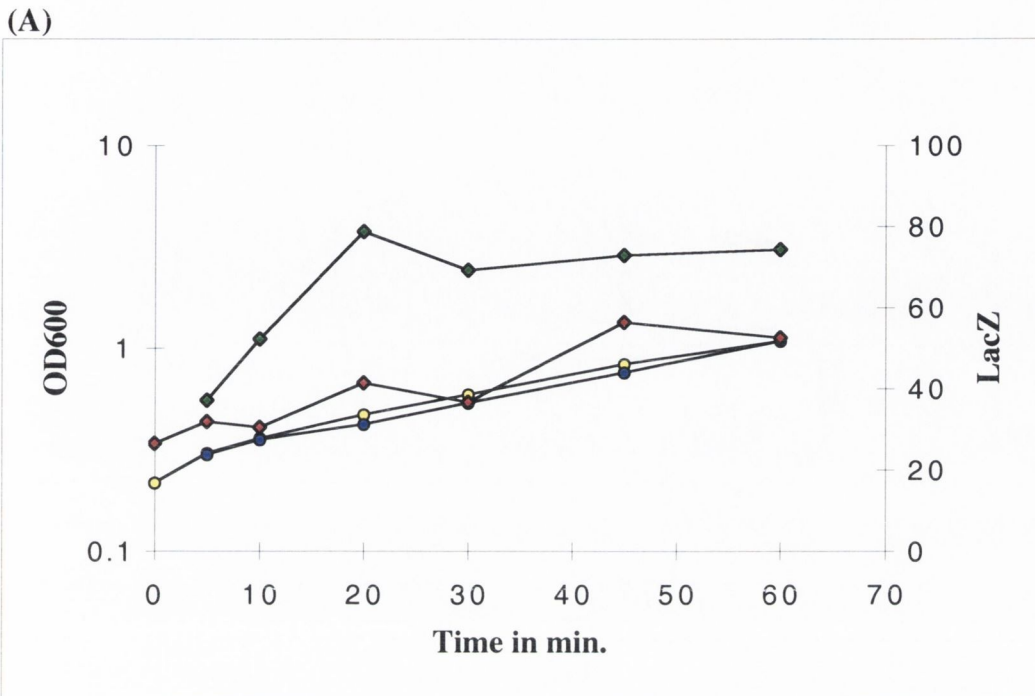
Figure 4.10 Growth and LacZ expression profiles after t-Butyl hydroperoxide induction. Growth and expression profiles of strains (A) ES993 (P*ahpCF* -*lacZ*) and (B) ES991 (P*katA* -*lacZ*) in LB medium with and without 100μM tBOOH added. Cell growth was monitored by measuring OD600 and is shown by yellow (non induced) and blue (tBOOH induced) circles. LacZ activity was measured as nmole ONPG hydrolysed/min/mg protein and is shown by red (non induced) and green (tBOOH induced) diamonds. Time is presented in minutes after tBOOH addition.

than would be expected after exposure to H_2O_2. This result implies that tBOOH and possibly other alkyl hydroperoxides do not derepress the genes of the Per regulon as efficiently as H_2O_2.

4.2.4 tBOOH sensitivity of a *sodA* mutant.

Growth and viability of the *sodA* mutant strain ES002 was examined after exposure to tBOOH in liquid medium. Both strain ES002 and the wild type were grown in LB to early exponential phase (OD_{600}~0.2). The culture was split and 500μM tBOOH was added to half the culture. For 120 minutes after splitting the culture the OD_{600} of both the induced and control halves of the culture was monitored. The results of this experiment are shown in Figure 4.11(A) and (B). There is no significant difference in growth between the wild type and the *sodA* mutant (ES002). The same experiment was carried out using 1mM tBOOH and no significant difference in growth was observed between the *sodA* mutant and the wild type (data not shown). To examine cell viability after exposure to tBOOH, cells in early exponential phase (OD_{600}~0.2) were grown in 100mM tBOOH for up to 30 minutes. This experiment was carried out with the wild type strain 168 and the *sodA* mutant strain, ES007, that contains a spectinomycin cassette inserted into *sodA*. At T_0 and 10, 20 and 30 minutes after exposure, cells were plated out on LB agar plates. Percentage survival was calculated as: (Cell count/Cell count at T_0)x100. The results are shown in Figure 4.11(B). In both the wild type and the *sodA* mutant strains 2 logs killing is achieved within 10 minutes of addition of tBOOH. The cell viability of the *sodA* mutant does not differ significantly from the wild type. Therefore it appears that the *sodA* mutant sensitivity to tBOOH is a phenotype specific to solid medium.

4.2.5 Paraquat induction of *sodA*.

Treatment of ES002 (*sodA'-lacZ*) with paraquat resulted in increased LacZ activity as described in 4.2.2. This induction was further examined in two *sodA+* strains: ES14 and ES22. Strain ES14 is a transcriptional fusion of the *sodA* promoter region (the 178 base pairs upstream of the *sodA* ATG) to *lacZ* at the amylase locus. Strain ES22 contains a translational fusion of the *sodA* promoter region to *lacZ*, also at the amylase locus. These strains are not *sodA* mutants, therefore any autoregulatory effect of SodA on its own expression could be ascertained. LacZ assays were carried out on these strains after paraquat induction. The LacZ expression profiles for ES14, ES22 and ES002 (the pMUTIN4 insertion into *sodA*) are shown in Figure 4.12. In strain ES22 containing the translational promoter fusion at the amylase locus the LacZ activity is high in both the
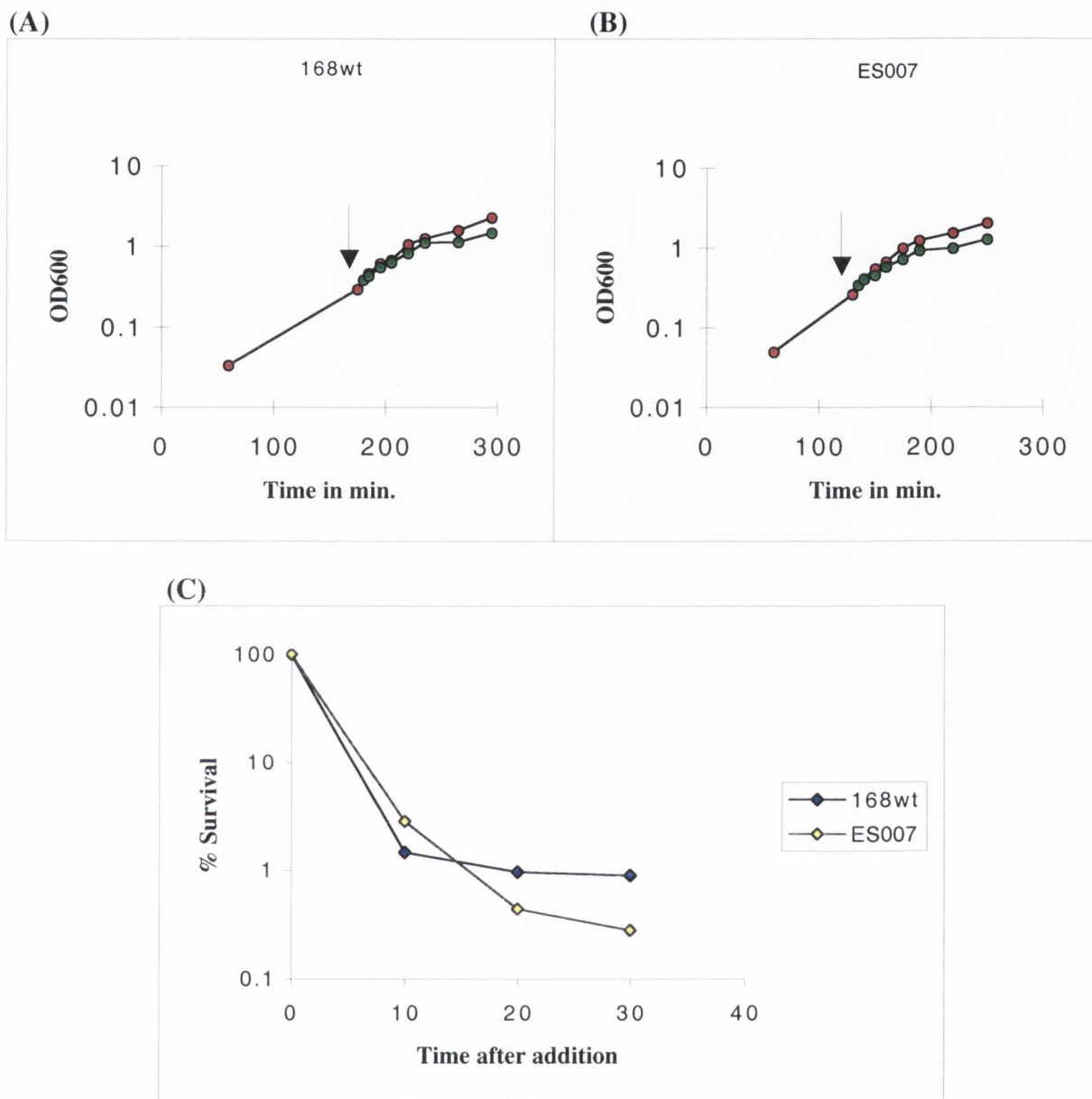
Figure 4.11 t-Butyl hydroperoxide sensitivity of *B. subtilis* strain 168 (wild type) and strain ES007 (*sodA* ::*spc*) in liquid LB medium.

Cell growth of strain(A)168 wild type and (B)the *sodA* null mutant strain ES007(*sodA* ::*spc)* with and without 500μM tBOOH added. Arrows indicate the point at which the culture was split. Cell growth was monitored by measuring OD_{600} and is shown by red (no addition) and green (500μM tBOOH added) circles. Time is presented in minutes after inoculation.

(C) Cell viability of strains 168 wild type and ES007 was measured after exposure to 100mM tBOOH. Percent survival was calculated 10, 20 and 30 minutes after tBOOH addition as described in Chapter 2 (2.7.3). Percent survival of strain 168 (wild type) is shown by blue diamonds. Percent survival of strain ES007 (*sodA* ::*spc*) is shown by yellow diamonds. Time is presented in minutes after tBOOH addition.
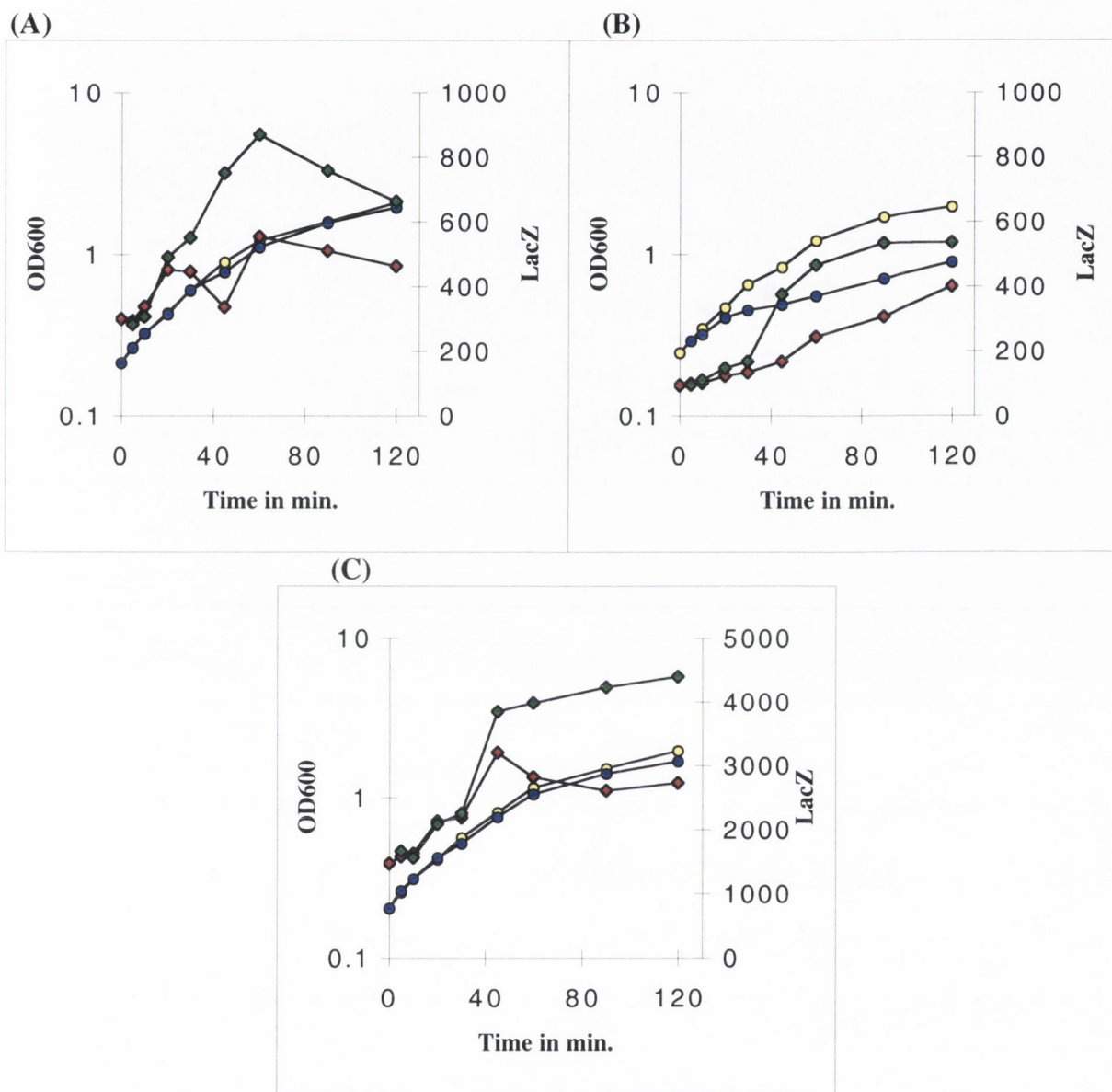
Figure 4.12 Expression profiles of *sodA* after paraquat induction.
Growth and LacZ expression profiles of strains (A)ES22 (translational P*sodA -lacZ* fusion)
(B)ES002 (sodA::pMUTIN4) and (C) ES14 (transcriptional P*sodA -lacZ* fusion) in LB
medium with and without 100µM paraquat addition. Growth was monitored by measuring
OD600 and is shown by yellow (non induced) and blue (paraquat induced) circles
LacZ activity was measured as nmole ONPG/min/mg and is shown by red
(noninduced) and green (paraquat induced) diamonds. Time is presented in minutes after
paraquat addition.

induced and non-induced cultures and the activity level does not differ significantly between the cultures until 30 minutes after paraquat addition. At this point and at 45 minutes after paraquat addition the activity levels diverge and the induction of activity in the culture containing paraquat becomes evident. Activity is approximately 2 fold greater in the induced culture than in the culture to which no addition was made. The induced culture maintains this higher level of expression up to 120 minutes after paraquat addition. This is presumably due to the continuing generation of superoxide radicals by redox cycling of paraquat within the cell. The LacZ activity levels in this strain (ES22) range between 200 and 900 units. In strain ES002 that contains the pMUTIN4 insertion into *sodA*, the pattern is similar. LacZ levels are high in both cultures, however at 45 minutes after paraquat addition the levels in the paraquat induced culture increase to approximately 2 fold greater than in the non-induced culture. The activity levels range between 200 and 700 units. In strain ES14, that contains the transcriptional promoter fusion at the amylase locus, the same pattern is observed. LacZ activity in the paraquat induced culture increases at 45 minutes after paraquat addition by approximately 1.5 fold. In ES14 however the LacZ activity levels range between 1,000 and 5,000 units, therefore though the pattern of expression is similar the overall level of LacZ activity is different between the strains.

To confirm that the observed *sodA* induction occurs at the transcriptional level and to elucidate the operon structure of the *sodA* gene, northern analysis was carried out. Samples were collected at time points after paraquat addition from control and induced cultures as indicated in Figure 4.13. RNA was prepared from the samples and northern analysis carried out as described in materials and methods (Chapter 2 section 2.5.2). The northern blot was hybridised with a *sodA* probe and the resultant autoradiogram is shown in Figure 4.13. The observed transcript size was between 623 and 955 base pairs. The *sodA* coding sequence is 606 bases pairs in length and the upstream untranslated region is 178 base pairs. This implies that *sodA* is monocistronic. Transcript of *sodA* clearly increases in the induced samples in comparison to the control. This increase is particularly significant 45 minutes after induction which correlates with the LacZ data. Transcript from the paraquat induced culture is still greater than that from the control 60 minutes after induction, but the difference is not as large as that seen in the 45 minute samples. This data confirms that *sodA* is monocistronic, that it can be induced by paraquat at the transcriptional level and that such induction takes 45 minutes to manifest itself under these conditions.
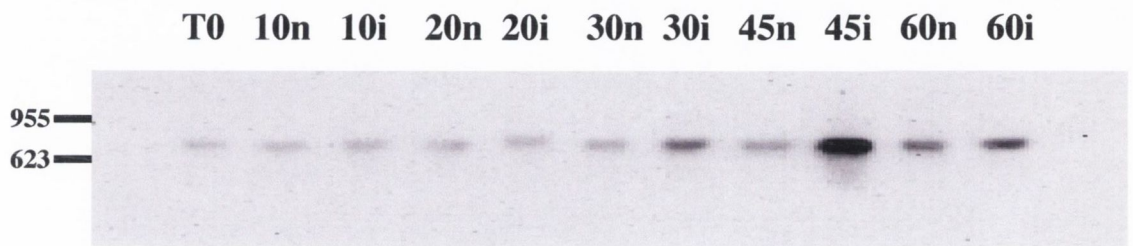
Figure 4.13 Northern blot analysis of *sodA*.
RNA was prepared from strain 168 (wild type) cells grown in LB medium before splitting the culture (T0) and 10, 20, 30, 45 and 60 minutes after from non induced cells (10n, 20n, 30n, 45n, 60n respectively) and from cells to which 100μM paraquat had been added (10i, 20i, 30i, 45i, 60i respectively). 25μg of total RNA was loaded onto each well. The positions to which the RNA size markers (623 and 955 bases) migrated are indicated. The northern blot was hybridised with a probe for *sodA*. The levels of *sodA* mRNA increase after paraquat addition.

4.2.6 Stress induction of *sodA*.

The expression of *sodA* in response to different stresses was examined using the P_{sodA}-*lacZ* transcriptional fusion in strain ES14. Addition of hydrogen peroxide did not lead to increased expression of *sodA* (data not shown). Superoxide was produced outside the cell by xanthine/xanthine oxidase (in contrast to the action of paraquat which produces superoxide intracellularly). Superoxide radicals produced in this manner did not result in increased expression of *sodA* (Figure 4.14 (A)). Cells were exposed to Nitric oxide (NO·) stress using GSNO (S-nitrosoglutathione) which is an NO· donor and SIN-1(3-morpholinosydnonimine) which produces NO· and superoxide radicals (O_2^-) that react rapidly together to produce peroxynitrite (OONO·). Neither GSNO or SIN-1 induced *sodA* transcription (Figure 4.14 (B) and (C)). The level of *sodA* transcription was increased after addition of ethanol (Figure 4.14 (D)). The activity of the *sodA* promoter rapidly increased between 5 and 10 minutes after ethanol addition. By 45 minutes after ethanol addition the activity of the induced culture is double that of the control. The level of activity in the induced culture decreases by the 60 and 90 minute time points, but is still greater than the activity of the control culture. By the 120 minute sample the level of activity of the induced culture has almost returned to that of the non-induced culture. The expression profile of *sodA* was examined in glucose replete and glucose limiting minimal medium to examine the effect of stress caused by glucose starvation (Figure 4.14 (E) and (F)). The pattern of expression differs depending on the level of glucose. Under glucose replete conditions expression is at its highest during exponential growth but falls dramatically as cells enter stationary phase. Under glucose starvation conditions the cells grow more slowly and enter the stationary phase at OD_{600}~1. The LacZ activity does not increase to levels as high as that seen in the exponential phase of glucose replete cells, probably because the cells have only a brief phase of exponential growth before they stop growing due to glucose limitation, therefore the *sodA* expression does not show the increase in expression typical of exponentially growing cells. The expression profiles show that *sodA* is not induced by glucose starvation, it is however maintained at high levels in glucose starved cells.

4.2.7 Regulation of *sodA* expression.

Primer extension analysis was used to determine the promoter structure of *sodA*. Northern analysis and transcriptional fusion data suggested that the promoter(s) for this gene lie in the 178 base pair region upstream of the *sodA* start codon. Two primers complementary to sequence within the 5' region of the *sodA* gene were used for primer extensions. Primer
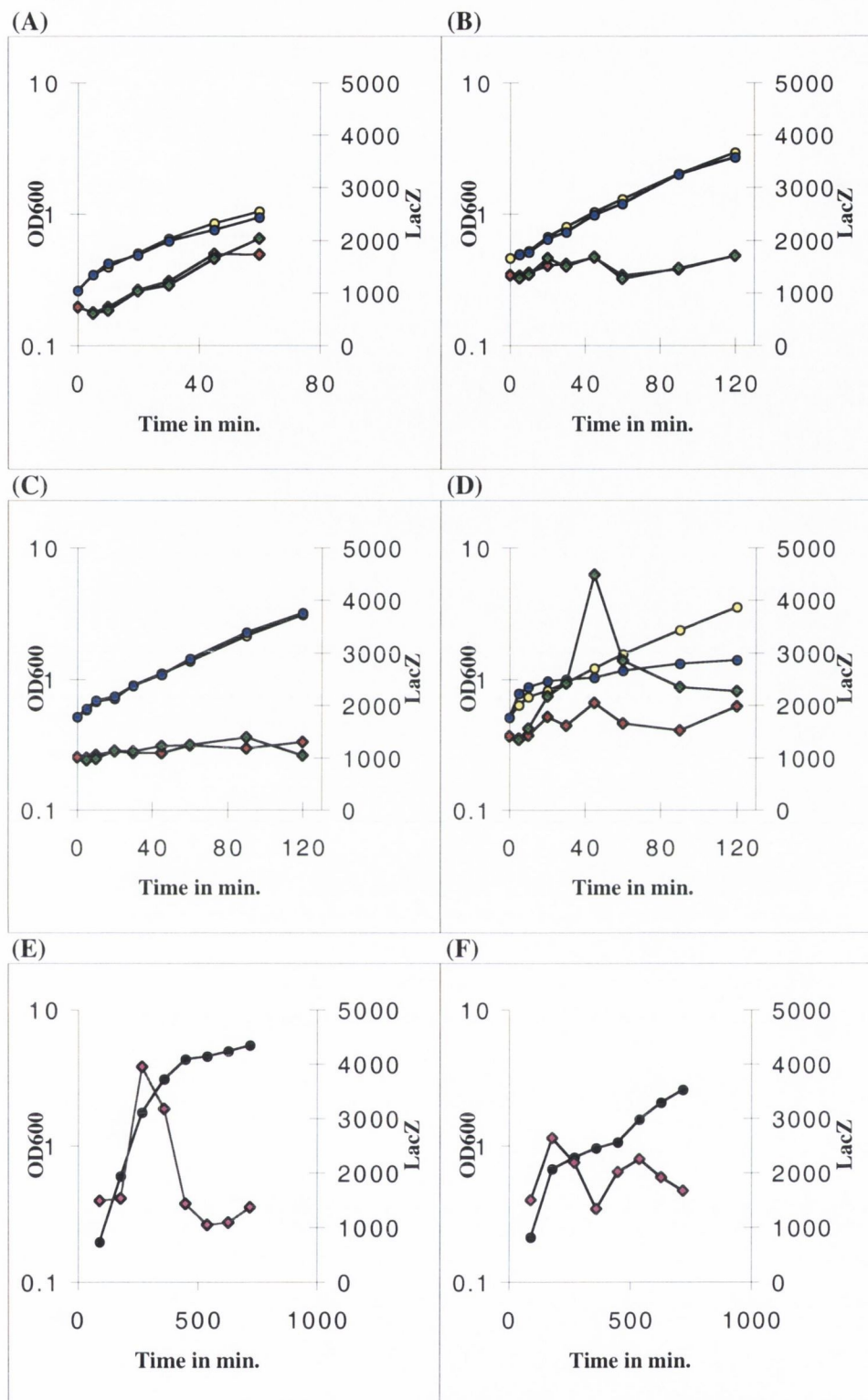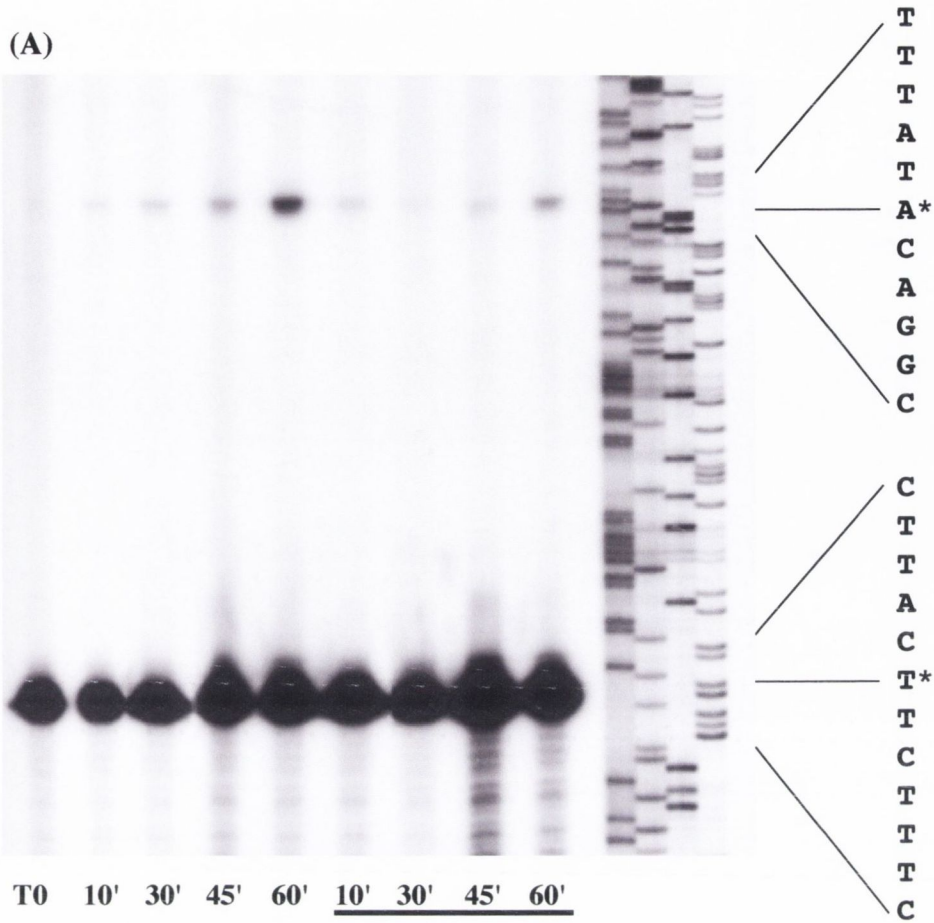
Figure 4.14 Expression profiles of *sodA-lacZ* **after exposure to various stresses.** Growth and LacZ expression profiles of strain ES14 (P*sodA* -*lacZ*) grown in the presence and absence of (A)superoxide radicals generated with xanthine/xanthine oxidase; nitric oxide stress generated by (B)GSNO and (C)SIN-1; and (D) ethanol stress. Growth was monitored by measuring OD_{600} and is shown by yellow (non induced) and blue (stress induced) circles. LacZ activity was measured as nmole ONPG/min/mg and is shown by red (non induced) and green (stress induced) diamonds. Time is presented in minutes after the addition of the stressor. (E) and (F) show growth curves and expression profiles of strain ES14 grown in (E) glucose replete and (F) glucose limiting medium. OD_{600} is shown by black circles. LacZ activity is shown by pink diamonds. Time is presented in minutes after inoculation.

extension analysis of RNA samples harvested from paraquat stressed and non-stressed wild type cells is shown in Figure 4.15. RNA was prepared from cells harvested just before the culture was divided (T_0). Paraquat was added to half the culture to a final concentration of 100μM and no addition was made to the other half of the culture. RNA samples were then prepared from cells harvested from both cultures 10, 30, 45 and 60 minutes after the culture was split. Two start sites are evident in all the samples of the primer extension (Figure 4.15). The majority of transcription initiates from a start site 56 bases upstream of the methionine of SodA. This start site is consistent with transcription beginning at a σ^A-type promoter. This promoter shows two mismatches with a consensus σ^A-type –35 motif (mismatches underlined: TTGA<u>TT</u>), has a conventional 17 base pair spacer region and has two mismatches with the consensus –10 sequence (mismatches underlined: TA<u>C</u>A<u>T</u>T). A second less intense transcript was also present in the primer extensions (Figure 4.15 (A)). This start site lies downstream of a σ^B-type consensus sequence. The –35 sequence is a perfect match to a σ^B consensus (AGGTTTAA), the spacer region between the –35 and –10 motifs is a conventional 14 bases and there are two mismatches to a consensus σ^B-type –10 motif (mismatches underlined GG<u>T</u>TA<u>A</u>). This primer extension analysis showed an increase in reverse transcript 45 minutes after paraquat addition driven by the σ^A–type promoter. There is no increase from the σ^B promoter. This indicates that the 45 minute increase in expression observed in northern and *lacZ* fusion data reflects increased transcription driven from the σ^A-type promoter. There is a very low level of transcription from the σ^B-type promoter that appears to increase in the later time points of the experiment in the 45 and 60 minute samples.

To confirm that expression of *sodA* is under σ^B- control, RNA was made from cells which had been exposed to ethanol, an inducer of the σ^B-dependent general stress response. Primer extension analysis of *sodA* transcription was performed on cells exposed to 4% ethanol stress. The results are shown in Figure 4.16 (A). At T_0, the point at which the culture was split , the σ^A directed transcript is evident but the σ^B transcript is not detectable. In the non-induced culture, transcript from both start sites increases after culture split. The majority of the transcript is from the σ^A-type promoter and transcript from the σ^B-type promoter remains low throughout the experiment. In the ethanol induced culture within 5 minutes of ethanol addition the σ^B transcript increases very significantly. Interestingly there is also a significant increase in the level of the σ^A transcript in the ethanol induced samples. To confirm that the ethanol induction was under σ^B control the

Figure 4.15 Primer extension analysis of *sodA* after paraquat stress.
(A) RNA was prepared from strain 168 wild type grown in LB medium before splitting the culture (T0) and 10, 30, 45 and 60 minutes after from non induced cells and from cells to which 100µM paraquat had been added (heavy underline). The signal is that obtained from 12.5µg of total RNA. The sequencing ladder is shown (A C G T) with the sequence indicated on the right. Two reverse transcripts were observed. The two bases at which transcription initiates are each indicated by an asterisk. (B) The sequence of the region is shown. The two bases at which transcription initiates are each shown by an asterisk. The putative σ^A motif (-35A; -10A) and the σ^B motif (-35B; -10B) are shown in bold uppercase letters. The first 20 amino acids of SodA are shown.

(A)


```
                                                    T
                                                    T
                                                    A
                                                    T
                                                    A*
                                                    C
                                                    A
                                                    G
                                                    C

                                                    C
                                                    T
                                                    T
                                                    A
                                                    C
                                                    T*
                                                    T
                                                    C
                                                    T
                                                    T
                                                    C
```

T0   10'   30'   45'   60'   10'   30'   45'   60'

(B)

```
                  -35 B                      -10 B                  *
taagaaatgctggcggcAGGTTTAAtggggcagattatcGGTTAAagtgaaatatgtccg
---------+---------+---------+---------+---------+---------+
attctttacgaccgccgtccaaattaccccgtctaatagccaatttcactttatacaggc

                           -35 A                        -10 A
ctgaaatgagccaaatcgtgagctttttcaTTGATTacaaagcattctttttgtTACATTg
---------+---------+---------+---------+---------+---------+
gactttactcggtttagcactcgaaaaagtaactaatgtttcgtaagaaaacaatgtaac

       *
aatgaagaaaggctcgctgtgaagcgtcatttcagtacatatatactaaggaggaattat
---------+---------+---------+---------+---------+---------+
ttacttctttccgagcgacacttcgcagtaaagtcatgtatatatgattcctccttaata

  M  A  Y  E  L  P  E  L  P  Y  A  Y  D  A  L  E  P  H  I  D
catggcttacgaacttccagaattaccttatgcgtacgatgctttagaaccgcatatcga
---------+---------+---------+---------+---------+---------+
gtaccgaatgcttgaaggtcttaatggaatacgcatgctacgaaatcttggcgtatagct
```

Figure 4.16 Primer extension analysis and LacZ expression profiles of *sodA* after ethanol stress in the wild type strain and a *sigB* null mutant strain.

RNA was prepared from cells of strain 168 (wild type) (A) or the *sigB* null mutant strain PB344 (*sigB::spc*) (B). Cells were grown in BFA minimal medium (section 2.1.2) and harvested before the culture was divided (T0) and then 5, 10 and 20 minutes afterwards from the non induced culture or after addition of 4% ethanol (heavy underline). The signal is that obtained from 12.5μg of total RNA. In wild type cells (A) ethanol causes induction at the σ^B promoter and the σ^A-type promoter as indicated to the left of the figure. In cells of strain PB344 (B) there is no reverse transcript detected from the σ^B promoter and transcription is induced by ethanol from the σ^A- type promoter. Figures (C) and (D) show the LacZ expression profiles after ethanol stress in BFA minimal medium for strain ES14 (P*sodA*-*lacZ*) and ES37 (P*sodA*-*lacZ sigB::spc*) respectively. Cells were sampled before the culture was divided and 5, 10, 20, 30, 45, 60, 90 and 120 minutes afterwards. 4% ethanol was added to one half of the culture and no addition was made to the non induced half of the culture. Cell growth was monitored by measuring OD_{600} and is shown by yellow (non induced) and blue (ethanol induced) circles. LacZ activity was measured as nmole ONPG hydrolysed/min/mg and is shown by red (non induced) and green (induced) diamonds. Time is presented in minutes after ethanol addition. The graph shown in (D) was previously shown in Figure 4.14 and is reproduced here for the purposes of comparison with primer extension data.

(A)

σB —

σA —

T0   5'   10'   20'   5'   10'   20'

(B)

T0   5'   10'   20'   5'   10'   20'

(C)

OD600

LacZ

10

5000

1

4000

3000

2000

1000

0

0   20   40   60   80   100   120

Time in min.

(D)

OD600

LacZ

10

5000

1

4000

3000

2000

1000

0

0   20   40   60   80   100   120

Time in min.
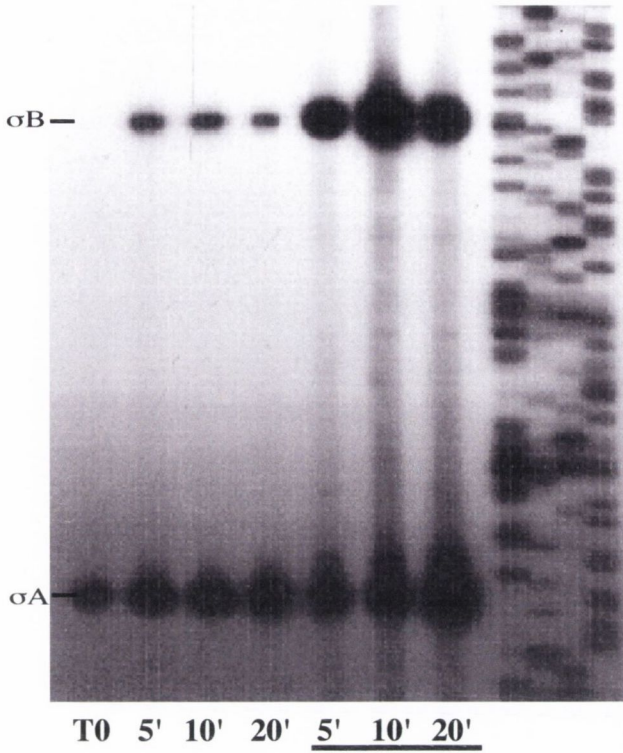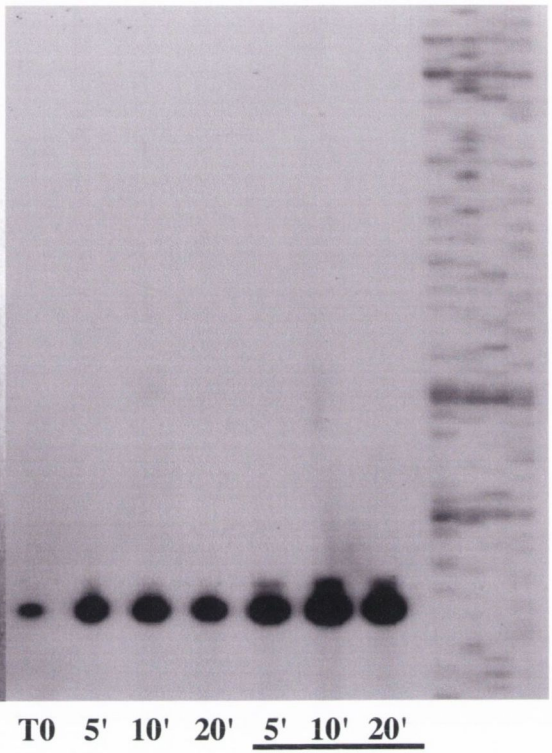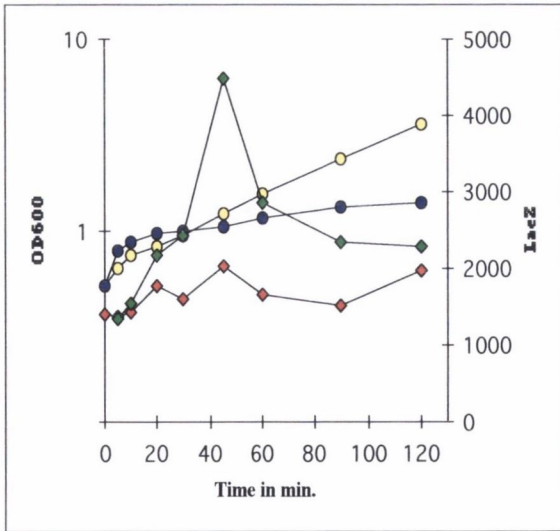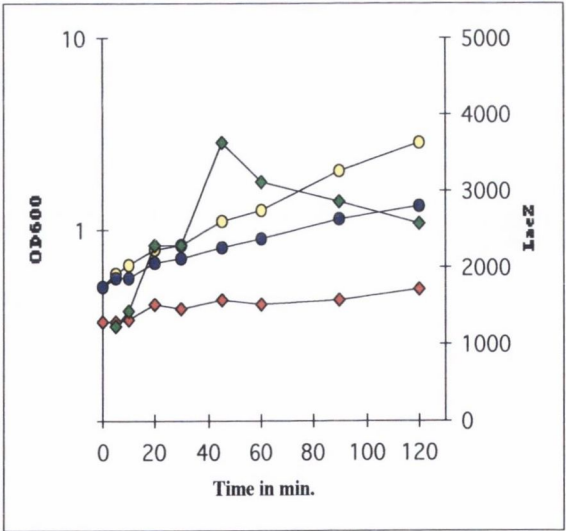
same experiment was carried out in σ^B mutant, PB344 (Boylan *et al.*, 1993). The result of this primer extension is shown in Figure 4.16 (B). The σ^B directed transcript disappears in these samples, confirming that the basal and induced transcript were both σ^B-dependent. However, the ethanol induction of transcription from the σ^A promoter is still evident in the σ^B mutant. LacZ assays on strain ES14 (amyE::P$_{sodA}$-*lacZ*) and ES37 (amyE::P$_{sodA}$-*lacZ sigB*::*spc*) confirmed the results of these primer extensions. These are pictured in Figure 4.16 (C) and (D). The graph shown in Figure 4.16 (C) is the same as shown previously in Figure 4.14 (D), it is reproduced here to allow comparison with the primer extension analysis. In strain ES14 the increase in expression from the *sodA* promoter was evident 10 minutes after ethanol addition. By 45 minutes the expression in the induced culture is over double that of the non-induced culture (4483 units in the induced culture and 2041 units in the non-induced culture). 120 minutes after the culture split the expression of the induced culture decreases to close to that of the non-induced culture. Figure 4.16 (D) shows the same experiment carried out for the *sodA* promoter in a σ^B mutant background (strain ES37). The expression profile is very similar to that in the wild type background with the greatest difference between the induced and non-induced samples evident at 45 minutes when the expression in the induced culture is over double that of the non-induced culture (3623 units in the induced culture and 1576 units in the non-induced culture). However expression levels do not reach the maximum achieved in the wild type cultures, reflecting the contribution of the σ^B-promoter to activity. This is an interesting result as it indicates that whatever regulator is involved in induction from this start site is capable of responding to both superoxide and ethanol stress.

To evaluate the contributions of the σ^A and σ^B promoters to *sodA* expression under glucose limiting and glucose replete growth conditions, primer extension analysis was performed (Figure 4.17 (A) and (B)). RNA was prepared from cells harvested at the times indicated (Figure 4.17 (C) and (D)). Figure 4.17 (A) shows the results of primer extension analysis carried out on RNA prepared from cells grown in glucose replete medium. Under these conditions expression is high during exponential growth and falls as cells enter stationary phase. The majority of transcript is from the σ^A promoter and expression from the σ^B promoter is very low at all the time points. Figure 4.17 (B) shows the results of primer extension analysis on RNA prepared from cells grown in glucose starvation conditions. The σ^A promoter directed expression is high during exponential growth and remains high throughout the glucose starvation induced stationary phase. Expression from the σ^B promoter varies slightly throughout growth but remains low. These analyses show that

Figure 4.17 Primer extension analysis of *sodA* after growth in glucose replete and glucose limiting media

RNA was prepared from cells of strain 168 (wild type) at eight time points during growth in glucose replete (0.4% glucose) and glucose limiting (0.05% glucose) medium. The growth curves from which cells were harvested are shown for cells grown in glucose replete medium (C) (timepoints indicated T1-T8) and glucose limiting medium(D) (time points indicated t1-t8). Cell growth was monitored by measuring OD_{600} and is shown by blue circles. Time is presented as minutes after inoculation. The results of primer extension analysis are shown for glucose replete conditions (A) and glucose limiting conditions (B). The signal is that obtained from 12.5μg of total RNA. Transcript initiates from a σ^B promoter and a σ^A- type promoter as indicated to the left of the figure.

(A) T1 T2 T3 T4 T5 T6 T7 T8

σB—

σA—

(B) t1 t2 t3 t4 t5 t6 t7 t8

(C)

OD600

10

1

T1 T2 T3 T4 T5 T6 T7 T8

0 200 400 600 800

Time

(D)

OD600

10

1

t1 t2 t3 t4 t5 t6 t7 t8

0 200 400 600 800

Time
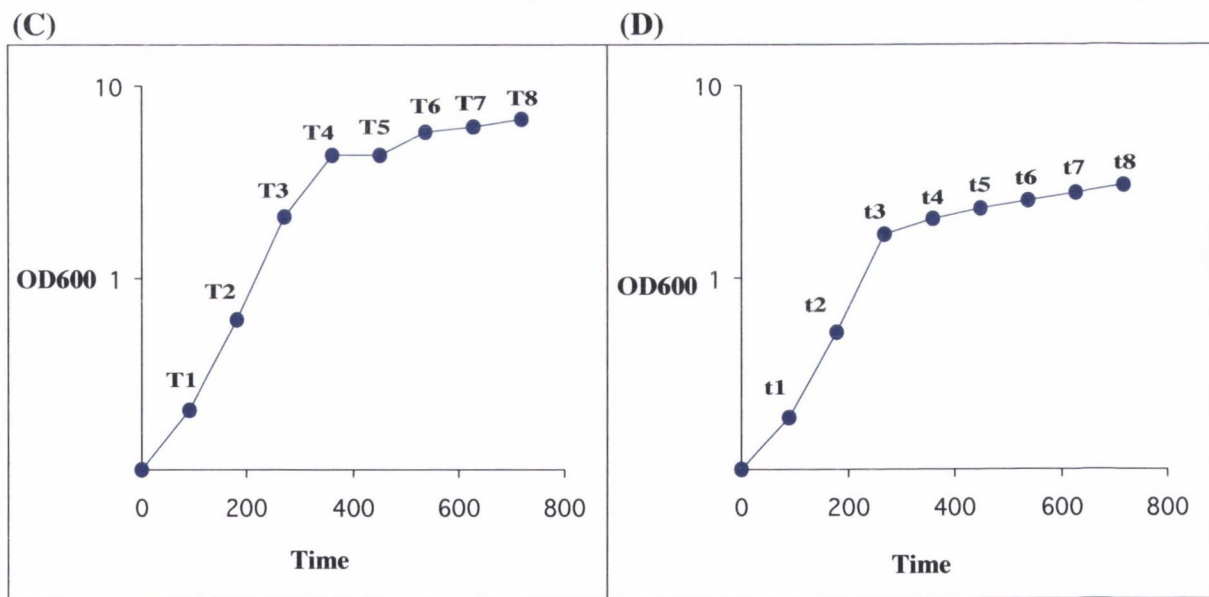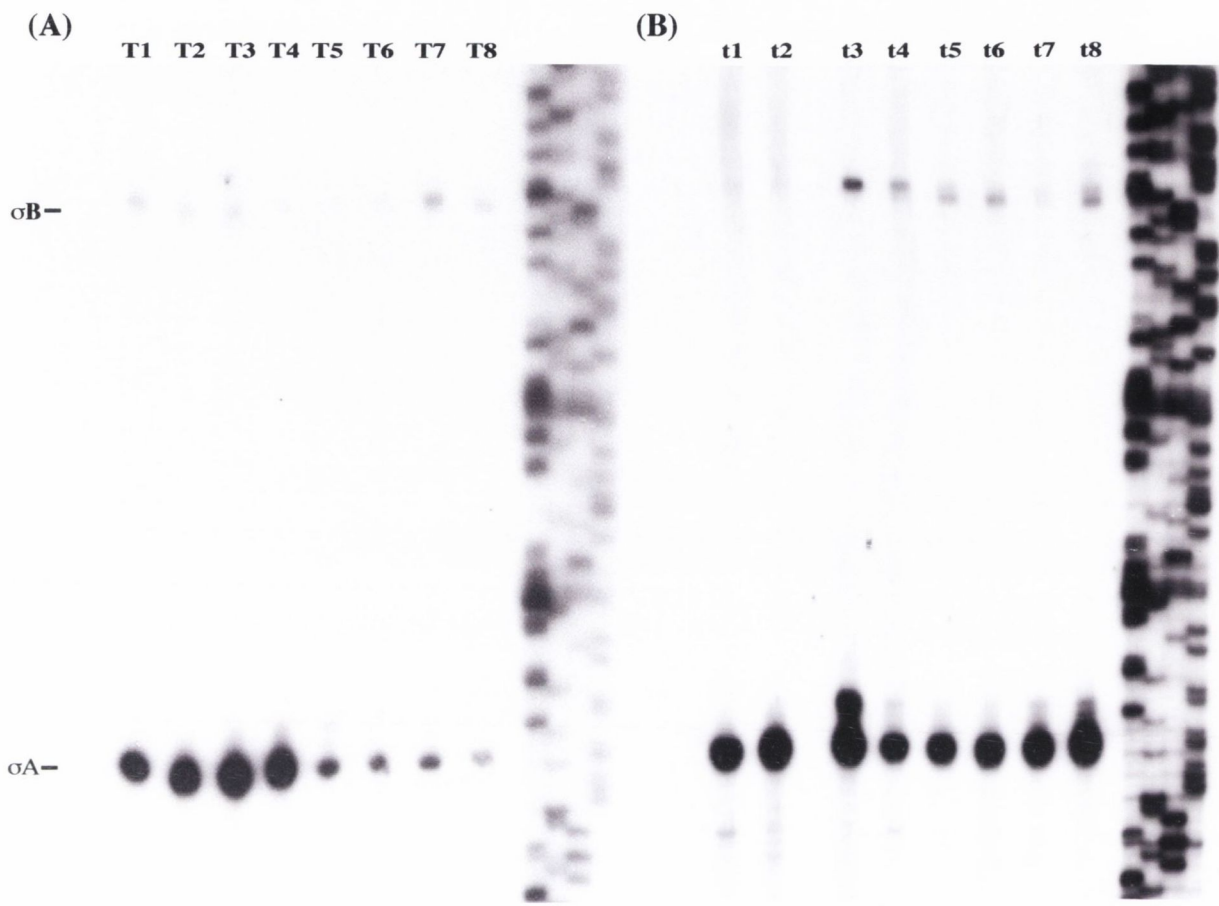
expression from the *sodA* promoter is not induced by energy stress caused by glucose starvation. However the level of *sodA* transcript persists during a glucose starvation induced stationary phase whereas it decreases dramatically during stationary phase after the cells have been grown in glucose replete medium.

4.2.8 Effect of null mutations in *yqgC*, *yqgB* and *fur* on *sodA* expression.

The primer extension results showed *sodA* is under dual control. It is a member of the σ^B-controlled general stress regulon and it is also induced by superoxide stress though a σ^A promoter in a manner not yet elucidated. To investigate the possibility that the *sodA* gene is regulated by a protein encoded by the upstream gene, *yqgC*, or the divergently transcribed gene, *yqgB*, strains carrying null mutations in these ORFs were generated using pMUTIN4XZ (Chapter 2, section 2.4.2) and the expression of *sodA* monitored through a P$_{sodA}$-*lacZ* fusion at the amylase locus (Figure 4.18 (A) and (B)). The strains ES20 and ES21 contain null mutations in the *yqgB* and *yqgC* genes respectively, generated by insertional inactivation. Results presented here show that inactivation of *yqgB* has no effect on basal *sodA* expression or on the level of expression after paraquat induction (Figure 4.18 (A)). However inactivation of *yqgC* leads to a lower level of *sodA* induction by paraquat than previously observed (Figure 4.18 (B), Figure 4.12 (C) for wild type levels). These data suggest that the protein encoded by *yqgC* may influence *sodA* expression.

Fur, the iron homeostasis regulator, has been shown to be involved in regulation of *sodA* (MnSOD), and *sodB* (FeSOD) in *E. coli* (Tardat and Touati, 1991; Dubrac and Touati, 2002). Since *B. subtilis* also contains a Fur regulator, (Bsat *et al.*, 1998) it was decided to see if *sodA* is regulated by Fur. A null *fur* mutant was constructed in *B. subtilis* by insertion of a plasmid carrying a kanamycin cassette into the *fur* gene. Strain ES25 contains the plasmid, pDG780, inserted into *fur* and the *sodA* promoter transcriptionally fused to *lacZ* at the amylase locus. When strain ES25 was subjected to paraquat stress the profile of expression differed substantially from that observed for the *sodA* promoter fusion alone (Figure 4.18 (C) and figure 4.12 (C) for wild type levels). For strain ES25 at 45 minutes after the culture was split the LacZ levels are almost 2,000 units, similar to the level observed in the wild type non-induced culture at this time point. However there is very little difference between expression levels in the non-induced and paraquat induced cultures. In contrast in the wild type expression is induced 2 fold at this time point. For strain ES25 expression in both cultures decreases after the 45 minute time point to less
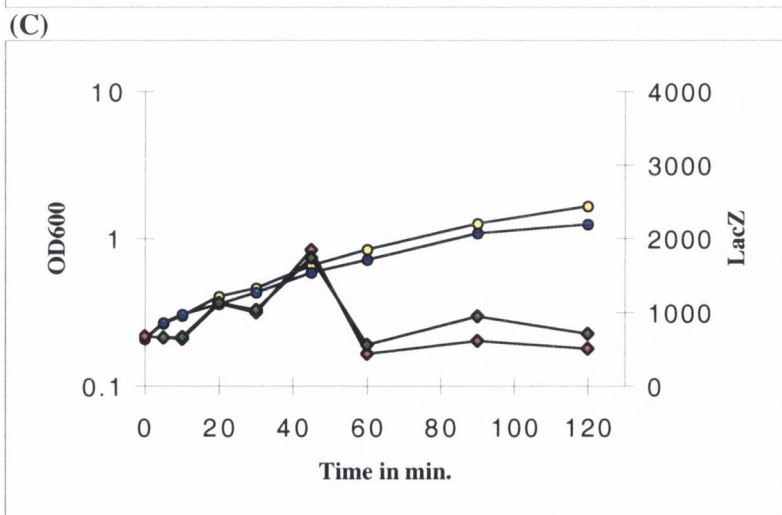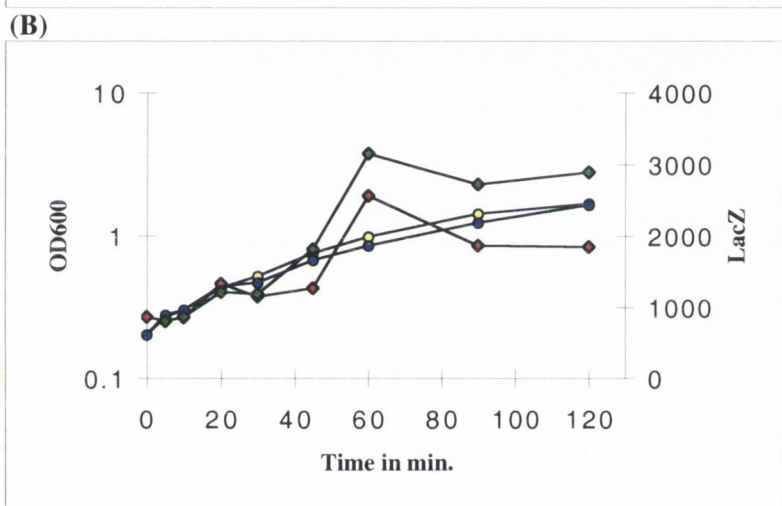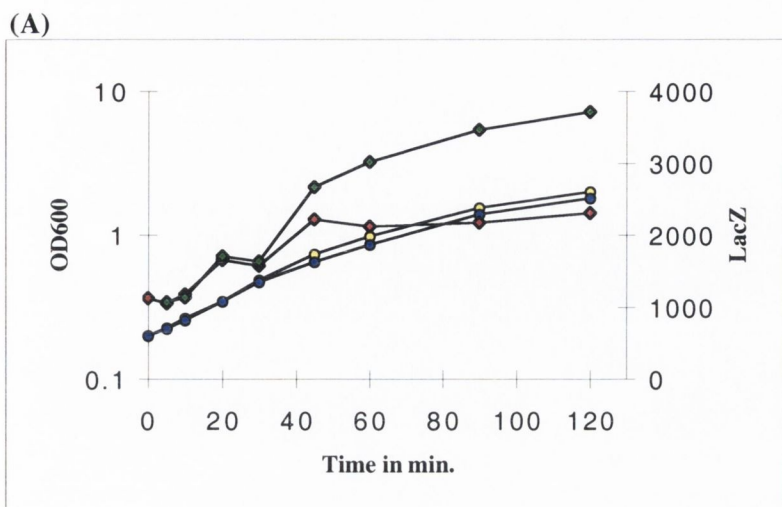
Figure 4.18 Expression from the *sodA* promoter in *yqgB*, *yqgC* and *fur* null mutant backgrounds. Growth and LacZ expression profiles of strains (A) ES20 (P*sodA*-*lacZ* *yqgB*::pMUTIN4XZ); (B) ES21(P*sodA*-*lacZ* *yqgC*::pMUTIN4XZ) and (C) ES25 (P*sodA*-*lacZ* fur::pDG780) grown in LB medium in the presence and absence of 100µM paraquat. Growth was monitored by measuring OD600 and is shown by yellow (non induced) and blue (paraquat induced) circles.LacZ activity was measured as nmole ONPG/mg/min and is shown as red (non induced) and green (paraquat induced) diamonds. Time is presented in minutes after paraquat addition.

than 1,000 units. This contrasts with the wild type where expression is maintained at ~2,000 units in the non-induced culture and ~4,000 units in the paraquat induced culture. In strain ES25 the *sodA* expression is only slightly increased in the paraquat induced culture at the 90 and 120 minute time points. In addition, it appears from the OD_{600} values that 100μM paraquat slightly inhibits growth of the *fur* mutant. These data clearly demonstrate that *sodA* expression is regulated either directly or indirectly by Fur in a manner that affects both basal and paraquat induced expression levels.

4.2.9 Expression of *sodF*.

The *sodF* gene encodes a protein that is homologous to the iron/manganese family of SODs but it is not obvious from the amino acid sequence of SodF whether it binds iron, manganese or both as its catalytic metal (Section 4.2.1). Strain ES001 contains a *sodF'-lacZ* transcriptional fusion in a null *sodF* background generated by pMUTIN4 insertion into the gene by homologous recombination. Expression of *sodF'-lacZ* was very low in this strain with less than 5 *lacZ* units of activity, and expression was not induced by paraquat (Figure 4.8 (A)) or by superoxide radicals generated extracellularly by xanthine/xanthine oxidase (data not shown). To determine if *sodF* expression could be activated by superoxide stress in the absence of *sodA,* strain ES0013 was constructed. This strain carries a pMUTIN4 insertion into the *sodF* gene generating a *sodF'-lacZ* fusion, and contains a spectinomycin cassette inserted into *sodA* to generate a *sodA* null mutant. As expected for a *sodA* null mutant this strain shows growth inhibition after exposure to paraquat (figure 4.19 (A)). However after superoxide induction, notwithstanding the null mutation in *sodA*, there is no induction of transcription of *sodF* (Figure 4.19 (A)). To determine whether this SOD is involved in the stress response to glucose starvation, strain ES0013 was grown in minimal medium under glucose replete and glucose starvation conditions. The results in figure 4.19 (B) and (C) show that accumulation of LacZ remains very low under glucose replete conditions with less than 2 units of activity. However a small increase in *sodF* expression is observed when cells enter a glucose starvation induced stationary phase, when levels rise to 6 units of activity. Nitric oxide stress mediated by SIN-1 does not induce *sodF* expression (data not shown). The *sodF-lacZ* expression was examined in a *fur* mutant background in strain ES23. These results show very low *sodF* expression throughout growth in this strain (data not shown).

To determine whether the *sodF* gene encoded a functional superoxide dismutase whose activity could be detected on activity gels the strain ES32 was constructed in which the
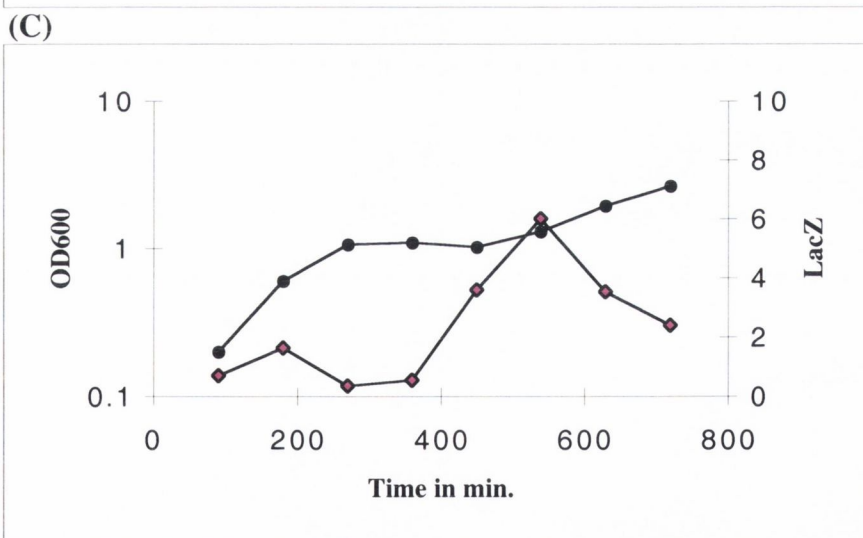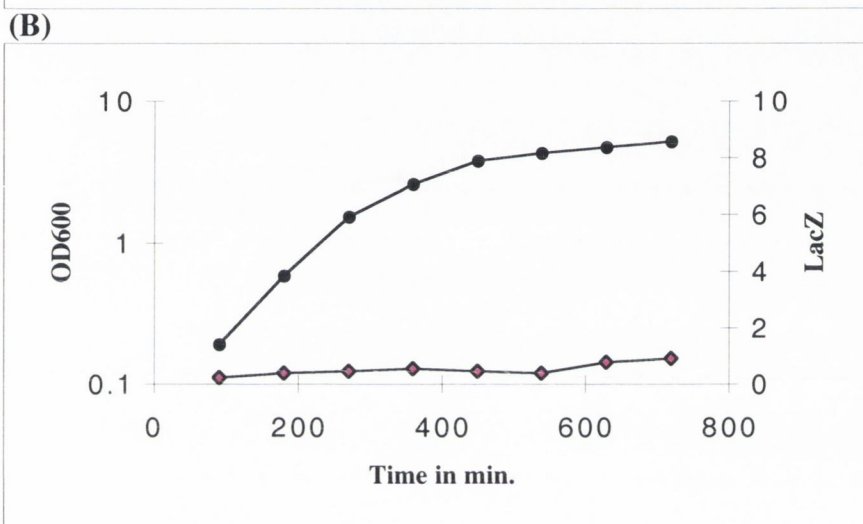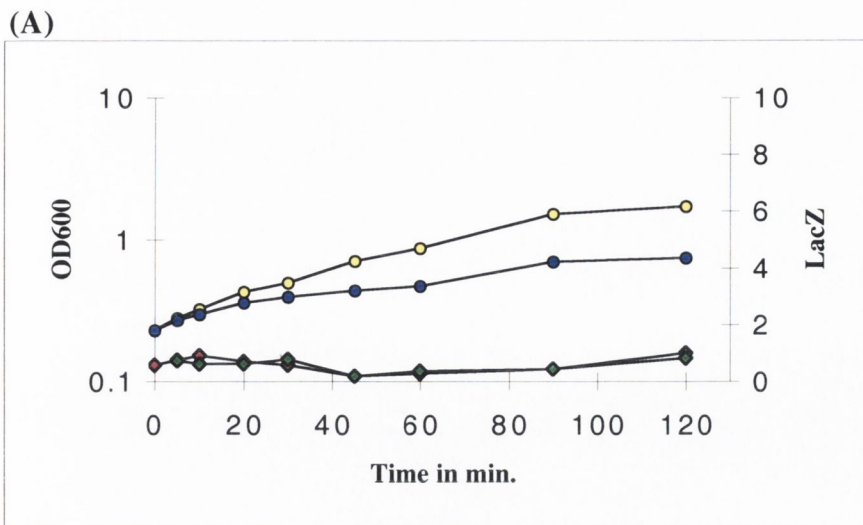
Figure 4.19 Expression of *sodF*.   Growth curves and LacZ expression profiles of strain (A) ES0013 (*sodF'-lacZ sodA* ::*spc*) grown in LB medium in the presence and absence of 100μM paraquat. Cell growth was monitored by measuring OD600 and is shown by yellow (non induced) and blue(paraquat induced) circles.LacZ activity was measured as nmole ONPG/mg/min and is shown by red (non induced) and green (paraquat induced) diamonds. (B) and (C) show growth curves and expression profiles of strain ES001 (*sodF'-lacZ*) grown in (B) glucose replete and (C) glucose limiting medium. OD600 is shown as black circles. LacZ activity is shown as pink triangles.

sodF gene is under the control of the IPTG inducible P_{SPAC} promoter. This strain is also a *sodA* mutant and therefore more sensitive to paraquat than the wild type. The sensitivity of this strain to paraquat was examined in the presence and absence of IPTG to ascertain whether it was affected by *sodF* expression. These experiments were carried out by a senior sophister student, Jennifer Chubb. The OD_{600} values of ES32 cultures grown in LB medium and exposed to paraquat in the presence and absence of IPTG were monitored. Results showed that the sensitivity of the cultures to paraquat was similar both in the presence and absence of IPTG (data not shown). Protein extracts were prepared from exponentially growing ES32 cells in LB medium in the presence and absence of IPTG. Protein extracts were also prepared from exponentially growing ES32 cells grown in basic limitation medium (Chapter 2 section 2.1.2) supplemented with 100μM $FeSO_4$. These samples were separated by non denaturing protein gel electrophoresis and the gel was stained for SOD activity (Figure 4.20). A band of superoxide dismutase activity was visible in extracts from wild type cells which was absent from protein extracts prepared from ES32 (a null *sodA* mutant), therefore this was probably SodA activity as it did not appear in *sodA* mutant strains. No SOD activity could be detected for strain ES32 even when *sodF* was fully expressed by addition of 1mM IPTG. This suggests that either SodF is produced but at undetectable levels or that the protein produced is not active under these conditions.

4.2.10 Expression of *yojM*.

The *yojM* gene is similar to copper, zinc superoxide dismutases (Section 4.2.1). Since the operon structure of the *yojM* region is not clear it was decided to examine the promoter activity of two DNA fragments: the 92 nucleotide region immediately upstream of *yojM* and the 381 nucleotide region upstream of *yojL*. Strain ES15 carries a transcriptional fusion in which the *yojM* upstream region is fused to *lacZ*. ES38 carries a transcriptional fusion in which the region upstream of *yojL* is fused to *lacZ*. Both fusions were placed in single copy at the amylase locus. LacZ activity under glucose starvation and glucose replete conditions was examined (Figure 4.21). During growth under these conditions LacZ activity is low for both strains. The ES38 strain is most highly expressed with activity levels reaching 18 lacZ units during growth in glucose replete medium and 16 units during growth in glucose limiting medium (Figure 4.21 (C) and (D)). The ES15 strain has very low activity with a maximum of 6 units in glucose limiting conditions and no expression in glucose replete medium (Figure 4.21(A) and (B)). Both strains were also tested for activity after nitric oxide stress (by addition of SIN-1), after paraquat stress in a
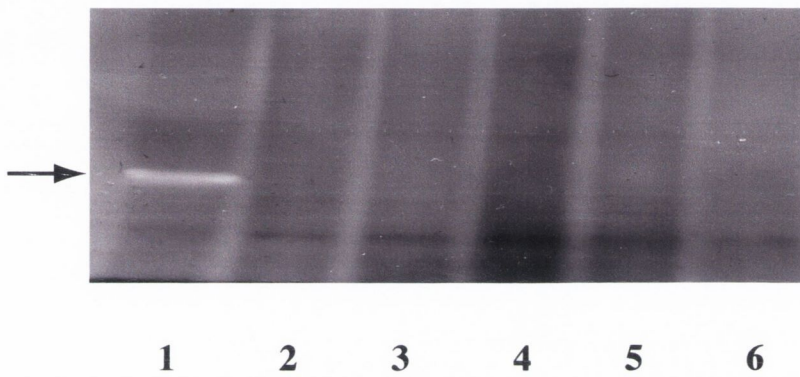
Figure 4.20 Superoxide dismutase activity gel.
Protein extracts were prepared from cells harvested during exponential growth,
seperated by non denaturing PAGE and stained for superoxide dismutase activity
as described in Chapter 2 (2.8.4). Protein extracts were prepared from strains
168 wild type (Lane 1); ES007 (*sodA::spc*) (Lane 2); ES31 grown with 1mM
IPTG (PSPAC-*yojM sodA::spc*) (Lane 3); ES31 grown without IPTG (Lane 4);
ES32 grown with 1mM IPTG *(PSPAC-sodF sodA::spc*) (Lane 5); ES32 grown
without IPTG (Lane 6). The achromatic band of SodA superoxide dismutase
activity is indicated with an arrow (Lane 1). No other superoxide dismutase
activity is detectable.
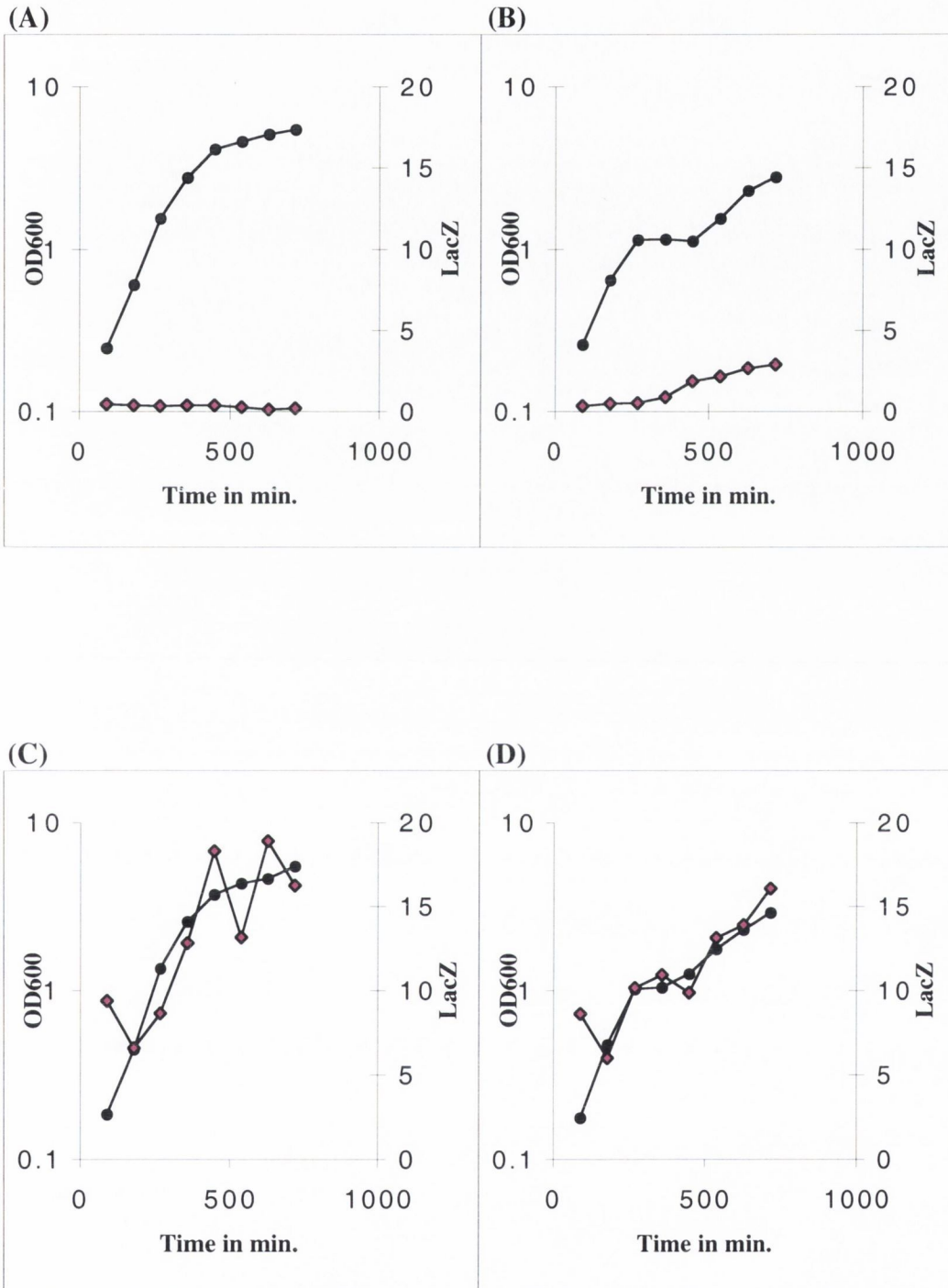
Figure 4.21 Expression profiles of putative *yojM* promoters.
Growth curves and LacZ expression data of strains (A) ES15 (intergenic region upstream of *yojM* fused to *lacZ*) in glucose replete and (B) ES15 in glucose limiting medium and (C) ES38 (intergenic region upstream of *yojL* fused to *lacZ*) in glucose replete and (D)ES38 in glucose limiting medium. Cell growth was monitored by measuring OD600 and is shown by black circles. LacZ activity was measured as nmole ONPG hydrolysed/min/mg and is shown by pink diamonds. Time is presented in minutes after inoculation.

sodA mutant background and after exposure to extracellular superoxide. Expression was not induced under these conditions and activity remained low (data not shown). Therefore *yojM* expression could not be induced either by superoxide or by glucose starvation. Expression levels remained low throughout the growth curve.

Since the gene positioned downstream of *yojM* is a putative nitric oxide reductase, it was decided to test the expression of *yojM* under conditions of nitric oxide stress by northern analysis. RNA was extracted from cells of wild type *B. subtilis* after nitric oxide stress generated by addition of either SIN-1 or SNAP (S-Nitroso-N-Acetylpenicillamine, an NO˙ donor). RNA samples were prepared from cells at the time the culture was split, from cultures to which SIN-1 or SNAP had been added from control cultures with no addition at 10, 20, 30 and 60 minutes after the culture was split. A northern blot of these samples was hybridised with a *yojM* probe. The resultant autoradiogram is shown in Figure 4.22 (A). There is a very faint transcript in the induced cultures 10 minutes after addition of either SIN-1 or SNAP. No bands were visible in any other samples. The size of this transcript was between 1049 and 1517 bases which does not fit the predicted transcript size of *yojM* alone or *yojM* and *yojN* (*yojM*:588 bases; *yojN*: 912 bases). However no increase in LacZ activity was observed at this time after induction in either strain ES15 or ES38. To confirm that SIN-1 and SNAP were inducing nitric oxide stress conditions the blot was stripped and then reprobed with an *hmp* probe. The *hmp* gene of *B. subtilis* encodes a flavohaemoglobin that is induced by nitric oxide stress (Nakano *et al.*, 2000). The control blot hybridised with the *hmp* probe clearly shows that the cells were undergoing nitric oxide stress in these cultures (Figure 4.22 (B)). Both SIN-1 (left hand samples) and SNAP (right hand samples) samples show increased levels of *hmp* transcript compared to their non induced controls. No *hmp* band is visible in the non-induced samples. After 10 minutes a band is visible in the culture which was induced by SNAP. After 20 minutes bands are visible in the induced samples of both SIN-1 and SNAP cultures. Transcript in both induced culture was higher after 30 minutes and induction of transcription was maximal 60 minutes after addition of SIN-1 or SNAP. The level of transcript was higher in the 60 minute SNAP induced sample than in the equivalent sample induced by SIN-1. The transcript ran alongside the marker band which was 1517 bases in length. The *hmp* Northern hybridisation clearly shows that the amount of SIN-1 and SNAP used in these experiments is sufficient to cause nitric oxide stress in the cell and to induce expression of at least one gene known to be induced by nitric oxide stress. However, the *yojM* gene is not clearly induced under these conditions (Figure 4.22).
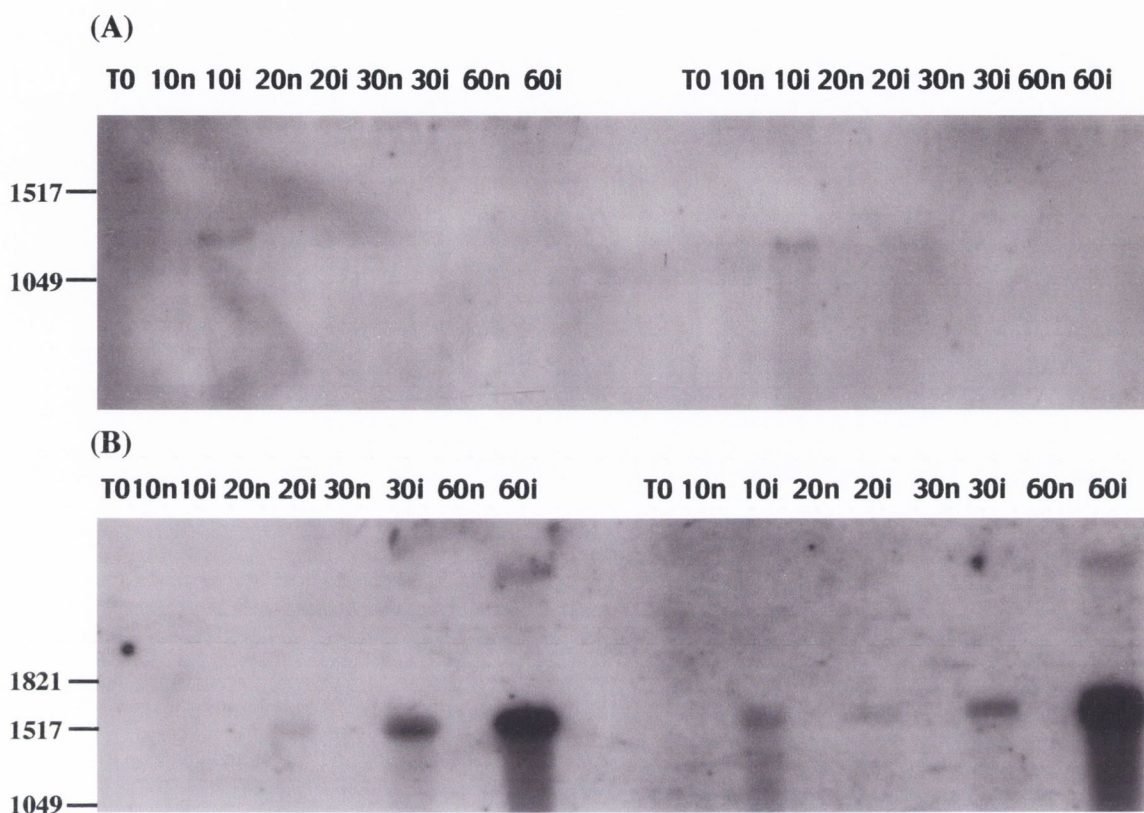
(A)

T0 10n 10i 20n 20i 30n 30i 60n 60i   T0 10n 10i 20n 20i 30n 30i 60n 60i

1517—

1049—

(B)

T010n10i 20n 20i 30n   30i 60n 60i   T0 10n 10i 20n 20i   30n 30i 60n 60i

1821 —

1517 —

1049—

Figure 4.22 Northern blot analysis of *yojM* and *hmp* after nitric oxide stress.
RNA was prepared from wild type strain 168 cells grown in BFA minimal medium
before the culture was divided (T0) and 10, 20, 30 and 60 minutes afterwards
from the non induced culture (indicated 10n, 20n, 30n, 60n) and the culture to which
SIN-1 (left side of blots) or SNAP (right side of blots) (indicated 10i, 20i, 30i, 60i)
had been added. Twenty five micrograms of total RNA was loaded onto each lane.
The positions to which the RNA size markers migrated are indicated to the left of
the figures. The Northern blots were hybridised with probes for (A) *yojM*. and
(B) *hmp*.

Strain ES31 contains a spectinomycin cassette inserted into *sodA* and a pMUTIN4 insertion at the *yojM* locus which results in the *yojM* gene being placed under P_{SPAC} control. The paraquat sensitivity of this strain in the presence and absence of IPTG was examined by monitoring growth turbidimetrically. These experiments were carried out by Jennifer Chubb. It was found that the expression of *yojM* by P_{SPAC} did not affect the paraquat sensitivity of this strain (data not shown). Exponentially growing ES31 cells grown in LB medium in the presence and absence of IPTG and cells were harvested and protein extracts prepared. In addition protein was also prepared from ES31 cells grown in basic limitation medium supplemented with copper (100µM $CuCl_2$) and zinc (100µM $ZnCl_2$). These protein extracts were separated on a non-denaturing protein gel and stained for SOD activity. No SOD activity bands were visible in protein extracts of ES31 grown with or without IPTG (Figure 4.20). These data show that *yojM* expression is very low, cannot be induced by superoxide, nitric oxide or glucose starvation stresses and that SOD activity cannot be detected after induction of the gene by a heterologous inducible expression system.

4.3 DISCUSSION

In initial experiments to characterise the superoxide stress response in *B. subtilis* several 'superoxide regulon' genes whose protein level increased after paraquat induced superoxide stress were studied. In 2D gel analysis studies it was found that paraquat exposure induced SodA, IlvC, LeuC, CysK and DhbB (Antelmann *et al.*, 1997). The *sodA* gene encodes the vegetative superoxide dismutase responsible for superoxide detoxification. The *ilvC* gene encodes ketol acid reductoisomerase and *leuC* encodes 3-isopropylmalate dehydratase. Both of these enzymes are involved in branched chain amino acid synthesis. The 3-isopropylmalate dehydratase contains an iron-sulfur cluster. The superoxide radical can oxidize the iron-sulfur clusters of enzymes leading to their inactivation (Flint *et al.*, 1993) which probably explains the upregulation of this protein under superoxide stress. It is not immediately obvious, however, why ketol acid reductoisomerase is induced, as this enzyme does not contain an iron-sulfur cluster. The *cysK* gene encodes cysteine synthetase A which is involved in cysteine biosynthesis. In *E. coli* cells, superoxide stress leads to a requirement for cysteine for optimal growth, due to a leakage of sulfite from the cell (Benov *et al.*, 1996). It is possible that a similar effect in *B. subtilis* cells causes the *cysK* upregulation. The *dhbB* gene encodes isochorismatase and is a member of the dihydroxybenzoate siderophore biosynthesis operon. This operon has been previously demonstrated to be under the control of Fur, the ferric uptake repressor protein (Bsat and Helmann, 1999). Null mutants and transcriptional *lacZ* fusions of these genes were generated with pMUTIN4 (Vagner *et al.*, 1998). Of these strains only the expression of the *sodA-lacZ* fusion was induced after addition of paraquat. The induction of *sodA* confirms the results of Antelmann *et al.* (Antelmann *et al.*, 1997) and demonstrates that the increase in SodA levels after paraquat stress is due to an increase in transcription from the *sodA* promoter. However the other fusions do not show any induction after addition of paraquat. This may reflect differences in the experiments. Antelmann *et al.* used minimal medium in their experiments and exposed cells to stress at $OD_{500}=0.4$ whereas in these experiments richer LB medium was used and cells were exposed to paraquat at $OD_{600}=0.2$. Different physiological effects of growth on rich medium compared to minimal medium could result in differences in expression. The use of different media is probably responsible for the time difference of the *sodA* induction in these two studies. Alternatively, the increased levels of protein observed by Antelmann *et al.* may be due to post-transcriptional regulation such as increase in the stability of the mRNA or the protein. These effects would not be detected using transcriptional *lacZ* fusions.

The sensitivities of the mutant stains listed above to various oxidants were examined. The *cysK* null mutant was sensitive to H_2O_2. The *cysK* gene encodes an enzyme involved in cysteine biosynthesis, cysteine synthetase A, and it is not clear why a mutation in this gene gives a H_2O_2 sensitive phenotype. The paraquat sensitivity phenotype of a *sodA* mutant has been reported previously (Casillas-Martinez and Setlow, 1997; Inaoka *et al.*, 1998) and it is unsurprising, given the high expression levels of SodA during growth, that this phenotype is also evident in liquid medium. The reason for the *sodA* mutant sensitivity to tBOOH on solid medium is less obvious. One possible explanation lies in the toxicity of iron to the cell. In the absence of SodA the increased level of superoxide radicals may react with metals, such as ferric iron, resulting in the formation of the reduced metal (ferrous iron). Superoxide may also cause the release of iron from the iron-sulfur clusters of susceptible enzymes. Such metals may then catalyse the Fenton reaction, where the reduced metal reacts with H_2O_2 or alkyl hydroperoxides and forms the toxic hydroxyl radical. So it is possible that the presence of hydroperoxides in the absence of SodA will lead to increased hydroxyl radical formation and this may be the cause of the tBOOH sensitivity of the *sodA* mutant. The fact that the *sodA* mutant is sensitive to tBOOH but not to CHP may be due to the different localisations of these hydroperoxides in the cell. tBOOH is located in the cytoplasm (Radi *et al.*, 1993) and *sodA* is probably a cytoplasmic enzyme. However CHP localises in the membrane (Klimek *et al.*, 1998). The *sodA* mutant is not sensitive to H_2O_2 in these experiments. This may be due to the upregulation of the genes of the PerR regulon by H_2O_2 and the resultant detoxification of H_2O_2 by the *katA* encoded catalase which is a member of the regulon. To examine whether members of the PerR regulon were also induced by tBOOH induction experiments were carried out using *ahpC* and *katA* promoter *lacZ* fusions. These promoters were induced after tBOOH exposure but not to a high level. This effect has been shown previously for *ahpC*. Quantification of the increase in *ahpC* mRNA after stress showed a significant increase after induction with cumene hydroperoxide but a greater increase was achieved after induction with H_2O_2 (Antelmann *et al.*, 1996). Consequently, the low level of induction of *katA* and *ahpC* after exposure to tBOOH may not be enough to protect the cell from oxidative damage in the absence of *sodA*. The fact that a *sodA* mutant and the wild type have the same sensitivity to H_2O_2 on solid medium was also reported by Casillas-Martinez and Setlow (Casillas-Martinez and Setlow, 1997). Interestingly Inaoka *et al.* observed that a *sodA* mutant is more sensitive than wild type to H_2O_2 when grown in liquid medium (Inaoka *et al.*, 1999). The difference between growth on liquid and solid medium is also

evident for the tBOOH sensitivity phenotype of a *sodA* mutant. This phenotype has only been observed when the strain is grown on solid and not in liquid medium.

The paraquat induction of *sodA* transcription was examined in three different *lacZ* fusion strains. These were strain ES002, a transcriptional *lacZ* fusion generated by pMUTIN4 insertion into *sodA*; ES22, a translational fusion of the *sodA* promoter to *lacZ* positioned at the amylase locus and ES14, a transcriptional *sodA* promoter fusion positioned at the amylase locus. The paraquat induction of *sodA* was evident in all three strains (Figure 4.12). The profile of expression was also similar in all three strains. The level of expression is high even in the non-induced cultures. The transcriptional induction due to superoxide stress does not occur until 30 to 45 minutes after addition of paraquat. This delay before induction of transcription was also observed in northern analysis (Figure 4.13). It is presumably due to the fact that paraquat itself is not the inducer. The cell takes up paraquat and upon redox cycling of this compound, the inducer molecule, the superoxide radical, is produced. This result appears inconsistent with the data of Antelmann *et al.* who see protein induction 20 minutes after paraquat addition (Antelmann *et al.*, 1997). This anomaly might be explained by the fact that Antelmann *et al.* use minimal medium in their paraquat experiments and richer LB medium was used in the experiments described here. It has been shown for *E. coli* that salts inhibit paraquat uptake (Kitzler and Fridovich, 1986). The high level of NaCl in LB may possibly inhibit paraquat uptake and may be responsible for the delay in induction of *sodA* seen in this study. The level of induction of transcription from the *sodA* promoter is approximately 2 fold. As *sodA* is a highly expressed gene under normal growth conditions, this level of induction is clearly sufficient to manage the detoxification requirements of superoxide stress conditions. However, as can be seen from Figure 4.12, there is a large discrepancy in the levels of LacZ activity observed between the transcriptional promoter *lacZ* fusion in strain ES14 and the levels seen in the other two strains. Whereas the lacZ levels range between approximately 200 to 800 units in strains ES22 and ES002, the levels of activity in strain ES14 range between approximately 1,000 and 4,000 units. The difference in *lacZ* expression between strains ES14 and ES22 may be due to differences in ribosome binding as strain ES14 uses the *spoVG* ribosome binding sequences and as ES22 is a translational fusion to *lacZ* it uses the *sodA* ribosome binding site. However the *spoVG* ribosome binding site and the *sodA* ribosome binding site both complement the 3' 16S rRNA sequence at ten out of fifteen bases, therefore another regulatory factor may be involved in the observed differences. ES14 also has differences in expression levels to ES002

(pMUTIN4 insertion into *sodA*), and both these *lacZ* fusions have *spoVG* ribosome binding sites. They are not directly comparable however as strain ES002 has a null mutation of *sodA*, whereas strain ES14 is wild type for *sodA*. In addition the *lacZ* fusion of strain ES002 is at the *sodA* locus whereas in strain ES14 (and ES22) the 187 base pair region upstream of *sodA* was fused to *lacZ* at the amylase locus. Primer extension results have shown that two promoters of *sodA* lie within this region (Section 4.2.7), however other factors involved in modulating expression may only be effective at the *sodA* locus and may not influence expression from the promoter fragment at the amylase locus. It has been shown that deletion of the gene upstream of *sodA, yqgC*, affects *sodA* expression even though it does not appear to be cotranscribed (Section 4.2.8). Such factors may explain the difference between strain ES002 and ES14, but does not account for the difference between ES22 and ES14, and the reason for this result is not clear.

Primer extension analysis revealed two start sites of transcription of *sodA*, the start sites are consistent with σ^A-type and σ^B-type consensus promoter motifs. The σ^B promoter was confirmed by showing that the reverse transcript from this promoter is inducible by ethanol and that the start site disappears in a σ^B mutant (Section 4.2.7). SodA had previously been classed as a σ^B -independent general stress protein (Bernhardt *et al.*, 1997) and has recently been recognised as a σ^B –dependent gene based on the presence of the σ^B promoter (Petersohn *et al.*, 2001). It is evident that the reason for the initial classification of the gene as σ^B –independent was due to the induction of transcription by stress in a σ^B mutant. Primer extension analysis has shown that stress induction occurs from two promoters and in the absence of σ^B induction still occurs from the σ^A-type promoter (Section 4.2.7). Analysis of expression under conditions of glucose starvation has shown that *sodA*, though highly expressed when grown in glucose limiting medium, is not induced by these conditions. It appears therefore that induction at the *sodA* σ^B promoter occurs via environmental stress (*e.g.* ethanol stress) and not energy stress (*e.g.* glucose starvation). These two types of stress are sensed by two different branches of the σ^B signal transduction pathway and the signal is transduced to the common regulators RsbV and RsbW. Clearly *sodA* responds only to the environmental type stress and not the energy type and it is regulated by another mechanism under glucose starvation conditions. The σ^B regulon is not induced by oxidative stress. The superoxide stress induction of *sodA* occurs from the σ^A-type promoter and is mediated by an unknown regulator. Induction of *sodA* expression by ethanol stress also occurs from this promoter in both the wild type and a σ^B mutant strain (Section 4.2.7). This is an interesting result as it indicates that whatever regulator is

involved in induction from this promoter is capable of responding to both superoxide and ethanol stress. It is possible that the environmental signal that leads to induction from this promoter occurs as a result of cellular damage that can be caused by superoxide or ethanol stress. Alternatively, more than one regulator may be involved in stress induction at this promoter. Expression of the major SOD of *E. coli* is governed by at least six different regulators (Compan *et al.*, 1993).

Null mutants in two genes altered the expression of *sodA* under superoxide stress conditions. A null mutant of *yqgC*, a gene of unknown function that lies upstream of *sodA*, affects the paraquat induction of *sodA* expression. Addition of paraquat to the *yqgC* mutant caused induction of transcription from the *sodA* promoter but the LacZ activity observed was lower than seen in the wild type background (Section 4.2.8). The *yqgC* gene encodes a protein that has similarities to other proteins of unknown function. The gene product may have a regulatory role in *sodA* expression. Alternatively the insertion of the pMUTIN4XZ plasmid so close to the *sodA* promoter region may have some polar effect on transcription from this locus.

A null mutation in the *fur* gene was also found to affect *sodA* induction by superoxide stress. Fur is a global regulator of genes involved in iron uptake, and it usually acts as a repressor. Mutation of this gene almost entirely eliminates paraquat induction at the *sodA* promoter (Section 4.2.8). It also causes a decrease in LacZ activity compared to the wild type, in both the induced and non-induced cultures. This implies that in a wild type background, Fur has a positive effect on *sodA* expression, but only at the later stages of the growth cycle. This is in contrast to the situation in *E. coli*, where Fur represses MnSOD at the transcriptional level, and the levels of *sodA* mRNA are increased in a *fur* mutant (Schrum and Hassan, 1994). However, Fur does activate the expression of *sodB*, the FeSOD, in *E. coli* by increasing mRNA stability and by transcriptional activation, although this may be an indirect effect rather than due to interaction with the Fur protein (Dubrac and Touati, 2000), (Dubrac and Touati, 2002). Cells of the Fur mutant presumably have much higher levels of iron than the wild type due to derepression of genes for siderophore synthesis. In *B. subtilis* Fur has been shown to bind the promoter region of the *dhb* operon for siderophore biosynthesis (Bsat and Helmann, 1999). Paraquat generated superoxide radicals may reduce the cellular iron so that it can participate in the Fenton reaction and generate toxic hydroxyl radicals. This may be the reason for the slight paraquat mediated growth inhibition in the *fur* mutant strain (Section 4.2.8). In addition, the superoxide

radical liberates iron from iron sulfur clusters of various enzymes. This may add to the iron overload of a Fur mutant. The amount of SodA in the cell of a *fur* mutant may not be sufficient to cope with the superoxide generated by paraquat and the superoxide radicals generated by paraquat may be more toxic than in the wild type due to the abundance of iron, hence the paraquat sensitivity of the strain. The promoter region of *sodA* does not contain a Fur box, therefore it seems unlikely that changes in expression are due to the Fur protein binding in this region. Instead changes in expression from the *sodA* promoter are probably due to indirect effects of derepression of the Fur operon. A protein of the Fur regulon may interact with a superoxide stress regulator of *sodA*, or the abundance of iron as a result of derepression of the regulon may down regulate *sodA* by some other mechanism. The oxidative stress regulator PerR is known to repress *fur* and this repression is not relieved by H_2O_2 (Herbig and Helmann, 2001). As Fur apparently has a positive effect on *sodA,* PerR may indirectly have a negative effect on *sodA*. Removal of the PerR repression of *fur* by a change in PerR caused by an undefined signal would, according to this model, result in an increase in *sodA* transcription. As described in chapter 1, elevated levels of manganese cause PerR to repress the genes of its regulon through the DNA binding of the manganous form of PerR. It is probably the manganous form of PerR which represses *fur* as iron does not mediate *fur* repression (Herbig and Helmann, 2001). Therefore, according to the model, high levels of manganese would cause a decrease in the level of transcription from the *sodA* promoter. This would be consistent with the idea that Mn(II) can catalyse the non-enzymic dismutaion of superoxide (Archibald and Fridovich, 1981), (Inaoka *et al.*, 1999), therefore higher levels of SodA would not be necessary and may even be damaging to the cell due to overproduction of H_2O_2.

The genome sequence of *B. subtilis* contains two further genes that encode superoxide dismutase homologues. Experiments carried out during this work have failed to reveal the role or regulation of these genes. The *sodF* gene appears to encode a superoxide dismutase of the iron/manganese binding family, however this gene was very lowly expressed, if at all, under the conditions examined. Expression was not induced by a wide variety of growth and stress conditions. Induced expression of the gene by P_{SPAC} did not affect the paraquat sensitivity of a *sodA* mutant and did not result in any superoxide dismutase activity that could be detected on a SOD activity gel. The role of SodF and the conditions under which it becomes expressed remain to be elucidated.

The *yojM* gene is very interesting as it encodes a potential Cu,ZnSOD. To date this type of SOD has been described in eukaryotes and Gram negative bacteria but not in Gram positive bacteria. It has, however, proven difficult to discover the conditions for expression of this gene and its role *in vivo*. Stress and growth conditions examined did not induce expression of *yojM*. Expression of the gene under the control of P$_{SPAC}$ appeared to have no effect on the paraquat sensitivity of a *sodA* mutant and did not produce superoxide dismutase activity that could be seen on a SOD activity gel. The *sodC* encoded Cu,ZnSOD of *E. coli* was not initially detected due to its thermal lability, periplasmic location and low expression compared to the FeSOD and MnSOD. It is possible that YojM and SodF of *B. subtilis* are produced at levels too low to be detected on a gel, or they may be inactivated in the preparation of protein extracts, a factor which was critical for the Cu,ZnSOD of *E. coli* (Benov and Fridovich, 1994). YojM may only ever be lowly and transiently expressed as seen in LacZ assays (15-20 LacZ units being the maximum observed). As both *sodF* and *yojM* appear to be in operons with other genes, determining when these genes are expressed may reveal when these alternative SODs are expressed.

The study of the superoxide dismutases of *B. subtilis* has shown that *sodA* encodes the most important SOD and its regulation is likely to be complex and involve more than one factor. Transcriptome and proteome work may reveal regulators of *sodA* and growth conditions under which regulation is altered. The expression of *sodA* may be affected by global regulators such as those involved in aerobic versus anaerobic growth: Fnr and ResDE, or possibly the regulator of manganese homeostasis MntR. Such analysis may identify the conditions of expression of SodF and YojM. It would be surprising if these SODs did not function in superoxide detoxification *in vivo*.

CHAPTER 5

TYPE 2 NADH DEHYDROGENASES OF *B. subtilis*

5.1 INTRODUCTION

NADH dehydrogenases function to reoxidise NADH to NAD$^+$ and to feed electrons to the quinone pool during respiration. Two NADH dehydrogenases have been described in bacteria: NDH-1 and NDH-2. NDH-1 is a complex multisubunit enzyme that is homologous to the mitochondrial NADH dehydrogenase complex 1. This enzyme couples oxidation of NADH with the generation of a proton motive force. NDH-2 is a single polypeptide which also oxidises NADH and reduces quinones but the activity of this enzyme is not coupled to a proton motive force. The *B. subtilis* genome does not contain genes for NDH-1, but does contain three genes encoding NDH-2 like proteins: *yjlD, yutJ* and *yumB*. These genes have not been previously characterised. However proteome and transcriptome analysis of anaerobic growth of *B. subtilis* showed that YjlD is repressed during anaerobic growth (Marino *et al.*, 2000; Ye *et al.*, 2000). Anaerobic repression of YjlD was not relieved by null mutations in *fnr* or *resDE* (Marino *et al.*, 2000). Transcription of *yjlCD* is also repressed during the stringent response of *B. subtilis* and this repression is mediated by RelA, a (p)ppGpp synthetase and regulator of the response (Eymann *et al.*, 2002). The *yutJ* and *yumB* encoded NADH dehydrogenases have not yet been characterised.

As part of the work for this thesis, experiments were carried out to investigate the role of the NDH-2 encoding genes of *B. subtilis*. These genes are of interest not only because they encode important enzymes of the respiratory chain, but also because of their potential role as sources of superoxide production during respiration. The NDH-2 of *E. coli* is a primary source of superoxide and H_2O_2 (Messner and Imlay, 1999). In addition, phenotypic analysis of null mutants of *yjlC* and *yjlD* during the *B. subtilis* functional analysis project showed that these mutants were resistant to the superoxide generator paraquat on solid medium which pointed to a role for YjlC and YjlD in the generation of superoxide in *B. subtilis* cells.

5.2 RESULTS

5.2.1 Bioinformatic analysis of the *B. subtilis* NADH dehydrogenases.

Phenotypic analysis of the genes of unknown function encoded in the *B. subtilis* genome revealed that strains carrying null mutations in *yjlC* and *yjlD* were more resistant than wild type to paraquat when grown on solid medium. BLAST searches revealed that *yjlD* is homologous to Type 2 NADH dehydrogenases from other organisms, while no homologues of *yjlC* have yet appeared in the databases. The *B. subtilis* genome contains two paralogues of *yjlD*, both of which are also genes of unknown function: *yumB* and *yutJ*.

The genes *yjlC* and *yjlD* form a dicistronic operon (Figure 5.1) The upstream gene, *yjlB,* is not part of the operon. The *yjlC* gene is preceded by a long untranslated region (388 base pairs). The *yjlC* gene encodes a protein of 140 amino acids that does not display significant homologies to any other proteins in current databases. There is a short untranslated region of 39 base pairs between the stop codon of *yjlC* and the start codon of *yjlD*. The *yjlD* gene encodes a protein of 392 amino acids that is homologous to Type 2 NADH dehydrogenases. There is a putative rho-independent terminator positioned after *yjlD* (Figure 5.1). Type 2 NADH dehydrogenases are NADH: quinone oxidoreductases that use the flavin cofactor FAD as their redox centre. The multiple alignment in Figure 5.2 shows the conserved amino acid residues of the FAD binding site and the NADH binding site. Both FAD and NADH binding sites consist of a β sheet- α helix- β sheet structure. This conserved structure contains a glycine-rich phosphate binding domain, a negatively charged residue at the end of the second β sheet and six conserved small hydrophobic residues which stabilize the secondary structure (Howitt *et al.*, 1999; Bellamacina, 1996). Figure 5.2 illustrates these domains which are conserved in YjlD, YutJ and YumB. The FAD-binding motif is located at the N-terminal end of each of the proteins. The six conserved hydrophobic residues are marked with diamonds in Figure 5.2. The positions of the conserved amino acids in YjlD are: Ileu-5, Ileu-7, Ala-18, Val-21, Val-32 and Val-34. The glycine motif is from position 9 to 14 in YjlD. The negatively charged residue is not apparent in YjlD and YumB. In YutJ it is Asp-35 and in the *E. coli* Ndh it is Asp-40. The NADH-binding motif is located in the centre of the proteins. Only four conserved hydrophobic residues are evident in the proteins in the multiple alignment. These are marked with triangles in Figure 5.2. In YjlD they are positioned at: Ileu-157, Ileu-159, Val-170 and Leu-173. The glycine motif is located between amino acids 161-166 in YjlD. The negatively charged residue is Asp-191 for YumB, Glu-190 for YjlD, Asp-165 for YutJ and Glu-207 for the *E. coli* Ndh.
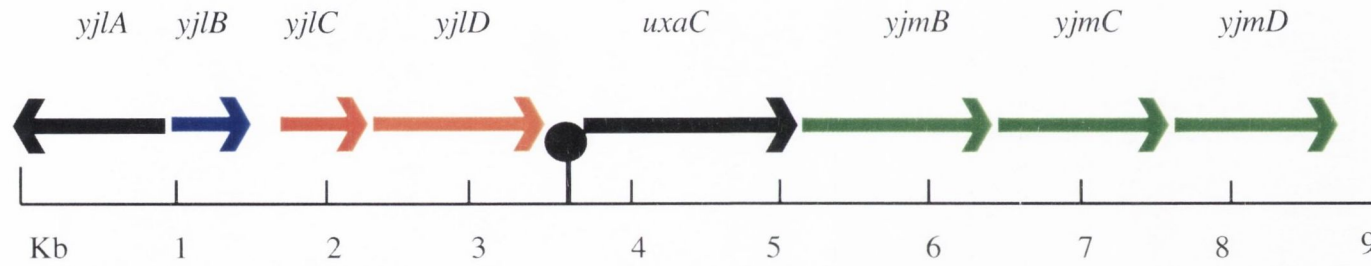
Figure 5.1 Organisation of the ORFs in the vicinity of the *yjlCD* genes of *B. subtilis.*
Arrows represent genes and are proportional to gene length. The lollipop represents a rho-independent terminator.
yjlC is a gene of unknown function. *yjlD* is a probable type 2 NADH dehydrogenase.

```
                              ♦ ♦   GXGXXG       ♦    ♦              ♦ ♦
B.subtilisYumB    1  MALNKPKIVILGAGYGGLTVTRLTKYVG-PNDADITLVNKHNYHYETTWIHEASAGTLHHDRCRYQIKDVINGSRVNFV
B.subtilisYjlD    1  --MSK-EIVILGAGYGGVLSALTVRKEYT-KEQARVTVVNKYPTHQIITELHRLAAGNVSEKAVAMPLEKIFGKDIDLK
B.subtilisYutJ    1  ----MKKIVLIGGGYGNMRVLHRLLPNQL-PDDVSITLIDRNPYHCLKTEYYALAAGTISDHHIRVSFP---EHPRIDVQ
E.coliNdh         1  MTTPLKKIVIVGGGAGGLEMATQLGHKLGRKKKARITLVDRNHSHLWKPLLHEVATGSLDEGVDALSYLAHARNEGFQFQ

B.subtilisYumB   80  QDTVKAIKIDEKKVVLANG-------------ELQYDYLVIGLGVPETFGIKGLKEYAFPIANINTSRLLREHGELQFAT
B.subtilisYjlD   77  IAEVSSFSVDKKEVALADGS------------TLTYDALVVGLGSVIAYFGIPGLEENSMVLKSAADANKVFQHVEDRVRE
B.subtilisYutJ   73  YGDITSIDIVQKQVLFQDRE-----------PISYDDAIIGLGCEDKYHNVPGAPEFTYSIQTIDQSRETYQKINN----
E.coliNdh        81  LGSVIDIDREAKTITIAELRDEKGELLVPERKIAYDTLVMALGSTSNDFNTPGVKENCIFLDNPHQARRFHQEILNLFLK

                              ▽ ▽   GXGXXG       ▽    ▽
B.subtilisYumB  148  YNTEAEKRPDRLTIVVGGAGFTGIEFLGELAARVPELCKE--YDVDRSLVRIICVEAAPTVLPGFDPELVDYAVHYLEEN
B.subtilisYjlD  146  YS--KTKNEADATILIGGGGLTGVELVGELADIMPNLAKK--YGVDKEIKLKLVEAGPKILPVLPDDLIERATASLEKR
B.subtilisYutJ  138  -------LSANATVAIVGAGLSGVELASELRE------------SRDDLNIIILFDRGNLILSSFPERLSKYVQKWFEEH
E.coliNdh       161  YS-ANLGANGKVNIAIVGGGATGVELSAELHNAVKQLHSYGYKGITNEALNVTLVEAGERILPALPPRISAAAHNELTKL

B.subtilisYumB  226  GVEFKIGTAVQECTPEGVRVGKKDEEPEQIKSQTVVWAAGVRGHPIVEAG-FENM-RGRVKVNPDLRAPGHDNVFIIGD
B.subtilisYjlD  222  GVEFLTGLPVTN--VEGNVIDLKDGS--KVVANTFVWTCGVQGNPIVGESG-LEVN-RGRATVNDFLQSTSHEDVVAGD
B.subtilisYutJ  198  GVRIINRANITK--VEEGVVYNEDDP---ISADAIVWTAGIQPNKVVRDLD-VEKDAQGRIVLTPHHNLPGDEHLYVVGD
E.coliNdh       240  GVRVLTQTMVTS--ADEGGLHTKDGE--YIEADLMVWAAGIKAPDFLKDIGGLETNRINQLVVEPTLQTTRDPDIYAIGD

B.subtilisYumB  304  SSLFMNEDTERPYPPTAQIAMQQGITVAKNLGRLIKGGELEEFKPDIKGTVASLGEHNAVGVVYG--RK----LKGTPAS
B.subtilisYjlD  296  SAVYFGPD-GRPYPPTAQIAWQMGELIGYNLFAYLEGKTLETFKPVNSGTLASLGRKDAVAIIGANSTP----LKGIPAS
B.subtilisYutJ  272  CAS-------LPHAPSAQLAEAQAEQIVQILQKRWNGEALPESMPQFK-------------------------LKGVLGS
E.coliNdh       316  CASCPRPE-GGFVPPRAQAAHQMATCAMNNILAQMNGKPLKNYQYKDGSLVSLSNFSTVGSLMGNLTRGSMMIEGRIAR

B.subtilisYumB  378  FMKKVIDNRSLFMIGGLGLTLKKGKFKFF-----------
B.subtilisYjlD  371  LMKEASNVRYLTHIKGL-FSLAY----------------
B.subtilisYutJ  320  TRKKSRLRACC----------------------------
E.coliNdh       395  FVYISLYRMHQIALHGYFKTGLMMLVGSINRVIRPRLKLH
```

Figure 5.2 Multiple alignment of the *E. coli ndh* encoded NADH dehydrogenase and its three homologues in *B. subtilis*: YumB, YjlD and YutJ. Where half or more of the residues in a column are identical/similar they are shaded. White letters in black boxes denote identical residues. Black letters in grey boxes denote similar residues. Diamonds and plain text denote the FAD-binding motif. Triangles and bold text denote the NADH-binding motif.

The *yumB* gene encodes a 406 amino acid protein that is homologous to *yjlD*. However unlike *yjlD* the *yumB* gene is monocistronic (Figure 5.3). The upstream gene (*yumC*) is divergently transcribed from *yumB* and the *yumB* ORF is followed by a rho-independent terminator. Interestingly, although *yumB* has been described as a stress inducible protein the control region does not contain a σ^B consensus motif(Petersohn *et al.*, 2001).

The *yutJ* gene encodes a 330 amino acid protein and is monocistronic. The gene is followed by a rho-independent terminator (Figure 5.4). The *yutJ* gene is a paralogue of *yjlD*. It lies 7,613 bases away from *yumB*.

5.2.2 Paraquat resistance of *yjlC* and *yjlD* null mutants.

As part of the *B. subtilis* functional analysis project a collection of insertional mutants were generated in genes of unknown function and each strain was subjected to a variety of phenotypic tests. The results of these tests can be viewed on the Micado website at http://locus.jouy.inra.fr/cgi-bin/genomic/madbase_home.pl Our laboratory tested this collection of strains for responses to the oxidising agents paraquat and hydrogen peroxide using gradient plates (Scanlan *et al.*, 2001). Of the 1,148 strains tested, 35 responded to these oxidising agents in a manner that differed from the wild type. Two strains, BFS841 and BFS842, were chosen for further investigation. Both strains were constructed using pMUTIN4 in D. Karamata's laboratory. These strains exhibited greater resistance to paraquat than the wild type strain grown on solid medium. Strain BFS841 contains a pMUTIN4 insertion into *yjlD* and strain BFS842 contains a pMUTIN4 insertion into *yjlC*. Bioinformatic analysis suggested that these two genes are cotranscribed and therefore constitute an operon. The first gene of the operon, *yjlC*, does not have significant homology to any genes in current databases and is therefore *B. subtilis* specific. The second gene of the operon, *yjlD*, is a probable type 2 NADH dehydrogenase. The operon structure and homologies are discussed in detail in section 5.2.1. In addition to their increased resistance to paraquat both BFS841 and BFS842 strains formed smaller colonies and grew more slowly than the wild type strain. After several days growth on agar plates containing X-gal, faster growing white colonies grew out from the existing small blue colonies. It is likely that these faster growing colonies are strains in which the pMUTIN4 plasmid has excised from the chromosome by a reverse Campbell-type event thereby restoring the wild type gene. It was concluded therefore, that the slow growth of these strains is due to *yjlC* and *yjlD* inactivation, implying that both these genes are required for wild type growth. Phenotypically, both strains were more resistant to paraquat than wild
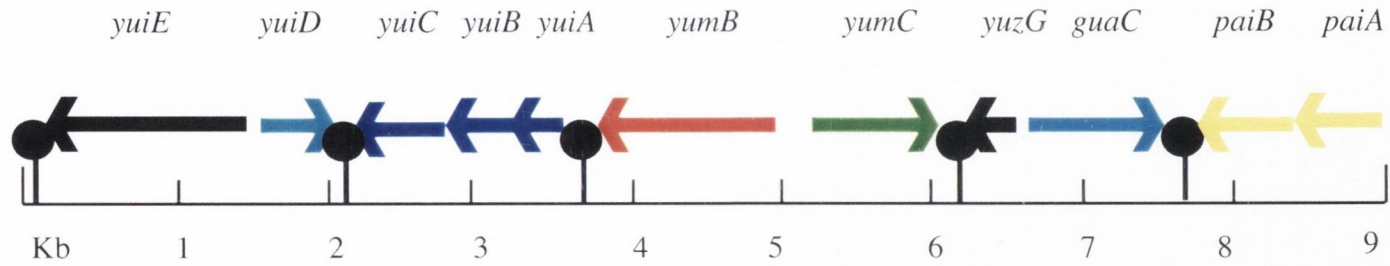
Figure 5.3 Organisation of the ORFs in the vicinity of the *yumB* gene of *B. subtilis*.
Arrows represent genes and are proportional to gene length. Lollipops represent
rho- independent terminators. *yumB* is similar to genes that encode type 2 NADH dehydrogenases.
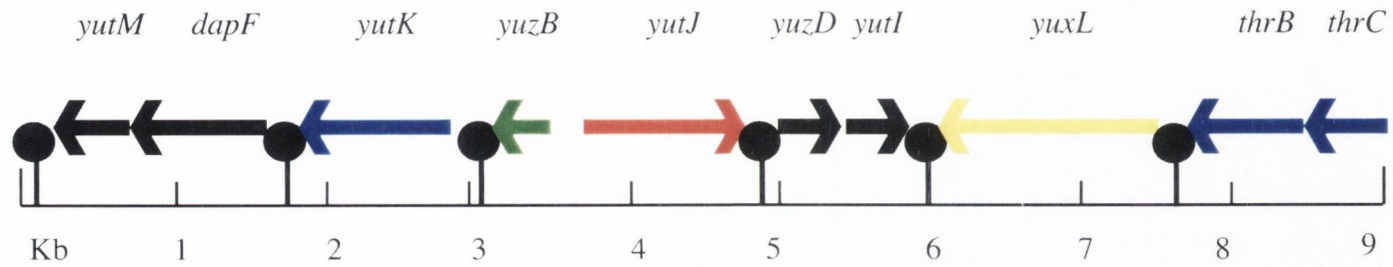
Figure 5.4 Organisation of the ORFs in the vicinity of the *yutJ* gene of *B. subtilis*.
Arrows represent genes and are proportional to gene length. Lollipops represent
rho- independent terminators. *yutJ* is similar to genes that encode type 2 NADH dehydrogenases.

type when grown on a gradient plate containing 0-875μM paraquat. The *yjlC* mutant is also more resistant than the wild type strain to paraquat on gradient plates even in the presence of IPTG which induces expression of *yjlD* which is downstream of *yjlC*. This suggests that the *yjlC* mutant phenotype is not simply a result of the polar effect of the *yjlC* insertion on *yjlD* expression.

To determine whether the increased paraquat resistance phenotype of the *yjlD* and *yjlC* mutant strains observed on gradient plates could be observed when cells were grown in liquid medium, cultures of strain 168 (wild type), strain BFS841 and strain BFS842 (in the presence and absence of IPTG) were grown to early exponential phase in LB medium. The cultures were split at an OD_{600} of ~0.2 and paraquat was added to half the culture to a final concentration of 75mM. Aliquots were removed 10, 20, 30, 40 and 50 minutes after the culture was split and cell viability was determined by plating dilutions on LB plates. Percentage survival was calculated by comparing the colony counts of the culture to which paraquat was added with the control samples to which no addition was made. The results of these experiments are shown in Figure 5.5 (A). The survival level of all the strains fell to between 10% and 30% 10 minutes after paraquat addition. After 50 minutes survival for all the strains was less than 20%. Differences seen in the graph are likely to be experimental deviations rather than a biological effect. It is evident therefore that there is not a significant difference in survival between the wild type and the mutant strains. In addition, the survival of strain BFS842 is not significantly affected by the presence of IPTG in the medium. It seemed likely that the paraquat resistance phenotype was specific for solid medium.

To quantitate the increase in paraquat resistance of strains BFS841 and BFS842 observed when the cells were grown on LB plates containing a paraquat gradient, exponentially growing cultures of the wild type strain, BFS841 and BFS842 were plated onto LB plates which contained 175μM, 350μM, 525μM and 700μM paraquat and onto LB plates without paraquat. Strain BFS842 was also grown in the presence of IPTG. Colonies were counted after 15 hours incubation at 37°C. The results are presented in Figure 5.5 (B). The colony counts of the wild type cells decrease almost 4 fold after growth on 175μM paraquat compared to colonies grown without paraquat, however colony counts fall by 3 logs when the paraquat concentration is 350μM and by 6 logs when the paraquat concentration is 525μM. At 700μM there is 7 logs difference between the colony counts with and without paraquat. In contrast, colony counts of strains BFS841 and BFS842 with/without IPTG
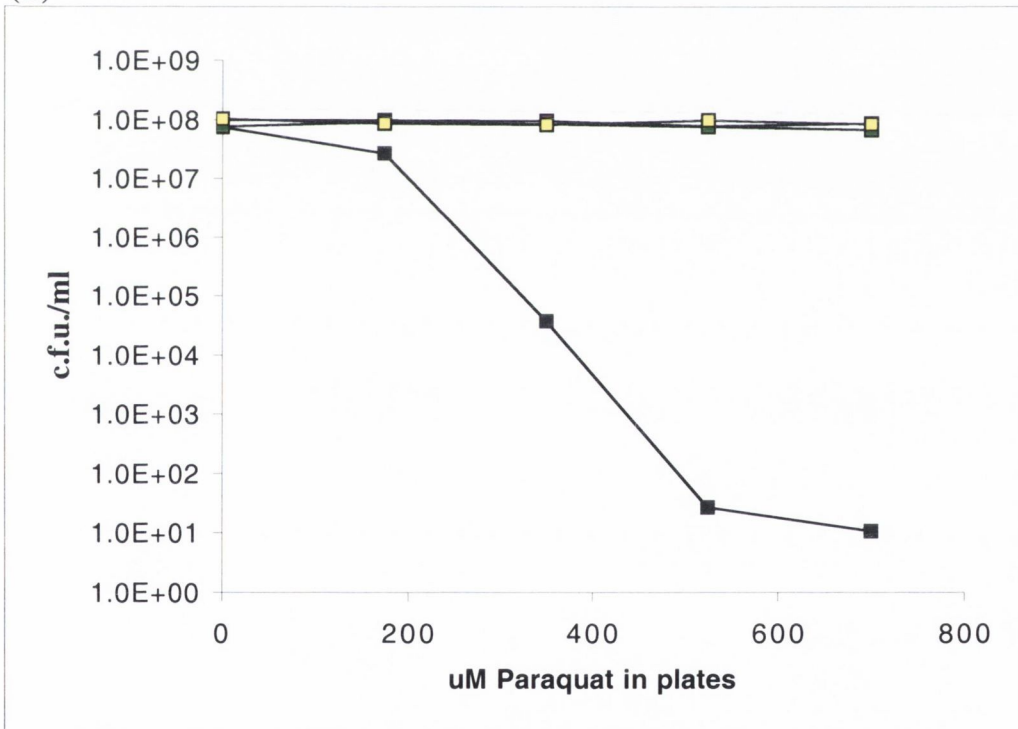
(A)
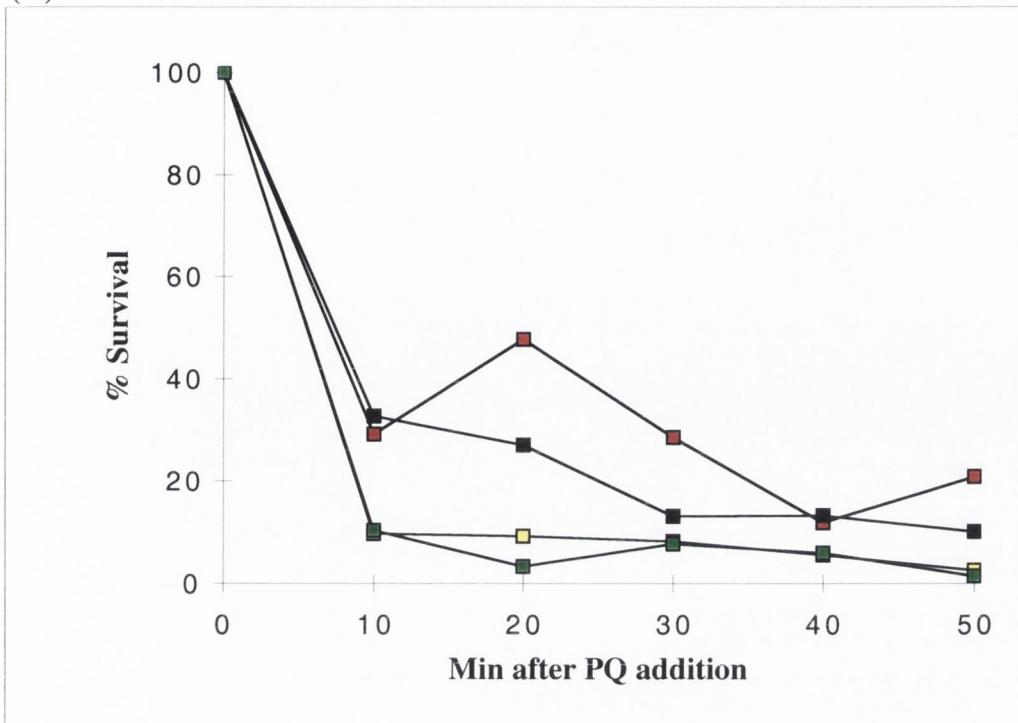
(B)

Figure 5.5 Paraquat resistance of *yjlCD* **mutant strains.**
(A) Survival of strains at time points indicated after exposure to 75mM
paraquat in liquid LB medium. Percent survival was calculated by comparison
with a non stressed culture as described in Chapter 2 (2.7.3). (B) Survival
of strains after growth on LB agar plates containing paraquat. Exponentially
growing cultures were plated on LB agar plates containing 175μM, 350μM,
525μM and 700μM paraquat. Survival of cells is measured by determining
colony counts in colony forming units/ml after overnight incubation.
For both graphs: black squares, strain 168 (wild type); red squares,
BFS841 (*yjlD* ::pMUTIN4); green squares, BFS842 (*yjlC* ::pMUTIN4);
yellow squares, BFS842 grown in the presence of 1mM IPTG.

decrease as the paraquat concentration increases and the colonies are smaller as the concentration of paraquat increases, nevertheless the drop in colony numbers compared to cells growing without paraquat is less than 1 log at the highest level of paraquat used (700μM). This is 6 logs less than the decrease in colony number seen for wild type cells. The *yjlC* and *yjlD* null mutant strains have significantly increased paraquat resistance compared to the wild type.

5.2.3 Expression and autoregulation of the *yjlCD* operon.

BFS841 and BFS842 contain pMUTIN4 insertions into *yjlD* and *yjlC* respectively which generate transcriptional *lacZ* fusions. The expression of these *lacZ* fusions was examined throughout growth in Schaeffer's medium (SM). An additional strain, ES9912, was generated in which the *yjlCD* upstream intergenic region was fused to *lacZ* and placed in single copy at the amylase locus (*amyE*::P$_{yjlCD}$-*lacZ*). The intergenic region between *yjlB* and *yjlC* is 388 bases in length and is likely to contain the promoter sequences for *yjlCD*. The LacZ activity for each strain was determined throughout the growth curve. The results are depicted in Figure 5.6. The expression profile of strain ES9912 (*amyE*::P$_{yjlCD}$-*lacZ*) is shown in Figure 5.6 (A). It is evident that expression is largely confined to exponential phase, where LacZ activity accumulates to a maximum of 65 units. Expression decreases at transition and LacZ levels fall to approximately 10 units. The expression profile of the *yjlCD* operon in a *yjlD* mutant (strain BFS841) is shown in Figure 5.6 (B). It is clear that the basal LacZ activity level during exponential growth is significantly higher than the maximum observed for strain ES9912. The activity continues to rise during exponential growth reaching a maximum of 470 units. The activity begins to fall as cells enter stationary phase. LacZ expression of the *yjlC* mutant was measured when cells were grown in the presence of IPTG (to induce expression of the downstream *yjlD* gene) and in the absence of IPTG. The profiles are shown in Figure 5.6 (C). The profiles for growth with and without IPTG are similar. Again the level of activity is significantly higher than seen for strain ES9912, though not as high as for BFS841. LacZ activity reaches a maximum of 395 units when cells are grown with IPTG and 327 when grown without IPTG. The differences in activity levels between the wild type (ES9912) and mutant (BFS841 and BFS842) strains may signify negative autoregulation at the *yjlCD* promoter by the gene products of *yjlCD*.

Figure 5.6 also shows the growth curves for these strains and the difference in growth rates between strains carrying the *yjlC* and *yjlD* mutants and ES9912, which is wild type for
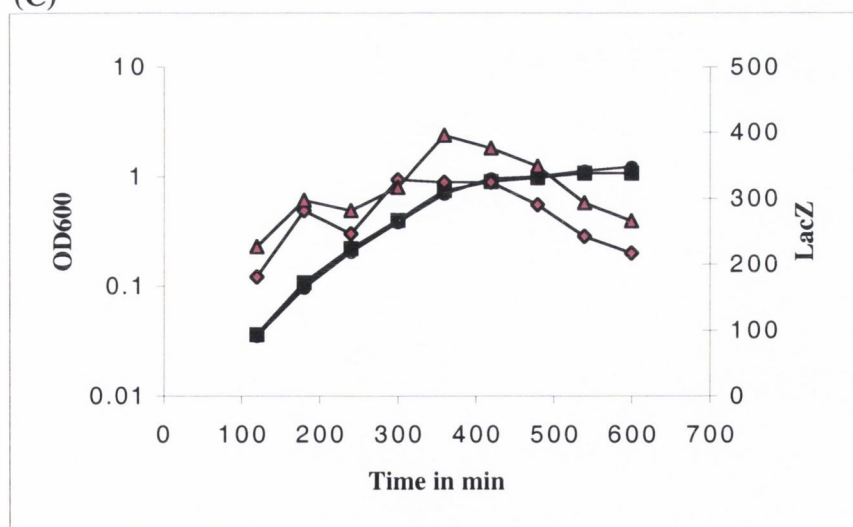
(A)
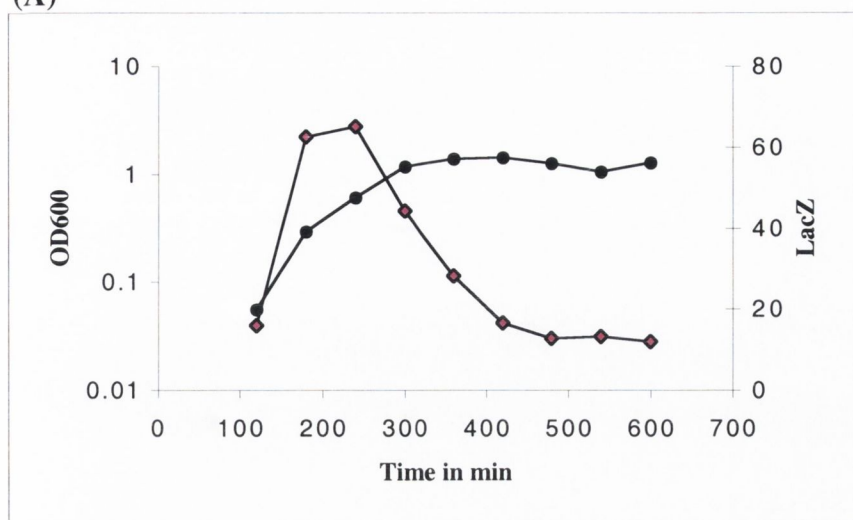
(B)
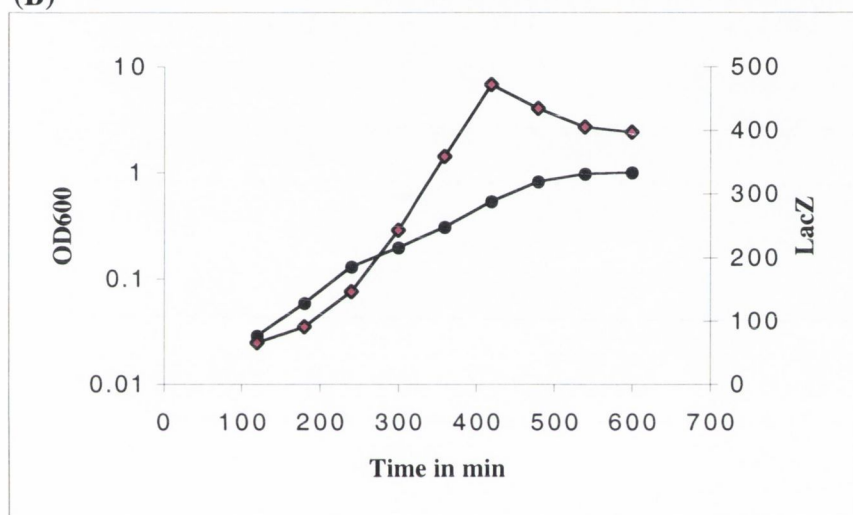
(C)

Figure 5.6 Growth and expression profiles for *yjlCD*.
Growth curves and expression profiles in SM medium for strains (A) ES9912
(*amyE* ::P*yjlCD* -*lacZ)* (B) BFS841 (*yjlD* '-*lacZ*) and (C) BFS842 (*yjlC* '-*lacZ*).
Growth was measured by monitoring OD_{600} and is shown by black circles,
except for BFS842 grown with 1mM IPTG where it is shown by black squares.
LacZ activity was measured as nmole ONPG/min/mg and is shown by pink
diamonds except for BFS842 grown with 1mM ONPG where it is shown by pink
triangles. Time is presented in minutes after inoculation.

yjlCD, is clear. It is especially evident that BFS841 and BFS842 have a very pronounced lag phase and take several hours before exponential growth begins. In contrast there is almost no lag phase when strain ES9912 is grown in this rich medium. These data show that (i) mutation of *yjlC* and *yjlD* leads to slow growth, (ii) *yjlCD* expression is confined to exponential growth and (iii) there is evidence of negative autoregulation of the *yjlCD* operon.

5.2.4 Northern blot analysis of the *yjlCD* operon.

Northern analysis was carried out to determine the operon structure of the *yjlCD* locus. RNA was prepared from cells harvested during mid exponential growth (T_1), transition phase (T_2) and stationary phase (T_3). RNA was separated by electrophoresis on formaldehyde gels and northern analysis was carried out. The northern blots were probed separately with *yjlC* and *yjlD* probes and the results are shown in Figure 5.7. It is evident that the number, sizes and levels of transcript are very similar in the northern blots probed with *yjlC* and *yjlD* probes indicating that the genes are co-transcribed. Both northerns show two transcripts. The larger, more intense transcript is between 1821 and 2661 and is sufficient to encode both the *yjlC* and *yjlD* cistrons. The smaller transcript is just under the 1517 size marker and may be a processed form of the large transcript. A northern blot containing the same RNA samples was also hybridised with a probe for the upstream gene, *yjlB*. Transcript could not be detected on this blot showing that this gene is not in the same operon as *yjlCD*.

5.2.5 Primer extension analysis of the *yjlCD* operon.

To determine the promoter structure of *yjlCD* and to investigate the possible negative autoregulation effect (Section 5.2.3), RNA was prepared from cells of strain BFS841 and the wild type strain in exponential (T_1: OD600~ 0.3), transition (T_2: OD600~ 1.0) and early stationary phase (T_3: OD600~ 1.3). Primer extension analysis was carried out on these samples to determine the start point of transcription. The results are shown in Figure 5.8. A single major reverse transcript is observed in RNA preparations from exponentially growing cells, showing the initiation point of transcription at the C marked with an asterisk (Figure 5.8 (A)). There is a potential σ^A-type promoter located at the appropriate distance upstream of this start. The potential σ^A promoter motif differs by two bases from the consensus in the –35 box (mismatches underlined TT<u>C</u>A<u>A</u>A), has the conventional 17 base spacer and has one mismatch from the consensus –10 sequence (mismatch underlined TAT<u>C</u>AT). A second transcript 5 bases smaller is also observed though the significance of
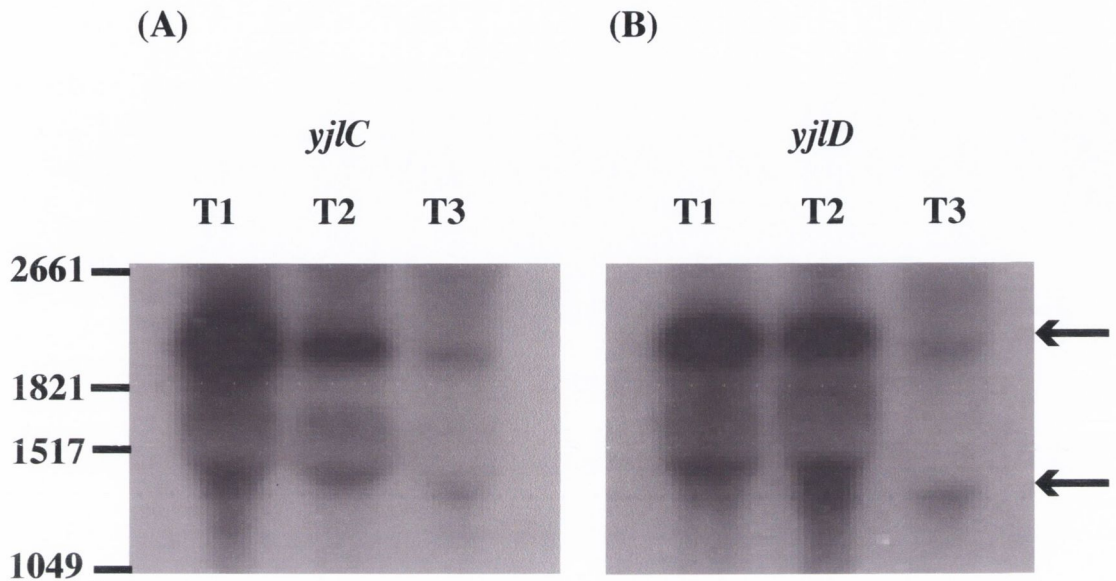
Figure 5.7 Northern blot analysis of *yjlC* and *yjlD*.
RNA was prepared from cells of strain 168 (wild type) that were harvested
from exponential (T1), transition (T2) and early stationary (T3) phases of growth
in SM medium. Twenty five micrograms of total RNA was loaded onto each well.
The positions to which the RNA size markers (1,049, 1,517, 1,821 and 2,661 bases)
migrated are indicated to the left of the figure. The northern blots were hybridised
with probes for (A) *yjlC* and (B) *yjlD*. The two transcripts detected are indicated
with arrows to the right of the figure.

Figure 5.8 Primer extension analysis of *yjlCD*.
(A) RNA was prepared from cells of strain 168 (wild type) and BFS841 (*yjlD*::pMUTIN4)
that were grown in SM medium and harvested at exponential (T1), transition (T2) and
early stationary phase (T3). The signal is that obtained from 12.5μg of total RNA. The
sequencing ladder is shown (A C G T) with the sequence indicated on the left of the figure.
Two reverse transcripts were observed. The two bases at which transcription initiates are
each indicated by an asterisk. (B) The sequence of the region is shown. The two bases at
which transcription initiates are each shown by an asterisk and a bold letter. The putative
σ^A motif (-35, -10) is shown in bold uppercase letters. The first 5 amino acids of YjlC are
shown.

(A)

T
A
T
A
T
C
T
T
C*
C
A
T
T
A
G*

T1 T2 T3   T1 T2 T3

BFS841   168wt

(B)
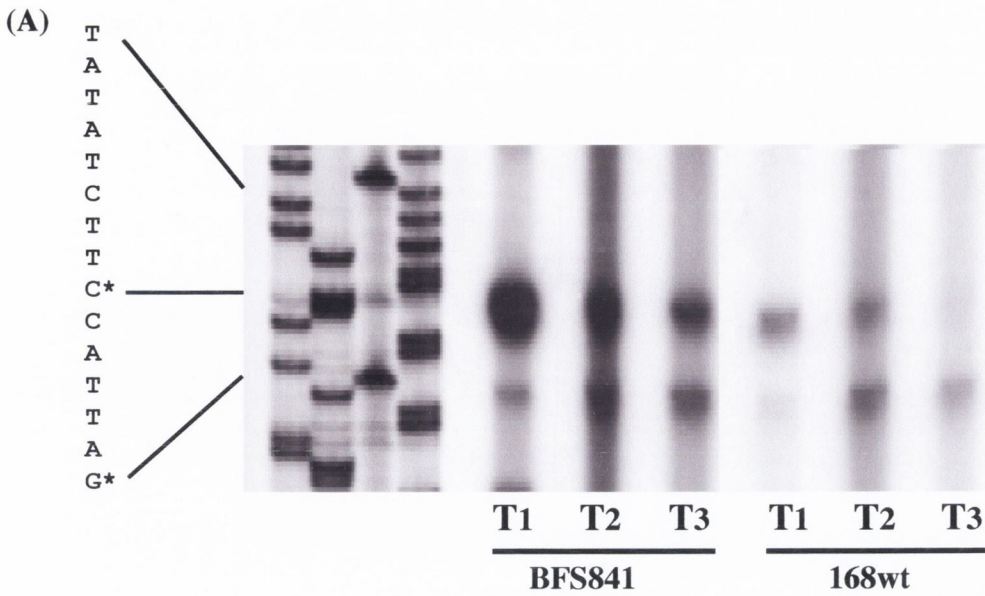
```
  E   I   W   V   K   * End yjlB
gaaatatgggtaaagtgaaggggagaaattgctggattgtcaataactagtgccgcagaa
---------+---------+---------+---------+---------+---------+
ctttatacccatttcacttcccctctttaacgacctaacagttattgatcacggcgtctt

                                                          -35
acgatatgacatgagctccatataaattaatttcataatcgaagcaattcacgTTCAAAt
---------+---------+---------+---------+---------+---------+
tgctatactgtactcgaggtatatttaattaaagtattagcttcgttaagtgcaagttta

      -10           *         *
tttgttcaaaaagtgtTATCATatagaaggtaatcgaattggagttttttgtgtttgcta
---------+---------+---------+---------+---------+---------+
aaacaagtttttcacaatagtatatcttccattagcttaacctcaaaaaacacaaacgat

ttttcagagagagtaaaggattttgacttattcaattgtttattttttagataaaaataat
---------+---------+---------+---------+---------+---------+
aaaagtctctctcatttcctaaaactgaataagttaacaaataaaaatctattttttatta

actgaatatggcaagatgaaaaactgggtagtaaacatttattttttgtaaatgattgtgc
---------+---------+---------+---------+---------+---------+
tgacttataccgttctactttttgacccatcatttgtaaataaaaacatttactaacacg

atgaattcacaatccttattcttatgccttaaaagcgtgagacagggagaggcactaccc
---------+---------+---------+---------+---------+---------+
tacttaagtgttaggaataagaatacggaattttcgcactctgtccctctccgtgatggg

                 Start yjlC            M   P   E   T   I
tttccaaaaaaatcagcaaactgcttatagtaaaggagaattggatatgccagaaacaat
---------+---------+---------+---------+---------+---------+
aaaggttttttttagtcgtttgacgaatatcatttcctcttaacctatacggtctttgtta
```

this is unclear. The level of reverse transcript is higher for strain BFS841 than for wild type although both strains show the same start point of transcription. This implies that negative autoregulation occurs at the transcriptional level and from the same start point of transcription as in wild type cells. In summary, the primer extension data shows that transcription initiates from a σ^A-type promoter, and interestingly, there is a long leader transcript (257 bases) between the initiation point of transcription and the *yjlC* start codon. Transcription of the operon is highest in exponential and transition phases and negative autoregulation occurs at the transcriptional level.

5.2.6 Both YjlC and YjlD are necessary for activity of the NADH dehydrogenase.

The *yjlD* gene is highly homologous to Type 2 NADH dehydrogenases from other organisms and has conserved NADH and FAD binding domains (Section 5.2.1) strongly indicating that YjlD functions as an NADH dehydrogenase in *B. subtilis*. YjlC is not homologous to proteins in the databases and its function is unknown. Experiments were carried out to determine the functions of both YjlC and YjlD and to investigate the relationship between them.

As described in 5.1.2 mutations in either *yjlC* or *yjlD* result in a strain which has a slower growth rate than wild type. This implies that the products of this operon are important for normal aerobic growth rates. It was of interest therefore to establish whether *yjlC* and *yjlD* are both required for normal growth and if so, to establish the contribution made by the *yjlC* gene product. The strategy applied was to construct strains in which *yjlC* and *yjlD* could be separately expressed in a controlled manner. Two strains were constructed: strain ES29 contains a pMUTIN4 insertion into *yjlD* (therefore is a *yjlD* null mutant), but an intact copy of the *yjlD* gene is under xylose inducible control at the amylase locus. Since *yjlD* is the downstream gene of the operon, the pMUTIN4 insertion in strain ES29 does not eliminate expression of *yjlC*. The pMUTIN4 insertion increases expression of *yjlC* from the natural promoter via the negative autoregulation effect (Section 5.2.3 and 5.2.5). Therefore it is possible to grow strain ES29 expressing *yjlC* only (no addition to the medium) or expressing both *yjlC* and *yjlD* (xylose addition to the medium). The second strain ES30 contains a pMUTIN4 insertion into *yjlC* (therefore *yjlC* is inactivated) but the entire *yjlC* gene has been placed under the control of the xylose inducible promoter at the amylase locus. The pMUTIN4 insertion means that the natural copy of *yjlC* is not expressed but the downstream *yjlD* gene is under the control of the IPTG inducible P_{SPAC} promoter at this locus. It is therefore possible to grow strain ES30 so that it expresses

neither *yjlC* or *yjlD* (no addition to the medium), *yjlD* only (addition of IPTG), *yjlC* only (addition of xylose) or both *yjlC* and *yjlD* (addition of IPTG and xylose). The effect of growth conditions on *yjlCD* expression in these strains is summarised in table 5.1 below.

Table 5.1

Summary of the effect of additions to the growth medium on the expression of *yjlC* and *yjlD* in strains ES29 and ES30.

Medium:	No Addition	+ Xylose	+ IPTG	+IPTG and +Xylose
Effect in strain ES29:	*yjlC+ yjlD-*	*yjlC+ yjlD+*		
Effect in strain ES30:	*yjlC- yjlD-*	*yjlC+ yjlD-*	*yjlC- yjlD+*	*yjlC+ yjlD+*

Firstly, it was noted that the addition of xylose to the medium gives a growth advantage to cells, this was examined by monitoring growth of the wild type strain in SM medium in the presence and absence of xylose. These growth curves are shown in Figure 5.9 (A). Addition of xylose to the medium does not appear to offer any significant growth advantage to wild type cells during early and mid exponential growth. However, as growth begins to slow at the end of exponential phase the culture with xylose grows to a higher OD than the culture without xylose. This growth advantage does not affect interpretation of the growth experiments on strains ES29 and ES30 since growth differences for these strains should be observed mainly during exponential growth and the observed growth differences should be greater than those that resulted from xylose addition.

Strains ES29 and ES30 were grown in SM medium with addition of xylose and IPTG as appropriate to the medium as shown in Table 5.1 and cell growth was monitored by measuring OD_{600}. The growth curves of strain ES29 are shown in Figure 5.9 (B). The same overnight culture was used to inoculate both flasks. One flask had xylose to induce expression of *yjlD* and the other did not. Cells grown in the absence of xylose express *yjlC* only, whereas cells grown in the presence of xylose express both *yjlC* and *yjlD*. Addition of xylose to cultures of ES29 gives the cells a clear growth advantage (Figure 5.9 (B)). Cells where only *yjlC* is expressed grow significantly slower than, and do not reach the same OD_{600} as, those expressing both *yjlC* and *yjlD*. It is evident that the products of both genes are required for the improvement in growth rate. The growth curves of cultures of ES30 with appropriate additions are shown in Figure 5.9 (C). All cultures were initiated with the same inoculum. Media are either unsupplemented or supplemented with IPTG
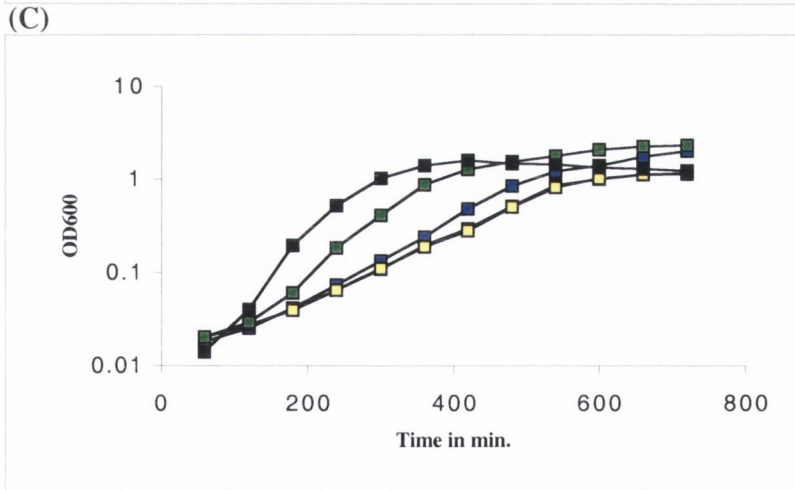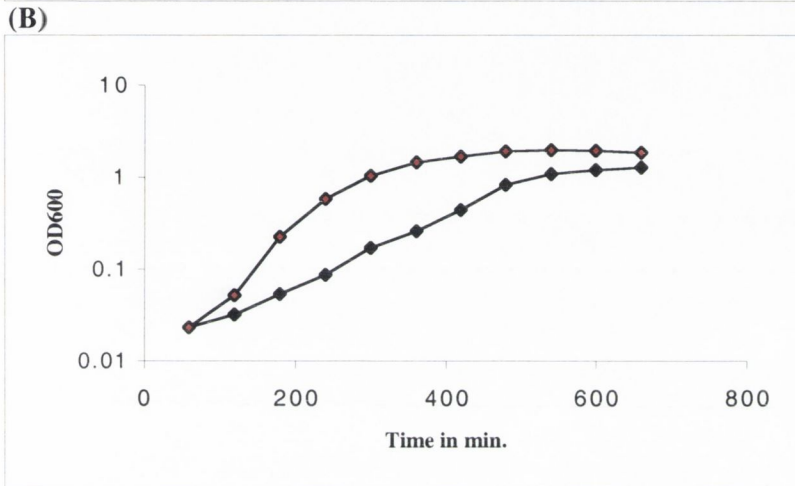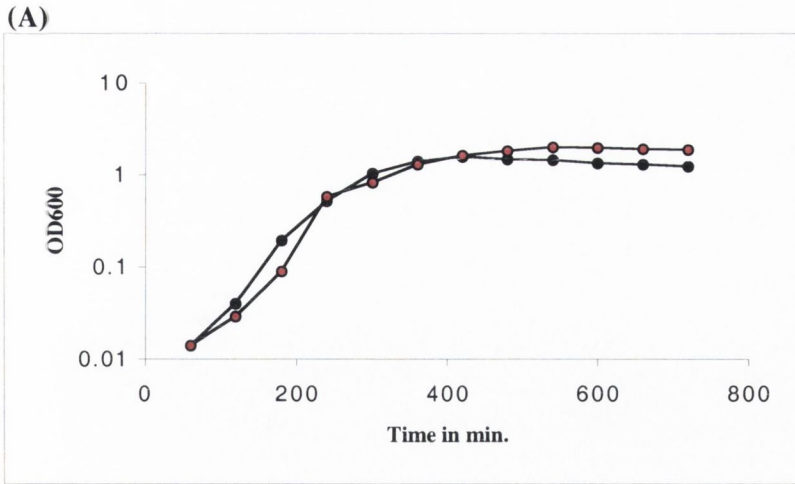
117

(A)

(B)

(C)

Figure 5.9 Growth curves for strains 168, ES29 and ES30.
Cells were grown in SM and growth was monitored by measuring OD600.
(A) Growth curves for strain 168 (wild type) with no addition (black circles)
and with 0.5% xylose added (red circles). (B) Growth curves for strain ES29
(P$_{yjlCD}$ -yjlC amyE ::P$_{xyl}$ -yjlD) with no addition (black diamonds) and with
0.5% xylose added (red diamonds). (C) Growth curves for strain 168 (wild type)
with no addition(black squares)and strain ES30 (P$_{SPAC}$-yjlD amyE ::P$_{xyl}$ -yjlC)
with no addition (red squares), 0.5% xylose added (blue squares), 1mM IPTG
added (yellow squares) and with both 0.5% xylose and 1mM IPTG added
(green squares). Time is presented in minutes after inoculation.

only, xylose only or both IPTG and xylose. The *yjlCD* expression status for each of these cultures is shown in Table 5.1. The growth curve for cells where *yjlD* is expressed (+IPTG) is almost identical to that in which neither gene is expressed (no addition). The growth curve for cells where *yjlC* (+xylose) is expressed shows a slight growth advantage over that in which neither gene is expressed (no addition), but this growth advantage could be due to the presence of xylose (as shown in Figure 5.9 (A)) and not the expression of *yjlC*. However in the culture where both *yjlC* and *yjlD* are expressed (+IPTG +xylose) it is evident that an increased growth rate is achieved, though the growth rate is still less than for the wild type strain. These data suggest that expression of both *yjlC* and *yjlD* is required for normal growth.

In view of the results just presented it was decided to determine whether NADH dehydrogenase activity could be detected in cells of strain ES30 grown in the presence of both xylose and IPTG. NADH dehydrogenase activity can be detected when cell extracts are separated on non-denaturing PAGE gels and stained with nitroblue tetrazolium in the presence of NADH (Chapter 2, section 2.8.3). Protein was prepared from cells harvested at an OD_{600} of 0.6 from ES30 cultures grown in SM with no addition (neither *yjlC* or *yjlD* expressed), with IPTG added (*yjlD* expressed), with xylose added (*yjlC* expressed) and with both xylose and IPTG added (both *yjlC* and *yjlD* expressed). Protein was also prepared from 168 wild type cells grown in SM. The samples were separated on non-denaturing PAGE gels and these gels were stained for NADH dehydrogenase activity. A representative gel is shown in Figure 5.10. Nine bands of NADH dehydrogenase activity appear for wild type cells. One band, (indicated by an arrow in Figure 5.10), was always absent in protein samples prepared from mutants of *yjlC* and *yjlD*. This band is likely to be the type 2 NADH dehydrogenase encoded by *yjlD*. However this gel also shows that even when both *yjlC* and *yjlD* are expressed by P_{xyl} and P_{SPAC}, a level of expression which apparently provides sufficient NADH dehydrogenase to partly relieve the growth defect of the *yjlCD* mutant strains, a band comigrating with the arrowed band in lane 5 of Figure 5.10 did not reappear on activity gels. It must be noted that, although growth of strain ES30 is improved by addition of xylose and IPTG, it is not restored to wild type levels. This suggests that for optimal growth rates and wild type levels of accumulation and activity of YjlCD the genes must be under the control of their natural promoter. These results show therefore that when *yjlC* and *yjlD* are expressed by P_{xyl} and P_{SPAC} promoters sufficient NADH dehydrogenase activity is produced to improve growth rate but the level
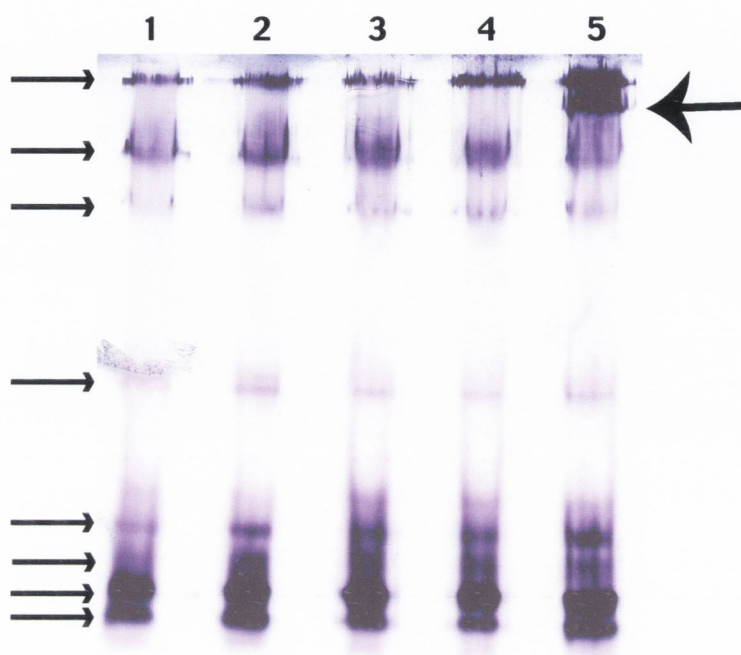
Figure 5.10 NADH dehydrogenase activity gel.
Protein extracts were prepared from cells grown in SM and harvested during mid
exponential growth, extracts were seperated by non denaturing PAGE and stained
for NADH dehydrogenase activity as described in Chapter 2 (2.8.3). Extracts were
prepared from strain ES30 (P$_{SPAC}$-*yjlD amyE*::P$_{xyl}$-*yjlC)* grown with no addition to the
medium (Lane 1), with 1mM IPTG added (Lane 2), with 0.5% xylose added (Lane 3)
and with both 1mM IPTG and 0.5% xylose added (Lane 4) and from strain 168
(wild type) with no addition (Lane 5). Small arrows at the left of the figure indicate
eight bands of NADH dehydrogenase activity. The large arrow at the right of the
figure indicates the ninth band which appears in protein samples prepared from wild
type cells (Lane 5) but not in *yjlC* or *yjlD* mutant protein samples, even when *yjlC* and
yjlD are induced by P$_{xyl}$ and P$_{SPAC}$ respectively (Lane 4).

of activity is not sufficient to restore wild type growth rates or to give an activity band on native protein gels.

When NADH dehydrogenase activity gels were carried out for *yutJ* and *yumB* null mutant strains and compared to wild type samples no YumB or YutJ-specific bands could be detected on the activity gels (data not shown). Perhaps these NADH dehydrogenases are too labile or, alternatively, do not accumulate to a level that is detectable on activity gels.

5.2.7 Expression profile and transcriptional analysis of *yutJ*.

There are three type 2 NADH dehydrogenases in the *B. subtilis* genome: *yjlD, yutJ* and *yumB* (Section 5.2.1). Expression profiles of *yutJ* and *yumB* were determined using transcriptional *lacZ* fusions. Strain ES998 contains a pMUTIN4 insertion into *yutJ* and is therefore a *yutJ* null mutant containing a transcriptional *lacZ* fusion to the *yutJ* promoter. A second strain, ES9911, contains a *lacZ* fusion with the 262 base pair putative *yutJ* promoter region fused to *lacZ* and placed at the amylase locus. The expression profiles for both *lacZ* fusions are shown in Figure 5.11. The levels of LacZ activity reach their maximum level of 25-30 units during exponential growth and expression decreases as cells enter stationary phase. It is evident that the expression profile of *yutJ* is very similar in the *yutJ* null mutant background to that in the wild type background (compare Figure 5.11 (A) and 5.11 (B)). This implies that YutJ does not autoregulate its expression. This expression profile was confirmed by northern analysis, shown in Figure 5.11 (C). RNA samples were prepared from cultures in mid exponential (T_1), transition (T_2) and stationary phase (T_3) of growth. Transcript is clearly visible in T_1 which is from exponentially growing cells. In transition, T_2, the level of transcript is greatly decreased and it is almost invisible in the T_3 stationary phase sample. The transcript migrates between the 1049 and 1517 nucleotide size markers. The *yutJ* open reading frame is predicted to be 990 bases in length and the upstream untranslated region to be 262 bases in length. The size of the transcript is consistent with the prediction that *yutJ* is monocistronic.

5.2.8 Expression profile and transcriptional analysis of *yumB*.

Strain ES999 contains a pMUTIN4 insertion into *yumB* that generates a *yumB* null mutant. Strain ES9910 carries a *lacZ* fusion to the 331 base pair promoter region of *yumB* placed at the amylase locus. The LacZ activity of both strains throughout the growth curve was determined and the results are shown in Figure 5.12 (A) and (B). The expression profiles of both transcriptional fusions are very similar. LacZ activity is low during exponential
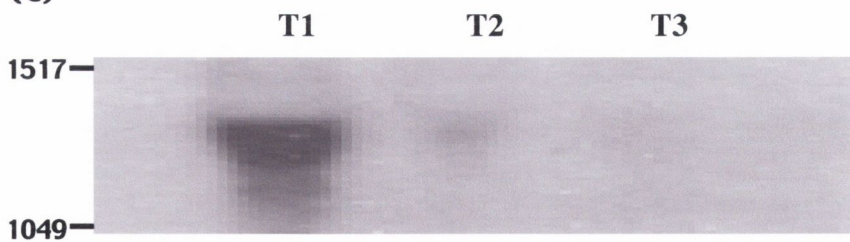
(A)
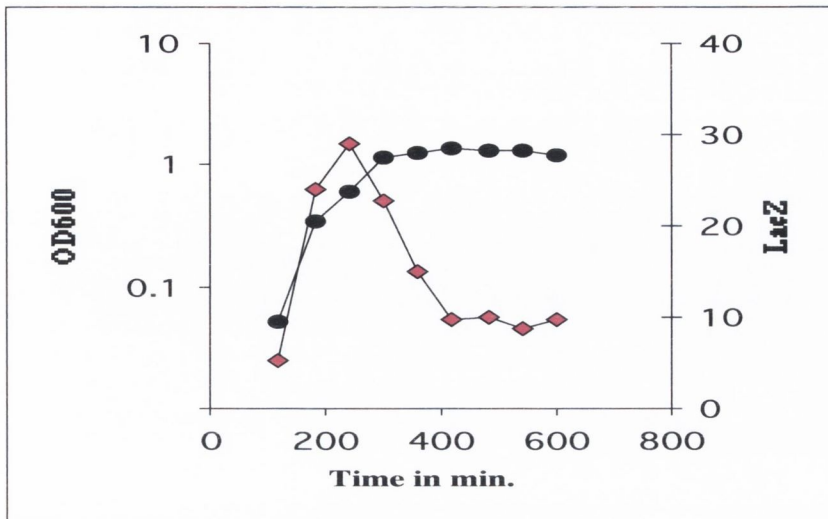
(B)
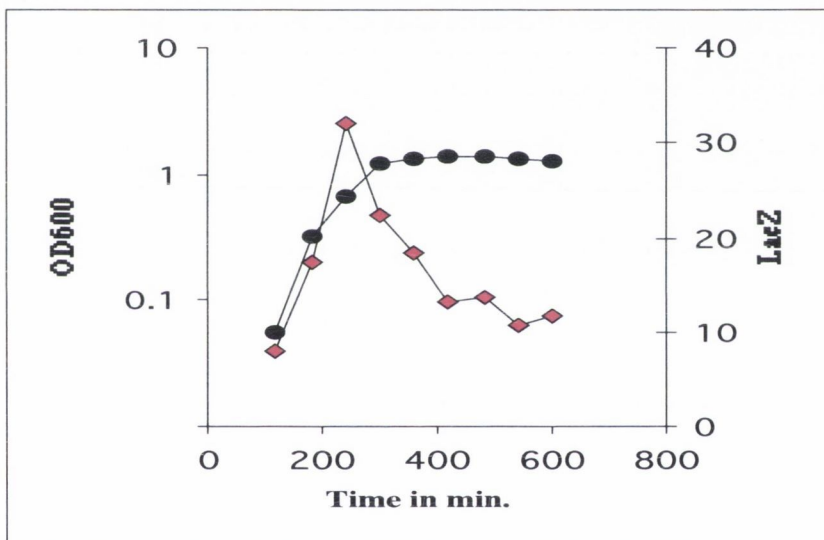
(C)

Figure 5.11 LacZ expression profiles and Northern blot analysis of *yutJ*.
Growth curves and LacZ expression profiles of strains (A) ES998 (*yutJ'-lacZ*)
and (B)ES9911 (*amyE::PyutJ-lacZ*) in SM medium. Growth was monitored
by measuring OD_{600} and is shown by black circles. LacZ activity was
measured as nmole ONPG/min/mg and is shown by pink diamonds. Time is
presented as minutes after inoculation. (C) RNA was prepared from cells of
strain 168 (wild type) harvested in exponential (T1), transition (T2) and early
stationary (T3) phases of growth. Twenty five micrograms of total RNA was
loaded onto each well. The positions to which the RNA size markers migrated
are indicated.The northern blot was hybridised with a probe for *yutJ*.

growth up to the transition phase. The level of expression rises after transition phase and continues to rise to a maximum of 15-17 units at the last point assayed. As the expression profiles for the *yumB* null background and the *yumB* wild type background are very similar, it is clear that YumB does not regulate its own expression (Figure 5.12 (A) and (B)). This expression profile is consistent with results obtained by northern analysis. RNA samples for the northern blot were prepared from cultures during mid exponential (T_1), transition (T_2) and stationary phases (T_3) of the growth cycle. Northern analysis was carried out and the blot probed for *yumB*. The resultant autoradiogram is shown in Figure 5.12 (C). Only a very faint transcript is visible in T_1 (exponential) and T_2 (transition). The level of transcript is maximal at T_3 (early stationary phase). The transcript migrates close to the 1,517 base pair marker. Since the *yumB* open reading frame is 1,218 bases in length this is consistent with *yumB* being a monocistronic operon.

### 5.2.9 Primer extension analysis of *yutJ* and *yumB*.

To determine the promoter structure for *yutJ* and *yumB*, primer extension analysis was carried out. RNA was prepared from wild type cells harvested at six points during growth in SM (t_1-t_6). The time points in the growth curve at which cells were harvested for preparation of RNA are indicated in Figure 5.13. RNA from cells harvested at t_1, t_2 and t_3 was used for primer extensions of *yutJ*. Cells are growing exponentially at t_1 and t_2 whereas by t_3 cells have reached transition phase. RNA from cells harvested at t_4, t_5 and t_6 was used for primer extensions of *yumB*. These time points are post transition phase. Two primers were used to determine the start point of *yutJ*. Consistent results were obtained with both primers. The autoradiogram is shown in Figure 5.14 (A). The initiation point of transcription lies 67 bases upstream of the first codon of *yutJ*. The initiation point is consistent with expression from a σ^A-type promoter (Figure 5.14 (B)). The promoter motif has two mismatches from the consensus σ^A sequence in the −35 region (mismatches underlined: TTG<u>TG</u>A), a conventional 17 base spacer and a perfect match to the consensus in the −10 region (TATAAT). The level of transcript visible in these primer extensions is highest in exponential growth (t_1) and decreases as cells approach transition (t_2). A low level of reverse transcript is seen in t_3, the transition phase sample, in agreement with northern and *lacZ* fusion data.

To determine the initiation point of transcription of *yumB* three different primers were used. Two initiation points of transcription were observed and each was confirmed by a second primer. The results are shown in Figure 5.15 (A) (the start point distal to *yumB*)
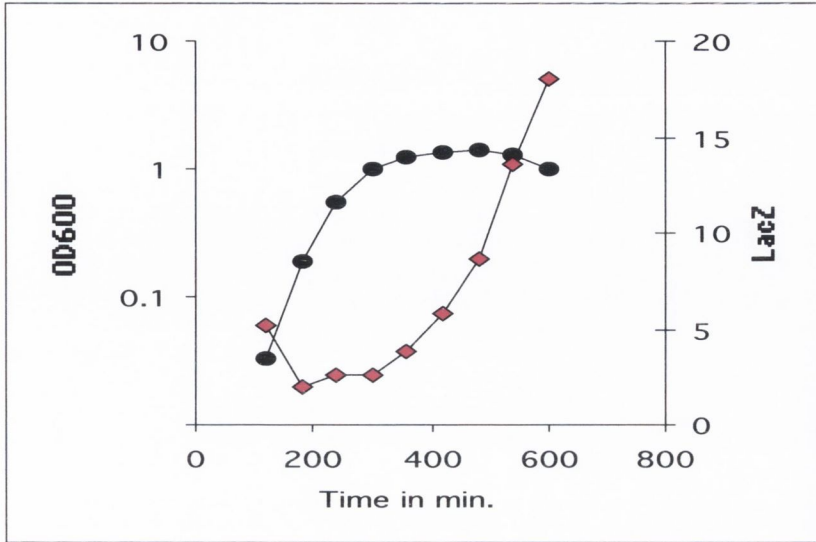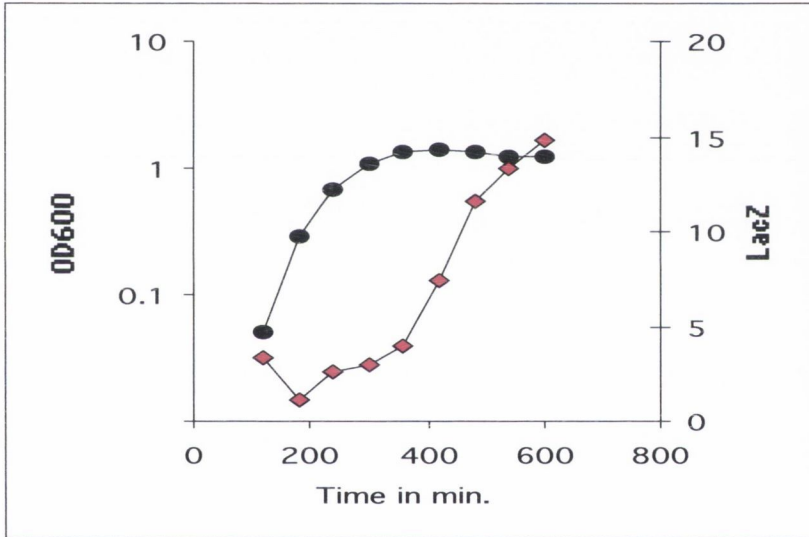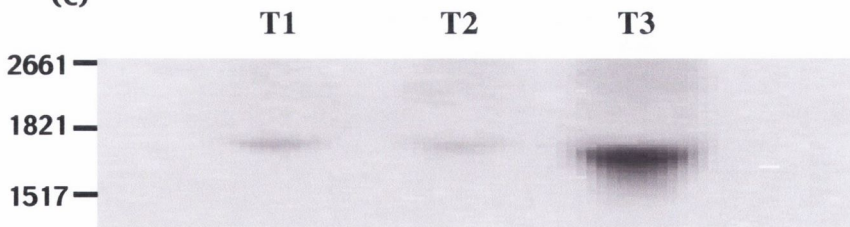
Figure 5.12 LacZ expression profiles and Northern blot analysis of *yumB*.
Growth curves in and LacZ expression profiles of strains (A) ES999 (*yumB'-lacZ*)
and (B)ES9910 (*amyE*::P*yumB-lacZ*) in SM medium Growth was monitored by
measuring OD_{600} and is shown by black circles. LacZ activity was measured as
nmole ONPG/min/mg and is shown by pink diamonds. Time is presented as
minutes after inoculation. (C) RNA was prepared from cells of strain 168 (wild type)
harvested in exponential (T1), transition (T2) and early stationary (T3) phases of growth.
Twenty five micrograms of total RNA was loaded onto each well. The positions to which
the RNA size markers migrated are indicated. The northern blot was hybridised with
a probe for *yumB*.
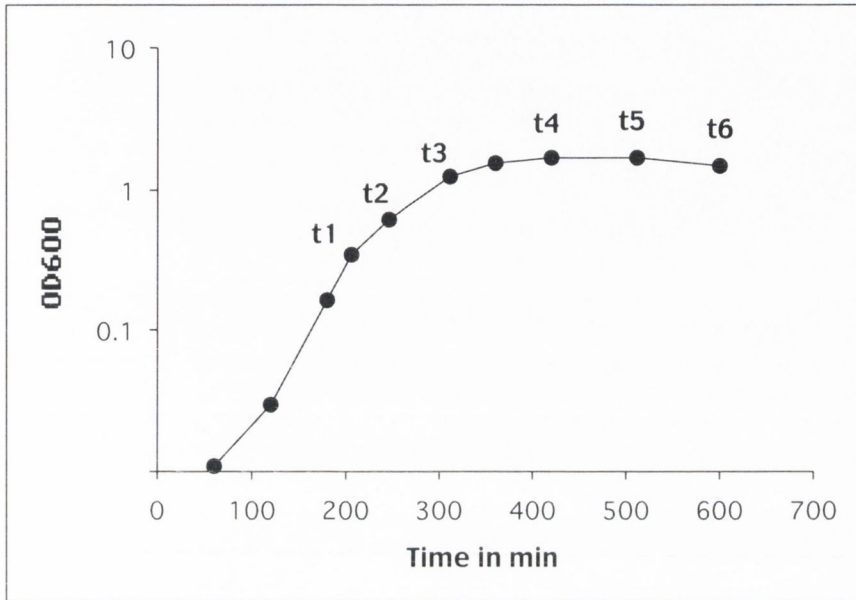
Figure 5.13 Growth curve for RNA samples used for primer extensions of
yutJ **and** *yumB*.
Growth curve of 168 wild type cells in SM medium indicating the timepoints
where cells were harvested for RNA preparation. RNA from the cells of this
growth curve was used to perform the primer extensions depicted in figure 5.14
and 5.15. RNA from timepoints t1, t2 and t3 above was used for primer
extensions with a *yutJ* primer. RNA from t4, t5 and t6 was used for primer
extensions with a *yumB* primer.

Figure 5.14 Primer extension analysis of *yutJ*.
(A)RNA was prepared from cells of strain 168 (wild type) at time points t1, t2 and t3 during growth as described in Figure 5.13. The signal is that obtained from 12.5μg of total RNA. The sequencing ladder is shown (A C G T) with the sequence indicated on the right of the figure. A single reverse transcript was observed. The base at which transcription initiates is indicated by an asterisk. (B) The sequence of the region is shown. The base at which transcription initiates is shown by an asterisk and a bold letter. The putative σ^A motif (-35, -10) is shown in bold uppercase letters. The first 6 amino acids of YutJ are shown.

(A)

(B)

```
attcaatcattggattcatgcggtttcacctcacatatgtccattatcgcgaattttttca
---------+---------+---------+---------+---------+---------+
taagttagtaacctaagtacgccaaagtggagtgtatacaggtaatagcgcttaaaaagt

   E  I  M  P  N  M  Start yuzB

ctattcgtcaaagacaacggatgattcaaaatttattatcgtaacatattatttctcaaa
---------+---------+---------+---------+---------+---------+
gataagcagtttctgttgcctactaagtttttaaataatagcattgtataataaagagttt

taacgttttactaatttgtgttttcaggctaaattgtaggttattttgaaactattaTTG
---------+---------+---------+---------+---------+---------+
attgcaaaatgattaaacacaaaagtccgatttaacatccaataaaactttgataataac

-35                      -10           *
TGAttttagatcaaattcgaTATAATtctagatgttatttttacatagttcgtattttctt
---------+---------+---------+---------+---------+---------+
actaaaatctagtttaagctatattaagatctacaataaaatgtatcaagcataaaagaa

                       Start yutJ      M  K  K  L  V  L
caatgtatctatcatactcttcggattaaaggggaattcaaatgaaaaaattggtcttga
---------+---------+---------+---------+---------+---------+
gttacatagatagtatgagaagcctaatttccccttaagtttacttttttaaccagaact
```

and (B) (the start point proximal to *yumB*). In Figure 5.15 (A) a low level of transcript is observed from the distal start point at t_4. The level increases in the later samples, t_5 and t_6. However transcript is not detected at t_4 from the proximal start point (Figure 5.15 (B)) but high levels of reverse transcript are observed at t_5 and t_6. There is a potential σ^F motif upstream of the distal start point. This motif has one mismatch from the σ^F consensus in the −35 region (mismatch underlined: GAATG), a consensus 15 base spacer and two mismatches from the consensus in the −10 region (mismatches underlined GGATAAACA). There is a potential σ^E motif upstream of the proximal start point. This motif is a perfect match to the σ^E consensus in the −35 region (TCATAAA), has a conventional 14 base spacer and four mismatches from the consensus in the −10 region (mismatches underlined: CTGAAGCT). The putative σ^F driven transcript appears before the σ^E driven transcript. These data suggest that *yumB* is expressed during sporulation in both the forespore and in the mother cell.

Figure 5.15 Primer extension analysis of *yumB*.
(A) and (B) RNA was prepared from cells of strain 168 (wild type) at time points t4, t5 and t6 during growth as described in Figure 5.13. The signal is that obtained from 12.5μg of total RNA. (A) and (B) are sections from the same autoradiogram. The sequencing ladder is shown (A C G T) with the sequence indicated on the right of the figures. Two reverse transcripts were observed. The bases at which transcription initiates are each indicated by an asterisk. (B) The sequence of the region is shown. The bases at which transcription initiates are each shown by an asterisk and a bold letter. The putative σ^F motif (-35F, -10F) and σ^E motif (-35E, -10E) are shown in bold uppercase letters. The first 18 amino acids of YumB are shown.

(A) t4 t5 t6 A C G T

A
T
T
A
T
C*
A
C
G
A

(B) t4 t5 t6 A C G T

T
T
G
T
G*
T
T
T
A
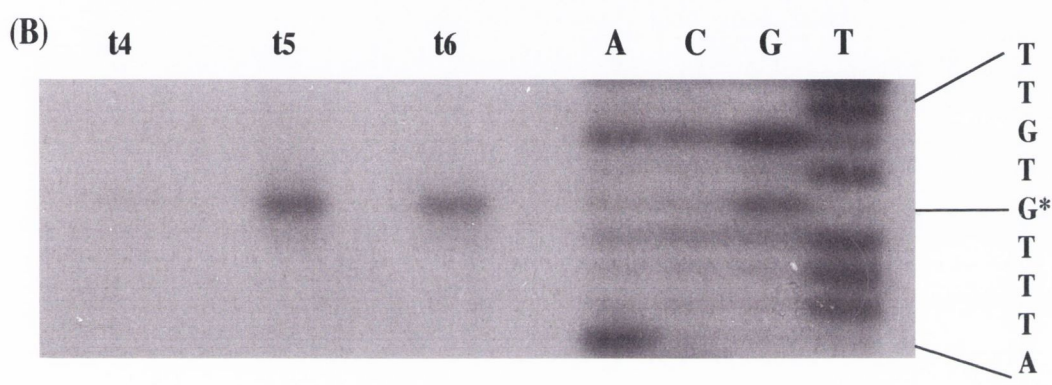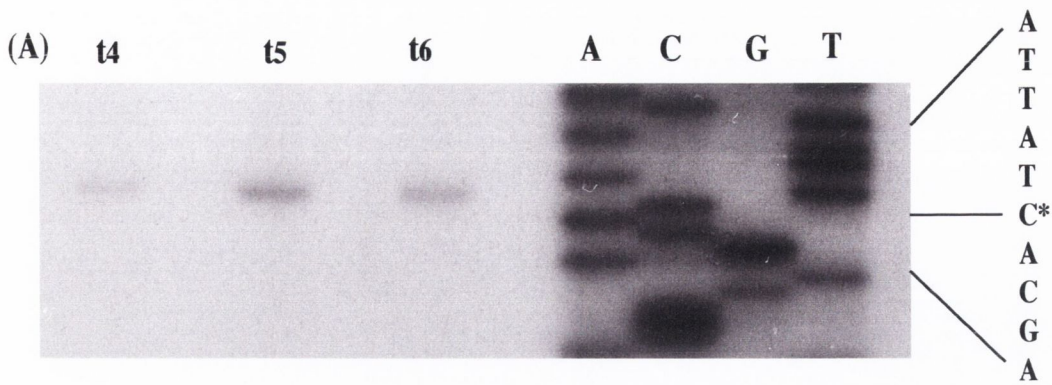
(C)

```
aattgtaatatcataaacctttgtatcctctcgcatacgaaaattgcctcctaaatgaaa
---------+---------+---------+---------+---------+---------+
ttaacattatagtatttggaaacataggagagcgtatgcttttaacggaggatttacttt
 I  T  I  D  Y  V  K  T  D  E  R  M  Start yumC

attgttattaattataccatattgccgttcaaattgacgagtcagcagaaaaaatcgtaa
---------+---------+---------+---------+---------+---------+
taacaataattaatatggtataacggcaagtttaactgctcagtcgtctttttttagcatt

ggatctaacgtgcccatccgttcttgaaaatatttccgcaggccgctacgataataacca
---------+---------+---------+---------+---------+---------+
cctagattgcacgggtaggcaagaacttttataaaggcgtccggcgatgctattattggt

                -35F                      -10F              *
aatgaaatgatcttttttctGAATGtttcttgggccagccGGATAAACAtgataatagtgc
---------+---------+---------+---------+---------+---------+
ttactttactagaaaaagacttacaaagaacccggtcggcctatttgtactattatcacg
                                        -35E
tacgggtaaggacttcacaaaaaaacttttcaatatattTCATAAAatcccgccaataagC
---------+---------+---------+---------+---------+---------+
atgcccattcctgaagtgttttttgaaaagttatataaagtattttagggcggttattcg

 -10E        *
TGAAGCTaacacaaatgatgattttattatgggtctaacaaaagcaaaaaaagagaaggt
---------+---------+---------+---------+---------+---------+
acttcgattgtgtttactactaaaataatacccagattgtttttcgttttttttctcttcca

Start yumB
      M  A  L  N  K  P  K  I  V  I  L  G  A  G  Y  G  G  L
ggatgtgatggcattgaataagcccaaaatcgtgattttaggtgcaggatatggcggatt
---------+---------+---------+---------+---------+---------+
cctacactaccgtaacttattcgggtttttagcactaaaatccacgtcctataccgcctaa
```

5.3 DISCUSSION

The *B. subtilis* genome contains three type 2 NADH dehydrogenases based on homology with these enzymes in other organisms. These are encoded by *yjlD, yutJ* and *yumB*. The proteins encoded by these genes have conserved NADH and FAD binding motifs and are homologous to the *ndh* encoded NDH-2 of *E. coli* (Figure 5.2). The genome does not contain genes for a type 1 NADH dehydrogenase. Why does *B. subtilis* require three NDH-2 enzymes? In the work described in the preceding sections expression analysis has given some indications as to why this is the case. Expression analysis of the three genes was carried out using *lacZ* transcriptional fusions, by northern analysis and by primer extension analysis. The results achieved by these three methods were consistent and gave some insight into why the cell has three NDH-2 genes.

Of the three genes the *yjlCD* operon is the most highly expressed. In LacZ assays the promoter fusion for this operon reaches 65 lacZ units in comparison to 32 units for the *yutJ* promoter and 15 units for the *yumB* promoter. This implies that the *yjlCD* encoded NDH-2 provides the major NADH dehydrogenase activity in the cell during exponential growth. Primer extension analysis showed the transcription of *yjlCD* initiates from a σ^A –type promoter that is consistent with its high expression during growth. It was also observed that the *yjlCD* operon is negatively autoregulated. In the absence of *yjlD* expression increases approximately 7 fold. In the absence of *yjlC* expression increases approximately 5 fold, but the profile of expression is different to that seen in the absence of *yjlD* (Figure 5.6). These results show that the level of expression of the major *yjlCD* encoded NDH-2 in the cell may be responsive to the level of *yjlCD* in the cell. This is not observed for the other two genes.

The *yjlCD* null mutant strains had notable phenotypes. The cells could still grow although at a slower rate than wild type and these strains formed small colonies. This result indicates that there must be a reserve NDH-2. Another enzyme in the cell must be capable of carrying out the function of the encoded NDH-2 albeit far less efficiently as is shown by the slow growing phenotype. The mutant strains also showed increased resistance to paraquat on solid medium. It has been shown in *E. coli* that the NDH-2 is one of the principal sources of superoxide generation during aerobic respiration. This occurs because the NDH-2 readily donates single electrons to oxygen via its exposed flavin cofactor (Messner and Imlay, 1999). It is likely that when the redox cycling compound paraquat is taken up into the cell it acts as a catalyst for this process sequestering electrons away from

122

the NDH-2 and donating them to oxygen to generate the toxic superoxide radical. However in the NDH-2 mutant background paraquat is likely to be less toxic to the cell as its most easily oxidisable source of electrons for superoxide generation is not present. In addition in the absence of NDH-2 it is likely that the cells are respiring less. Under these conditions the cells may grow fermentatively. This would mean less electrons flowing through the electron transport chain and therefore less electrons would be available for superoxide formation. These results again confirmed the important role of the *yjlCD* encoded NDH-2 as null mutants of *yutJ* and *yumB* do not have a paraquat resistant phenotype. In conclusion, these results imply that the *yjlCD* encoded NDH-2 is the most important NADH dehydrogenase during exponential growth. Its expression is autoregulated indicating that regulation of the operon is responsive to the $NAD^+/NADH$ balance in the cell.

Experiments that were carried out observing growth rate in the presence of *yjlC* or *yjlD* or both *yjlC* and *yjlD* are described in section 5.2.6. In strain ES30 that was used for these experiments both genes are expressed from heterologous promoters. The *yjlC* gene is expressed from P_{xyl} and the *yjlD* gene is expressed from P_{SPAC}. These experiments demonstrated that expression of both *yjlC* and *yjlD* is required to partially restore growth rate to wild type levels. When both genes are expressed from heterologous promoters growth rate is significantly improved in comparison to the growth rate of the *yjlCD* null background, however growth rate is still not completely restored to wild type levels. One of the most important unanswered questions about the *yjlCD* encoded NDH-2 regards the role of *yjlC*. Expression of this gene is required but it shows no homology to any genes in the current databases, therefore its role is unclear. In the growth experiments with strain ES30 both *yjlC* and *yjlD* are expressed from heterologous promoters. The *yjlC* gene is placed at the amylase locus. The *yjlC* gene product clearly carries out its function even though it is not expressed from its own promoter or at its natural locus. This suggests that *yjlC* has a role in NDH-2 function rather than a regulatory role. In strain ES30 the *yjlC* and *yjlD* transcripts are not the same as at their natural locus suggesting that YjlC is probably not involved in post transcriptional regulation although such a possibility cannot be ruled out. However these results point to a role for YjlC in activity, assembly or localisation of the NDH-2.

The *yutJ* gene is also expressed during exponential growth though at lower levels than *yjlCD*. The transcript initiates from a σ^A-type promoter. The expression of *yutJ* is not

responsive to the level of YutJ. It has not been determined whether *yutJ* expression is altered in a *yjlCD* null background. This will be determined as part of the future work on this gene. Expression of *yutJ* may increase in a *yjlCD* null background either through coordinate regulation with the *yjlCD* genes or in response to the alteration in the NAD$^+$/NADH balance in the cell. However in contrast to *yjlCD* the *yutJ* gene product is relatively dispensable. The *yutJ* null mutant does not have a growth defect or any alteration in paraquat sensitivity.

In contrast to *yjlCD* and *yutJ* the *yumB* encoded putative NDH-2 is expressed post transitionally and does not contribute to respiration during exponential growth. It has two start points of transcription that initiate from a potential σ^F promoter and a potential σ^E promoter. This profile of expression gives an explanation for the existence of this gene: it provides sporulation specific NADH dehydrogenase activity. The presence of two promoters imply that it is active in both the forespore (σ^F) and the mother cell (σ^E). It is known that σ^F is activated in the forespore before σ^E is activated in the mother cell and that σ^E activation requires prior activation of σ^F. The profile of the *yumB* transcript observed in primer extension analysis is consistent with this model. The σ^F-driven transcript appears before the σ^E- driven transcript, therefore *yumB* expression is sporulation specific and it is expressed in both compartments of the developing sporangium.

Future work on the NADH dehydrogenases will involve further investigation into the role of *yjlC*. Also, apart from a regulatory role of RelA during the stringent response, the regulator(s) of *yjlCD* that respond to the NAD$^+$/NADH balance and the availability of oxygen remain to be determined (Eymann *et al.*, 2002). In addition it will be interesting to determine whether there is overlap between the roles of *yjlCD* and *yutJ*. It is possible that these genes are coordinately regulated. Initial experiments will be carried out to see whether *yutJ* expression is altered in a *yjlCD* mutant. The sporulation specificity of *yumB* will be confirmed by the analysis of null mutants of *sig E* and *sigF*.

CHAPTER 6

DISCUSSION AND FUTURE WORK

DISCUSSION AND FUTURE WORK

The work for this thesis began as the *B. subtilis* genome sequencing project was drawing to a close. The complete genome sequence of 4,214,810 bases contains 4,100 genes 40% of which are genes of unknown function. The genome sequence is available on the SubtiList website (http://genolist.pasteur.fr/SubtiList/). The availability of the genome sequence has completely transformed *B. subtilis* research. It is now possible to identify all members of a gene family based on sequence similarities and therefore design experiments with this knowledge, for example generating double or multiple mutants of genes with similar functions. Very significantly the knowledge of the genome sequence paved the way for whole proteome and transcriptome analysis.

Upon the completion of the genome sequence the *B. subtilis* functional analysis project was undertaken by European and Japanese consortia. In a coordinated effort to assign function to many genes, 1,275 strains were constructed using pMUTIN vectors each containing a null mutation of a gene of unknown function. These strains comprise a large mutant collection that is a very useful resource which will certainly continue to yield information about these genes in the future. The collection was examined on three levels. Firstly, growth curves and LacZ assays were carried out on the strains to determine growth rates and expression profiles. Secondly, phenotypic tests were carried out in each lab to examine the response of the mutants to various stresses and growth media and to attempt to identify genes that may be significant in metabolism, competence, sporulation or germination. Thirdly, transcriptional analysis of large regions of the genome was carried out. The expression profiles and phenotypic data are all publicly available on the Micado website (http://locus.jouy.inra.fr/cgi-bin/genmic/madbase/progs/madbase.operl). This information provides the first steps in deducing the roles of the genes of unknown function. Phenotypic analysis of the mutant collection also identified phenotypes for many genes of unknown function that has initiated the further study of many genes, including *yjlC* and *yjlD* during the work for this thesis. It is clear that future research in the field of *B. subtilis* genetics will be either at the whole genome level or at the level of in depth study of a smaller number of genes. Transcriptome and proteome analysis will be carried out for all feasible stress responses and growth conditions and eventually the global effects of all the regulators in the genome will be analysed. These types of analyses will themselves provide an enormous amount of data to be exploited, comparable to the data generated from the genome sequencing. Much of the exploitation of data will involve new technologies, bioinformatic analysis or the use of older systems in the light of new

information (e.g. yeast two hybrid screens). A lot of information about specific gene function will also undoubtedly come from piecemeal studies carried out by researchers interested in a particular group of genes and their interactions.

The work carried out for this thesis focused on characterisation of genes involved in oxidative stress, and particularly superoxide stress. The NADH dehydrogenases were studied because of the increased paraquat resistance phenotype observed for null mutants of *yjlC* and *yjlD* and because this type of enzyme is a potential source of oxidative stress in the cell (Messner and Imlay, 1999). The expression profiles and operon structures of *yjlCD, yutJ* and *yumB* were deduced and some aspects of the regulation of these genes were characterised. One of the most interesting questions which results from this work is: what is the role of *yjlC*? Future work will involve further investigation of the role of this gene product in NDH-2 activity. The *yjlC* gene product may play a regulatory role or its role may be functional, it may optimise the NDH-2 activity by mediating localisation to the membrane for example. The role of this gene is very interesting because it does not have any significant homologues in the current databases.

The second aspect of oxidative stress that was studied was the SODs which are responsible for detoxification of the superoxide radical. The *sodA* gene encodes the probable MnSOD of *B. subtilis*. It is the principle SOD of *B. subtilis*. It is highly expressed during growth and induced in response to oxidative stress. It is also a member of the σ^B- dependant general stress regulon. The two alternative SODs encoded by *sodF* and *yojM* were very lowly expressed and expression was not increased under any of the conditions examined. Future work may reveal conditions under which these SODs are more highly expressed and reveal their role in the cell. The most interesting question to emerge from the work on SODs is: what is the regulator of the ethanol and paraquat induction at the σ^A-type promoter of *sodA*? If one regulator is responsible for this effect then it must respond to a signal which is common to both types of stress. Also what other genes are under the control of the same regulator? There may also be more than one regulator involved in these responses. The search for the regulator will be the subject of future work.

It is evident that while there are many similarities between *E. coli* and *B. subtilis* in terms of superoxide generation and detoxification the two species also have fundamental differences in these areas. It appears that the *yjlCD* encoded NDH-2 may be a source of oxidative stress as is the case for *E. coli*. However, while *B. subtilis* has three NDH-2

encoding genes, *E. coli* has one NDH-1 and one NDH-2. In addition the *soxRS* superoxide stress regulon is well elucidated in *E. coli* there does not appear to be an analogous regulon in *B. subtilis*. Therefore it will be interesting to study what regulators are involved in the superoxide stress response in *B. subtilis* and to what extent this response overlaps with the general stress response and other stress responses in the cell.

REFERENCES

REFERENCES

Adak, S., Aulak, K.S., and Stuehr, D.J. (2002) Direct evidence for nitric oxide production by a nitric oxide synthase-like protein from *Bacillus subtilis*. *J Biol Chem* **20**: 20.

Almiron, M., Link, A.J., Furlong, D., Kolter, R., Jenkins, D.E., Schultz, J.E., and Matin, A. (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev* **6**: 2646-2654.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Altuvia, S., Almiron, M., Huisman, G., Kolter, R., and Storz, G. (1994) The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol* **13**: 265-272.

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., and Storz, G. (1997) A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* **90**: 43-53.

Anagnostopolous, C., and Spizizen, J. (1961) Requirements for transformation in *Bacillus subtilis*. *J Bacteriol* **81**: 741-746.

Antelmann, H., Engelmann, S., Schmid, R., and Hecker, M. (1996) General and oxidative stress responses in *Bacillus subtilis*: cloning, expression, and mutation of the alkyl hydroperoxide reductase operon. *J Bacteriol* **178**: 6571-6578.

Antelmann, H., Engelmann, S., Schmid, R., Sorokin, A., Lapidus, A., and Hecker, M. (1997a) Expression of a stress- and starvation-induced *dps/pexB*-homologous gene is controlled by the alternative sigma factor sigmaB in *Bacillus subtilis*. *J Bacteriol* **179**: 7251-7256.

Antelmann, H., Bernhardt, J., Schmid, R., Mach, H., Volker, U., and Hecker, M. (1997b) First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. *Electrophoresis* **18**: 1451-1463.

Antoniewski, C., Savelli, B., and Stragier, P. (1990) The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. *J Bacteriol* **172**: 86-93.

Archibald, F.S., and Fridovich, I. (1981) Manganese and defenses against oxygen toxicity in *Lactobacillus plantarum*. *J Bacteriol* **145**: 442-451.

Argaman, L., and Altuvia, S. (2000) *fhlA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *J Mol Biol* **300**: 1101-1112.

Ariza, R.R., Li, Z., Ringstad, N., and Demple, B. (1995) Activation of multiple antibiotic resistance and binding of stress-inducible promoters by *Escherichia coli* Rob protein. *J Bacteriol* **177**: 1655-1661.

Ariza, R.R., Cohen, S.P., Bachhawat, N., Levy, S.B., and Demple, B. (1994) Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J Bacteriol* **176**: 143-148.

Aslund, F., and Beckwith, J. (1999) The thioredoxin superfamily: redundancy, specificity, and gray-area genomics. *J Bacteriol* **181**: 1375-1379.

Atichartpongkul, S., Loprasert, S., Vattanaviboon, P., Whangsuk, W., Helmann, J.D., and Mongkolsuk, S. (2001) Bacterial Ohr and OsmC paralogues define two protein families with distinct functions and patterns of expression. *Microbiology* **147**: 1775-1782.

Bagyan, I., Casillas-Martinez, L., Setlow, P., Jenkins, D.E., Schultz, J.E., and Matin, A. (1998) The *katX* gene, which codes for the catalase in spores of Bacillus subtilis, is a forespore-specific gene controlled by sigmaF, and KatX is essential for hydrogen peroxide resistance of the germinating spore. *J Bacteriol* **180**: 2057-2062.

Bannister, W.H., Bannister, J.V., Barra, D., Bond, J., and Bossa, F. (1991) Evolutionary aspects of superoxide dismutase: the copper/zinc enzyme. *Free Radic Res Commun* **12-13**: 349-361.

Barbosa, T.M., and Levy, S.B. (2000) Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. *J Bacteriol* **182**: 3467-3474.

Barker, M.M., Gaal, T., and Gourse, R.L. (2001a) Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. *J Mol Biol* **305**: 689-702.

Barker, M.M., Gaal, T., Josaitis, C.A., and Gourse, R.L. (2001b) Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation *in vivo* and *in vitro*. *J Mol Biol* **305**: 673-688.

Bauer, A.W., Kirby, W.M., Sherris, J.C., and Turck, M. (1966) Antibiotic susceptibility testing by a standardized single disk method. *Tech Bull Regist Med Technol* **36**: 49-52.

Beauchamp, C., and Fridovich, I. (1971) Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. *Anal Biochem* **44**: 276-287.

Becker-Hapak, M., and Eisenstark, A. (1995) Role of *rpoS* in the regulation of glutathione oxidoreductase *(gor)* in *Escherichia coli*. *FEMS Microbiol Lett* **134**: 39-44.

Bellamacina, C.R. (1996) The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J* **10**: 1257-1269.

Benov, L., and Fridovich, I. (1995) Superoxide dismutase protects against aerobic heat shock in *Escherichia coli*. *J Bacteriol* **177**: 3344-3346.

Benov, L., Kredich, N.M., and Fridovich, I. (1996) The mechanism of the auxotrophy for sulfur-containing amino acids imposed upon *Escherichia coli* by superoxide. *J Biol Chem* **271**: 21037-21040.

Benov, L.T., and Fridovich, I. (1994) *Escherichia coli* expresses a copper- and zinc-containing superoxide dismutase. *J Biol Chem* **269**: 25310-25314.

Bergsma, J., Van Dongen, M.B., and Konings, W.N. (1982) Purification and characterization of NADH dehydrogenase from *Bacillus subtilis*. *Eur J Biochem* **128**: 151-157.

Bergsma, J., Strijker, R., Alkema, J.Y., Seijen, H.G., Konings, W.N., Spiro, S., Roberts, R.E., and Guest, J.R. (1981) NADH dehydrogenase and NADH oxidation in membrane vesicles from *Bacillus subtilis*. *Eur J Biochem* **120**: 599-606.

Bernhardt, J., Buttner, K., Scharf, C., and Hecker, M. (1999) Dual channel imaging of two-dimensional electropherograms in *Bacillus subtilis*. *Electrophoresis* **20**: 2225-2240.

Bernhardt, J., Volker, U., Volker, A., Antelmann, H., Schmid, R., Mach, H., and Hecker, M. (1997) Specific and general stress proteins in Bacillus subtilis--a two-dimensional protein electrophoresis study. *Microbiology* **143**: 999-1017.

Bessette, P.H., Aslund, F., Beckwith, J., and Georgiou, G. (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci U S A* **96**: 13703-13708.

Bolhuis, A., Venema, G., Quax, W.J., Bron, S., and van Dijl, J.M. (1999) Functional analysis of paralogous thiol-disulfide oxidoreductases in *Bacillus subtilis*. *J Biol Chem* **274**: 24531-24538.

Bollag, D.M., and Edelstein, S.J. (1991) *Protein Methods*. New York: Wiley-Liss.

Bongaerts, J., Zoske, S., Weidner, U., and Unden, G. (1995) Transcriptional regulation of the proton translocating NADH dehydrogenase genes (*nuoA-N*) of *Escherichia coli* by electron acceptors, electron donors and gene regulators. *Mol Microbiol* **16**: 521-534.

Boylan, S.A., Redfield, A.R., Brody, M.S., and Price, C.W. (1993) Stress-induced activation of the sigma B transcription factor of *Bacillus subtilis*. *J Bacteriol* **175**: 7931-7937.

Bsat, N., and Helmann, J.D. (1999) Interaction of *Bacillus subtilis* Fur (ferric uptake repressor) with the *dhb* operator in vitro and in vivo. *J Bacteriol* **181**: 4299-4307.

Bsat, N., Herbig, A., Casillas-Martinez, L., Setlow, P., and Helmann, J.D. (1998) *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. *Mol Microbiol* **29**: 189-198.

Calhoun, M.W., Gennis, R.B., Glaser, P., and Jahn, D. (1993) Demonstration of separate genetic loci encoding distinct membrane-bound respiratory NADH dehydrogenases in *Escherichia coli*. *J Bacteriol* **175**: 3013-3019.

Casillas-Martinez, L., and Setlow, P. (1997) Alkyl hydroperoxide reductase, catalase, MrgA, and superoxide dismutase are not involved in resistance of *Bacillus subtilis* spores to heat or oxidizing agents. *J Bacteriol* **179**: 7420-7425.

Casillas-Martinez, L., Driks, A., Setlow, B., and Setlow, P. (2000) Lack of a significant role for the PerR regulator in *Bacillus subtilis* spore resistance. *FEMS Microbiol Lett* **188**: 203-208.

Chatterji, D., Ojha, A.K., Gentry, D.R., Hernandez, V.J., Nguyen, L.H., Jensen, D.B., Cashel, M., Bongaerts, J., Zoske, S., Weidner, U., and Unden, G. (2001) Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* **4**: 160-165.

Chen, L., Keramati, L., and Helmann, J.D. (1995) Coordinate regulation of *Bacillus subtilis* peroxide stress genes by hydrogen peroxide and metal ions. *Proc Natl Acad Sci U S A* **92**: 8190-8194.

Chen, L., Xie, Q.W., and Nathan, C. (1998) Alkyl hydroperoxide reductase subunit C (AhpC) protects bacterial and human cells against reactive nitrogen intermediates. *Mol Cell* **1**: 795-805.

Clark, J.M., and Beardsley, G.P. (1987) Functional effects of cis-thymine glycol lesions on DNA synthesis in vitro. *Biochemistry* **26**: 5398-5403.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B.G., and et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537-544.

Compan, I., Touati, D., Nunoshiba, T., Hidalgo, E., Li, Z., and Demple, B. (1993) Interaction of six global transcription regulators in expression of manganese superoxide dismutase in *Escherichia coli* K-12. *J Bacteriol* **175**: 1687-1696.

Costa Seaver, L., Imlay, J.A., Tanaka, K., Handel, K., Loewen, P.C., and Takahashi, H. (2001) Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol* **183**: 7173-7181.

Crawford, M.J., and Goldberg, D.E. (1998) Regulation of the *Salmonella typhimurium* flavohemoglobin gene. A new pathway for bacterial gene expression in response to nitric oxide. *J Biol Chem* **273**: 34028-34032.

Cruz Ramos, H., Boursier, L., Moszer, I., Kunst, F., Danchin, A., and Glaser, P. (1995) Anaerobic transcription activation in *Bacillus subtilis*: identification of distinct FNR-dependent and -independent regulatory mechanisms. *EMBO J* **14**: 5984-5994.

Cruz Ramos, H., Hoffmann, T., Marino, M., Nedjari, H., Presecan-Siedel, E., Dreesen, O., Glaser, P., and Jahn, D. (2000) Fermentative metabolism of *Bacillus subtilis*: physiology and regulation of gene expression. *J Bacteriol* **182**: 3072-3080.

Davidson, J.F., and Schiestl, R.H. (2001) Cytotoxic and genotoxic consequences of heat stress are dependent on the presence of oxygen in *Saccharomyces cerevisiae*. *J Bacteriol* **183**: 4580-4587.

Davidson, J.F., Whyte, B., Bissinger, P.H., and Schiestl, R.H. (1996) Oxidative stress is involved in heat-induced cell death in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **93**: 5116-5121.

Davies, K.J. (1987) Protein damage and degradation by oxygen radicals. I. general aspects. *J Biol Chem* **262**: 9895-9901.

De Groote, M.A., Testerman, T., Xu, Y., Stauffer, G., and Fang, F.C. (1996) Homocysteine antagonism of nitric oxide-related cytostasis in *Salmonella typhimurium*. *Science* **272**: 414-417.

De Groote, M.A., Granger, D., Xu, Y., Campbell, G., Prince, R., and Fang, F.C. (1995) Genetic and redox determinants of nitric oxide cytotoxicity in a *Salmonella typhimurium* model. *Proc Natl Acad Sci U S A* **92**: 6399-6403.

De Groote, M.A., Ochsner, U.A., Shiloh, M.U., Nathan, C., McCord, J.M., Dinauer, M.C., Libby, S.J., Vazquez-Torres, A., Xu, Y., and Fang, F.C. (1997) Periplasmic superoxide dismutase protects *Salmonella* from products of phagocyte NADPH-oxidase and nitric oxide synthase. *Proc Natl Acad Sci U S A* **94**: 13997-14001.

Demple, B., and Linn, S. (1982) 5,6-Saturated thymine lesions in DNA: production by ultraviolet light or hydrogen peroxide. *Nucleic Acids Res* **10**: 3781-3789.

Ding, H., and Demple, B. (1997) In vivo kinetics of a redox-regulated transcriptional switch. *Proc Natl Acad Sci U S A* **94**: 8445-8449.

Ding, H., and Demple, B. (1998) Thiol-mediated disassembly and reassembly of [2Fe-2S] clusters in the redox-regulated transcription factor SoxR. *Biochemistry* **37**: 17280-17286.

Ding, H., and Demple, B. (2000) Direct nitric oxide signal transduction via nitrosylation of iron-sulfur centers in the SoxR transcription activator. *Proc Natl Acad Sci U S A* **97**: 5146-5150.

Ding, H., Hidalgo, E., and Demple, B. (1996) The redox state of the [2Fe-2S] clusters in SoxR protein regulates its activity as a transcription factor. *J Biol Chem* **271**: 33173-33175.

Djinovic, K., Coda, A., Antolini, L., Pelosi, G., Desideri, A., Falconi, M., Rotilio, G., and Bolognesi, M. (1992a) Crystal structure solution and refinement of the semisynthetic cobalt-substituted bovine erythrocyte superoxide dismutase at 2.0 A resolution. *J Mol Biol* **226**: 227-238.

Djinovic, K., Gatti, G., Coda, A., Antolini, L., Pelosi, G., Desideri, A., Falconi, M., Marmocchi, F., Rotilio, G., and Bolognesi, M. (1992b) Crystal structure of yeast Cu,Zn superoxide dismutase. Crystallographic refinement at 2.5 A resolution. *J Mol Biol* **225**: 791-809.

Dowds, B.C. (1994) The oxidative stress response in *Bacillus subtilis*. *FEMS Microbiol Lett* **124**: 255-263.

Dubrac, S., and Touati, D. (2000) Fur positive regulation of iron superoxide dismutase in *Escherichia coli*: functional analysis of the *sodB* promoter. *J Bacteriol* **182**: 3802-3808.

Dubrac, S., and Touati, D. (2002) Fur-mediated transcriptional and post-transcriptional regulation of FeSOD expression in *Escherichia coli*. *Microbiology* **148**: 147-156.

Dukan, S., and Nystrom, T. (1998) Bacterial senescence: stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. *Genes Dev* **12**: 3431-3441.

Engelmann, S., and Hecker, M. (1996) Impaired oxidative stress resistance of *Bacillus subtilis sigB* mutants and the role of *katA* and *katE*. *FEMS Microbiol Lett* **145**: 63-69.

Engelmann, S., Lindner, C., and Hecker, M. (1995) Cloning, nucleotide sequence, and regulation of *katE* encoding a sigma B-dependent catalase in *Bacillus subtilis*. *J Bacteriol* **177**: 5598-5605.

Eymann, C., Homuth, G., Scharf, C., and Hecker, M. (2002) *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. *J Bacteriol* **184**: 2500-2520.

Farr, S.B., D'Ari, R., and Touati, D. (1986) Oxygen-dependent mutagenesis in *Escherichia coli* lacking superoxide dismutase. *Proc Natl Acad Sci U S A* **83**: 8268-8272.

Fawcett, P., Eichenberger, P., Losick, R., and Youngman, P. (2000) The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **97**: 8063-8068.

Ferrari, E., Howard, S.M., and Hoch, J.A. (1986) Effect of stage 0 sporulation mutations on subtilisin expression. *J Bacteriol* **166**: 173-179.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., and et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496-512.

Flint, D.H., Tuminello, J.F., and Emptage, M.H. (1993) The inactivation of Fe-S cluster containing hydro-lyases by superoxide. *J Biol Chem* **268**: 22369-22376.

Fuangthong, M., Atichartpongkul, S., Mongkolsuk, S., and Helmann, J.D. (2001) OhrR is a repressor of *ohrA*, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J Bacteriol* **183**: 4134-4141.

Furchgott, R.F., and Zawadzki, J.V. (1980) The obligatory role of endothelial cells in the relaxation of arterial smooth muscle by acetylcholine. *Nature* **288**: 373-376.

Furchgott, R.F., and Jothianandan, D. (1991) Endothelium-dependent and -independent vasodilation involving cyclic GMP: relaxation induced by nitric oxide, carbon monoxide and light. *Blood Vessels* **28**: 52-61.

Gardner, P.R., and Fridovich, I. (1991) Superoxide sensitivity of the *Escherichia coli* aconitase. *J Biol Chem* **266**: 19328-19333.

Gardner, P.R., Costantino, G., Szabo, C., and Salzman, A.L. (1997) Nitric oxide sensitivity of the aconitases. *J Biol Chem* **272**: 25071-25076.

Gardner, P.R., Gardner, A.M., Martin, L.A., and Salzman, A.L. (1998) Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proc Natl Acad Sci U S A* **95**: 10378-10383.

Gardner, P.R., Gardner, A.M., Martin, L.A., Dou, Y., Li, T., Olson, J.S., Zhu, H., and Riggs, A.F. (2000) Nitric-oxide dioxygenase activity and function of flavohemoglobins. sensitivity to nitric oxide and carbon monoxide inhibition. *J Biol Chem* **275**: 31581-31587.

Gennis, R.B., and Stewart, V. (1996) Respiration. In *Escherichia coli and Salmonella*: *cellular and molecular biology*. Neidhardt, F.C. (ed). Washington D.C.: ASM Press, pp. 217-260.

Gentry, D.R., Hernandez, V.J., Nguyen, L.H., Jensen, D.B., Cashel, M., Bongaerts, J., Zoske, S., Weidner, U., and Unden, G. (1993) Synthesis of the stationary-phase sigma factor sigma S is positively regulated by ppGpp. *J Bacteriol* **175**: 7982-7989.

Glaser, P., Kunst, F., Arnaud, M., Coudart, M.P., Gonzales, W., Hullo, M.F., Ionescu, M., Lubochinsky, B., Marcelino, L., Moszer, I. *et al.,* (1993) *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees. *Mol Microbiol* **10**: 371-384.

Gort, A.S., Ferber, D.M., and Imlay, J.A. (1999) The regulation and role of the periplasmic copper, zinc superoxide dismutase of *Escherichia coli*. *Mol Microbiol* **32**: 179-191.

Green, J., and Guest, J.R. (1994) Regulation of transcription at the *ndh* promoter of *Escherichia coli* by FNR and novel factors. *Mol Microbiol* **12**: 433-444.

Green, J., Anjum, M.F., and Guest, J.R. (1996) The ndh-binding protein (Nbp) regulates the *ndh* gene of *Escherichia coli* in response to growth phase and is identical to Fis. *Mol Microbiol* **20**: 1043-1055.

Green, J., Anjum, M.F., and Guest, J.R. (1997) Regulation of the *ndh* gene of *Escherichia coli* by integration host factor and a novel regulator, Arr. *Microbiology* **143**: 2865-2875.

Greenberg, J.T., Chou, J.H., Monach, P.A., and Demple, B. (1991) Activation of oxidative stress genes by mutations at the *soxQ/cfxB/marA* locus of *Escherichia coli*. *J Bacteriol* **173**: 4433-4439.

Gregory, E.M., and Dapper, C.H. (1983) Isolation of iron-containing superoxide dismutase from *Bacteroides fragilis*: reconstitution as a Mn-containing enzyme. *Arch Biochem Biophys* **220**: 293-300.

Grimshaw, C.E., Huang, S., Hanstein, C.G., Strauch, M.A., Burbulys, D., Wang, L., Hoch, J.A., and Whiteley, J.M. (1998) Synergistic kinetic interactions between components of the phosphorelay controlling sporulation in *Bacillus subtilis*. *Biochemistry* **37**: 1365-1375.

Guerout-Fleury, A.M., Shazand, K., Frandsen, N., and Stragier, P. (1995) Antibiotic-resistance cassettes for *Bacillus subtilis*. *Gene* **167**: 335-336.

Haldenwang W.G. (1995) The sigma factors of *Bacillus subtilis*. *Microbiol Rev* **59**: 1-30.

Hartford, O.M., and Dowds, B.C. (1994) Isolation and characterization of a hydrogen peroxide resistant mutant of *Bacillus subtilis*. *Microbiology* **140**: 297-304.

Hausladen, A., Privalle, C.T., Keng, T., DeAngelo, J., and Stamler, J.S. (1996) Nitrosative stress: activation of the transcription factor OxyR. *Cell* **86**: 719-729.

Hecker, M., and Volker, U. (2001) General stress response of *Bacillus subtilis* and other bacteria. *Adv Microb Physiol* **44**: 35-91.

Helmann, J.D., Wu, M.F., Kobel, P.A., Gamo, F.J., Wilson, M., Morshedi, M.M., Navre, M., and Paddon, C. (2001) Global transcriptional response of *Bacillus subtilis* to heat shock. *J Bacteriol* **183**: 7318-7328.

Hengge-Aronis, R., Lange, R., Henneberg, N., Fischer, D., Jenkins, D.E., Schultz, J.E., and Matin, A. (1993) Osmotic regulation of rpoS-dependent genes in *Escherichia coli*. *J Bacteriol* **175**: 259-265.

Henriques, A.O., Melsen, L.R., and Moran, C.P., Jr. (1998) Involvement of superoxide dismutase in spore coat assembly in *Bacillus subtilis*. *J Bacteriol* **180**: 2285-2291.

Herbig, A.F., and Helmann, J.D. (2001) Roles of metal ions and hydrogen peroxide in modulating the interaction of the *Bacillus subtilis* PerR peroxide regulon repressor with operator DNA. *Mol Microbiol* **41**: 849-859.

Herbig, A.F., and Helmann, J.D. (2002) Metal ion uptake and oxidative stress. In *Bacillus subtilis* and its closest relatives: from genes to cells. Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds.). Washington D.C.: ASM Press, pp.405-414.

Heym, B., Saint-Joanis, B., and Cole, S.T. (1999) The molecular basis of isoniazid resistance in *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**: 267-271.

Hidalgo, E., Leautaud, V. and Demple, B. (1998) The redox-regulated SoxR protein acts from a single DNA site as a repressor and an allosteric activator. *EMBO J* **17**: 2629-2636.

Hoa, T.T., Tortosa, P., Albano, M., and Dubnau, D. (2002) Rok (YkuW) regulates genetic competence in *Bacillus subtilis* by directly repressing *comK*. *Mol Microbiol* **43**: 15-26.

Hoffmann, T., Frankenberg, N., Marino, M., and Jahn, D. (1998) Ammonification in *Bacillus subtilis* utilizing dissimilatory nitrite reductase is dependent on *resDE*. *J Bacteriol* **180**: 186-189.

Holmes, D.S., and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* **114**: 193-197.

Homuth, G., Rompf, A., Schumann, W., and Jahn, D. (1999) Transcriptional control of *Bacillus subtilis hemN* and *hemZ*. *J Bacteriol* **181**: 5922-5929.

Horsburgh, M.J., Clements, M.O., Crossley, H., Ingham, E., and Foster, S.J. (2001) PerR controls oxidative stress resistance and iron storage proteins and is required for virulence in *Staphylococcus aureus*. *Infect Immun* **69**: 3744-3754.

Howitt, C.A., Udall, P.K., and Vermaas, W.F. (1999) Type 2 NADH dehydrogenases in the cyanobacterium *Synechocystis* sp. strain PCC 6803 are involved in regulation rather than respiration. *J Bacteriol* **181**: 3994-4003.

Hu, Y., Butcher, P.D., Mangan, J.A., Rajandream, M.A., and Coates, A.R. (1999) Regulation of *hmp* gene transcription in *Mycobacterium tuberculosis*: effects of oxygen limitation and nitrosative and oxidative stress. *J Bacteriol* **181**: 3486-3493.

Ignarro, L.J., Buga, G.M., Wood, K.S., Byrns, R.E., and Chaudhuri, G. (1987) Endothelium-derived relaxing factor produced and released from artery and vein is nitric oxide. *Proc Natl Acad Sci U S A* **84**: 9265-9269.

Imlay, K.R., and Imlay, J.A. (1996) Cloning and analysis of *sodC*, encoding the copper-zinc superoxide dismutase of *Escherichia coli*. *J Bacteriol* **178**: 2564-2571.

Inaoka, T., Matsumura, Y., and Tsuchido, T. (1998) Molecular cloning and nucleotide sequence of the superoxide dismutase gene and characterization of its product from *Bacillus subtilis*. *J Bacteriol* **180**: 3697-3703.

Inaoka, T., Matsumura, Y., and Tsuchido, T. (1999) SodA and manganese are essential for resistance to oxidative stress in growing and sporulating cells of *Bacillus subtilis*. *J Bacteriol* **181**: 1939-1943.

Ishige, K., Nagasawa, S., Tokishita, S., Mizuno, T., Glaser, P., and Jahn, D. (1994) A novel device of bacterial signal transducers. *EMBO J* **13**: 5195-5202.

Ivanova, A., Miller, C., Glinsky, G., and Eisenstark, A. (1994) Role of *rpoS* (*katF*) in *oxyR*-independent regulation of hydroperoxidase I in *Escherichia coli*. *Mol Microbiol* **12**: 571-578.

Jair, K.W., Martin, R.G., Rosner, J.L., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1995) Purification and regulatory properties of MarA protein, a transcriptional activator of *Escherichia coli* multiple antibiotic and superoxide resistance promoters. *J Bacteriol* **177**: 7100-7104.

Jakob, U., Eser, M., and Bardwell, J.C. (2000) Redox switch of hsp33 has a novel zinc-binding motif. *J Biol Chem* **275**: 38302-38310.

Jakob, U., Muse, W., Eser, M., and Bardwell, J.C. (1999) Chaperone activity with a redox switch. *Cell* **96**: 341-352.

Jenkins, D.E., Schultz, J.E., and Matin, A. (1988) Starvation-induced cross protection against heat or H_2O_2 challenge in *Escherichia coli*. *J Bacteriol* **170**: 3910-3914.

Jenkins, D.E., Chaisson, S.A., and Matin, A. (1990) Starvation-induced cross protection against osmotic challenge in *Escherichia coli*. *J Bacteriol* **172**: 2779-2781.

Jourlin-Castelli, C., Mani, N., Nakano, M.M., and Sonenshein, A.L. (2000) CcpC, a novel regulator of the LysR family required for glucose repression of the *citB* gene in *Bacillus subtilis*. *J Mol Biol* **295**: 865-878.

Jovanovic, S.V., and Simic, M.G. (1989) The DNA guanyl radical: kinetics and mechanisms of generation and repair. *Biochim Biophys Acta* **1008**: 39-44.

Karatan, E., Saulmon, M.M., Bunn, M.W., and Ordal, G.W. (2001) Phosphorylation of the response regulator CheV is required for adaptation to attractants during *Bacillus subtilis* chemotaxis. *J Biol Chem* **276**: 43618-43626.

Keyer, K., and Imlay, J.A. (1996) Superoxide accelerates DNA damage by elevating free-iron levels. *Proc Natl Acad Sci U S A* **93**: 13635-13640.

Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., and Kiley, P.J. (1997) Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O_2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci U S A* **94**: 6087-6092.

Kim, L., Mogk, A., and Schumann, W. (1996) A xylose-inducible *Bacillus subtilis* integration vector and its application. *Gene* **181**: 71-76.

Kitzler, J., and Fridovich, I. (1986) Effects of salts on the lethality of paraquat. *J Bacteriol* **167**: 346-349.

Klimek, J., Wozniak, M., Szymanska, G., and Zelewski, L. (1998) Inhibitory effect of free radicals derived from organic hydroperoxide on progesterone synthesis in human term placental mitochondria. *Free Radic Biol Med* **24**: 1168-1175.

Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K., Ogasawara, N., Tanaka, T., and Fujita, Y. (2001) Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *J Bacteriol* **183**: 7365-7370.

Kodama, T., Takamatsu, H., Asai, K., Ogasawara, N., Sadaie, Y., and Watabe, K. (2000) Synthesis and characterization of the spore proteins of *Bacillus subtilis* YdhD, YkuD, and YkvP, which carry a motif conserved among cell wall binding proteins. *J Biochem (Tokyo)* **128**: 655-663.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Danchin, A. *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249-256.

LaCelle, M., Kumano, M., Kurita, K., Yamane, K., Zuber, P., and Nakano, M.M. (1996) Oxygen-controlled regulation of the flavohemoglobin gene in *Bacillus subtilis*. *J Bacteriol* **178**: 3803-3808.

Lee, A.S., Teo, A.S., and Wong, S.Y. (2001) Novel mutations in *ndh* in isoniazid-resistant *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother* **45**: 2157-2159.

Li, Z., and Demple, B. (1994) SoxS, an activator of superoxide stress genes in *Escherichia coli*. Purification and interaction with DNA. *J Biol Chem* **269**: 18371-18377.

Liochev, S.I., and Fridovich, I. (1994) Paraquat diaphorases in *Escherichia coli*. *Free Radic Biol Med* **16**: 555-559.

Liochev, S.I., Hausladen, A., Beyer, W.F., Jr., and Fridovich, I. (1994) NADPH: ferredoxin oxidoreductase acts as a paraquat diaphorase and is a member of the *soxRS* regulon. *Proc Natl Acad Sci U S A* **91**: 1328-1331.

Liu, X., and Taber, H.W. (1998) Catabolite regulation of the *Bacillus subtilis ctaBCDEF* gene cluster. *J Bacteriol* **180**: 6154-6163.

Marino, M., Hoffmann, T., Schmid, R., Mobitz, H., and Jahn, D. (2000) Changes in protein synthesis during the adaptation of *Bacillus subtilis* to anaerobic growth conditions. *Microbiology* **146**: 97-105.

Martin, M.E., Byers, B.R., Olson, M.O., Salin, M.L., Arceneaux, J.E., and Tolbert, C. (1986) A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. *J Biol Chem* **261**: 9361-9367.

Martin, R.G., Gillette, W.K., Martin, N.I., and Rosner, J.L. (2002) Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli*. *Mol Microbiol* **43**: 355-370.

Martin-Verstraete, I., Debarbouille, M., Klier, A., and Rapoport, G. (1992) Mutagenesis of the *Bacillus subtilis* "-12, -24" promoter of the levanase operon and evidence for the existence of an upstream activating sequence. *J Mol Biol* **226**: 85-99.

Martinac, B. (2001) Mechanosensitive channels in prokaryotes. *Cell Physiol Biochem* **11**: 61-76.

McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981) Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* beta-lactamase gene. *J Biol Chem* **256**: 11283-11291.

Meier, B., Barra, D., Bossa, F., Calabrese, L., and Rotilio, G. (1982) Synthesis of either Fe- or Mn-superoxide dismutase with an apparently identical protein moiety by an anaerobic bacterium dependent on the metal supplied. *J Biol Chem* **257**: 13977-13980.

Meima, R., Eschevins, C., Fillinger, S., Bolhuis, A., Hamoen, L.W., Dorenbos, R., Quax, W.J., van Dijl, J.M., Provvedi, R., Chen, I., Dubnau, D., and Bron, S. (2002) The *bdbDC* operon of *Bacillus subtilis* encodes thiol-disulfide oxidoreductases required for competence development. *J Biol Chem* **277**: 6994-7001.

Membrillo-Hernandez, J., Coopamah, M.D., Anjum, M.F., Stevanin, T.M., Kelly, A., Hughes, M.N., and Poole, R.K. (1999) The flavohemoglobin of *Escherichia coli* confers resistance to a nitrosating agent, a "Nitric oxide Releaser," and paraquat and is essential for transcriptional responses to oxidative stress. *J Biol Chem* **274**: 748-754.

Meng, W., Green, J., and Guest, J.R. (1997) FNR-dependent repression of *ndh* gene expression requires two upstream FNR-binding sites. *Microbiology* **143**: 1521-1532.

Messner, K.R., and Imlay, J.A. (1999) The identification of primary sites of superoxide and hydrogen peroxide formation in the aerobic respiratory chain and sulfite reductase complex of *Escherichia coli*. *J Biol Chem* **274**: 10119-10128.

Miesel, L., Weisbrod, T.R., Marcinkeviciene, J.A., Bittman, R., and Jacobs, W.R., Jr. (1998) NADH dehydrogenase defects confer isoniazid resistance and conditional lethality in *Mycobacterium smegmatis*. *J Bacteriol* **180**: 2459-2467.

Miller, J.H. (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Miller, P.F., Gambino, L.F., Sulavik, M.C., and Gracheck, S.J. (1994) Genetic relationship between *soxRS* and *mar* loci in promoting multiple antibiotic resistance in *Escherichia coli*. *Antimicrob Agents Chemother* **38**: 1773-1779.

Mongkolsuk, S., Praituan, W., Loprasert, S., Fuangthong, M., and Chamnongpol, S. (1998) Identification and characterization of a new organic hydroperoxide resistance (*ohr*) gene

with a novel pattern of oxidative stress regulation from *Xanthomonas campestris* pv. phaseoli. *J Bacteriol* **180**: 2636-2643.

Moreno, M.S., Schneider, B.L., Maile, R.R., Weyler, W., and Saier, M.H., Jr. (2001) Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* **39**: 1366-1381.

Muffler, A., Fischer, D., and Hengge-Aronis, R. (1996) The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for rpoS translation in *Escherichia coli. Genes Dev* **10**: 1143-1151.

Musser, J.M., Kapur, V., Williams, D.L., Kreiswirth, B.N., van Soolingen, D., and van Embden, J.D. (1996) Characterization of the catalase-peroxidase gene (*katG*) and *inhA* locus in isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance. *J Infect Dis* **173**: 196-202.

Nakano, M.M. (2002) Induction of ResDE-Dependent Gene Expression in *Bacillus subtilis* in Response to Nitric Oxide and Nitrosative Stress. *J Bacteriol* **184**: 1783-1787.

Nakano, M.M., and Zuber, P. (1998) Anaerobic growth of a "strict aerobe" (*Bacillus subtilis*). *Annu Rev Microbiol* **52**: 165-190.

Nakano, M.M., and Zuber, P. (2002) Anaerobiosis. In *Bacillus subtilis and its closest relatives, From genes to cells*. Sonenshein, A.L., Hoch, J.A. and Losick, R. (eds). Washington D.C.: ASM Press, pp. 393-404.

Nakano, M.M., Dailly, Y.P., Zuber, P., and Clark, D.P. (1997) Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. *J Bacteriol* **179**: 6749-6755.

Nakano, M.M., Hoffmann, T., Zhu, Y., and Jahn, D. (1998) Nitrogen and oxygen regulation of *Bacillus subtilis nasDEF* encoding NADH-dependent nitrite reductase by TnrA and ResDE. *J Bacteriol* **180**: 5344-5350.

Nakano, M.M., Zuber, P., Glaser, P., Danchin, A., and Hulett, F.M. (1996) Two-component regulatory proteins ResD-ResE are required for transcriptional activation of *fnr* upon oxygen limitation in *Bacillus subtilis*. *J Bacteriol* **178**: 3796-3802.

Nakano, M.M., Zhu, Y., Lacelle, M., Zhang, X., and Hulett, F.M. (2000a) Interaction of ResD with regulatory regions of anaerobically induced genes in *Bacillus subtilis*. *Mol Microbiol* **37**: 1198-1207.

Nakano, M.M., Zheng, G., Zuber, P., Cruz Ramos, H., Hoffmann, T., Marino, M., Nedjari, H., Presecan-Siedel, E., Dreesen, O., Glaser, P., and Jahn, D. (2000b) Dual control of *sbo-alb* operon expression by the Spo0 and ResDE systems of signal transduction under anaerobic conditions in *Bacillus subtilis*. *J Bacteriol* **182**: 3274-3277.

Nakano, M.M., Zhu, Y., Haga, K., Yoshikawa, H., Sonenshein, A.L., Zuber, P., Marino, M., Nedjari, H., Presecan-Siedel, E., Dreesen, O., Glaser, P., and Jahn, D. (1999) A mutation in the 3-phosphoglycerate kinase gene allows anaerobic growth of *Bacillus subtilis* in the absence of ResE kinase. *J Bacteriol* **181**: 7087-7097.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**: 1-6.

Nunoshiba, T., Hidalgo, E., Li, Z., and Demple, B. (1993a) Negative autoregulation by the *Escherichia coli* SoxS protein: a dampening mechanism for the *soxRS* redox stress response. *J Bacteriol* **175**: 7492-7494.

Nunoshiba, T., deRojas-Walker, T., Wishnok, J.S., Tannenbaum, S.R., and Demple, B. (1993b) Activation by nitric oxide of an oxidative-stress response that defends *Escherichia coli* against activated macrophages. *Proc Natl Acad Sci U S A* **90**: 9993-9997.

O'Dell, T.J., Hawkins, R.D., Kandel, E.R., and Arancio, O. (1991) Tests of the roles of two diffusible substances in long-term potentiation: evidence for nitric oxide as a possible early retrograde messenger. *Proc Natl Acad Sci U S A* **88**: 11285-11289.

Ochsner, U.A., Hassett, D.J., and Vasil, M.L. (2001) Genetic and physiological characterization of *ohr*, encoding a protein involved in organic hydroperoxide resistance in *Pseudomonas aeruginosa*. *J Bacteriol* **183**: 773-778.

Ogawa, K., Akagawa, E., Yamane, K., Sun, Z.W., LaCelle, M., Zuber, P., and Nakano, M.M. (1995) The *nasB* operon and *nasA* gene are required for nitrate/nitrite assimilation in *Bacillus subtilis*. *J Bacteriol* **177**: 1409-1413.

Ogura, M., Yamaguchi, H., Yoshida, K., Fujita, Y., and Tanaka, T. (2001) DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B.subtilis* two-component regulatory systems. *Nucleic Acids Res* **29**: 3804-3813.

Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y., and Tanaka, T. (2002) Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol* **184**: 2344-2351.

Paidhungat, M., Setlow, B., Driks, A., and Setlow, P. (2000) Characterization of spores of *Bacillus subtilis* which lack dipicolinic acid. *J Bacteriol* **182**: 5505-5512.

Palmer, R.M., Ferrige, A.G., and Moncada, S. (1987) Nitric oxide release accounts for the biological activity of endothelium-derived relaxing factor. *Nature* **327**: 524-526.

Parge, H.E., Hallewell, R.A., and Tainer, J.A. (1992) Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn superoxide dismutase. *Proc Natl Acad Sci U S A* **89**: 6109-6113.

Parker, M.W., and Blake, C.C. (1988) Iron- and manganese-containing superoxide dismutases can be distinguished by analysis of their primary structures. *FEBS Lett* **229**: 377-382.

Petersohn, A., Engelmann, S., Setlow, P., and Hecker, M. (1999) The *katX* gene of *Bacillus subtilis* is under dual control of sigmaB and sigmaF. *Mol Gen Genet* **262**: 173-179.

Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J.D., Volker, U., and Hecker, M. (2001) Global analysis of the general stress response of *Bacillus subtilis*. *J Bacteriol* **183**: 5617-5631.

Pomposiello, P.J., Bennik, M.H., Demple, B., Nunoshiba, T., Hidalgo, E., and Li, Z. (2001) Genome-wide transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *J Bacteriol* **183**: 3890-3902.

Poole, R.K., and Hughes, M.N. (2000) New functions for the ancient globin family: bacterial responses to nitric oxide and nitrosative stress. *Mol Microbiol* **36**: 775-783.

Price, C.W. (2002) General stress response. In *Bacillus subtilis and its closest relatives, From genes to cells*. Sonenshein, A.L., Hoch, J.A. and Losick, R. (eds). Washington D.C.: ASM Press, pp. 369-384.

Price, C.W., Fawcett, P., Ceremonie, H., Su, N., Murphy, C.K., and Youngman, P. (2001) Genome-wide analysis of the general stress response in *Bacillus subtilis*. *Mol Microbiol* **41**: 757-774.

Prinz, W.A., Aslund, F., Holmgren, A., and Beckwith, J. (1997) The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J Biol Chem* **272**: 15661-15667.

Que, Q., and Helmann, J.D. (2000) Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a bifunctional regulator related to the diphtheria toxin repressor family of proteins. *Mol Microbiol* **35**: 1454-1468.

Radi, R., Bush, K.M., and Freeman, B.A. (1993) The role of cytochrome c and mitochondrial catalase in hydroperoxide-induced heart mitochondrial lipid peroxidation. *Arch Biochem Biophys* **300**: 409-415.

Rietsch, A., and Beckwith, J. (1998) The genetics of disulfide bond metabolism. *Annu Rev Genet* **32**: 163-184.

Riley, P.A. (1994) Free radicals in biology: oxidative stress and the effects of ionizing radiation. *Int J Radiat Biol* **65**: 27-33.

Ritz, D., Lim, J., Reynolds, C.M., Poole, L.B., and Beckwith, J. (2001) Conversion of a peroxiredoxin into a disulfide reductase by a triplet repeat expansion. *Science* **294**: 158-160.

Ritz, D., Patel, H., Doan, B., Zheng, M., Aslund, F., Storz, G., and Beckwith, J. (2000) Thioredoxin 2 is involved in the oxidative stress response in *Escherichia coli*. *J Biol Chem* **275**: 2505-2512.

Rosario, M.M., Fredrick, K.L., Ordal, G.W., and Helmann, J.D. (1994) Chemotaxis in *Bacillus subtilis* requires either of two functionally redundant CheW homologs. *J Bacteriol* **176**: 2736-2739.

Rosner, J.L., and Slonczewski, J.L. (1994) Dual regulation of *inaA* by the multiple antibiotic resistance (mar) and superoxide (soxRS) stress response systems of *Escherichia coli. J Bacteriol* **176**: 6262-6269.

Rozwarski, D.A., Grant, G.A., Barton, D.H., Jacobs, W.R., Jr., and Sacchettini, J.C. (1998) Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis. Science* **279**: 98-102.

Ruddock, L.W., and Klappa, P. (1999) Oxidative stress: Protein folding with a novel redox switch. *Curr Biol* **9**: R400-402.

Salvemini, F., Franze, A., Iervolino, A., Filosa, S., Salzano, S., Ursini, M.V., Nunoshiba, T., Hidalgo, E., Li, Z., and Demple, B. (1999) Enhanced glutathione levels and oxidoresistance mediated by increased glucose-6-phosphate dehydrogenase expression. *J Biol Chem* **274**: 2750-2757.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual.* NY: Cold Spring Harbour Laboratory Press, Cold Spring Harbour.

Scanlan, E., O'Reilly, M., Andersen, K.K., Noone, D., and Devine, K.M. (2001) Screening for increased sensitivity or resistance and level of *lacZ* reporter gene expression in response to the oxidising agents hydrogen peroxide and methyl viologen. In *Functional Analysis of Bacterial Genes; A Practical Manual.* Schumann, W., Ehrlich, S.D. and Ogasawara, N. (eds). Chichester, UK: Wiley, pp. 227-230.

Schaeffer, P., Miller, J., and Aubert, J. (1965) Catabolic repression of bacterial sporulation. *Proc Natl Acad Sci U S A* **54**: 701-711.

Scharf, C., Riethdorf, S., Ernst, H., Engelmann, S., Volker, U., and Hecker, M. (1998) Thioredoxin is an essential protein induced by multiple stresses in *Bacillus subtilis. J Bacteriol* **180**: 1869-1877.

Schrum, L.W., and Hassan, H.M. (1994) The effects of fur on the transcriptional and post-transcriptional regulation of MnSOD gene (*sodA*) in *Escherichia coli. Arch Biochem Biophys* **309**: 288-292.

Setlow, B., Setlow, P., Jenkins, D.E., Schultz, J.E., and Matin, A. (1993) Binding of small, acid-soluble spore proteins to DNA plays a significant role in the resistance of *Bacillus subtilis* spores to hydrogen peroxide. *Appl Environ Microbiol* **59**: 3418-3423.

Shibutani, S., Takeshita, M., and Grollman, A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**: 431-434.

Smith, M.A., Finel, M., Korolik, V., and Mendz, G.L. (2000) Characteristics of the aerobic respiratory chains of the microaerophiles *Campylobacter jejuni* and *Helicobacter pylori*. *Arch Microbiol* **174**: 1-10.

Spiro, S., Roberts, R.E., and Guest, J.R. (1989) FNR-dependent repression of the *ndh* gene of *Escherichia coli* and metal ion requirement for FNR-regulated gene expression. *Mol Microbiol* **3**: 601-608.

Steinman, H.M., Weinstein, L., and Brenowitz, M. (1994) The manganese superoxide dismutase of *Escherichia coli* K-12 associates with DNA. *J Biol Chem* **269**: 28629-28634.

Stephenson, K., and Hoch, J.A. (2001) PAS-A domain of phosphorelay sensor kinase A: a catalytic ATP-binding domain involved in the initiation of development in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **98**: 15251-15256.

Storz, G., Tartaglia, L.A., and Ames, B.N. (1990) Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science* **248**: 189-194.

Storz, G., Imlay, J.A., and Slonczewski, J.L. (1999) Oxidative stress. *Curr Opin Microbiol* **2**: 188-194.

Stulke, J., Hanschke, R., and Hecker, M. (1993) Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J Gen Microbiol* **139**: 2041-2045.

Sun, G., Birkey, S.M., and Hulett, F.M. (1996a) Three two-component signal-transduction systems interact for Pho regulation in *Bacillus subtilis*. *Mol Microbiol* **19**: 941-948.

Sun, G., Sharkova, E., Chesnut, R., Birkey, S., Duggan, M.F., Sorokin, A., Pujic, P., Ehrlich, S.D., and Hulett, F.M. (1996b) Regulators of aerobic and anaerobic respiration in *Bacillus subtilis*. *J Bacteriol* **178**: 1374-1385.

Tainer, J.A., Getzoff, E.D., Beem, K.M., Richardson, J.S., and Richardson, D.C. (1982) Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase. *J Mol Biol* **160**: 181-217.

Takamatsu, H., Chikahiro, Y., Kodama, T., Koide, H., Kozuka, S., Tochikubo, K., and Watabe, K. (1998) A spore coat protein, CotS, of *Bacillus subtilis* is synthesized under the regulation of sigmaK and GerE during development and is located in the inner coat layer of spores. *J Bacteriol* **180**: 2968-2974.

Tanaka, K., Handel, K., Loewen, P.C., and Takahashi, H. (1997) Identification and analysis of the rpoS-dependent promoter of *katE*, encoding catalase HPII in *Escherichia coli*. *Biochim Biophys Acta* **1352**: 161-166.

Tardat, B., and Touati, D. (1991) Two global regulators repress the anaerobic expression of MnSOD in *Escherichia coli*::Fur (ferric uptake regulation) and Arc (aerobic respiration control). *Mol Microbiol* **5**: 455-465.

Telenti, A., Honore, N., Bernasconi, C., March, J., Ortega, A., Heym, B., Takiff, H.E., and Cole, S.T. (1997) Genotypic assessment of isoniazid and rifampin resistance in *Mycobacterium tuberculosis*: a blind study at reference laboratory level. *J Clin Microbiol* **35**: 719-723.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680.

Toledano, M.B., Kullik, I., Trinh, F., Baird, P.T., Schneider, T.D., and Storz, G. (1994) Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* **78**: 897-909.

Touati, D. (2000) Sensing and protecting against superoxide stress in *Escherichia coli*-- how many ways are there to trigger *soxRS* response? *Redox Rep* **5**: 287-293.

Toulokhonov, II, Shulgina, I., and Hernandez, V.J. (2001) Binding of the transcription effector ppGpp to *Escherichia coli* RNA polymerase is allosteric, modular, and occurs near the N terminus of the beta'-subunit. *J Biol Chem* **276**: 1220-1225.

Tseng, H.J., Srikhanta, Y., McEwan, A.G., and Jennings, M.P. (2001) Accumulation of manganese in *Neisseria gonorrhoeae* correlates with resistance to oxidative killing by superoxide anion and is independent of superoxide dismutase activity. *Mol Microbiol* **40**: 1175-1186.

Tsuzuki, M., Ishige, K., Mizuno, T., Glaser, P., and Jahn, D. (1995) Phosphotransfer circuitry of the putative multi-signal transducer, ArcB, of *Escherichia coli*: in vitro studies with mutants. *Mol Microbiol* **18**: 953-962.

Vagner, V., Dervyn, E., and Ehrlich, S.D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**: 3097-3104.

Volker, U., Andersen, K.K., Antelmann, H., Devine, K.M., and Hecker, M. (1998) One of two *osmC* homologs in *Bacillus subtilis* is part of the sigmaB-dependent general stress regulon. *J Bacteriol* **180**: 4212-4218.

Wackwitz, B., Bongaerts, J., Goodman, S.D., and Unden, G. (1999) Growth phase-dependent regulation of *nuoA-N* expression in *Escherichia coli* K-12 by the Fis protein: upstream binding sites and bioenergetic significance. *Mol Gen Genet* **262**: 876-883.

Wakao, H., Wakagi, T., and Oshima, T. (1987) Purification and properties of NADH dehydrogenase from a thermoacidophilic archaebacterium, *Sulfolobus acidocaldarius*. *J Biochem (Tokyo)* **102**: 255-262.

Wendrich, T.M., and Marahiel, M.A. (1997) Cloning and characterization of a *relA/spoT* homologue from *Bacillus subtilis*. *Mol Microbiol* **26**: 65-79.

Wink, D.A., Kasprzak, K.S., Maragos, C.M., Elespuru, R.K., Misra, M., Dunams, T.M., Cebula, T.A., Koch, W.H., Andrews, A.W., Allen, J.S., *et al.* (1991) DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. *Science* **254**: 1001-1003.

Winstedt, L., and von Wachenfeldt, C. (2000) Terminal oxidases of *Bacillus subtilis* strain 168: one quinol oxidase, cytochrome aa(3) or cytochrome bd, is required for aerobic growth. *J Bacteriol* **182**: 6557-6564.

Wray, L.V., Jr., Ferson, A.E., Rohrer, K., and Fisher, S.H. (1996) TnrA, a transcription factor required for global nitrogen regulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **93**: 8841-8845.

Yagi, T. (1993) The bacterial energy-transducing NADH-quinone oxidoreductases. *Biochim Biophys Acta* **1141**: 1-17.

Yano, T., and Yagi, T. (1999) H(+)-translocating NADH-quinone oxidoreductase (NDH-1) of *Paracoccus denitrificans*. Studies on topology and stoichiometry of the peripheral subunits. *J Biol Chem* **274**: 28606-28611.

Yansura, D.G., and Henner, D.J. (1984) Use of the *Escherichia coli* lac repressor and operator to control gene expression in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **81**: 439-443.

Ye, R.W., Tao, W., Bedzyk, L., Young, T., Chen, M., and Li, L. (2000) Global gene expression profiles of *Bacillus subtilis* grown under anaerobic conditions. *J Bacteriol* **182**: 4458-4465.

Yoshida, K., Ishio, I., Nagakawa, E., Yamamoto, Y., Yamamoto, M., and Fujita, Y. (2000) Systematic study of gene expression and transcription organization in the *gntZ-ywaA* region of the *Bacillus subtilis* genome. *Microbiology* **146**: 573-579.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G. (1998) The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. *EMBO J* **17**: 6061-6068.

Zheng, M., Aslund, F., and Storz, G. (1998) Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**: 1718-1721.

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

Zheng, M., Wang, X., Templeton, L.J., Smulski, D.R., LaRossa, R.A., and Storz, G. (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J Bacteriol* **183**: 4562-4570.