Dans les champs de l'observation, le hazard ne favorise que les esprits prepares.
- Louis Pasteur

I have never let my schooling interfere with my education.
- Samuel Clemens

# University of Alberta

**Regulation of the general stress response and characterization of the novel dihemic cytochrome *c* catalase of *Azotobacter vinelandii.***

by

James Robert Sandercock 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Microbiology and Biotechnology
Department of Biological Sciences

Edmonton, Alberta
Spring, 2008

*Your file Votre référence*
*ISBN: 978-0-494-45592-0*
*Our file Notre référence*
*ISBN: 978-0-494-45592-0*

# ABSTRACT

The general stress response mediated by the sigma factor RpoS is important for the survival of bacteria in adverse environments. A mutant unable to produce RpoS was constructed in the diazotrophic bacterium *Azotobacter vinelandii* strain UW. Under non-desiccating solid-medium growth conditions the wild-type was culturable for 16.5 yr, but the *rpoS* mutant for only 10 mo. The *rpoS* mutant also exhibited reduced survival compared to the wild-type following $H_2O_2$ stress, with stationary phase cells being killed by 15 mM $H_2O_2$. RpoS expression was induced 26- and 28-fold during carbon and nitrogen diauxic shifts, declining to basal levels upon growth recovery. However *rpoS* mRNA levels did not correlate with RpoS levels following nitrogen diauxie, suggestive of post-transcriptional regulation. Evidence using an RpoS-dependent reporter suggests there is regulation of the RNAP:RpoS holoenzyme at the level of formation or activity. Thus we have evidence for RpoS regulation at the transcriptional, post-transcriptional and formation/activity levels. No evidence of proteolytic regulation was found.

Three catalases were produced in the wild-type, but only two were observed in the *rpoS* mutant. The RpoS-dependent catalase was purified and identified as a novel BCCP-like protein, and was designated an *A. vinelandii* cytochrome *c* catalase (*Av*CCC). *Av*CCC was purified 183-fold from whole cell lysates to a specific activity of 38,654 U/mg. Two $K_M$ of 8.4 and 67.3 mM were observed. The enzyme was thermally- and chemically-resistant, exhibiting residual activity after treatments as high as 92°C and retaining tetrameric organization in the presence of 5.6 M urea at 70°C, or SDS and _ME at 100°C. The monomer mass was determined to be 50,606.81 ± 3.90 Da by ESI-MS,

consistent with re-assigned ORF limits and the presence of two hemes. Tertiary-folding prediction software was used to determine the possible structure of *Av*CCC. The overall predicted structure was similar to that of *Ne*CCP. Both hemes are predicted to have His-Phe coordination, similar to that observed for the class I cytochrome *c'* proteins. Optical, resonance Raman and electron paramagnetic resonance studies suggest a 5c QS at the hemes, as observed previously for cytochrome *c'* (i.e. His-Phe) hemes, supporting the computationally predicted structures.

## Acknowledgements

Thank you Dr. Page for welcoming me into your lab group, and for supporting the multi-headed beast that was my research. I appreciate the degree of freedom you allowed me as I attempted to pioneer new directions in *Azotobacter vinelandii* research. Your additional financial support allowed me to reduce my teaching commitments over several semesters, for which I am greatly indebted.

Thank you also to my lab mates in the Page lab: Tony for taking the time to show me the ropes on the Bioflow IIIs and your ready whit and easy banter which made the intellectual grind bearable; Trevor Samis for help with many late night experiments (Fermentor buddies indeed!); Angela Scott for help in the early days as we started the fermentation studies and were still struggling to break the buggers open.

The success of this research would never have been possible without the help of our collaborators, who supplied access to highly specialized devices, advanced skills, and often both. I express my gratitude to Dr. Peter Loewen (University of Manitoba), Dr. Richard Rothery (University of Alberta) and Dr. Glen Loppnow (University of Alberta) for allowing me access to Oxygraph, EPR and RR units respectively. I also appreciate the assistance of several service unit technicians, including Mr. Troy Lock (FPLC), Dr. Amanda Doherty-Kirby (proteinase K LC-MS/MS, ESI-MS), Mr. Lorne Burke (trypsin LC-MS/MS), and technicians at the Johns Hopkins University (N-terminal sequencing).

And most importantly, thank you Melanie for your support through many years of effort. Your patience during many late night (and over-night, and over-week) experiments is greatly appreciated, as I could not have finished this work without them. Your love, combined with very strong coffee, fueled my resolve in the face of setbacks and challenges that only you can appreciate.

# Table *of* Contents

# List *of* Tables

# List *of* Figures

**CHAPTER 1**

**CHAPTER 3**



**CHAPTER 4**



**CHAPTER 5**

# Abbreviations

| | |
|---|---|
| 5c | 5-axial coordination |
| 6c | 6-axial coordination |
| Å | ångstrom, $1.0 \times 10^{-10}$ meters |
| βME | β-mercaptoethanol |
| λ | wavelength, nm |
| σ | sigma factor |
| Amp | ampicillin |
| BB | Burk's buffer with salts |
| BBGN | Burk's buffer containing glucose and ammonium acetate |
| BCCP | bacterial cytochrome *c* peroxidase |
| bp | base pair |
| BP-1 | barley grain peroxidase |
| cAMP | cyclic adenosine monophosphate |
| CAPS | 3-cyclohexamino-1-propanesolfonic acid |
| CCC | cytochrome *c* catalase |
| CCP | cytochrome *c* peroxidase |
| CFU | colony forming units |
| CFX | cell-free extract |
| CPx | catalase peroxidase |
| CRP | catabolic repression protein |
| CT1 | charge transfer band |
| Da | Dalton, protein mass unit |
| DAB | 3,3' diaminobenzidine |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytosine triphosphate |
| DNA | deoxyribonucleic acid |
| $dO_2$ | dissolved molecular oxygen |

| | |
|---|---|
| DTT | 1,4-dithio-DL-threitol |
| E | enzyme |
| Eσ | RNA polymerase holoenzyme |
| ECF | extracytoplasmic function sigma factor |
| EDTA | ethylenediamine tetraacetic acid |
| EGTA | ethyleneglycol tetraacetic acid |
| $E_M$ | electropotential midpoint |
| EPR | electron paramagnetic resonance |
| ESI-MS | electrospray ionization mass spectroscopy |
| FPLC | fast protein liquid chromatography |
| g | gravity |
| GC | gas chromatography |
| Gm | gentamycin |
| h | hour(s) |
| HP | high potential |
| HPI | hydroperoxidase I, KatG |
| HPII | hyderoperoxidase II, KatE |
| HPLC | high pressure liquid chromatography |
| HRP | horse radish peroxidase |
| HS | high spin, S = 5/2 |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| IS | intermediate spin, S = 3/2 |
| kbp | kilobase pair |
| kDa | kiloDalton |
| Km | kanamycin |
| $K_M$ | substrate binding affinity |
| LB | Luria-Bertani media |
| LC | liquid chromatography |
| LP | low potential |
| LS | low spin, S = 1/2 |
| min | minute(s) |

| | |
|---|---|
| mRNA | messenger ribonucleic acid |
| MS | mass spectroscopy |
| MWCO | molecular weight cut off |
| NADH | nicotinamide adenine dinucleotide |
| NADPH | nicotinamide adenine dinucleotide phosphate |
| $OD_{600}$ | optical density at 600 nm |
| ORF | open reading frame |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| PHB | poly-β-hydroxybutyrate |
| PMSF | phenylmethylsulphonyl flouride |
| ppGpp | guanosine tetraphosphate |
| PVDF | polyvinylidene difluoride |
| QS | quantum admixed spin, S = 3/2 – 5/2 |
| RBS | ribosome binding site |
| RF | relative fluorescence |
| RFU | relative fluorescence unit |
| RMSD | root mean square deviation |
| RNA | ribonucleic acid |
| RNAP | RNA polymerase |
| ROI | reactive oxygen intermediate |
| ROS | reactive oxygen species |
| rpm | rotations per minute |
| rRNA | ribosomal ribonucleic acid |
| SDS | sodium dodecyl sulfate |
| sec | second(s) |
| SOD | superoxide dismutase |
| sRNA | small (untranslated) ribonucleic acid |
| Tc | tetracycline |
| TEMED | tetramethylethylenediamine |
| TMBZ | 3,3',5,5,' tetramethylbenzidine |

| | |
|---|---|
| Tris | 2-amino-2-hydroxymethyl-1,3-propanediol |
| UV | ultraviolet light |
| vol | volume |
| v/m | volume per minute |
| v/v | volume per volume |
| $V_{max}$ | maximum reaction velocity |
| X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |

# CHAPTER 1

## Review of the Literature

## 1.1 General Biology of *Azotobacter vinelandii.*

Originally described in 1903 by Lipman, the bacterium *Azotobacter vinelandii* is a free-living diazotrophic obligate aerobe that has been isolated from both soil and fresh water samples (Winogradsky 1938; Claus and Hempel, 1970). Primarily considered a soil organism, *A. vinelandii* is likely to experience extreme changes in environmental conditions, including but not limited to changes in temperature, pH, osmotic pressure, UV exposure, nutrient and mineral availability, even within the same soil type. Though well-studied organisms such as the enteric bacterium *Escherichia coli* have the capacity to respond to changes within their environment, they tend to experience less extreme fluctuations in temperature, pH and nutrient concentration as compared to soil organisms. The order *Azotobacteraceae* is very closely related to *Pseudomonadaceae*, and so shares many traits with well-characterized bacteria such as *Pseudomonas aeruginosa* and *Pseudomonas putida.* As a result much of the research done on *A. vinelandii* may be compared to that done on *P. aeruginosa.* However, many biological systems have only been elucidated for *E. coli* and so despite being a poor model for soil bacteria, much of the *A. vinelandii* research must be compared to *E. coli* out of necessity.

*A. vinelandii* has a number of characteristics that allow it to endure various environmental stress conditions. Some of these characteristics are shared by other soil organisms, while some are unique to the *Azotobacteraceae.* This thesis will focus on how the bacterium *A. vinelandii* regulates the central regulator of the general stress response, RpoS, under various stress conditions. Further analysis of a unique RpoS-dependent gene product, a cytochrome *c* catalase, will follow.

### 1.1.1 *A. vinelandii* is a Diazotrophic Bacterium.

Nitrogen is a very common element, found at very high levels as a diatomic gas ($N_2$) in the Earth's atmosphere. However relatively few organisms are capable of converting nitrogen from the diatomic state to the 'fixed' state of ammonium ($NH_3$).

All nitrogen-fixing organisms are bacteria, and fall in three major groups. Some fix nitrogen in intra or inter-genic symbiotic relationships, as is seen with various cyanobacteria or various *Rhizobium* spp. and legumes, respectively. Other organisms, such as *Clostridium pasteurianum* are able to fix nitrogen under anaerobic conditions. To

date only two genera, *Azomonas* and *Azotobacter*, are known to fix nitrogen aerobically as free-living bacteria (Bingle et al, 1984). Three different nitrogenases have been identified in *A. vinelandii*, each requiring a different set of metallic cofactors. It is believed that the presence of multiple nitrogenases, each with a unique metallic cofactor requirement (Kennedy and Toukdarian, 1987; Pau *et al.*, 1993; Bishop, 1993), increases the number of environments in which *Azotobacter* can grow. All three enzymes catalyze the same reaction:

$$10H^+ + 8e^- + {}_nATP + N_2 \rightarrow {}_nADP + {}_nP_i + 2\,NH_4^+ + H_2$$

$n \geq 16$ (Halbleib and Ludden, 2000)

Though energetically expensive, this reaction is invaluable to organisms attempting to grow in environments lacking available nitrogen-sources. Interestingly, while *Azotobacter* spp. are able to utilize exogenous ammonium, they do not appear to have amino acid transporters (Thompson and Skerman, 1979). This, combined with the relatively slow growth rate of *A. vinelandii*, suggests that the organism may be optimized for growth in nitrogen poor environments where little competition is likely to exist.

Nitrogen fixation is extremely sensitive to oxygen, as nitrogenases are irreversibly inactivated by $O_2$ (Robson and Postgate, 1980). However, several solutions to oxygen inactivation exist. Many nitrogen-fixing bacteria avoid oxygen inactivation by living in symbiotic relationships with legumes. Once within the legume nodule, the bacterium produces nitrogenases and fixes nitrogen, while the legume provides energy via carbon-sources and produces an anoxic environment by binding free $O_2$ to leghaemoglobin (Kuzma *et al.*, 1993). Several free-living obligate anaerobes are capable of fixing nitrogen. Because of the lack of $O_2$ in their environment, they are free from needing a mechanism of nitrogenase protection.

In contrast, the free-living *Azotobacteraceae* are obligate aerobes, so can not avoid oxygen, and are incapable of producing oxygen-binding molecules such as leghaemoglobin. These free-living bacteria have acquired three unique traits that are believed to protect their nitrogenases from oxygen inactivation. First, under low carbon-source conditions *A. vinelandii* protects its nitrogenases by conformational protection.

This requires the production of the FeSII protein, which binds and protects nitrogenase in the presence of $O_2$, but inhibits nitrogen fixation (Moshiri et al, 1994). Unlike oxygen inhibition, conformational protection is reversible, allowing the organism to resume nitrogen fixation under conditions of lower oxygen tension.

Secondly the production of thick polysaccharide capsules, particularly those made of alginate, can act as diffusional barriers to oxygen (Hassett, 1996). As a result the alginate capsule may help protect nitrogenases by reducing the rate of oxygen diffusion below the rate of metabolism, thereby reducing the cytoplasmic oxygen concentration to all but trace levels (Sabra *et al.*, 2000). However non-mucoid mutants can still fix nitrogen, weakening this argument.

Third, *Azotobacter* spp. can utilize a branched electron transport chain (cytochrome *bd*), allowing it to shunt electrons directly to the terminal electron acceptor, oxygen, resulting in the very rapid elimination of $O_2$ and the production of $H_2O$. This process, referred to as respiratory protection, is believed to lower the oxygen tension within the cell to levels which allow for the resumption of nitrogenase activity (Robson and Postgate, 1980). However, use of the alternative pathway further increases the energy expenditure associated with nitrogen fixation, as the cytochrome *bd* system only translocates 1 proton per electron rather than the typical three protons per electron by the standard cyt *bc1*-cyt *c4/c5*-cyt *o* transport chain (Bertsova et al, 1997). This additional energy cost is unavoidable since mutations of cytochrome *bd* result in a reduced capacity to fix nitrogen under high carbon-source and oxygen conditions (Kelly et al, 1990), suggesting that rapid oxygen depletion is absolutely required for nitrogen fixation.

### 1.1.2 The Expanded Iron Requirements of Free-Living Nitrogen Fixing Bacteria.

With the exception of *Lactobacillus* spp., which utilize manganese and cobalt, most organisms have a requirement for iron for various enzymological functions (Archibald, 1983; Gureinot, 1994). Iron is required as a cofactor for many enzymes including those within the electron transport chain (cytochromes, succinate dehydrogenase, ferredoxin), those required for protection from toxic oxygen products (catalases, peroxidases, superoxide dismutase) and for the coordination of oxygen binding by heme groups.

Nitrogen fixation by a free-living bacterium can increase the iron requirements considerably. Iron is needed as a cofactor for nitrogenases, for the conformational protection enzyme, FeSII, and for the production of additional cytochromes for respiratory protection. Additionally, the extremely rapid respiration rate of *Azotobacter* spp. (Jurtshuk and Yang, 1980) could conceivably result in the production of large amounts of reactive oxygen species (ROS) such as superoxide free radicals ($\bullet O_2^-$) and hydrogen peroxide ($H_2O_2$). These oxygen products can damage cellular components, leading to aging, mutation and eventually the death of the cell (Wai et al, 1996). In response the cell produces superoxide dismutase and various catalases or peroxidases to reduce $\bullet O_2^-$ and $H_2O_2$ to water, thereby minimizing the amount of damage done to the cell (see Section 1.5 for a thorough discussion of these enzymes). These "anti-oxidant" enzymes frequently require iron as a cofactor, increasing the iron requirements of the cell further. As a result, *Azotobacter* species accumulate substantial quantities of iron, which have the potential to generate large quantities of ROS, particularly the hydroxyl radical which can be produced following superoxide radical-free $Fe^{3+}$ interactions. The production of hydroxyl radicals can be mitigated by ferritin-type proteins, which sequester iron in a biologically accessible but non-reactive form (Briat, 1992).

### 1.1.3 *A. vinelandii* Forms Dormant Cysts.

Like several other soil organisms, *A. vinelandii* is able to form dormant-state cells during periods of extreme stress. These cells, referred to as cysts, are desiccation and UV resistant (Socolofsky and Wyss, 1962; Moreno et al, 1986) and can be resuscitated up to 24 years after encystment if stored in dry conditions (Moreno et al, 1986). Cysts, being nearly metabolically inactive (Socolofsky and Wyss, 1962) are also able to wait out nutrient limitations, which confers a definite advantage over organisms that only exist in a vegetative state. Even the closely related *Pseudomonas* spp. are less able to weather such profound stresses despite their ability to form biofilm layers. Finally, vegetative *A. vinelandii* cells are stable at 4 °C for up to 12 months, but are not stable at –4 °C over this time frame. It is possible that cyst formation is also necessary for the survival of the bacterium in temperate zones, where temperatures can drop below freezing for greater than 4 months of the year.

Cyst formation is induced under nutrient limitation when carbon-sources are in excess (Lin and Sadoff, 1968; Stevenson and Socolofsky, 1972). Excess carbon-source is converted into the storage polymer poly-β-hydroxybutyrate (PHB) and can then be turned over to fuel the encystment process. PHB also has potential as an industrially useful polymer that can be used to manufacture biodegradable plastics. The mature cyst structure consists of a central body (the vegetative cell), surrounded by a thick, viscous layer referred to as the "intine layer" which is in turn surrounded by the rigid "exine layer". Both layers contain significant amounts of alginate, protein, and in the case of the exine, alkyl resorcinols (Lin and Sadoff, 1969; Reusch and Sadoff, 1979). Calcium is a required element for successful encystment as it allows rigid packing of the alginate polymers (Page and Sadoff, 1975).

### 1.2 RpoS/$\sigma^S$ is the Central Regulator of the General Stress Response.

Bacteria must mount appropriate responses to deal with environmental stresses, such as nutrient limitation, pH changes, osmotic fluctuations, oxidative stresses and temperature extremes. Most studies to date focus on the responses of enteric bacteria to such stresses, but it should be noted that the appropriate response will differ not only by the type of stress experienced, but also by the context in which the organism lives. Being highly specialized for a particular niche, most bacteria have a narrow set of responses that are optimized for survival in their "normal" environment. In some cases relatively few enzymes are required to protect the cell from specific damages (i.e. oxidative stress), so the general stress response is not invoked, and is in some cases even suppressed. In other cases, such as during nutrient limitation, numerous enzymes and protective proteins are required to help the cell survive the challenge, so many genes are induced. In Gram negative bacteria many of these stress signals are integrated by a single regulator, the sigma factor $\sigma^S$ (also known as RpoS and $\sigma^{38}$). Once transcribed and translated, RpoS binds to the core RNA polymerase (E, RNAP), allowing transcription from a specific subset of promoters that are normally silent in non-stressed cells. However, the RpoS-mediated response induces more genes than are needed to deal with the specific initiating stress: the general stress response regulates nearly 360 genes (10% of the annotated genome) in *E. coli* (Patten *et al.*, 2004) and approximately 772 genes (14% of the

genome; Schuster *et al.*, 2004) in *P. aeruginosa*, regardless of which initial stress was experienced by the cell. Because of the number of genes involved, this response could be likened to a morphological differentiation event such as sporulation.

The majority of the genes regulated by RpoS are up regulated, though the expression of some is down regulated. In most cases it is believed that down regulation is due to sigma factor competition, where a limited number of RNAP are bound by a variety of available sigma factors. It could be expected that as the amount of $\sigma^S$ increases there should be a concomitant increase in the amount of RNAP:$\sigma^S$ ($E\sigma^S$) and a decrease in the amount of $E\sigma^{70}$. This would theoretically result in an increase in the amount of transcription at $\sigma^s$-dependent promoters and a decrease in the transcription of some $\sigma^{70}$ regulated genes (Maeda *et al.*, 2000). In reality several factors have been shown to modulate the efficiency of Eσ complex formation (see Section 1.4.4.1).

The regulation of RpoS concentrations in the cell is one of the most complex bacterial regulation circuits known to date. Numerous global stresses, including changes in growth rate, shifts in temperature, osmolarity, starvation and acidification of the growth medium are able to elevate RpoS concentrations (Fig. 1.1). The RpoS response itself is modulated at the transcriptional, translational, post-translation, and activity levels. More than 20 (molecular) regulatory systems have been identified that regulate RpoS in *E. coli*. These are summarized in Fig. 1.2 and discussed below (Section 1.4). Before exploring the regulation of this sigma factor, a more in-depth discussion of the function and role of sigma factors is in order.

### 1.3 Structures and Classes of Sigma Factors.

Sigma factors are transcriptional accessory proteins that bind the core RNA polymerase complex. Once bound, the holoenzyme (Eσ) is able to bind DNA in a sequence specific manner at the gene promoter. Typically, the sigma factor allows the polymerase to recognize a hexameric sequence 10 bases upstream of the transcription start site, referred to as the –10 region, and also recognize a hexamer 35 basepairs upstream, referred to as the –35 region of the promoter. Four major classes of sigma factor exist, the Primary sigma factors, the Alternative sigma factors and the

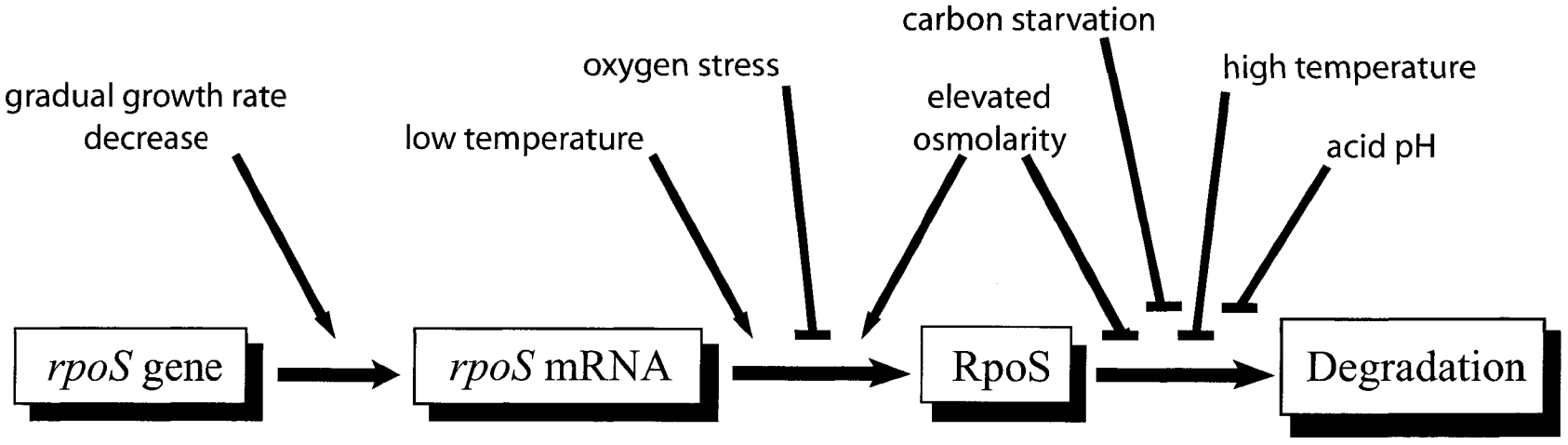



**Figure 1.1 The Environmental Stress Inputs Regulating RpoS Synthesis and Activity.** Positive regulation is indicated by arrows whereas negative regulation is indicated by blunted lines. Modified from Hengge-Aronis (2002).

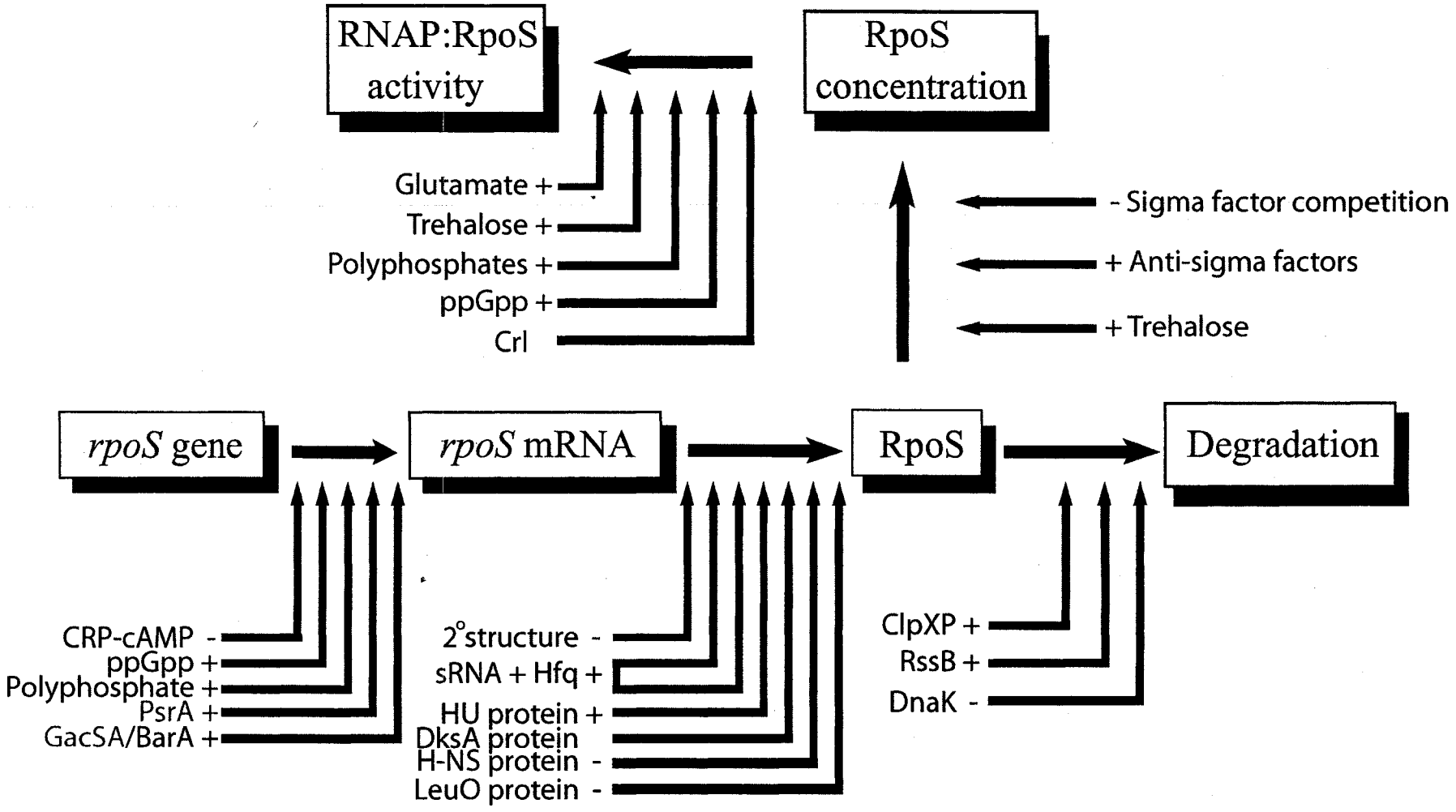



**Figure 1.2 The Regulatory Network Responsible for the Modulation of Synthesis and Activity of RpoS at the Molecular Level.** Known regulators in enteric bacteria are indicated in black, those of pseudomonads are indicated in green. Positive regulators are indicated by the + sign, whereas negative regulators are indicated by the - sign.

Extracytoplasmic Function (ECF) sigma factors. The final sigma factor, $\sigma^{54}$ (RpoN), shares little homology with other sigma factors so has been assigned to a separate class.

### 1.3.1 RpoD/$\sigma^{70}$ Structure and Domains.

The *rpoD* product, alternatively referred to as $\sigma^{D}$ or $\sigma^{70}$, is considered the "house keeping" sigma factor of most bacteria. Once bound to the core RNAP, RpoD directs the transcription of genes that are required for the normal functions of the cell under moderate to high growth rate conditions. The major domains of RpoD have been identified (Fig. 1.3). Region 1 is a unique and conserved feature of $\sigma^{70}$ that acts as an inhibitor of DNA binding. This region blocks the recognition of promoter sequences by $\sigma^{70}$ until it is complexed with the RNAP core enzyme (Dombroski *et al.*, 1992) and also promotes binding to the core RNAP (Nagai and Shimamoto, 1997). Region 2 is responsible for binding the RNAP core enzyme and for recognition of the –10 promoter region. Region 2 has been further divided into 5 subsections. Of these, region 2.2 plays a role in binding the RNAP, while regions 2.4 and 2.3 are responsible for the binding and the melting of the –10 region, respectively (Gross *et al.*, 1998). Region 2.5 is responsible for binding the extended –10 promoter region (Ponnambalam *et al.*, 1986). Region 3 is involved in abortive transcription events at several "suboptimal" promoters. Crystallographic studies have shown that region 3.2 sterically hinders the elongation process at these promoters by blocking the extrusion of newly synthesized RNA from the transcribing RNA Polymerase, once the promoter open complex is formed (Murakami *et al.*, 2002; Vassylyev *et al.*, 2002). Region 3 may also play a role in the stringent response (Cashel, *et al.*, 2003; see also Section 1.4.4.5). Region 4.2 is responsible for sequence recognition at the –35 region (Siegele et al, 1989) and is targeted by the anti-sigma factors Rsd (Jishage and Ishihama, 1998) and AsiA (Severinov and Muir, 1998).

The consensus –10 and –35 sequences recognized by the $\sigma^{70}$ of *E. coli* are TATAAT and TTGACA respectively, with an average spacing of 17 ±1 bp between them (Harley and Reynolds, 1987). Individual amino acid:nucleotide interactions between the sigma factor and the promoter sequences have been identified, but fall outside the scope of this thesis.

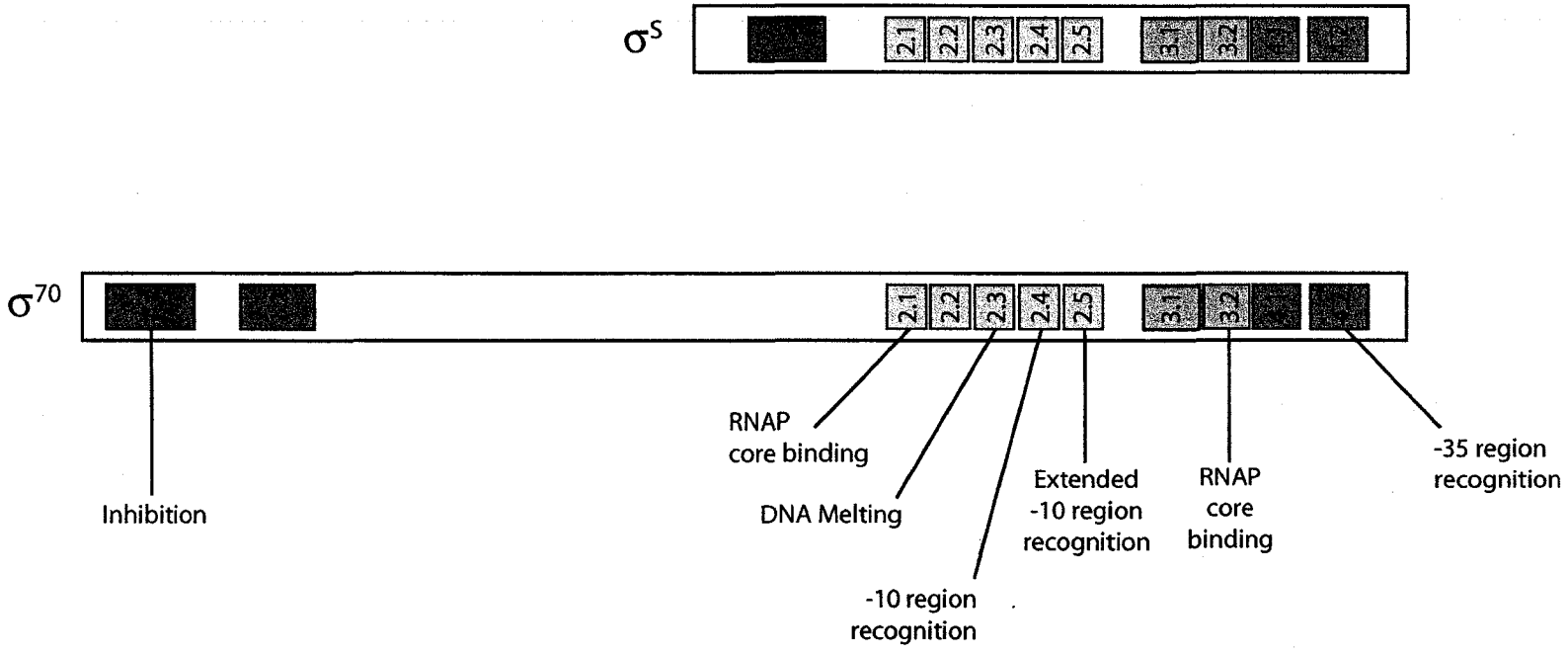



**Figure 1.3 Comparison of the Functional Domains of RpoD (σ70) and RpoS (σS) of *E. coli*.** Modified from Colland et al. (1999). Regions are defined by the system of Lonetto et al. (1992).

### 1.3.2 RpoS/$\sigma^S$ is an Alternative Sigma Factor.

Besides the primary sigma factor, most organisms produce a number of additional sigma factors. The first group of these, $\sigma^E$, $\sigma^F$, $\sigma^G$, and $\sigma^K$, was discovered in *Bacillus subtilis*, and was found to regulate the formation of spores through the "sigma cascade" (Losick and Pero, 1981). These and other sigma factors were discovered to share a high degree of identity with $\sigma^{70}$, and so were termed the "alternative" sigma factors. In most cases these sigma factors control regulons responsible for specific global functions, such as those involved in flagellar synthesis ($\sigma^{F/FliA}$), the heat shock response ($\sigma^E$ and $\sigma^H$), and the general stress response ($\sigma^S$) in *E. coli*, or sporulation ($\sigma^E$, $\sigma^F$, $\sigma^G$, and $\sigma^K$) in *Bacillus* spp. (Lonetto *et al.*, 1994). For a comprehensive list of sigma factors by bacterial species see Gross *et al.* (1998).

*E. coli* RpoS protein is 330 amino acids in length with a mass of 38 kDa (Loewen and Hengge-Aronis, 1994). The homologue in *A. vinelandii* is 334 amino acids long, and shares 76% identity with RpoS of *E. coli* K-12, and 87% identity with the RpoS of *P. aeruginosa*.

Though *E. coli* $\sigma^S$ shares considerable homology with $\sigma^{70}$ there are some notable differences. First, $\sigma^S$ has only region 1.2 at the N-terminal end of the protein, so is unlikely to have the inhibitory domain found in the primary sigma factor (Lonetto *et al.*, 1992). Region 1.2 may promote binding of $\sigma^S$ to the core enzyme. Second, a high degree of similarity to $\sigma^{70}$ throughout region 2 allows RpoS to recognize the consensus sequence CTATACT (Fig. 1.4; Espinosa-Urgel *et al.*, 1996a), which is very similar to the typical $\sigma^{70}$ –10 promoter sequence, TATAAT. While the consensus sequences are very similar, deviations from consensus are differentially tolerated by the two sigma factors due to minor amino acid substitutions. Notably, E$\sigma^S$ binds the extended -10 region best when the sequence is CT, rather than the GT sequence which is preferred by E$\sigma^{70}$ (Becker and Hengge-Aronis, 2001; Burr *et al.*, 2000). Region 2.5, centred on L173, is critical to the proteolytic regulation of RpoS *via* the chaperone/anti-sigma factor RssB and the protease ClpXP (Muffler *et al.*, 1996a; see also Section 1.5.3.1). Finally, there is considerable divergence between the $\sigma^{70}$ and $\sigma^S$ sequence in region 4. While $\sigma^{70}$ recognizes the –35 sequence TTGACA, a –35 consensus sequence for $\sigma^S$ has yet to be determined. At this


**Figure 1.4 Comparison of the Sequences in RpoS-Dependent Promoters.** The -10 region is boxed and the consensus sequence is indicated below. The spacing between the -10 region and the transcription start site (+1) is 6 basepairs on average. From Espinosa-Urgel et al. (1996a).

time the only requirement seems to be that the sequence upstream of the –10 box contains a degree of curvature. Such curvature may increase $E\sigma^S$ activity by binding the α subunit and thereby stabilizing the $E\sigma^S$:promoter complex (Espinosa-Urgel *et al.*, 1996b) especially if the region contains a distal half-UP element (an A-T rich sequence 10-20 residues upstream of the -35 hexamer) (Ballesteros et al, 1998). Alternatively the curvature may promote the binding of nucleoid proteins such as H-NS and IHF, which would in turn affect transcriptional activity at these promoters (Colland *et al.*, 2000).

### 1.3.3 The Extracytoplasmic Function (ECF) Sigma Factor Subfamily.

Lonetto and co-workers (1994) first coined the term Extracytoplasmic Function (ECF) when they compared the *Streptomyces coelicolor* gene product SigE to CarQ (*Myxococcus xanthus*), AlgU (*P. aeruginosa*), HrpL (*Pseudomonas syringae*), $\sigma^E$ (*E. coli*), CnrH (*Alcaligenes eutrophus*), FecI (*E. coli*) and SigX (*B. subtilis*) and found that they were all based upon the same modular design as $\sigma^{70}$. In each case the product was shown to regulate extracytoplasmic functions in response to extracellular stimuli (Lonetto *et al.*, 1994). The ECF are considered a subfamily of sigma factors, distinct from the primary and alternative sigma factors (Fig. 1.5; Lonetto *et al.*, 1994). These sigmas usually mediate highly specific cell responses rather than global responses, as is seen with the primary and alternative factors. For example, in *P. aeruginosa* and *A. vinelandii* the ECF sigma factor AlgU mediates only the synthesis of alginate (Hershberger *et al.*, 1995; Martinez-Salazar *et al.*, 1996) while $\sigma^{Fec}$ of *E. coli* (PvdS in *P. aeruginosa*) is exclusively involved with iron uptake (Leoni et al,. 2000). It has been observed that "generalist" organisms such as *S. coelicolor* and *Pseudomonas putida* have more ECF factors (49 and 19 ECFs respectively) than specialist organisms like *E. coli* (2 ECFs) and *Mycoplasma genitalium* (no ECFs) (Bentley *et al.*, 2002; Martinez-Bueno *et al.*, 2002; Blattner *et al.*, 1997; Fraser *et al.*, 1995). It is interesting to note that a cursory examination of the *A. vinelandii* genome revealed 16 putative ECF factors (Sandercock, unpublished). As was found for *P. aeruginosa* and *P. putida* the majority of the *A. vinelandii* ECFs shared significant similarity with *E. coli* FecI (Table A, Appendix). Though it has been suggested that these may all be involved in iron regulation in the pseudomonads

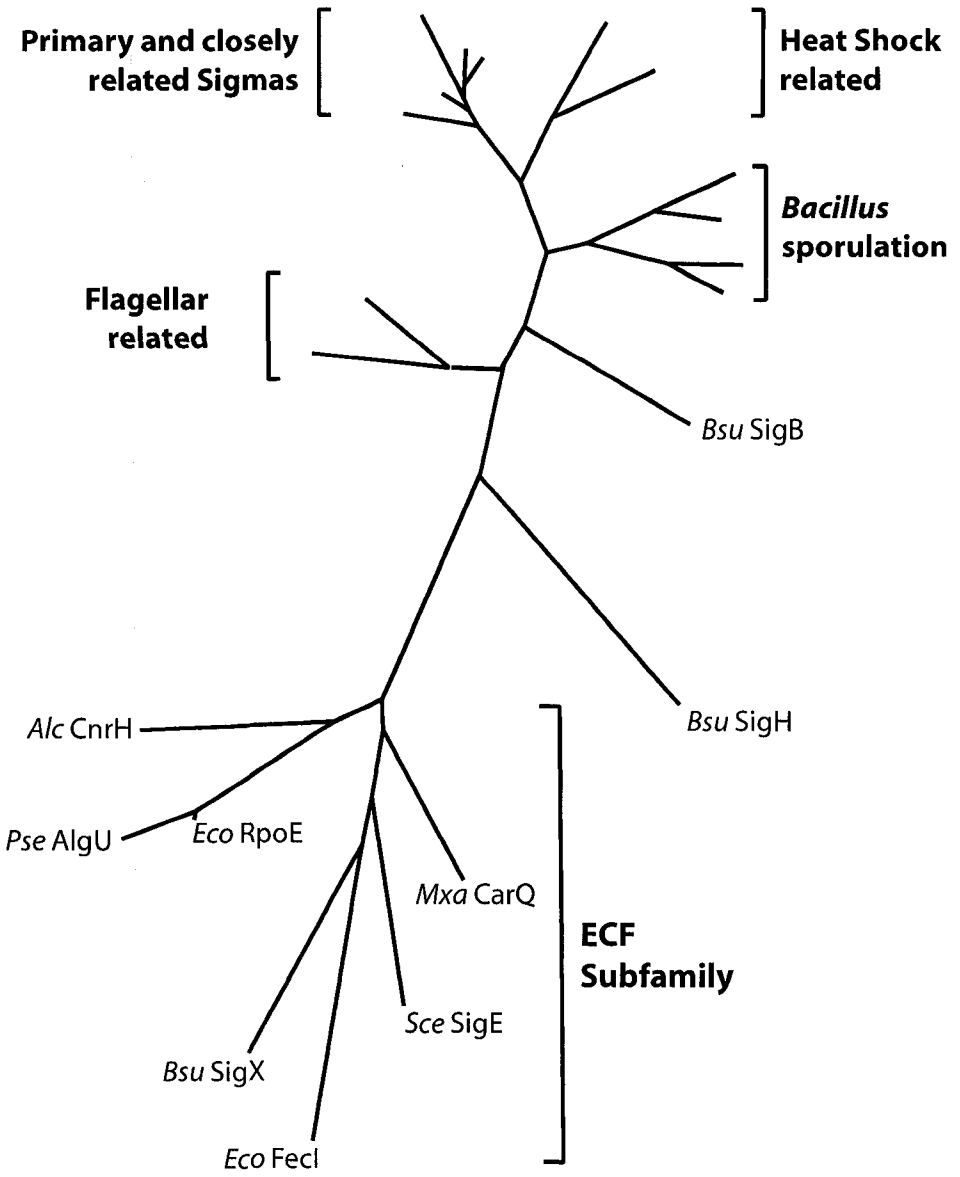



**Figure 1.5 Early Phylogenetic Tree of the Major Sigma Factors from Several Bacteria,** highlighting the divergence of the ECF sigma factors from other classes. From Lonetto et al. (1994).

(Martinez-Bueno et al, 2002), it is important to note that function does not necessarily follow homology.

Structurally, the ECF sigma factors diverge from $\sigma^{70}$ in regions 2.3, 2.4 and 3 (Gross *et al.*, 1998), resulting in the recognition of unique –10 sequences for each of these sigma factors. Because region 3 may be responsive to the stringent response, individual ECFs may also have varied activity profiles in the presence of the small regulating molecule, guanosine tetraphosphate (ppGpp). Conversely, region 4 of the ECF subfamily shares a high degree of similarity to $\sigma^{70}$, resulting in the recognition of sequences that are similar to the typical –35 sequence (Visca *et al.*, 2002).

The activity of most ECFs is regulated by a closely associated membrane bound anti-sigma factor, where external inducing cues cause the dissociation of the sequestered σ-factor from the anti-sigma factor (Fig. 1.6). As an example, the AlgU sigma factor of *P. aeruginosa* and *A. vinelandii* is a member of a tri-partite membrane-anchored system. The sigma factor AlgU is sequestered by the membrane-bound anti-sigma factor, MucA, which in turn is bound by the periplasmic modulator, MucB. Removal of MucB in response to envelope stress or other unidentified signals result in conformational changes within the MucA protein structure, liberating the sigma factor. Free AlgU can then bind free RNAP, directing transcription at the *algD* promoter, resulting in alginate synthesis (Rowen and Deretic, 2000; Hughes and Mathee, 1998).

### 1.4 Regulation of RpoS.

RpoS activity, defined as transcriptional activity by the holoenzyme $E\sigma^S$, is regulated at the transcriptional, translational, post-translational and activity levels (Fig. 1.2). Transcription of *E. coli rpoS* is induced 5 to 10-fold when cells experience a gradual reduction in growth rate, such as when they enter stationary phase when grown on minimal or defined media, or when they are grown under nutrient limiting chemostatic conditions (Lange and Hengge-Aronis, 1991; Mulvey *et al.*, 1990; Zgurskaya *et al.*, 1997). Stresses, such as shifts to low temperature, high osmolarity, low pH, or increased cell density can increase the efficiency of translation of available *rpoS* mRNA (Sledjeski *et al.*, 1996a; Muffler *et al.*, 1996a; Lange and Hengge-Aronis, 1994a). Conversely oxygen stress can decrease the efficiency of *rpoS* translation (Zhang *et al.*, 1998). Other

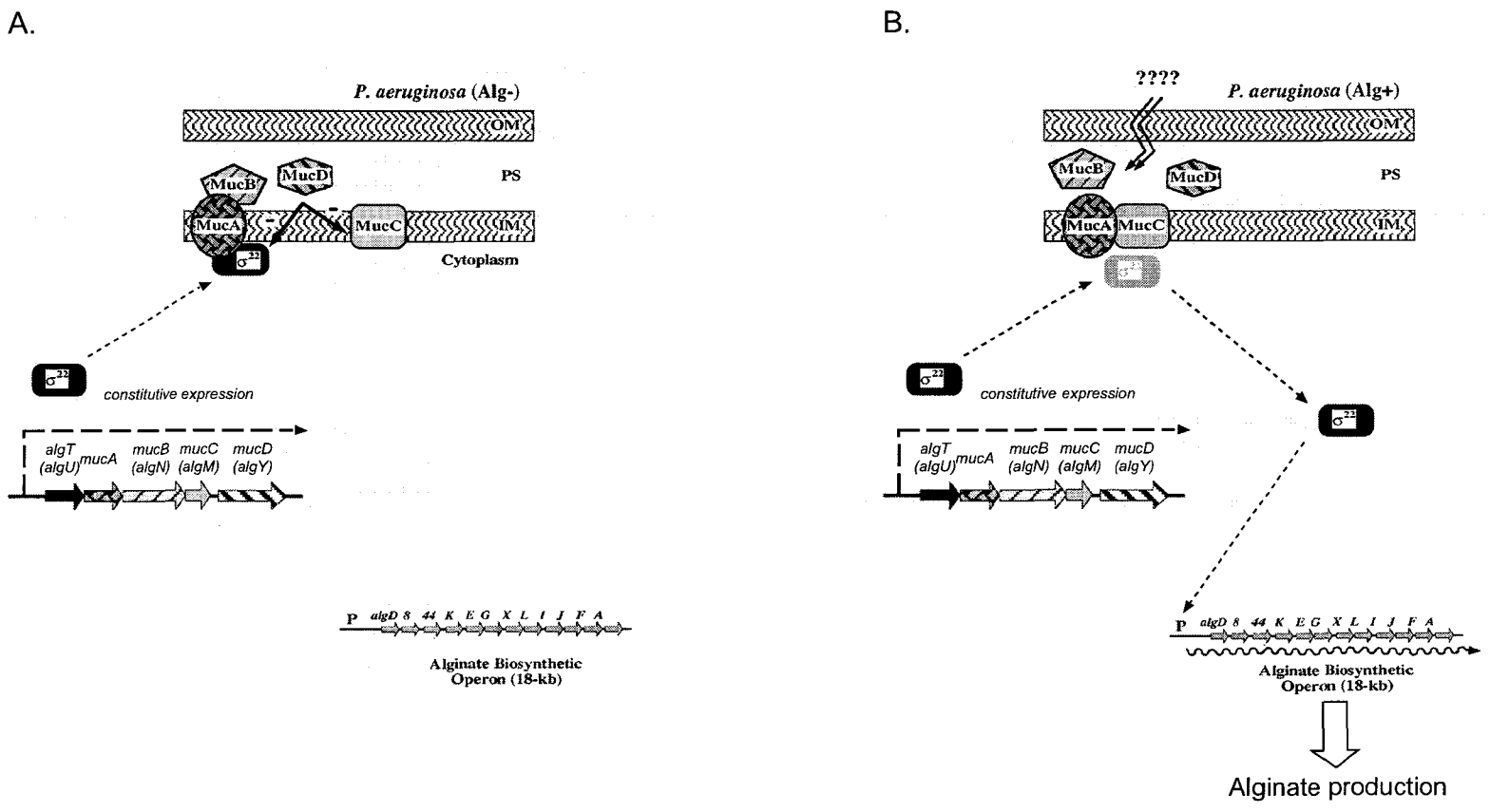



**Fig. 1.6 A Schematic of the AlgU/$\sigma^{22}$ Regulatory Circuit of *P. aeruginosa* as an Example of ECF Regulation.** Under non-alginate producing conditions the sigma factor $\sigma^{22}$ is sequestered by the membrane bound anti-sigma factor MucA, so it cannot activate transcription at the alginate biosynthetic operon (A). Following activation by an unknown signal MucC interacts with MucA, causing the release of the sigma factor. Once partitioned to the cytoplasm, $\sigma^{22}$ can activate transcription at the alginate biosynthetic operon, resulting in the production of alginate (B). From Hughes and Matthee (1998).

stresses, such as carbon starvation, high temperature shock, shifts to low pH and high osmolarity can cause a reduction in the efficiency of $\sigma^S$ proteolysis (Lange and Hengge-Aronis, 1994a; Muffler *et al.*, 1997; Muffler *et al.*, 1996a; Bearson *et al.*, 1996; Muffler *et al.*, 1996b). Additionally, several stress responses cause changes in the efficiency of $\sigma$:RNAP formation, while others affect the efficiency of $E\sigma^S$ promoter recognition. For most of these stresses, individual molecular regulators have been identified that account for the observed responses.

### 1.4.1 Transcriptional Regulation of *rpoS*.

Transcriptional regulation of *rpoS* occurs primarily as a response to reduced growth rate. This is seen when organisms exhaust a carbon-source or other nutrient and either enter stationary phase or undergo a diauxic shift, thereby utilizing an alternative nutrient source. The associated molecular signals that regulate transcription include CRP-cAMP (or equivalent catabolite repression system), the two-component signal transduction system GacS/A, protein PsrA and the polyphosphates, including ppGpp. Quorum sensing systems are known to affect transcription of *rpoS*, but since there is no evidence for quorum sensing in *A. vinelandii*, these systems are not discussed further.

#### 1.4.1.1 *rpoS* Expression is Regulated by Catabolite Repression Systems.

Transcription of the *rpoS* gene occurs from three major promoters. The first two are promoters for the upstream gene, *nlpD*, which is expressed at basal levels regardless of growth rate or growth phase (Lange and Hengge-Aronis, 1994b). Occasional read-through of the *nlpD* gene results in low levels of *rpoS* transcription even during exponential growth phase. The inducible *rpoS* promoter has been mapped to the *nlpD* structural gene and is flanked by two putative CRP boxes, CRP I and CRP II (Fig 1.7; Lange and Hengge-Aronis, 1994b; Hengge-Aronis, 2002). Lange and Hengge-Aronis observed that CRP-cAMP is a negative regulator of *rpoS* transcription, and that *cya* or *crp* mutants were derepressed for *rpoS* transcription during exponential phase (Lange and Hengge-Aronis, 1994a). However it was also observed that upon entering stationary phase, when cAMP levels are expected to increase, CRP-cAMP became a positive regulator of *rpoS* transcription (Hengge-Aronis, 2002). This dichotomy may be

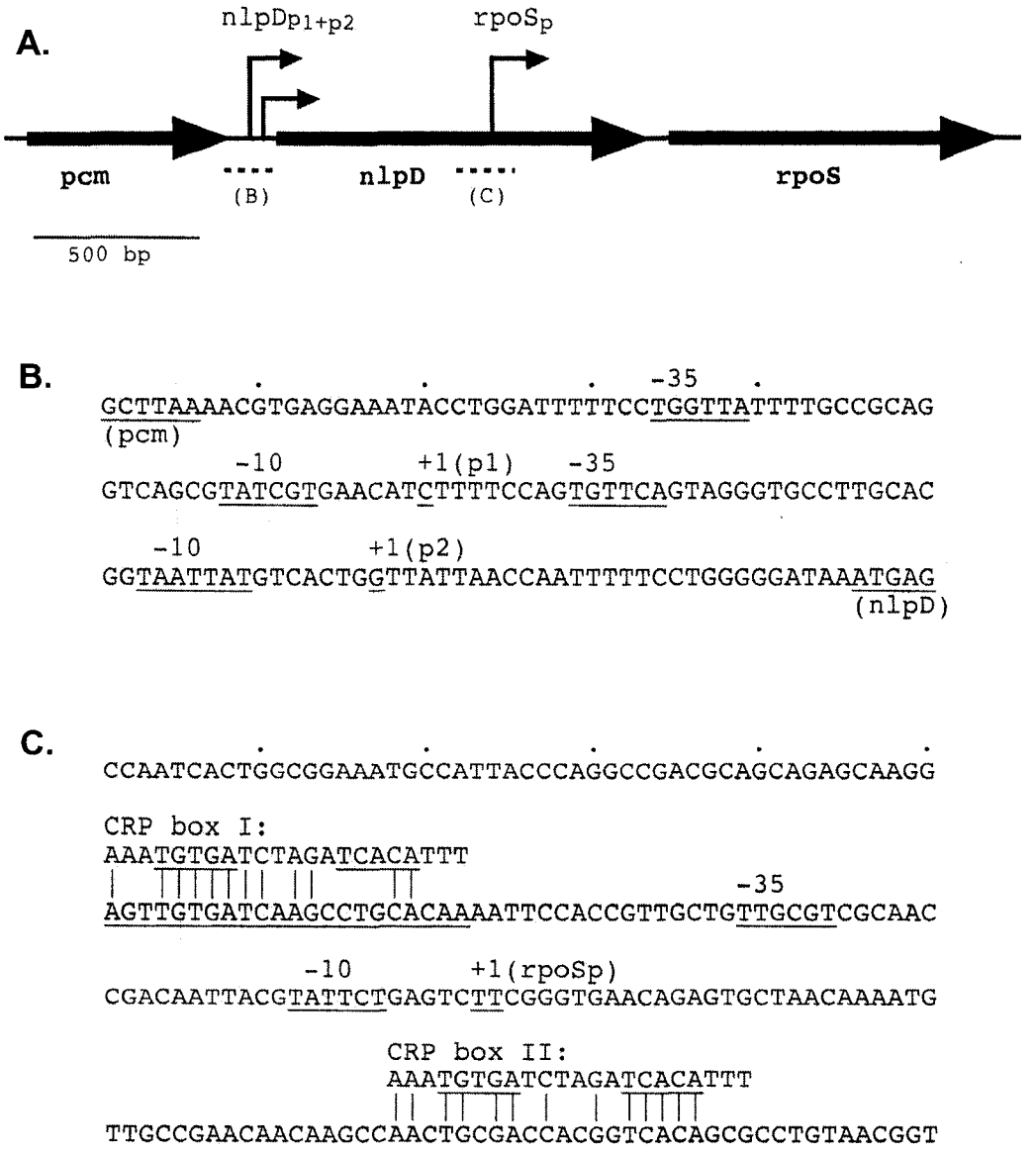



**Figure 1.7 A Schematic Representation of the Three Promoters Known to Result in *rpoS* mRNA Transcription.** Two vegetative promoters (*nlpD*P1, *nlpD*P2) for the upstream open reading frame *nlpD* are known to occasionally read-through their cognate terminators, resulting in the transcription of *rpoS*. A stress regulated promoter (*rpoS*P) is located within the coding region of *nlpD* (A). The -10 and -35 sequences of the *nlpD* promoters are shown in panel B and largely conform to the RpoD consensus sequence. The stress-responsive promoter (C) contains two CRP boxes. Interestingly the -10 and -35 sequences conform to the RpoD consensus sequence. From Hengge-Aronis (2002).

explained by the differential affinity of cAMP-CRP for the two binding sites. If true, the few molecules of cAMP-CRP present during exponential phase would bind the high affinity site (II) and cause transcriptional repression. During stationary phase, when significantly more cAMP is present, the low affinity site (I) would be saturated, and the over-riding activating function would dominate. However the affinity of cAMP-CRP for these sites has not been confirmed experimentally, so this hypothesis remains theoretical.

Many soil organisms do not employ CRP-cAMP for catablolite repression, and in many cases glucose does not cause a significant degree of catabolite repression. For example, *Bacillus* spp. utilize the catabolite control protein (CcpA; Kruger, 1993) and the protein Hpr (Deutscher *et al.*, 1994) instead of CRP, and are not glucose repressible. Hpr, a member of the PEP phosphotransferase system, undergoes autophosphorylation under high-energy conditions, and is then able to form a complex with CcpA. The complex then binds promoters at CRE sites, resulting in promoter repression (Wray *et al.*, 1994).

Likewise, in pseudomonads the presence of glucose does not cause catabolite repression. Rather the organisms preferentially utilize acetate and tricarboxylic acids, with relatively little glucose metabolism until the organic acids become exhausted. Furthermore cAMP pools do not fluctuate with growth rate or with variations in sugar availability (Suh *et al.*, 2002). Instead, in the pseudomonads (and so likely *A. vinelandii*) the Crc protein is responsible for catabolite exclusion. However, at this time the molecular mechanism of Crc is unclear, as it is not a DNA binding protein.

#### 1.4.1.2 *rpoS* Expression is Regulated by GacA/S (BarA/UvrY).

*rpoS* mutant strains of *P. fluorescens*, in addition to being stress sensitive, were found to produce elevated levels of the anti-fungal secondary metabolites pyoluteorin and 2,4-diacetyl-phloroglucinol but were unable to produce pyrrolnitrin, an additional stationary phase specific metabolite (Sarniguet *et al.*, 1995). Additional studies identified non-*rpoS* mutations that also resulted in the modification of secondary metabolite expression, and increased sensitivity to oxidative or osmotic stresses. One mutation was localized to *gacA*, the product of which falls in the FixJ family of response regulators. The other mutation was localized to the *gacS* gene; who's product is a membrane bound sensor

kinase (Whistler *et al.*, 1998). Widely considered members of the same two-component regulatory system, GacS/A were shown to regulate *rpoS* through mutational studies. Mutants *gacS*::Tn*5*, *gacA*(V203), and wild-type *P. fluorescens* all exhibited equal expression rates of *rpoS* mRNA throughout exponential phase when grown on a rich medium. As the culture entered stationary phase, the wild-type strain exhibited a rapid increase in *rpoS* transcription while the *gacS*::Tn*5* and *gacA*(V203) mutants showed only moderate induction. By late stationary phase the mutants had about half as much *rpoS* expression as the wild-type strain. Because the DNA binding sites of GacA have not been described yet, it is unknown if the response regulator binds the *rpoS* promoter directly or acts through intermediates (Whistler et al, 1998).

In *E. coli* the GacS homologue, BarA, was identified for its importance in iron-starvation (Zhang and Normark, 1996) and oxygen stress responses (Mukhopadhyay and Schellhorn, 1997). Mutants of *barA* produced significantly less *rpoS* mRNA in both exponential and stationary phase when grown on LB media (Mukhopadhyay et al., 2000), which differs somewhat from the *P. fluorescens* data, where the *gacS* mutant saw a reduction in *rpoS* transcription only in stationary phase. However most of the *E. coli* work was done in strain MC4100, a *relA* mutant that produces reduced levels of ppGpp. Unfortunately ppGpp is also an inducer of *rpoS* expression, making it difficult to differentiate between the effects of the *barA* mutation and the effects of the *relA* mutation (see also Section 1.4.4.5). It is possible that the *barA* mutation causes the depressed expression of *rpoS* in stationary phase growth, but that the *relA* mutation is responsible for the lowered *rpoS* mRNA concentrations during exponential growth. When the MC4100 *barA* mutant was grown on a minimal media, no such decrease in the stationary phase expression of *rpoS* was observed. So while the signal for BarA kinase activity has yet to be identified, it appears that the signal is associated with growth phase transition when *E. coli* is grown on rich, but not minimal, media (Mukhopadhyay *et al.*, 2000).

More difficulty was experienced when attempting to identify the response regulator in *E. coli*. BarA has been shown to donate phosphoryl groups to the response regulator UvrY (the *yecB* product) efficiently and with a higher degree of specificity versus other known response regulators (Pernestig *et al.*, 2001). Furthermore UvrY shares a high degree of protein identity (60%) with the GacA of *P. aeruginosa*. However mutations in

*yecB/uvrY* failed to modify *rpoS* expression either positively or negatively in *E. coli* (Hengge-Aronis, 2002). Therefore, it is likely that in *E. coli*, BarA is able to recognize at least one other response regulator, which then regulates *rpoS* expression.

Equivalent homologues of both GacS and GacA have been identified in *A. vinelandii* (Castañeda *et al.*, 2000). Antibiotic cassette disruptions of *gacA* resulted in the total loss of *rpoS* transcription in both exponential phase and stationary phase when growth occurred on minimal media (Castañeda *et al.*, 2001). This differs significantly from the systems discussed above, where minimal *rpoS* expression was observed at exponential phase growth, and mutations only resulted in a 2-fold reduction in *rpoS* expression during stationary phase. Furthermore, the drastic decline in *rpoS* expression in *P. fluorescens* and *E. coli* was only observed during growth on rich media. Finally, Castañeda and co-workers observed no *rpoS* transcripts in exponential phase in the wild-type strain, which may be a unique feature of *A. vinelandii* (or more likely an indication that their probe was not sensitive enough to detect low levels of *rpoS* mRNA).

#### 1.4.1.3 PsrA is a Positive Regulator of *rpoS* Transcription in Pseudomonads.

PsrA (*Pseudomonas* sigma regulator) is a 26.3 kDa protein that falls within the TetR family of DNA binding regulators. Originally identified by Tn*5* mutagenesis in *P. putida*, homologues have since been identified in *P. aeruginosa* (90% identity; Kojic and Venturi, 2001) and *A. vinelandii* (67% identity; Sandercock, unpublished). The *P. putida psrA* mutant displayed a 90% lower *rpoS* transcription rate and an associated decrease in stress resistance upon entering stationary phase. By 24 hours however the total *rpoS* expression was approximately 50% of that of wild-type cells, suggesting that regulation by PsrA is more important in early stationary phase than in late stationary phase (Kojic and Venturi, 2001). The regulatory protein recognizes the minimum palindromic sequence C/GAAACN$_{2-4}$GTTTG/C and is capable of causing supershifts (i.e. oligomerization) in gel-shift experiments (Kojic et al, 2002).

Apparently PsrA controls a relatively small regulon, involving only 7 promoters in the *P. aeruginosa* genome (Kojic *et al.*, 2002). However, since *rpoS* is one of these targets, PsrA expression has a broad regulatory effect on the physiology of the cell. PsrA

both positively and negatively regulates gene expression; the different effects are due to the placement of the binding site relative to the -10 and +1 elements of the promoter.

### 1.4.2 Translational Regulation of RpoS.

RpoS activity is also regulated at the level of translational control. Takayanagi and co-workers (1994) identified a 567 nucleotide untranslated region (UTR) at the 5' end of the *rpoS* mRNA of *E. coli*. Though the secondary structure has not been determined experimentally, a difficult task given the extreme length of the UTR, several likely structures have been proposed (Lease *et al.*, 1998; Cunning *et al.*, 1998; Hirsch and Elliott., 2005) as shown in Fig. 1.8. Most notably, the ribosome binding site (RBS) is expected to be partially occluded due to the secondary structure within the mRNA, blocking translation (Brown and Elliott, 1997). Effective translation is dependent upon a number of small RNA species and is completely dependent upon the nucleoid protein, Hfq. A summary of the regulatory web controlling RpoS translation can be seen in Fig. 1.9.

#### 1.4.2.1 Three Nucleoid Proteins that Regulate *rpoS* Translation.

Hfq (HF-I protein) is an 11.2 kDa, ribosomally associated RNA binding protein that is present at about 60,000 copies per cell in early exponential phase, dropping to about 20,000 copies per cell in stationary and late stationary phase (Azam *et al.*, 2000; Azam *et al.*, 1999). It was observed that *hfq* mutant strains have a pleiotropic phenotype that closely resembles that of *rpoS* mutants and that the Hfq protein is required for the effective translation of *rpoS* mRNA in both *Salmonella typhimurium* and *E. coli* (Brown and Elliott, 1996; Muffler *et al.*, 1996b). It has since been shown that one of the roles of Hfq protein is to bind small untranslated RNA (sRNA) species and modulate RNA:RNA interactions (Wassarman *et al.*, 2001). Hfq specifically recognizes and binds structures containing an AU rich single stranded region immediately bordered by a stem-loop structure (Brescia et al 2003). Interestingly RNase E recognizes a very similar sequence/structure (Kaberdin et al, 2000), which led to the hypothesis that the primary function of Hfq is to stabilize RNA species *in vivo* (Moll *et al.*, 2003). Indeed, Hfq protects many sRNAs from RNase E degradation *in vitro*, and the phenotype of some

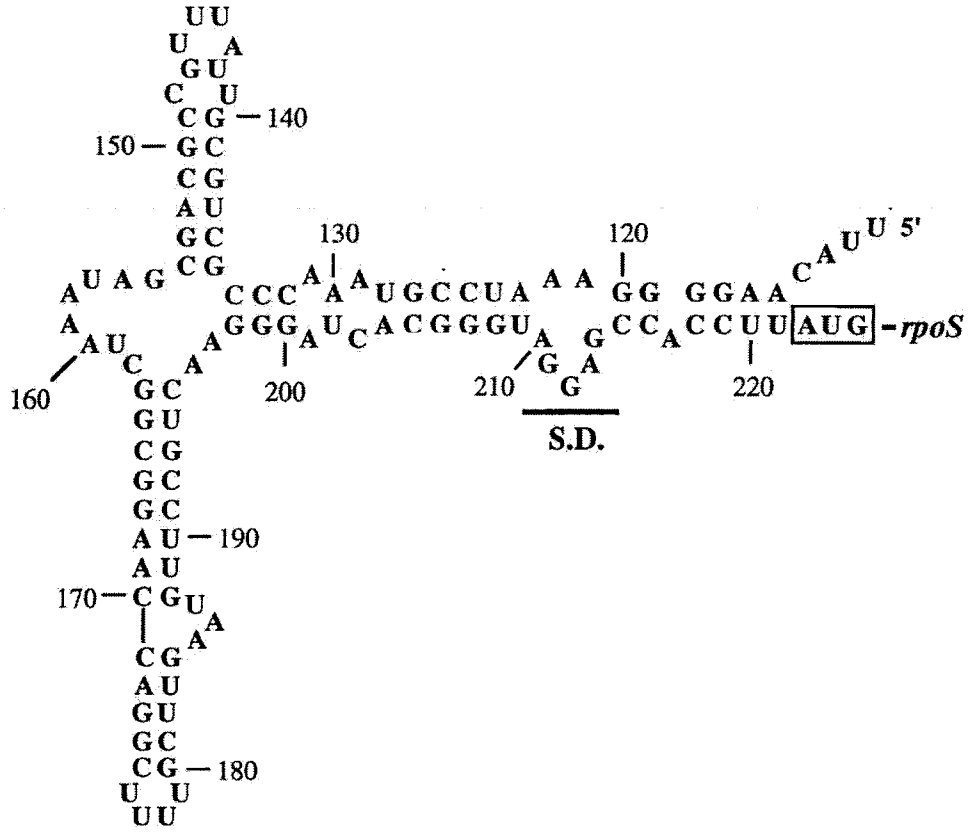



**Figure 1.8 The Proposed Secondary Structure of the 5' UTR of *rpoS* mRNA.** The Shine-Delgarno/ribosome binding site (S.D.) is proposed to be inaccessible due to secondary structure constraints, limiting the translational efficiency. The translation start site (AUG) is indicated with a box. From Cunning et al. (1998).

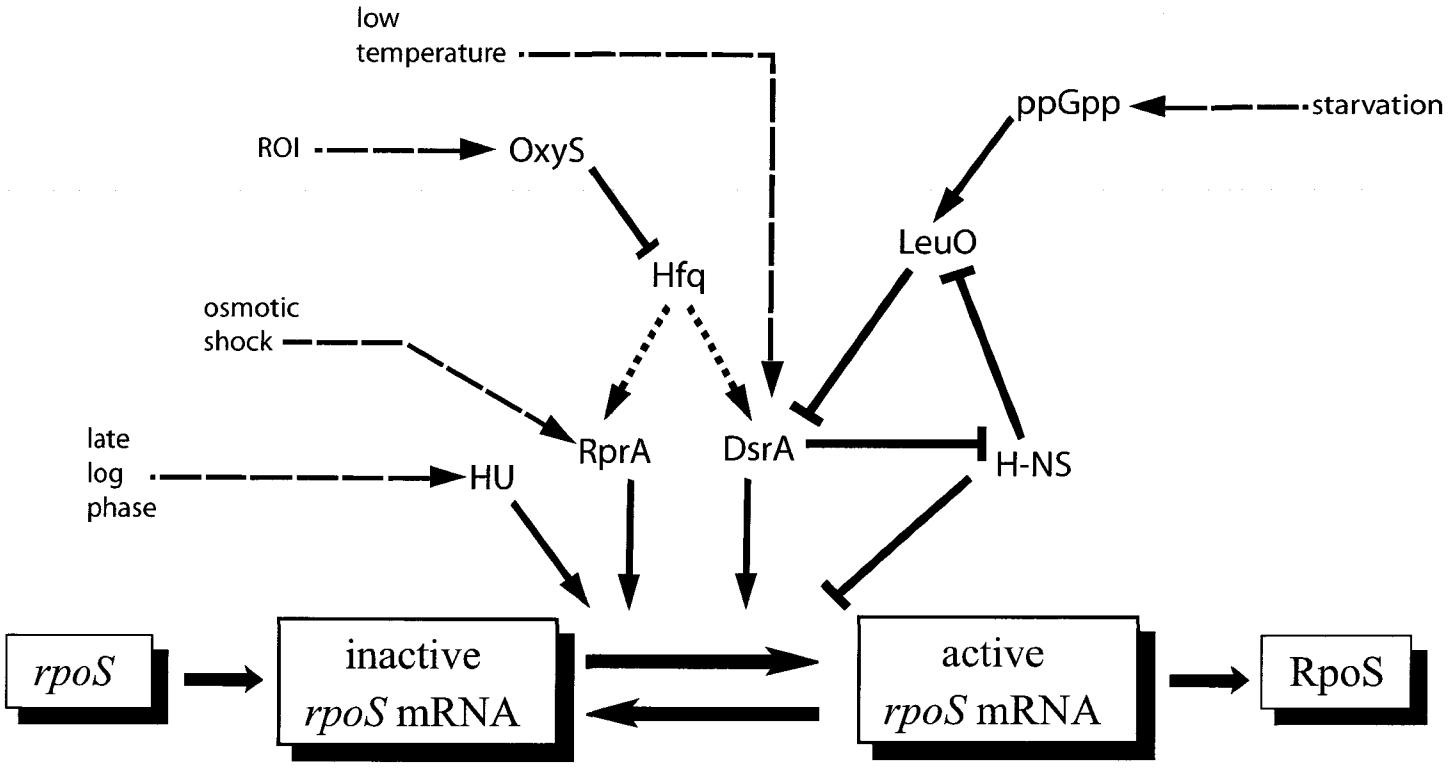



**Figure 1.9 The Regulatory Network Controlling the Transition of *rpoS* mRNA from an Inactive to Active Form.** Stress inputs are indicated by the broken lines, activation is indicated by arrow heads, and inactivation is indicated by blunted lines. Hfq acts by promoting RNA:mRNA duplex formation, which is indicated by the dotted line. Based on Hengge-Aronis (2002).

sRNA in an *hfq* mutant can be rescued by overexpression of the sRNA gene (Gottesman 2004).

Relative to the enterics, little molecular work regarding the role of Hfq in sRNA regulation has been done in the pseudomonads, primarily because few sRNA species have been identified at this time. *hfq* mutants exhibit reduced expression of the pathogenicity determinants elastase, catalase, and pyocyanin, as well as reduced expression of RpoS (Sonnleitner *et al.*, 2003). It has been suggested that the method of RpoS regulation by *Pa*Hfq may be analogous to that of the enterics, as the *P. aeruginosa* gene can complement *E. coli hfq* mutants (Sonnleitner *et al.*, 2002). More recently Hfq interactions with an archetypical Hfq/RNase E binding site on the quorum sensing sRNA, RsmY, have been demonstrated (Sorger-Domenigg *et al.* 2007), further supporting arguments for the homologous action of the pseudomonad Hfq protein.

HU protein is a heterodimeric protein composed of two subunits, HUα and HUβ. During exponential phase HU is primarily found as $HU\alpha_2$ homodimers, while in stationary phase the heterodimer HUαβ predominates (Claret and Rouviere-Yaniv, 1997). The protein is required for prolonged starvation survival, and double mutants (*hupA*$^-$ and *hupB*$^-$) accumulate less RpoS in stationary phase because of a decreased efficiency of translation (Claret and Rouviere-Yaniv, 1997; Balandina *et al.*, 2001). Unfortunately single mutants have not been studied for their effect on RpoS accumulation, so it is unclear which of the HU moieties, the heterodimer or the homodimer, is required for this effect. HU has been shown to bind the 5' UTR of *rpoS* mRNA up to 1000 times more strongly than linear dsDNA, which is comparable to the 1000-fold preference of HU for highly structured DNA (i.e. cruciform structures) over linear dsDNA (Balandina *et al.*, 2001). It is possible that HU binding results in a conformation change that makes the *rpoS* RBS more accessible. *A. vinelandii* has three putative *hup* genes (http://azotobacter.org), but no information regarding their expression is available at this time.

H-NS protein is a global regulator that controls more than 35 genes and operons (Atlung and Ingmer, 1997). Widely referred to as the "gene silencing" factor, H-NS

binds many different gene promoters, universally acting by repression. Oligomerized H-NS also plays a role in bacterial nucleoid structure and organization, sequestering large stretches of DNA. Barth and co-workers found that *hns* mutants produced exceptionally large amounts of RpoS protein, even during exponential phase (Barth *et al.*, 1995). However this effect was not exclusively due to derepression of the *rpoS* promoters. Rather H-NS regulation was imposed *via* translational regulation of *rpoS* mRNA, and by modifying the stability of the *rpoS* transcript and protein (Barth *et al.*, 1995; Brescia *et al.*, 2004; Yamashino *et al.*, 1995).

Originally considered a strict DNA binding protein, H-NS has since been shown to bind both *rpoS* and DsrA RNA *in vitro* and *in vivo* with high affinity (Brescia *et al.*, 2004). Following binding, H-NS causes these target species to refold, increasing their susceptibility to degradation by RNase I. As such, the mechanism by which H-NS modulates *rpoS* mRNA stability has been fully described. The mechanism by which H-NS regulates RpoS protein stability is still unclear, but involves the modulation of RssB activity (but not RssB expression; Zhou and Gottesman, 2006).

Though it is uncertain at this time how H-NS modulates the efficiency of RpoS translation, two mechanisms are suggested. First, it is possible that when H-NS binds *rpoS* mRNA it occludes or refolds the 5'UTR such that sRNA cannot bind and activate translation. Second, some experimental evidence suggests that H-NS translational control may act through gene silencing of specific activators, such as HU or Hfq proteins (Muffler *et al.*, 1996b; Hengge-Aronis, 2002).

#### 1.4.2.2 Several sRNA Species Regulate *rpoS* Translation.

Quite recently an entirely new level of regulation has been uncovered in bacteria involving small untranslated RNA (sRNA, also referred to as non-coding RNA). As of 2004, the expression of over 50 sRNA species had been confirmed in *E. coli*, though many more had been predicted by computational means (Gottesman 2004). Small scale studies indicate that a very high proportion of these regulate *rpoS* mRNA translation: sRNA that positively regulate *rpoS* translation include DsrA, RprA, RybB, and RyhA, whereas OxyS, RydB, RyeE, and RyhB negatively regulate *rpoS* translation (Sledjeski and Gottesman, 1995, Argaman *et al.*, 2001, Altuvia *et al.*, 1997, Wassarman *et al.*,

2001). At this time sRNA have been studied in only a few other bacteria. Initial research efforts in *P. aeruginosa* uncovered two sRNA, PrrF1 and PrrF2, both of which were involved with iron regulation (Wilderman *et al.*, 2004). Recent computational efforts identified 36 additional sRNA candidates in *P. aeruginosa*, of which 17 were confirmed by Northern analysis (Livny *et al.*, 2006). However their functions, and so the possibility that any are involved in *rpoS* regulation, have not yet been determined. Considering the importance of Hfq in pseudomonas RpoS expression (Sonnleitner *et al.*, 2003), it seems likely that at least one *P. aeruginosa* sRNA is involved with *rpoS* regulation. The most fully described sRNAs from *E. coli* and their role in *rpoS* regulation are discussed below.

DsrA is an 85 nucleotide (nt) untranslated RNA that plays a negative regulatory role in the accumulation of H-NS protein, and a positive role in the translation of RpoS (Sledjeski and Gottesman, 1995; Sledjeski *et al.*, 1996). DsrA is a required regulator for translation of *rpoS* mRNA in either exponential or stationary phase at 25 °C, but is not required for translation at 37 °C (Repoila and Gottesman, 2001). It is not surprising to see that DsrA is itself temperature regulated. The rate of transcription of DsrA is 25-fold higher at 25 °C than at 37° C, and the half-life of the transcript increases from 4 min to 23 min following temperature shifts from 37 °C to 25 °C (Repoila and Gottesman, 2001). As a result, DsrA increases the efficiency of translation of RpoS at temperatures at 25 °C, but has little effect upon translation at temperatures at or above 37 °C.

The structure of the RNA, shown in Fig. 1.10A, includes 3 stem-loops with a 12 nt A/U rich single stranded spacer located between stem-loop I and stem-loop II (Worhunsky *et al.*, 2003). DsrA interacts with the *rpoS* 5' UTR directly, forming RNA:RNA duplexes. Deletion analysis of the stem-loop and spacer structures indicated that stem-loop I is the most important determinant for RpoS translation. It is believed that this stem-loop structure binds the complementary sequence in the *rpoS* 5' UTR, causing a conformational change within the UTR, thereby making the RBS more accessible to ribosomes (Majdalani *et al.*, 1998). Gel shift experiments with Hfq have identified the poly A/U spacer and stem-loop II as being the site of binding. However, no unwinding of the stem-loops was observed by circular dichroism following binding (Brescia *et al.*, 2003) which is somewhat unusual for proteins that mediate RNA:RNA

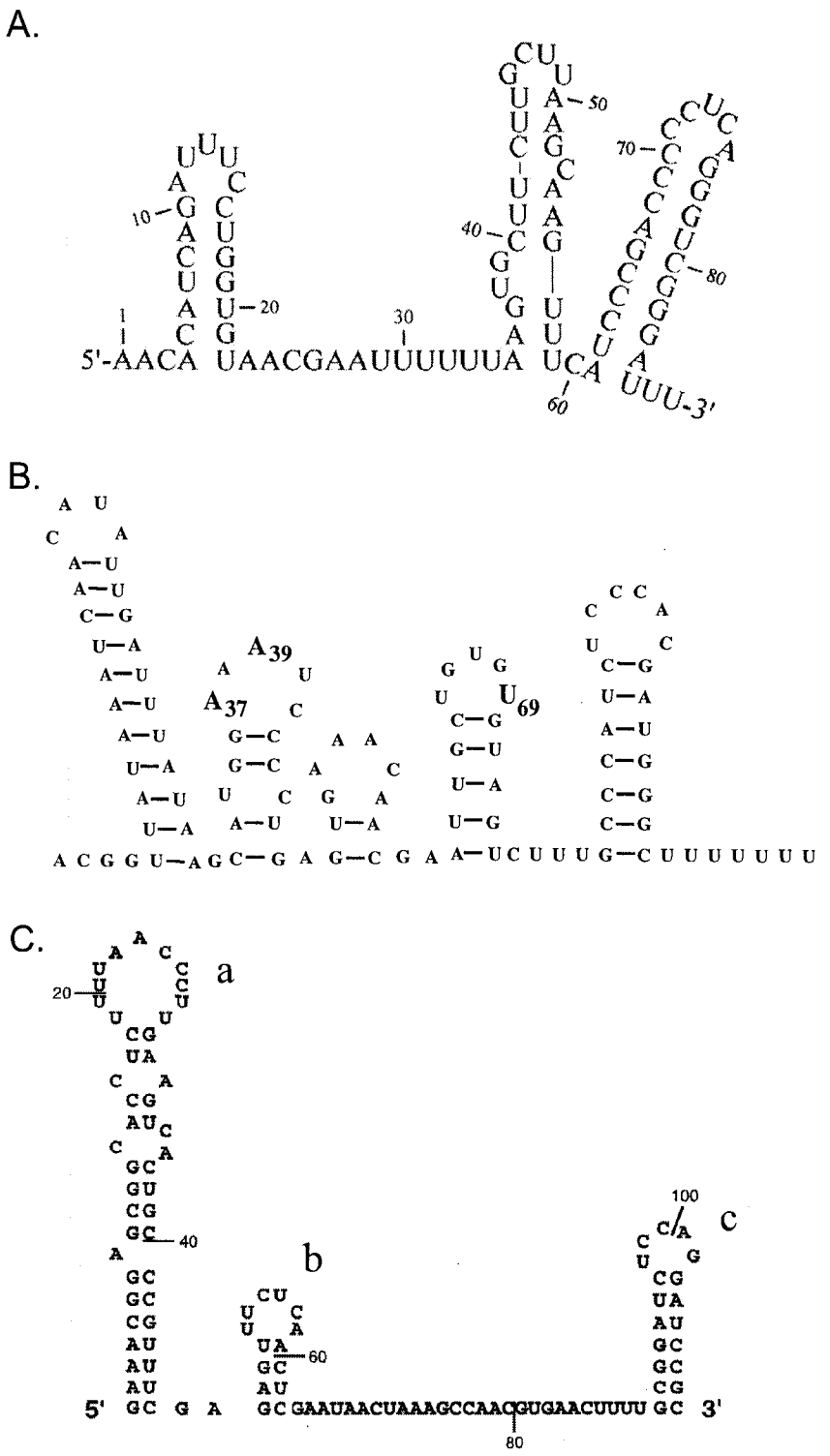



**Figure 1.10 Proposed Structures of sRNA Species Known to Modulate rpoS mRNA Translation Efficiency.** DsrA, RprA, and OxyS are shown in panels A, B, and C respectively. Stem-loops a through c are shown for OxyS.

interactions.

RprA is a 105 nt untranslated RNA (Argaman *et al.*, 2001) with a highly structured putative secondary structure (Fig. 1.10B; Majdalani *et al.*, 2002). Concentrations of this product are approximately 8-fold higher in late stationary phase than in mid-exponential phase (Majdalani *et al.*, 2002). Majdalani and co-workers showed that rpoS translation under osmotic stress is dependent upon RprA in *dsrA* mutants (Majdalani *et al.*, 2001) suggesting that the primary role of RprA is to induce the general stress response under conditions of osmotic challenge. Expression of the RNA is regulated by the two-component signal transduction system, RcsC/B, which also regulates capsule synthesis, and *osmC* and *ftsZ* expression. RcsC has been found to regulate a number of promoters in response to assaults on the cell surface (Parker *et al.*, 1992; Kelley and Georgopoulous, 1997) suggesting that RprA may be part of a stress response network that includes not only osmotic shock, but also general surface shock. The predicted RprA:*rpoS* mRNA duplex structure is believed to "open" the *rpoS* 5' UTR structure in a manner analogous to that of DsrA (Majdalani *et al.*, 2002). Hfq is known to bind RprA, and is required for RprA activity on *rpoS* mRNA (Wassarman *et al.*, 2001).

OxyS, a 109 nt RNA product with 3 stem-loop structures (Fig. 1.10C; Altuvia *et al.*, 1997), is known to play a regulatory role in the expression of approximately 40 different genes in *E. coli* in response to oxidative stress (Altuvia *et al.*, 1998). OxyS is known to inhibit translation from the *flhA* mRNA by forming an RNA:RNA duplex that occludes the RBS (Altuvia *et al.*, 1998). It is also known to inhibit RpoS translation in an Hfq dependent manner (Zhang *et al.*, 1998). Gel shift assays indicated that OxyS binds Hfq in two complexes; complex I has an apparent stoicheiometric ratio of 1 OxyS RNA to 1 Hfq multimer whereas complex II has a ratio of 1 OxyS: 2 Hfq multimers (Zhang *et al.*, 2002). Subsequent to Hfq binding, stem-loop b is unwound and the protein rests over the single stranded spacer region. Since Hfq protein is absolutely required for OxyS activity against RpoS translation, it is possible that the stem-loop unwinding event results in a refolding of OxyS such that it can bind complementary regions within the *rpoS* mRNA (Zhang, *et al.*, 2002). Although it is clear that RNA duplex formation occurs between

OxyS and the *flhA* target RNA, the exact nature of a putative OxyS:*rpoS* mRNA complex is not known at this time. It is possible that, as with *flhA* mRNA, OxyS binds the *rpoS* mRNA in such a manner that the RBS is directly occluded. Alternatively OxyS may bind elsewhere within the *rpoS* mRNA 5' UTR, causing a conformational change that blocks activation by DsrA and/or RprA. It has also been proposed that high concentrations of OxyS simply acts to titrate Hfq, temporarily blocking sRNA activation of the *rpoS* mRNA.

A question remains, why would OxyS, a regulator of oxidative stress, act to reduce the accumulation of the general stress response protein, RpoS? It is important to remember that expression of the RpoS regulon represents a significant developmental commitment on the part of the cell (Ruiz *et al.*, 2001; Patten *et al.*, 2004). However upon oxygen stress, a minimal number of stressors ($H_2O_2$, $\bullet O_2^-$) are present, which can be readily removed by a limited number of protein products (catalases, peroxidases, superoxide dismutases). As a result a comprehensive general stress response is unnecessary. That said, following oxygen stress, RpoS concentrations do increase, due to a concomitant decrease in the rate of protein turnover. This additional level of regulation is believed to be due to the increase in ppGpp concentrations following $H_2O_2$ stress, though the mechanism remains unclear.

#### 1.4.2.3 LeuO Protein Decreases the Translational Efficiency of *rpoS* mRNA *via* DsrA Repression.

LeuO protein, a member of the LysR-like regulators causes a decrease in RpoS translation when overexpressed, especially at lower temperatures (Klauck *et al.*, 1997). This hinted that perhaps LeuO acts by antagonizing the positive regulatory effect of the temperature responsive sRNA, DsrA. The decrease in RpoS translation was shown to be entirely DsrA dependent, as *dsrA* mutants displayed no greater decreases in RpoS translation under conditions of wild-type level or overexpressed LeuO (Klauck *et al.*, 1997). It is possible that LeuO expression helps the cell avoid run-away activation of RpoS translation by decreasing the level of *dsrA* transcription at lower temperatures.

Both H-NS and ppGpp regulate *leuO* transcription. In growing cells, H-NS acts as a repressor of *leuO* transcription, nearly silencing expression (Klauck *et al.*, 1997).

However, as cells enter stationary phase, H-NS concentrations decrease moderately (Azam *et al.*, 1999) allowing ppGpp to induce *leuO* expression (Fang *et al.*, 2000). The increased LeuO concentration upon entry into stationary phase causes a decrease in the amount of DsrA RNA within the cell, and so acts to reduce the efficiency of RpoS translation. This may act as a brake on the amount of RpoS produced in stationary phase, helping the cell to avoid unchecked production of the sigma factor in a manner similar to that seen with OxyS during oxidative stress.

### 1.4.3 Post Translational Regulation of RpoS.

RpoS is also regulated at the level of protein stability in the enteric bacteria. During exponential phase the protein has a half-life of about 1.5 min (Muffler *et al.*, 1997) but this can increase substantially in response to various stresses such as nutrient starvation, high temperature, high osmolarity or low pH (Muffler *et al.*, 1996a and references within). Proteolysis of RpoS is mediated by the serine protease ClpXP and the chaperone RssB, which binds $\sigma^S$ and directs it to the protease. While RssB is regulated to some degree by stress responses (particularily carbon-source starvation), ClpXP concentrations and activities do not fluctuate significantly. The most recent evidence suggests that the RssB/ClpXP system primarily regulates RpoS in response to carbon limitation, as stability changes are less extreme during nitrogen or phosphate limitation (Mandel and Silhavy, 2005). The novel "anti-adaptor" protein, IraP, protects RpoS from RssB-mediated degradation and is likely responsible for the moderately improved stability of RpoS during nitrogen and phosphate limitation (Bougdour *et al.*, 2006).

#### 1.4.3.1 ClpXP and the Chaperone RssB Regulate the Degradation of RpoS.

ClpXP is a serine protease that contains two protein complexes. ClpP is found as a hexameric complex and functions as a protease. Specificity of the protease is driven by its association with an accessory chaperone protein, such as ClpA or ClpX, which mediate target binding. The accessory chaperone then unfolds and feeds the substrate protein into the ClpP portion of the complex for degradation (Singh *et al.*, 2000). It was found that the ClpXP protease is responsible for RpoS degradation during the exponential growth phase, but that proteolysis decreased greatly during nutrient limiting conditions

(Schweder *et al.*, 1996). The rapidity of the response suggested that ClpXP activity was modified upon nutrient limitation (rather than *clpX* or *clpP* gene regulation). However ClpXP activity levels were unaffected by changes in growth phase (Zhou and Gottesman, 1998), suggesting the presence of an additional factor.

Several groups independently identified a response regulator, known variably as SprE, MviA, ExpM and RssB, which regulates RpoS stabilities in a ClpXP dependent manner (Pratt and Silhavy, 1996; Bearson et al, 1996; Andersson *et al.*, 1999; Muffler *et al.*, 1996a). Two RpoS sites are critical for RssB-mediated proteolytic regulation. The first is the RssB recognition domain centred around K173 which overlaps the portion of RpoS responsible for promoter recognition at the -10 hexamer (Becker *et al.*, 1999; Stüdemann *et al.*, 2003). K173E mutants are not bound by RssB, and are protease resistant (Becker *et al.*, 1999). The second RpoS site, responsible for ClpXP recognition, is located within the N-terminus between amino acids 7 and 35. ClpXP is only able to utilize this site if RssB is bound to the sigma factor (Stüdemann *et al.*, 2003) which suggests that RssB causes some degree of conformational change in RpoS following binding. Interestingly RssB is not degraded by ClpXP, but is instead recycled to the cytoplasm, potentially binding free RpoS to start the process again.

Some controversy regarding the importance of RssB phosphorylation exists (Peterson *et al.*, 2004; Mika and Hengge, 2005). Initial reports indicated that high functioning RssB was phosphorylated, and had a much higher binding efficiency for RpoS than did the unphosphorylated form (Bouche *et al.*, 1998). The *in vivo* phosphoryl donor has not been identified to date, but one candidate, acetyl~P, has been shown to donate a phosphate group to residue Asp58 of the RssB protein *in vitro* (Bouche *et al.*, 1998). It is likely that during logarithmic growth, energy conditions within the cell are very high, and so high-energy phosphate donors such as acetyl~P are more available. As a result more RssB~P is available to bind and direct RpoS to the ClpXP complex. When the cell is at a lower energy state however (i.e. stationary phase), less RssB~P is likely to be formed, reducing the amount of RpoS degradation.

The *rssB* gene is a member of the *rssA-rssB* operon. Two promoters exist for the transcription of *rssB*: the polycistronic *rssA* promoter at the front of the operon, and an *rssB* specific promoter, found within the intergenic region separating *rssA* and *rssB* (Ruiz

*et al.*, 2001). Though the conditions of the *rssA* expression are not clearly defined, it was found that the *rssB* promoter conforms to the consensus sequence recognized by $\sigma^S$, and that *rpoS* mutants produce substantially less RssB than do wild-type cells (Ruiz *et al.*, 2001). Ruiz and coworkers concluded that RssB is regulated by RpoS *via* the *rssB* promoter, and possibly by unknown factors *via* the *rssA* promoter. Although an RssA homologue has been identified in several pseudomonads, no RssB homologues have been found. Nor has any evidence of proteolytic regulation of RpoS been demonstrated in these organisms (Bertani *et al.*, 2003).

#### 1.4.3.2 DnaK: The Heat Shock Response and RpoS Stability.

The protein chaperone DnaK is involved in the disaggregation of mis-folded protein complexes during heat shock in bacteria (Hengge and Bukau, 2003). It was originally observed that *dnaK* mutants have pleiotropic phenotypes, which included elevated $H_2O_2$ sensitivity and reduced starvation survival (Rockabrand *et al.*, 1995), so was likely to play a role in the general stress response. Further work showed that RpoS concentrations were induced 3 to 5-fold in exponentially growing cells when shifted from 30 °C to 42 °C, and that the half life of the sigma factor increased nearly 5-fold (Muffler *et al.*, 1997). RpoS levels did not increase during heat shock in the *dnaK* mutant, suggesting that DnaK is absolutely required for RpoS induction during thermal up-shifts. Furthermore RpoS levels were slightly elevated in the mutant during exponential phase, but were only induced 2-fold rather than 20-fold when the cells entered stationary phase. This suggested that the trace amounts of DnaK present in non-thermally stressed cells played an important role in RpoS regulation, even when the cells encountered other stresses (Muffler *et al.*, 1997). Studies using *rssB* and *clpP* mutants indicated a link between DnaK and RssB function, as these mutations completely abolished the DnaK phenotype (as it relates to RpoS stability, but not as it relates to protein disaggregation during heat stress; Muffler *et al.*, 1997).

### 1.4.4 Control of RpoS Activity.

Finally, the activity of $\sigma^S$ is regulated at the level of association with free RNA polymerase core enzyme (E) and at the level of function (initiation of transcription by the

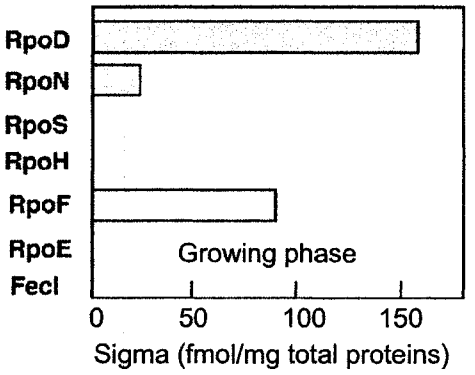
$E\sigma^S$ holoenzyme). Two major dynamics regulate the efficiency of $E\sigma^S$ formation; sigma factor competition and anti-sigma interference. However the presence of $E\sigma^S$ within the cell does not guarantee the high expression of the $\sigma^S$ dependent genes, as several small molecules (trehalose, glutamate, polyphosphate, and ppGpp) regulate the efficiency of the $E\sigma^S$ complex at the target promoters.

### 1.4.4.1 Sigma Factor Competition.

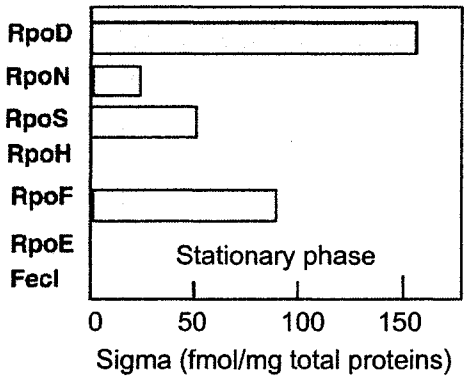
Sigma factor switching was first shown to occur during infection by the bacteriophage T4. The phage enzyme factor gp55 was shown *in vivo* to out-compete $\sigma^{70}$ for available E, resulting in a decreased rate of transcription from *E. coli* housekeeping genes, and an increase in the amount of viral transcription (Malik *et al.*, 1987). However *in vitro* $\sigma^{70}$ demonstrated a higher affinity for RNAP than did gp55, successfully out-competing gp55 for available core enzyme. The authors suggested that another factor or factors besides sigma affinity for E must be involved in regulating $E\sigma$ formation (Malik *et al.*, 1987).

More recently sigma factor competition has been demonstrated between $\sigma^{70}$ and $\sigma^S$. During stationary phase, *rpoS* mutants were observed to express stationary phase induced genes much less than in wild-type cells. However it was also observed that there was a coincidental increase in the level of expression at $\sigma^{70}$-dependent genes in the *rpoS* mutants during stationary phase (Farewell *et al.*, 1998). Similarly, overproduction of $\sigma^{70}$ in a wild-type background resulted in a very similar phenotype; decreased expression of $\sigma^S$-dependent genes and an increase in the expression of $\sigma^{70}$-dependent genes during stationary phase (Farewell *et al.*, 1998). Such competition for available core RNAP is not surprising, as there are only about 2000 RNAP molecules available per cell during exponential growth in *E. coli.* Of these, roughly 70% are transcribing genes, so are not available for sigma factors, leaving approximately 30% or 600 RNAP molecules available for sigma factors (Ishihama, 2000). However in balanced, exponentially growing *E. coli* there are approximately 1200 sigma factor molecules ($\sigma^{70}$, $\sigma^N$, $\sigma^S$, $\sigma^H$, $\sigma^F$, $\sigma^E$ and $\sigma^{FecI}$) available to bind these free RNAP (Fig. 1.11A; Ishihama, 2000) 60% of which are $\sigma^{70}$ (700 molecules; Jishage and Ishihama, 1995; Jishage *et al.*, 1996; Maeda *et*

A.




B.




C.

| Sigma Factor | $K_d$ |
|---|---|
| RpoD, $\sigma^D$ | 0.26 |
| RpoN, $\sigma^N$ | 0.30 |
| RpoS, $\sigma^S$ | 4.26 |
| RpoH, $\sigma^H$ | 1.24 |
| RpoF, $\sigma^F$ | 0.74 |
| RpoE, $\sigma^E$ | 2.43 |
| FecI, $\sigma^{FecI}$ | 1.73 |

**Figure 1.11 Relative Quantities of the Seven Sigma Factors of *E. coli* During Exponential (A) and Stationary Phase Growth (B).** The dissociation constant ($K_d$) of each sigma factor, determined for free RNAP *in vitro*, is indicated in panel C. From Ishihama (2002) and Maeda et al. (2000).

*al.*, 2000). It has been observed that several organisms express many more ECF sigma factors than *E. coli*, so competition for available RNAP in these organisms may be more intense. As growth conditions in *E. coli* change, the amount of $\sigma^S$ increases from undetectable levels to approximately 16% (~230 molecules) of the total sigma factor in the cell (Fig. 1.11B; values estimated from Ishihama, 2000).

Though it would be convenient to suggest that sigma factor usage is directly proportional to the relative concentration of the σ factor in the cell, two additional factors also mediate sigma factor competition: affinity of σ factors for the available RNAP and sequestration of sigma factors by anti-sigma factors. The relative dissociation constants ($K_d$) of the seven major sigma factors of *E. coli*, determined *in vitro* using (presumably) exponential phase derived proteins, are shown in Fig. 1.11C. Most significantly $\sigma^{70}$ has the lowest dissociation constant, with a $K_d$ of 0.26 nM, while RpoS has the highest $K_d$ of 4.26 nM. This would suggest that the amount of $E\sigma^S$ in stationary phase cells would be far less than 16% of the total holoenzyme, while the amount of $E\sigma^{70}$ should exceed 70% of the total Eσ within the cell (Ishihama, 2000). A nuanced discussion of these factors and their effect on Eσ partitioning is discussed below.

### 1.4.4.2 Anti-Sigma Factors and Sigma Factor Competition.

Several sigma factors have associated anti-sigma factors that bind and store the σ in an inactive form. As a result the estimates of sigma availability, as reported above (Maeda *et al.*, 2000; Ishihama 2000), are likely to be somewhat inaccurate as a fraction of the total σ in the cell is likely to be in a sequestered state. Most notably, the highly expressed $\sigma^{70}$ can be trapped by the anti-sigma factor Rsd. The flagellar sigma factor $\sigma^F$ is constitutively expressed throughout all growth phases, and can be sequestered by the anti-sigma factor FlgM.

Rsd is not detectable in exponentially growing *E. coli*, allowing nearly all of the $\sigma^{70}$ present to bind free RNAP core enzyme (Jishage and Ishihama, 1999). During stationary phase elevated levels of Rsd are expressed, which sequester a portion of the functional $\sigma^{70}$ causing a decline in the amount of $E\sigma^{70}$ directed transcription. As with most anti-sigma factors, Rsd binds $\sigma^{70}$ within region 4. Because region 4 varies significantly

between sigma factors, most anti-sigma factors are highly specific for their targets (Ishihama, 2000).

Sigma F, responsible for the transcription of the flagellar genes (but not the basal hook assembly) is expressed in large quantities in a constitutive fashion in *E. coli* (Kutsukake and Iino, 1994) so could potentially compete for RNAP. Large quantities of the anti-sigma factor, FlgM, are produced in the cell however, severely limiting the amount of unsequestered $\sigma^F$ in the cell. Only when flagella are initially constructed, or following shearing of a flagellum from the basal body, does the FlgM concentration drop in the cell (it is exported from the cell through a specific mechanism associated with flagellar basal bodies; Hughes *et al.*, 1993) allowing $\sigma^F$ to bind RNAP core enzymes. So despite the fact that sigma F accounts for up to 28% of the $\sigma$ in stationary phase cells, and has a relatively low $K_d$ *in vitro* (0.74 nM; Ishihama, 2000), it is unlikely to interfere with the formation of $E\sigma^S$ under normal stationary conditions.

Similarly, the ECF are unlikely to interfere with $E\sigma^S$ formation, as they are typically sequestered by their membrane bound anti-sigma factors unless an appropriate stress is experienced.

#### 1.4.4.3 The Effect of Small Molecules on $E\sigma^S$ Activity.

Trehalose, an osmoprotectant disaccharide that accumulates in stationary phase, has been shown to affect the ability of RNAP holoenzyme to initiate transcription at promoters. Maximal $E\sigma^{70}$ activity has been observed at trehalose concentrations of 0.5 M, whereas maximal $E\sigma^S$ transcriptional activity has been seen at concentrations of 1.0 M (Kusano and Ishihama, 1997a). These concentrations are representative of normal trehalose concentrations in moderately osmotically shocked cells and severely osmotically shocked (or stationary phase) cells respectively (Cayley *et al.*, 1991; Welsh *et al.*, 1991). Trehalose was also shown to affect the efficiency of $E\sigma^S$ formation by about 40 percent *in vitro*, suggesting that trehalose may effect $\sigma^S$ efficiency by two methods *in vivo* (Kusano and Ishihama, 1997a).

Polyphosphate is widely considered a storage molecule for inorganic phosphate (Pi), accumulated most often when cells experience stress. Tens to hundreds of Pi molecules have been seen to form polyphosphate strands, polymerized by the *ppk* product,

polyphosphate kinase. Mutants of *ppk* are unable to form polyphosphate, and have a drastically lower survival rate after a few days in stationary phase (Kornberg *et al.*, 1999). Biochemical tests show that polyphosphates are able to irreversibly associate with Eσ complexes, reducing the transcriptional efficiency of all RNAP holoenzymes at low salt concentrations. However at high salt concentrations and in the presence of the osmoprotectant molecule glutamate, $E\sigma^S$ activity is restored, while $E\sigma^{70}$ activity is not (Kusano and Ishihama, 1997b). The region of RpoS responsible for this response has been mapped to the C-terminal 16 amino acids, which has been designated the CTE (C-terminal extension; Ohnuma et al, 2000). The authors also demonstrated that the CTE is required for improved transcriptional activity by $E\sigma^S$ under high salt conditions, but is not required for the formation of the $E\sigma^S$ holoenzyme (Ohnuma *et al.*, 2000). *ppk* mutants generated in *P. aeruginosa* displayed a slight increase in RpoS levels at high cell densities (Bertani et al, 2003), but the effect upon Eσ activities has not yet been studied.

Another osmoprotectant, glutamate, has also been shown to modify Eσ activity at promoters. Concentrations above 0.5 M have been shown to completely inactivate transcription by $E\sigma^{70}$ but only moderately reduce $E\sigma^S$ *in vitro* (Ding *et al.*, 1995). However all tests were done *in vitro*, and some controversy exists regarding the results (see Hengge-Aronis, 2002). Perhaps more importantly, glutamate works synergistically with the small molecules trehalose and polyphosphate. Kusano and Ishihama (1997a) demonstrated that increasing concentrations of trehalose exacerbated the inhibition of $E\sigma^{70}$ by 0.3 M glutamate, but alleviated the mild inhibitory effects on $E\sigma^S$ in the presence of glutamate. Additionally, glutamate was shown to reduce the inhibition of $E\sigma^S$ by polyphosphate under high salt conditions, but not to alleviate the inhibition of $E\sigma^{70}$ under these same conditions (Kusano and Ishihama, 1997b).

#### 1.4.4.4 The Stringent Response Affects the Formation and Activity of $E\sigma^S$.

The stringent response occurs when cells become starved for carbon-source, or when ribosomes stall during translation due to a paucity of amino charged tRNA (often due to nitrogen-source limitation). The result is an accumulation of the small regulator molecule guanosine tetraphosphate (ppGpp), and to a lesser degree guanosine pentaphosphate (pppGpp), by the action of the proteins RelA and SpoT. *relA spoT*

double mutants, unable to make ppGpp (phenotypically designated ppGpp$^{O}$), were found to produce significantly less RpoS than the parental strain (Xiao *et al.*, 1991). Conversely over production of ppGpp caused an increase in the amount of RpoS in both exponentially growing cells and those in stationary phase (Gentry *et al.*, 1993). It is now known that ppGpp affects RpoS levels and activities at at least four levels: through transcriptional and post-transcriptional regulation of the *rpoS* gene, as well as affecting the efficiencies of Eσ formation and by modulating Eσ activity at promoters.

It was shown that ppGpp regulates *rpoS* transcription since ppGpp$^{O}$ mutants produced no *rpoS* mRNA during exponential phase and substantially less during early stationary phase (Gentry *et al.*, 1993; Lange *et al.*, 1995; Zgurskaya *et al.*, 1997). This indicates that *rpoS* expression is dependent upon ppGpp during rapid growth, but that other factors can partially rescue expression during stationary phase. The discovery poses a significant problem for the study of the general stress response in *E. coli*, as much of the work in the field employed (and continues to use) the *relA* strain, MC4100. The mechanism by which ppGpp regulates *rpoS* transcription has not been fully described at this time, though it may have to do with transcriptional efficiencies of Eσ at the *rpoS* promoter.

Several communications highlight the inhibitory effects of ppGpp on the efficiency of Eσ formation. It has been noted that ppGpp can bind $\sigma^{70}$ at region 3.2 (Reddy *et al.*, 1995; Cashel *et al.*, 2003) and the resultant $\sigma^{70}$ppGpp complex is less able to bind available E (Jishage *et al.*, 2002). Consequently $\sigma^{H}$, $\sigma^{N}$ and $\sigma^{S}$ compete more efficiently for RNAP core enzyme in the presence of ppGpp, than in its absence (Jishage *et al.*, 2002; Laurie *et al.*, 2003). The ECF sigma factor, $\sigma^{E}$, is also activated by elevated ppGpp levels independent of its normal signal, envelope stress (Costanzo and Ades, 2006), skewing sigma competition in the presence of ppGpp even further from $\sigma^{70}$. The binding sites of ppGpp for various sigma factors have not yet been determined.

It is interesting to note that ppGpp can also regulate gene expression independently of σ factor binding, as the β subunit of the RNAP core enzyme is also a ppGpp target (Zhou and Jin, 1998; Barker et al, 2001a; Barker et al, 2001b). The resulting complex EppGpp:$\sigma^{70}$ has a higher dissociation constant at many promoters. This acts to decrease waste, as stringent promoters (i.e. rRNA promoters) typically make frequent abortive transcripts before successfully escaping to make full length RNA products. As a result of

ppGpp modification the holoenzyme transcribes RNA less often and produces fewer abortive products at stringent promoters (Barker *et al.*, 2001a; Barker *et al.*, 2001b). Conversely many $\sigma^S$ dependent genes absolutely require the presence of a ppGpp modified core polymerases for expression (Kvint *et al.*, 2000). An additional protein factor, DksA, acts to amplify the ppGpp effect on $E\sigma^{70}$ activity. The observation that DksA has its most profound effect during periods of low ppGpp concentration, combined with structural work that suggests DksA binds next to ppGpp near the RNAP active site (Perederina *et al.*, 2004) has led to a model in which DksA stabilizes ppGpp-RNAP interactions, thereby modifying $E\sigma^{70}$ activities even when ppGpp levels are low (Bernardo *et al.*, 2006). A role for DksA in $E\sigma^S$ activity has not yet been studied, but a positive DksA-linked effect on $E\sigma^N$ activity has been observed. However the authors suggest that the $E\sigma^N$ effect is actually due to the suppression of $E\sigma^{70}$ activities in the presence of DksA (Bernardo *et al.*, 2006). Similar ppGpp-dependent DksA effects have also been observed in *P. aeruginosa* (Perron *et al.*, 2005).

Finally, ppGpp also modifies the expression of the $\sigma^{70}$ anti-sigma factor, Rsd. Jishage and Ishihama demonstrated that over expression of ppGpp resulted in a coincidental increase in the amount of Rsd, which encouraged the formation of $E\sigma^S$ rather than $E\sigma^{70}$ when cells experienced the stringent response (Jishage and Ishihama, 1999).

Relatively little work has been published regarding the effect of ppGpp on the stress responses of pseudomonads. Furthermore, no efforts to examine the regulatory roles of the stringent response in diazotrophic organisms during nitrogen limitation and subsequent fixation have been made. At this time it is known that the *relA* mutation is sufficient to abolish ppGpp production in *P. aeruginosa* (Erickson *et al.*, 2004) despite the presence of a *spoT* homologue. The $ppGpp^O$ phenotype results in significantly lower RpoS levels, reduced virulence, reduced antibiotic resistance, and modified quorum sensing in *P. aeruginosa* (Erickson *et al.*, 2004; Viducic *et al.*, 2006; van Delden *et al.*, 2001).

### 1.5 Responses to Reactive Oxygen Species.

Aerobic metabolism allows for the efficient capture of energy during nutrient oxidation (due to the large potential between the $NAD^+ + 2H^+/NADH + H^+$ and the $2H^+ +$

$½O_2/H_2O$ half-reactions) but also results in the production of a large number of damaging reactive oxygen intermediates. Of particular biological interest are superoxide ($\cdot O_2^-$), hydrogen peroxide ($H_2O_2$) and the hydroxyl radical ($\cdot OH^-$).

The primary metabolic generator of superoxide radicals is NADH dehydrogenase II (Messner and Imlay, 1999), though the degree to which $O_2$ will be radicalized differs between Kingdoms (and presumably even at the species level) due to the efficiency of the enzyme; only 0.1% of oxygen consumed by *E. coli* becomes superoxide (Imlay and Fridovich, 1991), whereas mitochondria are less efficient, allowing 2-4% of $O_2$ to be converted to $\cdot O_2^-$ (Scandalios *et al.*, 1997). Because of its greater overall respiration rate however, *E. coli* produces more ROS in total (Imlay and Fridovich, 1991). It is possible that this is even more true of *A. vinelandii* as its respiration rate far exceeds that of *E. coli* (Kelly *et al.*, 1990). Superoxide is dismutated by superoxide dismutases (SODs) to hydrogen peroxide, which is in turn reduced by peroxidases to $H_2O$, or disproportionated by catalases to $H_2O$ and $O_2$.

Superoxide radicals that are not removed can react with free iron ($Fe^{3+}$), thereby forming the hydroxyl radical:

$$\cdot O_2^- + Fe^{3+} \rightarrow O_2 + Fe^{2+}$$
$$H_2O_2 + Fe^{2+} \rightarrow Fe^{3+} + OH^- + \cdot OH$$

This radical is the most reactive of the toxic oxidizing agents (Neilands, 1981), and is doubly dangerous because, unlike superoxide and hydrogen peroxide, no enzyme exists to fully eliminate it. Only two preventative mechanisms exist to mitigate hydroxyl radical accumulation: the expression of SODs and the sequestration of free iron by ferritin-type proteins.

Because of their importance to cell survival, most bacteria produce a set of constitutively expressed anti-oxidant proteins during aerobic growth. During periods of elevated oxidative stress, many organisms also produce a supplemental set of enzymes, most frequently regulated by the proteins SoxRS and OxyR, which measure superoxide or $H_2O_2$ levels directly (reviewed in Storz and Imlay, 1999). Additional anti-oxidant proteins can be induced during the general stress response (reviewed in Chelikani *et al.*,

2004) following the transition to non-optimal growth conditions or varied (i.e. not specifically oxidative) stresses.

### 1.5.1 Superoxide Dismutases.

Three major classes of superoxide dismutase have been described to date: the Fe/Mn-, Ni-, and CuZn-types. The overall reaction mediated by each type of enzyme is identical, with the dismutation of two $\cdot O_2^-$ to $O_2$ and $H_2O_2$. The growth phase expression and localization differ between the classes of SOD, and in most cases SODs of only one or two of the classes are produced by a particular organism. *A. vinelandii* expresses both an FeSOD and a CuZnSOD, whereas *P. aeruginosa* produces an Fe- and Mn-type SOD but no CuZnSOD homologue (Hassett *et al.*, 1993).

*A. vinelandii* growing in exponential phase exhibits 5-fold more superoxide dismutase activity than *E. coli* during exponential phase growth (Jurtshuk *et al.*, 1984), with moderate induction with increasing oxygen concentration (Dingler and Oelze, 1987) and greater induction during stationary phase (Qurollo *et al.*, 2001). The nitrogen-fixing state of the culture does not affect SOD expression in *A. vinelandii*.

#### 1.5.1.1 Fe/Mn Superoxide Dismutases.

X-ray crystal studies indicate that the Fe- (*sodB*) and Mn-type (*sodA*) superoxide dismutases have considerable structure similarity (Miller, 2004), though the mechanics of superoxide-metal interactions differ somewhat between these SOD types. The metal co-factor is coordinated by three His and a single Asp residue in both types.

The Fe- and Mn-type SODs are cytoplasmically localized (Qurollo *et al.*, 2001) and can be distinguished from one another by the sensitivity of FeSOD to $H_2O_2$ (Asada *et al.*, 1975). FeSODs are generally expressed constitutively in bacteria (Nettleton *et al.*, 1984), as was found for *A. vinelandii* (Pagani *et al.*, 1995; Qurollo *et al.*, 2001), whereas the MnSODs tend to be oxygen inducible (Nettleton *et al.*, 1984).

#### 1.5.1.2 Cu,Zn Superoxide Dismutases.

The CuZn superoxide dismutases (*sodC*) constitute a distinct enzyme class, being unrelated to the Fe- and Mn-type SODs. The mammalian CuZnSODs, and by extension

the bacterial homologues, have received particular attention because of a possible roll in the hereditary disease, amyotrophic lateral sclerosis (ALS). There is considerable divergence between the prokaryotic and eukaryotic CuZnSODs by sequence, yet the prokaryotic and eukaryotic varieties share significant structural similarities (Kroll *et al.*, 1995). The prokaryotic CuZnSODs are more sensitive to chelation than the eukaryotic form, but exhibit greater resistance to $H_2O_2$ inactivation than the eukaryotic CuZn- (Gabbianelli *et al.*, 2004) and Fe-type SODs. The enzymatic center is located at the Cu ion; superoxide dismutation involves the oxidation of one $\cdot O_2^-$ radical at the resting state Cu(II) core, followed by the reduction of a second radical at the resultant Cu(I) core. The ligands at the active site include six His and one Asp residue, which coordinate the Cu and Zn ions (reviewed in Livesay *et al.*, 2003). The turn-over rate of CuZnSOD is very high, exceeding the theoretical diffusion limit for enzyme-substrate interaction. This accelerated rate is due to cationic funnelling of the negatively charged superoxide radicals to the active site (Lepock *et al.*, 1985).

CuZn superoxide dismutase is a virulence factor for several organisms and *sodC* mutants exhibit reduced infectivity (Battistoni, 2003). Given (i) their periplasmic localization (Steinman, 1993) (ii) the inability of endogenously produced $\cdot O_2^-$ to cross the inner membrane (Hassan and Fridovich, 1979) and (iii) the fact that more virulent pathogens tend to express multiple CuZn superoxide dismutase genes (Fang *et al.*, 1999), it seems likely that these SODs are primarily involved with the neutralization of exogenously produced (e.g. host-derived) superoxide (Battistoni, 2003; Desideri and Falconi, 2003). Interestingly the CuZnSODs of the enteric bacteria are RpoS-dependent (Fang *et al.*, 1999), though other regulatory proteins also modulate their expression.

*A. vinelandii* and many other non-pathogenic organisms also produce a CuZnSOD. The *A. vinelandii* CuZnSOD was originally reported as a MnSOD due to its insensitivity to hydrogen peroxide (Pagani *et al.*, 1995). However cloning and mutational experiments have proven it to be a periplasmically localized CuZnSOD, which is most strongly expressed during stationary phase (Qurollo *et al.*, 2001). It is unknown at this time if the expression of the *Av*CuZnSOD is strictly RpoS-dependent, but initial reports suggest the presence of a typical $\sigma^{70}$ promoter sequence (Qurollo *et al.*, 2001).

### 1.5.2 Peroxidases.

Peroxidases degrade the ROS hydrogen peroxide by one of two apparently simple reactions:

$$H_2O_2 + 2\ AH + E \rightarrow 2\ H_2O + 2\ A + E$$
$$H_2O_2 + AH_2 + E \rightarrow 2\ H_2O + A + E$$

Where E is the peroxide enzyme, AH is a single-electron organic donor molecule, and $AH_2$ is a two-electron organic donor molecule. Thus hydrogen peroxide degradation by peroxidatic activity is metabolically dependent since a donor must be regenerated, whereas catalatic activity is independent of the metabolic state of the bacterium because no donor is required (see below). While the metabolic load would not be considered burdensome to rapidly metabolizing cells, it could be significant to organisms that enter metabolically quiescent states (i.e. spores and cysts) for years or decades.

#### 1.5.2.1 The Eukaryotic Peroxidase Superfamily.

Heme-peroxidases are ubiquitous, and are produced by all kingdoms of life. Within the "plant-peroxidase" superfamily several peroxidases have been thoroughly studied because of their ease of purification. Two model enzymes, yeast cytochrome *c* peroxidase (CCP, class I) and horseradish peroxidase (HRP, class III) have been studied spectroscopically and *via* crystal structure determinations. Despite low sequence homologies, several invariant structures are common to all members of this superfamily. First, their enzyme cores contain the distal triad Arg48 His52 Trp51 and the proximal triad Trp191 His175 and Asp235 (*Mi*CCP numbering; Wang *et al.*, 1990), second the heme groups are localized to the protein surface to allow heme-substrate interactions, and third the enzymes are relatively non-specific, oxidizing a broad range of one-electron organic donor substrates. The eukaryotic CCP are unique in that they preferably accept reductive power from (ferro)cytochrome *c* proteins (Yonetani 1976).

Peroxidases belonging to the plant-peroxidase superfamily undergo very similar peroxide reactions, owing to their highly conserved nature. The catalytic core of the resting state enzymes contains a ferric heme, typically in the 5-coordinate state with a

proximal His residue and an unbonded water molecule on the immediate distal side (HRP Fig. 1.12A, CCP Fig 1.12B; Hersleth *et al.*, 2006). Upon $H_2O_2$ entry, the intermediate "compound 0" is believed to form (Fig. 1.13; Matsui *et al.*, 1999), though it should be stated that the unstable nature of this compound has frustrated attempts to characterize it definitively. This conceptual intermediate is formed when the incoming peroxide is deprotonated by an acid/base reaction with the distal His (Poulos and Kraut, 1980) and the resultant $OOH^-$ ion binds the ferric iron. Heterolytic cleavage between the oxygen atoms then occurs (Schonbaum and Lo, 1972), with the hydroxide ion stabilized by the nearby distal Arg (Poulos and Kraut, 1980) until the $H^+$ (held by the distal His) can react with the cleaved hydroxide ion. The resulting water then exits the catalytic core, leaving compound I, which contains a cation radical (Yonetani *et al.*, 1966) associated with either the porphyrin ring (Coulson *et al.*, 1971) or a distal amino acid residue (Everse 1998), and an oxygen bound ferryl (FeIV=O) heme centre. Compound I then reacts with a donor molecule, resulting in the formation of compound II, which still contains the oxoferryl heme, but is no longer radicalized. A second donor molecule is then oxidized to regenerate the resting state enzyme with the release of a second water molecule.

The oxygen carrying enzyme myoglobin (Mb) is known to have weak peroxidatic and catalatic functions, with an active core structure very similar to that of the plant-peroxidases (Fig. 1.14; Matsui *et al.*, 1999). These anti-oxidant activities are remarkably improved by substitutive mutations which change the spacing between the distal His residue and the heme iron. Activities increase 5.1- to 46-fold when the distal His spacing is increased to 6.6 or 5.4 Å, versus the wild-type distance of 4.3 Å (Matsui *et al.*, 1999). Thus there is an apparent optimal distance requirement (between 4.3 and 6.6 Å) if histidine is to act as an effective acid/base catalyst. This distal His spacing is in agreement with that typically found in the "plant-peroxidases." However, while the plant peroxidases exhibit no catalatic activity, wild-type Mb does. The improved catalatic nature of certain Mb mutants may be explained by their planar His:heme arrangement; the distal His of peroxidatic enzymes tend to sit directly over the iron with a perpendicular orientation relative to the heme (5.5 and 6.0 Å distant for HRP and CCP; Phillips *et al.*, 1990, Gajhede *et al.*, 1997), whereas the His residues of catalases are off-

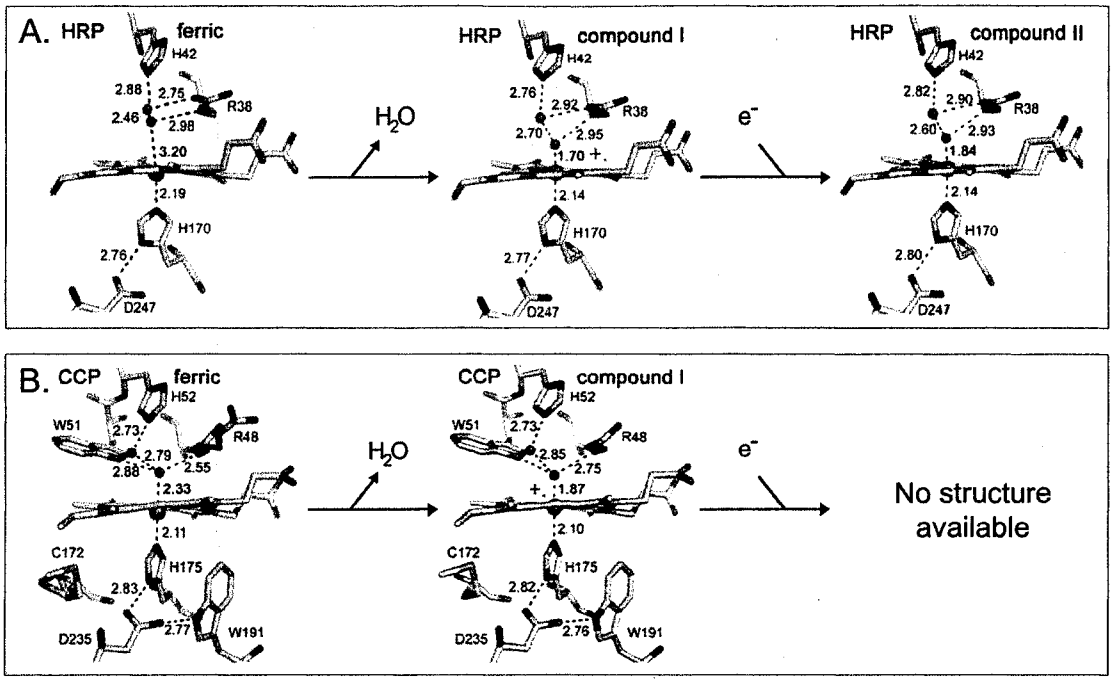



**Figure 1.12 Crystal Structures of Active Sites of the Peroxidases HRP (A) and Yeast CCP (B).** Peroxidatic reaction sites contain a heme group with a conserved proximal histidine shown below the heme iron. The distal axial ligands are shown above the heme group. Conversion of the resting state enzyme and hydrogen peroxide to compound I results in the release of one water. The conversion of compound I to compound II requires the input of one electron. From Hersleth et al. (2006).

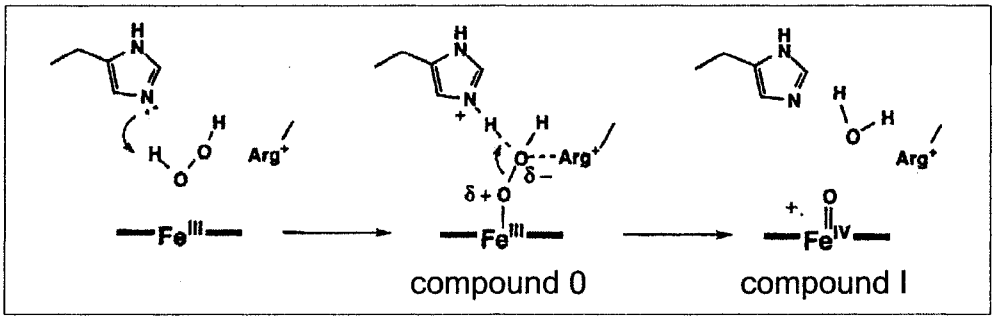



**Figure 1.13 The Proposed Scheme for the Heterolytic Cleavage of Hydrogen Peroxide** through the Hypothetical Intermediate, Compound 0. The distally located His deprotonates the incoming peroxide (indicated by a dashed line) and the resultant OOH- ion binds the ferric heme resulting in the formation of compound 0. Self-cleavage of the ferriperoxide is mediated by the stabilizing interaction of the distal Arg residue. The product, compound I has a cation radical localized to the porphyrin ring. From Matsui et al. (1999).

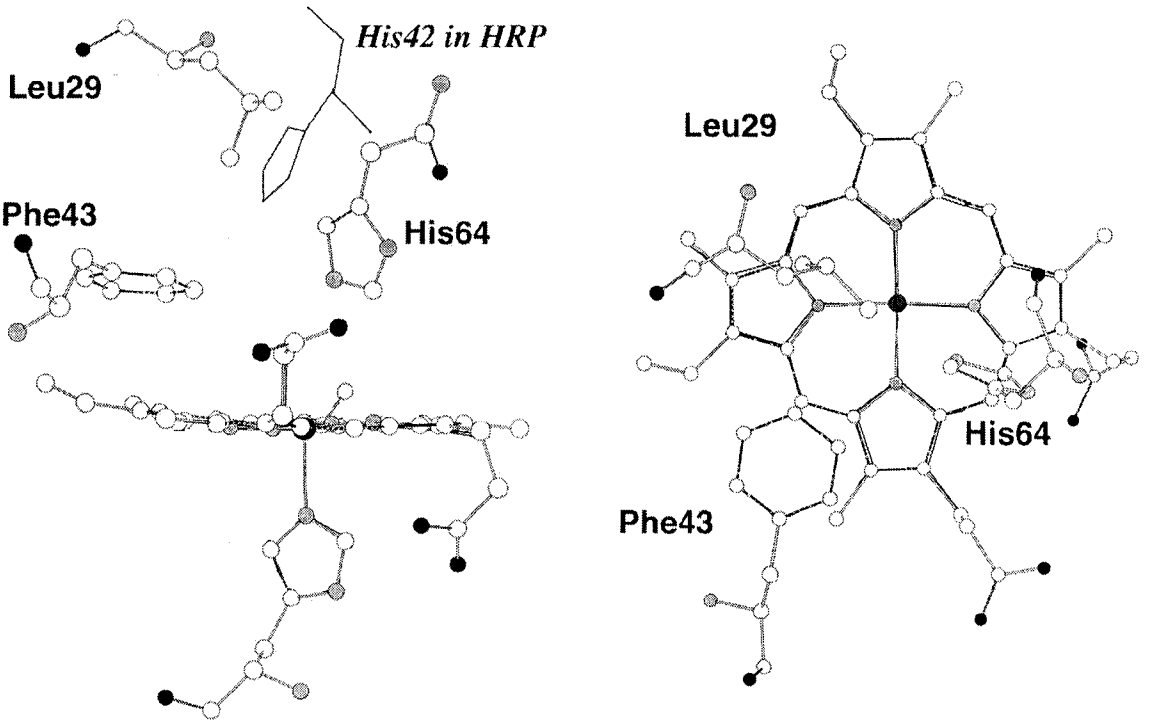



**Figure 1.14 Side and Top Views of the Axial Ligands of the Heme of Myoglobin.** Leu29His and Phe43His mutants displayed improved peroxidatic and catalatic activities, suggesting that optimal His spacing is critical for hydrogen peroxide degradation. Matsui et al. (1999).

set from the iron, oriented parallel to the heme (at ~ 3.5 Å). This may constitute the key structural difference which defines peroxidatic *versus* catalatic function, though surprisingly little additional substitutive work has been done. It should be noted that the peroxidatic and catalatic activities of various Mb mutants, though greater than their wild-type form, are substantially less than those of the typical peroxidases and catalases. This may be attributed to the relatively high oxidation-reduction midpoint ($E_M$) of the Mb mutants (+50 mV; Matsui *et al.*, 1999) as compared to those of dedicated anti-oxidizing enzymes such as HRP (-271 mV; Harbury, 1957), *Synechocystis* KatG (-226 mV; Bellei *et al.*, 2006) and *Pa*CCP (-330 mV; Ellfolk *et al.*, 1983).

### 1.5.2.2 Bacterial Di-hemic Cytochrome *c* Peroxidases.

The bacterial di-hemic cytochrome *c* peroxidases (BCCP) constitute a unique group of peroxidases, quite dissimilar to their eukaryotic namesakes (Zamocky, 2004). Unlike the yeast CCP, the bacterial forms contain two heme groups which are fused *via* thioether bonds to the protein backbone, whereas the former contain a single type-*b* heme. The typical type-*c* binding motif CXXCH is conserved in all members of the BCCP group, and due to heme *c* maturation requirements all members are localized to the periplasmic or membrane fraction (Thöny-Meyer, 2002; Goodhew *et al.*, 1990; Turner *et al.*, 2003). The C-terminal heme is coordinated with 6 axial ligands; the 5$^{th}$ ligand is the proximal His residue contained within the CXXCH motif, and the 6$^{th}$ ligand is a distal Met, typically located 74 residues downstream of the proximal His. As a result of this His-Met ligation, the C-terminal heme centres have high redox potentials (HP), ranging between +226 and +450 mV (Elliott *et al.*, 2007) and exhibit a low spin (LS) heme signature. Unlike other peroxidases, the BCCPs do not accept electrons from small organic substrates. Rather the BCCPs dock with specific proteins, such as cytochrome *c* or the cupric proteins azurin and pseudoazurin, accepting electrons from them singly (Foote *et al.*, 1992). The electrons are received by the HP heme group, and are then transferred to the N-terminal low potential (LP) heme, which is generally believed to constitute the peroxidatic active site. Redox potentials range between -100 and -320 mV at the LP heme (Elliot *et al.*, 2007), with potential difference of about 600 mV between the heme pairs of most BCCPs.

Interestingly a sizeable space exists between the hemes of any given BCCP, with the iron cores separated by approximately 20 Å (the propionate groups are 9.6 Å apart; Fülöp *et al.*, 1995). As a result, effective communication of electrons is proposed to occur *via* a highly conserved Trp residue which sits directly between the two hemes (Fig. 1.15) reducing the distance traveled to only 2.7 Å (propionate to Trp). Although electron tunnelling has been observed in several biological systems over distances greater than 20 Å, W97A and W97F substitutions in *Rhodobacter capsulatus* CCP resulted in the complete loss or 100-fold decrease in peroxidatic functions, respectively (DeSmet et al, 2006). This likely indicates that tunnelling does not occur at a rate sufficient to support peroxidatic function in the BCCPs (De Smet *et al.*, 2006).

In all but two cases (*Nitrosomonas europaea* CCP and *Methylococcus capsulatus* CCP) purified BCCP is found in a functionally inactive, fully oxidized state (Anciero *et al.*, 1994; Zahn *et al.*, 1997; Shimizu *et al.*, 2001). As crystal structures of BCCPs were published in numerous redox states, it became apparent that the majority utilize an inhibitory protein loop. This domain swings into the active site when the enzyme is fully oxidized, resulting in bis-His axial ligation to the heme iron, thereby blocking solvent and so peroxide access (Fig. 1.16). Following partial or complete reduction, the loop moves out of the active site, leaving a peroxide-accessible low potential heme in a 5c state, a 6c state (where the $6^{th}$ ligand is solvent/water) or in a mixed-coordinate state (Fülöp *et al.*, 1995; Dias *et al.*, 2004; Echalier *et al.*, 2006; DeSmet *et al.*, 2006). Interestingly *Ne*CCP lacks this inhibitory loop. Instead the LP heme is perpetually 5c HS, so readily interacts with $H_2O_2$ even in the fully oxidized state (Shimizu *et al.*, 2001).

Several residues have been identified in the semi-reduced structures that may play a role in the dismutation of $H_2O_2$. Interestingly no distal His, which is so strongly conserved in the eukaryotic peroxidases, is present in any of the BCCP active sites. Rather, conserved Met, Gln and Glu are present in the active site, all of which could potentially act as acid/base catalysts. Glu102, positioned in the *Ne*CCP active site is reminiscent of the catalytic Glu of the chloroperoxidases, and is predicted to promote heterolytic cleavage, leading to the formation of compound 0 (Shimizu *et al.*, 2001). However acceptable positioning of the conserved Glu is not present in other BCCPs (*Pa*CCP, Fülöp *et al.*, 1995; *Pn*CCP, Dias *et al.*, 2004; *Rc*CCP, DeSmet *et al.*, 2006), so

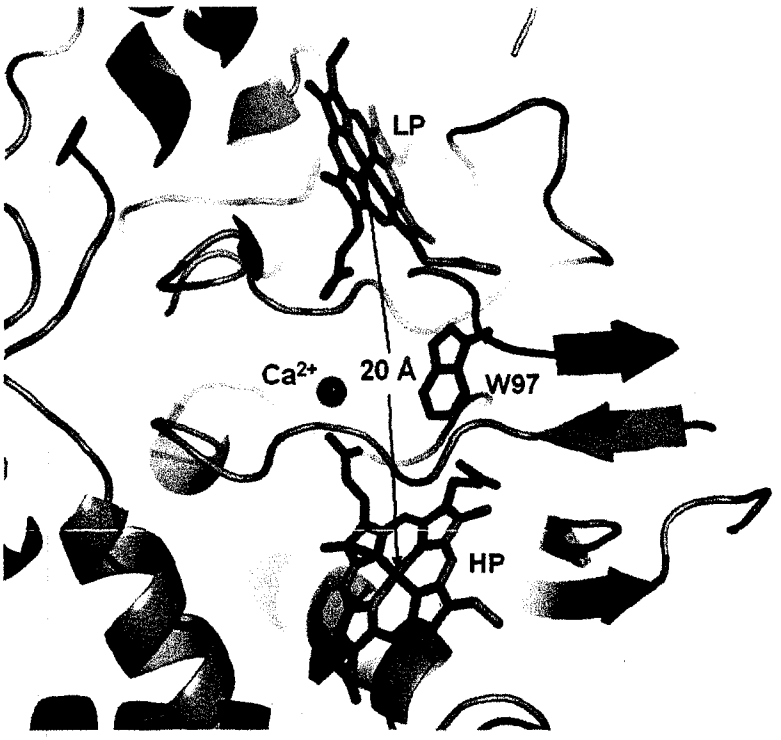



**Figure 1.15 Crystal Structure of the Inter-Hemic Region of *Rc*CCP.** The distance between the high potential (HP) electron-donating heme iron and the low potential (LP) active-site heme iron is 20 Å in free space. Electron traveling distance is reduced by the presence of the residue, Trp97, which bridges the "inter-hemic cleft". DeSmet et al. (2006).

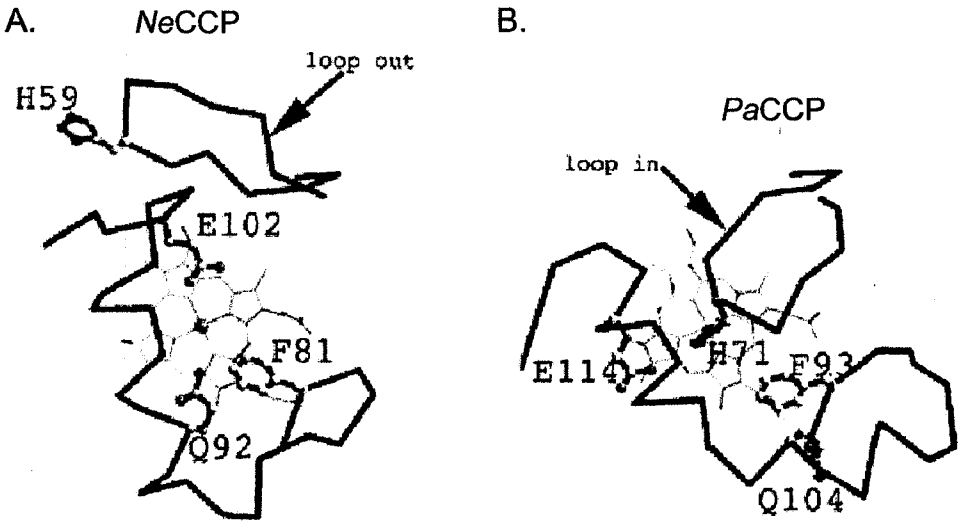



**Figure 1.16 The Inhibitory Distal Protein-Loop of Bacterial Cytochrome *c* Peroxidases.** The loop is permanently in the OUT position in *Ne*CCP, allowing solvent access at the active-site under oxidizing conditions (A). In other BCCPs the loop is in the IN position, resulting in a bis-His blockage of the heme (B). Reduction gives the OUT form, allowing solvent access. Shimizu et al. (2001).

residues such as Phe and Arg may alternatively act to stabilize heterolytic cleavage in these enzymes. No mutational studies of the active site have been undertaken in the BCCPs however, largely due to cloning difficulties, so it is unclear at this time what residues are critical for peroxidatic activity. In the favoured model the resting state enzyme is of mixed valence with a 5c HS N-terminal ferrous heme and a 6c LS C-terminal ferric heme (Fig 1.17; DeSmet *et al.*, 2006). The presence of $H_2O_2$, and subsequent transition through Compound 0, would result in the formation of a Compound I-like intermediate, with the oxoferryl compound forming at the LP N-terminal heme (Arciero and Hooper, 1994). Unlike the eukaryotic peroxidases, it should be possible for an electron to be immediately donated from the HP heme (Fülöp *et al.*, 1995), thereby avoiding the radicalization of the porphyrin ring at the active site and potentially increasing the longevity of the enzyme (Arciero and Hooper, 1994; Fülöp *et al.*, 1995). This non-radicalized "compound I-like" product would be very stable, and could be reduced by sequential electron donations from cytochrome *c* proteins as electrons became available. In the case where the cellular metabolic state is depressed (i.e. during starvation) electron donors may not be available to the HP heme. In the case of *Ne*CCP, which does not have an inactivating protein loop, this could leave the active site heme in a "true" Compound I state, complete with cation radicals localized to the porphyrin ring (or nearby amino acid). Subsequent devolution to Compound II could result in the inactivation of the enzyme if not resolved in a timely manner (Arciero and Hooper, 1994) and may explain why some BCCPs utilize the "inactivating" loop when fully oxidized.

Despite years of extensive biochemical study, very little has been published regarding the regulation of BCCP expression. Ironically, the BCCPs are most frequently expressed during anaerobic growth. The most studied BCCP-related regulatory system is that of *N. gonorrhoeae*, for which the peroxidase is a critical pathogenicity determinant in the face of host derived peroxide challenge (Seib *et al.*, 2004; Turner *et al.*, 2003). *Ng*CCP is positively regulated under low oxygen conditions by the oxygen sensing global regulator, FNR (Lissenden *et al.*, 2000; Turner *et al.*, 2003). Induction of *Ng*CCP in response to peroxide stress has also been observed, though the regulator has not yet been identified (Stohl *et al.*, 2005). Whereas anaerobic regulation makes sense for a pathogen that has colonized a target organism and is waiting in a "prepared state" for potential ROS-based

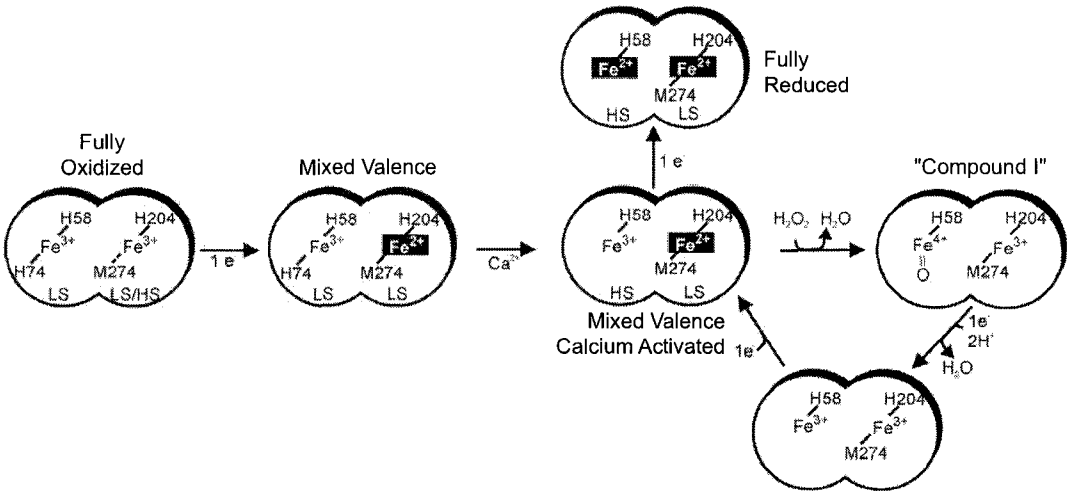



**Figure 1.17 The Proposed Redox and Axial Coordination Schema of the BCCPs.** Two heme domains exist in the BCCPs. The N-terminal heme is believed to constitute the active site, and is coordinated by H58 and H74 in the fully oxidized state. The C-terminal heme, which is believed to be the electron accepting/donating heme, is coordinated by H204 and M274 under all redox states and so is invariantly 6c LS in nature. Under physiological conditions (mixed-valence, $Ca^{2+}$activated) the inactivating protein loop adopts the "OUT" conformation, removing the coordinating residue H74 such that the active site heme exhibits the 5c HS state. Solvent access to the active site heme allows reaction with $H_2O_2$, yielding an oxyferryl moiety and water. Immediate oxidation of the C-terminal heme allows the enzyme to avoid cation radicalization of the porphyrin ring. Two-stage electron donation from small protein donors via the C-terminal heme regenerates the mixed-valence $Ca^{2+}$activated enzyme state. The fully reduced form of the enzyme is inactive, and is not observed *in vivo*. DeSmet et al. (2006).

host responses, satisfactory explanations for the anaerobic expression of BCCPs by non-pathogens are still lacking.

### 1.5.3 Catalases.

Catalases act to degrade hydrogen peroxide into water and oxygen by the reaction:

$$2\ H_2O_2 + E \rightarrow 2\ H_2O + O_2 + E$$

Three major classes of catalases have been identified to date: the monofunctional heme catalases first described in 1900 (Loew, 1900), the bifunctional catalase-peroxidases (often abbreviated CPx) first characterized in 1979 (Clairborne and Fridovich 1979), and the pseudocatalases which contain a manganese reactive centre, first purified in 1983 (Kono and Fridovich, 1983). Each catalase class catalyses the catalatic cleavage of hydrogen peroxide characteristically.

To date the catalases of *A. vinelandii* have not been identified. Surprisingly the catalase specific activity of the exponentially growing *A. vinelandii* culture (Dingler and Oelze, 1987; Curatti *et al.*, 2005) is roughly equivalent to that of exponentially growing *E. coli* (Meir and Yagil, 1984) despite having a much higher respiration rate (Kelly *et al.*, 1990).

#### 1.5.3.1 The Monofunctional Hemic Catalases.

Hemic monofunctional catalases have been classified in three clades: clades I and III constitute small subunit catalases of mainly plant and non-plant origin respectively, while clade II catalases contain large subunits of bacterial or fungal origin only. Some evidence for horizontal transfer, particularly of the KatE/HPII type catalases of clade II exists (Klotz *et al.*, 1997). Clade II catalases also tend to be more resistant to heat and chemical abuses, which may be the major driving force for their maintenance following interspecies transfer. In the case of the clade II catalase of the mesophilic bacterium *E. coli* (*Ec*HPII) the protein is both heat resistant (with a half-life of 12 h at 65 $^{o}$C; Switala and Loewen, 2002) and resistant to protease treatment (Chelikani *et al.*, 2003). Clade I and III catalases contain heme *b* catalytic cores, whereas clade II catalases utilize type-*d* hemes.

Enzymatic degradation of $H_2O_2$ by the monofunctional hemic catalases proceeds in two stages. In the first stage $H_2O_2$ is oxidized to water and the enzyme is converted to compound I, which contains a cation-radicalized porphyrin ring and an oxoferryl heme centre (Fig 1.18). Unlike the peroxidases, the radical does not associate with an axial ligand, but is strictly limited to the porphyrin ring (Andreoletti *et al.*, 2003). The ferric hydroperoxo-intermediate compound 0 is also believed to exist for the catalases, but as with the peroxidases it has yet to be captured due to its unstable nature. Unlike the peroxidases, compound I is directly reduced by a second $H_2O_2$ rather than proceeding to compound II. As a result the second $H_2O_2$ is reduced to water and $O_2$ and the enzymatic core is regenerated by the following reaction:

$$H_2O_2 + Por + Fe(III) \rightarrow Por + Fe(III)\text{-}O\text{-}O\text{-}H^- \rightarrow H_2O + \bullet Por^+ + Fe(IV)=O$$

$$H_2O_2 + \bullet Por^+ + Fe(IV)=O \rightarrow H_2O + O_2 + Por + Fe(III)$$

From Chelikani *et al.*, 2004 and Hersleth *et al.*, 2006

Under low peroxide conditions, as might be experienced *in vivo*, the catalase core can be trapped as compound I, which can spontaneously degrade to compound II, permanently inactivating the enzyme. As a result, several catalases have associated NADPH that can rescue compound II by directly reducing the heme to the resting ferric state (Andreoletti *et al.*, 1995).

Several crystal structures have been determined for the monomeric catalases. Being of type-*b* or -*d*, the hemic moieties interact with the protein non-covalently in a "binding pocket" which is typically buried deeply within the multimeric protein. Many of the residues that constitute the heme pocket are necessary for enzymatic function. For example, the proximal residue is invariably Tyr, located directly next to the heme iron. On the distal side, a highly conserved His residue is oriented such that it is off-set from the heme iron, being directly above and parallel to either ring III or IV of the porphyrin structure (Fig. 1.19). This His residue likely undergoes an acid-base reaction with $H_2O_2$ in a manner analogous to the conserved distal His of the class I peroxidases. The

Fig. 1.18 pp 55

Fig 1.19 pp 55

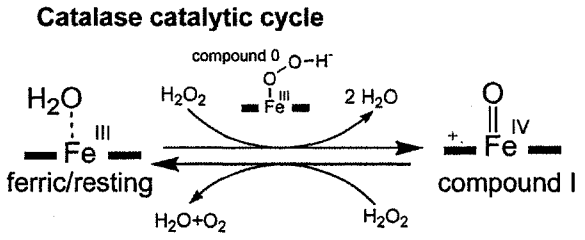



**Figure 1.18 The Catalase Reaction Cycle with $H_2O_2$.** The heme macrocycle is represented by the thick black lines on either side of the iron. The cation radical is indicated in pink. Hersleth et al. (2006).

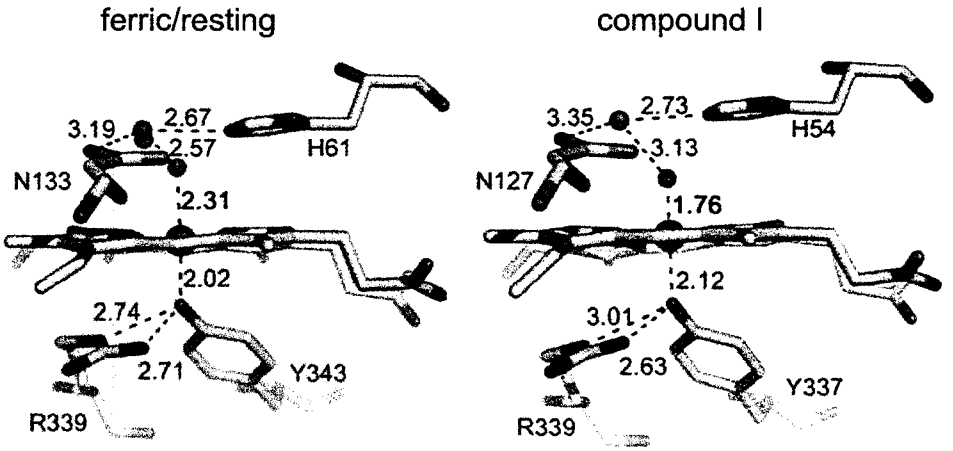



**Figure 1.19 Crystal-Derived Structures of the KatE Enzymatic Core in the Resting and Compound I States** (from *Micrococcus lysodeikticus* and *Proteus mirabilis*, respectively). The invariant His residue is aligned parallel to the heme plane, whereas the invariant proximal Tyr residue is oriented perpendicularly. Hersleth et al. (2006).

conserved distal Asn in the catalases probably acts by stabilizing the hydroxyl ion following the heterolytic cleavage reaction, which is functionally equivalent to the conserved Arg of the peroxidases. A conserved Phe, stacked parallel to the heme group, is also found within the active site on the distal side (Maj *et al.*, 1998). It is unclear at this time what structural elements truly differentiate the catalatic reaction from the peroxidatic reaction, allowing the enzyme to by-pass compound II during the enzymatic cycle. The protein structure also forms access channels to the catalatic core. These channels tend to be more restrictive than in the peroxidases, resulting in sharply defined substrate specificities as well as reduced susceptibility to various inhibitors (Switala and Loewen, 2002).

### 1.5.3.2 The Bifunctional Hemic Catalases.

The bifunctional catalase-peroxidases (variably referred to as CPx, KatG, or HPI) are remarkably similar in structure to the class I peroxidases, leading to the proposal that they share a recent phylogenetic lineage (Klotz and Loewen, 2003). The KatG protein has two highly similar heme binding sites which may have been generated by a gene duplication event (Welinder, 1991). The C-terminal site is apparently inactive, so both the catalase and peroxidase functions are mediated at the N-terminal heme. KatG proteins utilize a type-*b* heme, which often undergoes chemical modification (Carpena *et al.*, 2003).

The structure and even residue conservation between the type I peroxidases and KatG active sites is very significant (Fig. 1.20A), yet only the latter is capable of catalatic functions. Unlike the type I peroxidases, the KatG enzymes contain a unique covalent linkage between the highly conserved Trp95, Tyr218 and Met244 residues (all numbering adjusted to that of *Haloarcula marismortui* KatG) on the distal side of the heme (Fig. 1.20B). Of these, only the Trp residue functions within the enzymatic core, though mutational replacement of any of the triad members results in the complete loss of catalatic, but not peroxidatic, function (Chelikani *et al.*, 2004). Replacement of the conserved Trp with a Phe residue actually increases peroxidatic activity nearly 3-fold while essentially eliminating catalase activity (reduced 1000-fold; Hillar *et al.*, 2000). The distal pocket also contains conserved arginine (Arg92) and histidine (His96) residues, both of which are required for optimal peroxidatic and catalatic functions

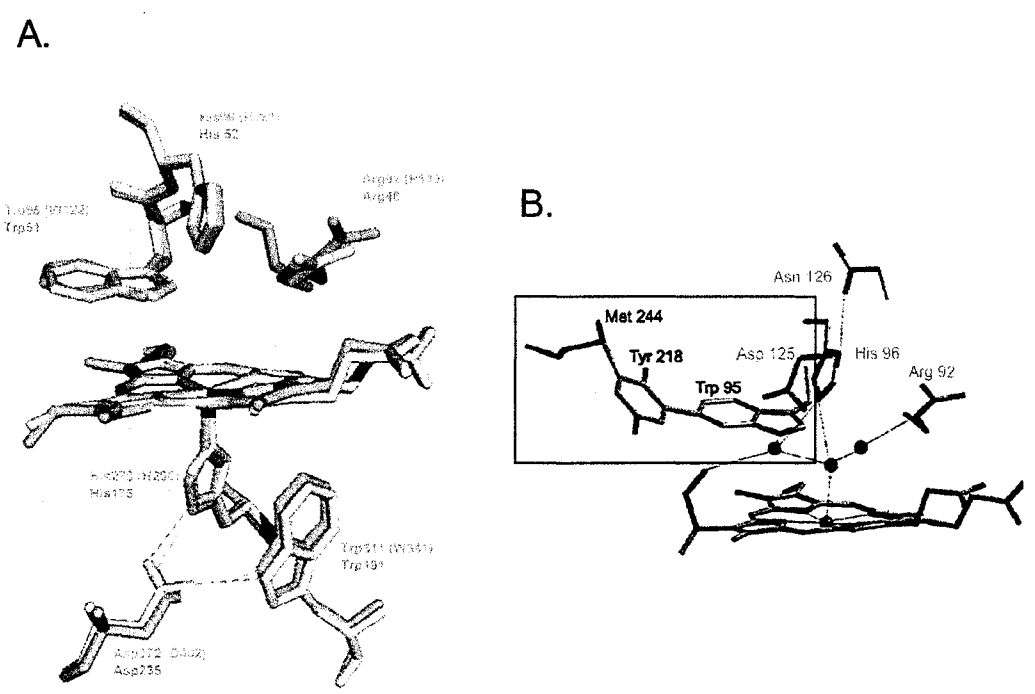



**Figure 1.20 The Structural Alignment of the Bifunctional Catalase, KatG, of *Haloarcula marismortui* (yellow) and Yeast CCP (gray)** demonstrates the high degree of structural homology between the enzymes (A). The KatG-specific covalent adduct between Trp95, Tyr218 and Met244 is highlighted by a box in panel B. Although only Trp95 is present within the active domain, substitutions at any of these residues result in the loss of catalase activity. From Smulevich et al. (2006) and Chelikani et al. (2004).

(Hillar *et al.*, 2000). These residues are predicted to be critical for the acid/base stabilization of peroxide during degradation as demonstrated for the class I peroxidases (Santoni *et al.*, 2004). KatG proteins have an additional conserved Asp residue in the distal heme pocket, which is required for strong catalase activity; D125N, D125S and D1225W mutants exhibit 18- to 175-fold lower catalatic activity but 2.3- to 6.7-fold greater peroxidatic activity (Santoni *et al.*, 2004). Unfortunately the mechanism that differentiates between peroxidatic and catalatic activities continues to elude research efforts. In contrast to yeast CCP, substitutions of the coordinating Trp and Asp residues in KatG often affect peroxidatic functions negatively (Santoni *et al.*, 2004). Catalatic activity is always negatively affected by proximal and distal substitutions, suggesting that the catalase functions are more restrictive than are peroxidase functions (Smulevich *et al.*, 2006).

Interestingly, the bifunctional catalase channels tend to approach the active site parallel to the heme-plane, whereas the monofunctional catalase channels enter the active site in a perpendicular fashion (Chelikani *et al.*, 2004). The KatG access channels are less restrictive than their monofunctional counterparts, resulting in a greater degree of sensitivity to inhibitory chemicals.

#### 1.5.3.3 The Pseudocatalases.

First isolated by Kono and Fridovich (1983) the manganese containing catalases constitute the most divergent family of catalases (Chelikani *et al.*, 2004). Primarily produced by thermo- and hyperthermophilic bacteria, such as *Thermoleophilum album* (Allgood and Perry, 1986) and *Pyrobaculum calidifontis* (Amo *et al.*, 2002), most of the Mn-catalases exhibit exceptional heat resistance. The sensitivity of these catalases to artificial ligands such as azide, cyanide and triazole is significantly lower than the hemic catalases, though hydroxylamine remains a significant inhibitor (Allgood and Perry, 1986).

The Mn-catalase subunit masses are typically 30 kDa, with reports of tetrameric, pentameric and hexameric organisation. The proteins are highly structured, being primarily α-helical in nature (Barynin *et al.*, 2001) with next to no β-sheet structure. The catalatic cores typically contain two Mn ions, coordinated by three highly conserved Glu

and one or two His residues, as well as water/hydroxyl ions (Barynin *et al.*, 1997; Barynin *et al.*, 2001). The Mn ions were determined to be in only the Mn(II)-Mn(II) or Mn(III)-Mn(III) redox states, even though alternative states are conceptually possible (Waldo and Penner-Hahn, 1995). The catalatic reaction cycles between these two enzyme states in the following reaction:

$$H_2O_2 + \text{Mn(II)-Mn(II)} + 2H^+ \rightarrow \text{Mn(III)-Mn(III)} + 2\ H_2O$$
$$H_2O_2 + \text{Mn(III)-Mn(III)} \rightarrow \text{Mn(II)-Mn(II)} + 2H^+ + O_2$$

The specific mechanism of these reactions remain somewhat unclear, though several models have been suggested (Boelrijk and Dismukes, 2000). This reaction scheme is complicated by the low-level peroxidatic activity of some Mn-catalases (Allgood and Perry, 1986).

**1.6 Thesis Objectives.**

The purpose of this thesis is to profile the stress responses of *Azotobacter vinelandii*, a significant soil organism and potentially industrially relevant organism. It will be important to identify what stresses result in the expression of RpoS and by what mechanisms this regulation is imposed. While many systems will be conserved amongst all Gram negative bacteria, *A. vinelandii* is likely to have a number of unique RpoS-regulatory systems not seen in the enteric bacteria. In particular this diazotoph may have a novel stress response to nitrogen-limitation not observed in the enterics. Furthermore the very rapid respiration rate employed during nitrogen-fixation may cause oxidative stresses not experienced by other bacterial species, necessitating novel stress response schemes.

The second area of study will investigate the hierarchy of RpoS regulation, that is, the degree of regulation at the transcriptional, post-transcriptional, post-translational, and activity levels. Although well studied in the enterics, relatively little is known about RpoS regulation in the pseudomonads. It is currently believed that *rpoS* is primarily regulated at the transcriptional level in the pseudomonads and that no regulation occurs at

the post-translational level. However no RpoS stability data is currently available in the literature for pseudomonad-like bacteria. Studies using *hfq* mutants suggest that there may be *rpoS* regulation at the post-transcriptional level. Unfortunately these studies have only examined protein levels, and have ignored transcript levels. As a result it is unclear if Hfq acts directly at the translational level as it does in the enterics, or if there is an indirect effect on RpoS expression. Furthermore, little attention has been directed to the study of $E\sigma^S$ formation and activity in the pseudomonads. Considering that most pseudomonads contain 3 to 4 times more sigma factor species than does *E. coli*, albeit mostly ECFs with likely cognate anti-sigma factors, sigma competition may play a much more significant role in RpoS regulation than has been seen previously.

The third area of study will focus on the discovery of an RpoS-dependent cytochrome *c*-type catalase produced by *A. vinelandii* during periods of nutrient depletion. Though originally studied as an indicator of $E\sigma^S$ activity, once the novelty of the enzyme was realized, enzymatic characterization was pursued. Additionally several methods were used to measure the degree of susceptibility of the catalase to chemical and heat damage, which could be considered proxies for enzyme stability in the encysted cell.

Given the reported difficulties in the cloning and transgenic expression of dihemic cytochrome *c* proteins and the difficulties inherent to *A. vinelandii* mutagenesis, structural work was confined to spectrometric methods and computational modeling. Interpretations of the spectrometric methods largely agree with each other, suggesting that the hemic ligand states differ significantly from those of other heme *c* antioxidant proteins. Instead the catalase protein demonstrates more similarities to group 1 cytochrome *c'* proteins, for which catalatic activities have never been described.

# CHAPTER 2

# Materials and Methods

**2.1 Bacterial Strains and Growth Conditions.** The parental *Azotobacter vinelandii* strains used were the the alginate non-producing strain, UW (ATCC 13705) and the PHB overproducing strain (UWD, ATCC 53799) derived from it. Mutant strains not previously described in the literature are summarized in Table 2.1. Strains were grown in Burk's complete medium (BBGN) composed of Burk's buffer salts (BB) supplemented with 18.0 μM ferric citrate, 55.5 mM glucose and 15 mM ammonium acetate, using the growth conditions described previously (Page *et al.*, 2001). In ammonium-augmentation experiments, 100 ml of 7 mM $(NH_4)_2SO_4$, pH 7.0, were added to 2.5 L of BBGN at 5 ml $min^{-1}$ just prior to ammonium limitation. The antibiotics kanamycin, gentamycin and ampicillin were used in liquid and solid media at 10 μg $mL^{-1}$, 500 ng $mL^{-1}$ and 100 μg $mL^{-1}$, respectively.

Shake flask cultures of *A. vinelandii* were inoculated from three day old BBGN agar slants to an initial $OD_{600}$ of 0.030 to 0.040 and grown with agitation at 200 rpm on a rotary shaker at 28 °C unless otherwise noted. Large batch cultures were grown in a 2.5 L BioFlowIII bioreactor (NewBrunswick Scientific) that was inoculated from liquid seed culture (as above) with no more than a 4% inoculum (v/v) and aerated with 1 v/m air, with a constant impeller speed of 200 rpm. Dissolved oxygen ($dO_2$) was monitored electronically and samples were removed manually at the time intervals noted in the results.

*Escherichia coli* DH5α was grown at 37 °C in liquid LB medium with shaking (200 rpm) and on solid LB medium containing 0.2 mM IPTG and 40 μg/mL X-gal.

**2.2 Quantification of Medium Components.** Glucose and ammonium concentrations in the culture fluid were determined by the Trinder assay (Sigma Diagnostics) and the sodium nitroprusside method (Strickland and Parsons, 1972). Acetate concentration was determined by gas chromatography (Hewlett Packard 5890) using a DB-FFAP megabore column (Agilent). All experiments were repeated at least twice and the values reported are means of duplicate or triplicate assays.

| Strain | Relevant Genotype or Phenotype | Reference |
|---|---|---|
| UW | *algU*::Tn [(1)] | Page et. al., 2001 |
| UWD | *algU*::Tn, $PHB^{++}$ [(2)] | Knosp and Page, 1989 ; Page et. al., 2001 |
| UWS | *algU*::Tn, *rpoS*::$Km^R$ | Kujat-Choy and Page, unpublished |
| UWDS | *algU*::Tn, *rpoS*:: $Km^R$ | Kujat-Choy and Page, unpublished |
| UWD *cccA*1 | *algU*::Tn, $PHB^{++}$, *cccA*::$Gm^R$ [(3)] | This work |
| UWD *cccA*2 | *algU*::Tn, $PHB^{++}$, *cccA*::$Gm^R$ [(4)] | This work |
| UWDS *cccA*1 | *algU*::Tn, $PHB^{++}$, *rpoS*::$Km^R$, *cccA*::$Gm^R$ | This work |
| UWDS *cccA*2 | *algU*::Tn, $PHB^{++}$, *rpoS*::$Km^R$, *cccA*::$Gm^R$ | This work |

[(1)] natural transposon interruption of the *algU* gene.

[(2)] PHB overproduction is due to an unmapped spontaneous mutation in strain UW when back-crossed with strain 113 chromosomal DNA.

[(3)] *cccA*1 mutations have the $Gm^R$ cassette inserted in the same orientation as the *cccA* ORF.

[(4)] *cccA*2 mutations have the $Gm^R$ cassette inserted in the opposite orientation as the *cccA* ORF.

**Table 2.1 Strains Used in this Work.** The relevant genotypes (or phenotypes where the mutation loci are uncertain) and references are listed.

**2.3 Cloning and Interruption of *rpoS*.** The *A. vinelandii rpoS* mutants UWS and UWDS were constructed by Dr. Sonja Kujat-Choi. A brief description of the methods used to make and validate the mutants is provided below.

*A. vinelandii* chromosomal DNA was cut by *Xho*I, and the products were cloned into pCR2.1-TOPO (Invitrogen). Primers WJP58 (5'ATCATGAACCAGACC(dI)GCAC CATC-3') and WJP59 (5'-CG(dI)TC(dI)CGGGTGAG(dI)CCGAT(dI)TC-3') were used for the amplification of the *rpoS* gene for subsequent screening of the library by Southern hybridization. These primers were designed using the highly conserved amino acid sequences IMNQTRTI and EIGLTRER found in the RpoS protein (GenBank P45684), according to *A. vinelandii* codon preference (www.kazusa.or.jp), with deoxyinosine (dI) substituted at points of nucleotide uncertainty. Primer WJP79 (5'-CCGCCTGGTTGAC AAGCTCT-3') hybridized to the 3'-end of the *A. vinelandii fdxA* gene (GenBank 63007). The plasmid containing the *rpoS* fragment was designated pRPOS and the sequence was confirmed using the DYEnamic ET kit (Molecular Dynamics). A 960 bp kanamycin resistance cassette was excised from p34S-km (Dennis and Zylstra, 1998) with *Sma*I which was ligated into the *Nru*I site of pRPOS to generate pRPOS::km. This plasmid was transformed into *A. vinelandii* (Page & von Tigerstrom, 1978) and homologous recombination produced the kanamycin resistant *rpoS* mutant UWS. Insertion of the $Km^R$ cassette into *rpoS* was confirmed by PCR and sequencing. The stability of the *rpoS::km* interruption was confirmed by PCR of 25 randomly chosen UWS colonies after 10 days of nutrient starvation. Strain UWDS was produced in a similar manner but used the parental PHB producing strain, UWD.

**2.4 Cloning and Interruption of *cccA*.** The *cccA* gene was amplified from *A. vinelandii* chromosomal DNA by Taq polymerase using the forward primer WJP352 (5'-GGCTTTCGCTTCCTGGTCCTTTC-3') and the reverse primer WJP353 (5'-GTACTGCAACTGATCCTCGCCCG-3'). An additional 10 min extension step was included for preferential addition of a single adenosine residue at the 3' ends. The 1.1 kb product was separated electrophoretically on a 0.8% agarose gel (OmniPur, EM Science) and extracted using a QIAquick kit (Qiagen). The product was ligated into pGEM-Teasy (Promega), producing the plasmid pCCC. Sequence analysis from the M13 forward

primer (5'-GTAAAACGACGGCCAGT-3') confirmed the insertion of the *cccA* gene. pCCC was digested by *Stu*I, and the linearized plasmid was purified from a 0.8% agarose gel. The gentamycin antibiotic cassette ($Gm^R$) was excised from the plasmid p34S-Gm (Dennis and Zylstra, 1998) by *Sma*I digestion, and the host vector was destroyed by cleavage with *Nde*I. The $Gm^R$ cassette was treated with shrimp alkaline phosphatase (Roche), purified from an agarose gel, and ligated into *Stu*I-linearized pCCC using a rapid ligation kit (Fermentas). Transformants were selected on LB gentamycin plates and plasmid purification was performed from cultures inoculated from isolated colonies. The orientation of the $Gm^R$ cassette was determined by restriction analysis and confirmed by sequencing (DYEnamic ET kit, Molecular Dynamics). Two plasmids were isolated, pCCC::Gm1 and pCCC::Gm2, with the $Gm^R$ cassette in the forward and reverse orientation respectively. These plasmids were separately transformed into *A. vinelandii* strains UWD and UWDS by natural transformation techniques (Page and von Tigerstrom, 1978). Following transformation the plasmids underwent homologous recombination with the genome, resulting in gentamycin selectable mutants. Due to genome polyploidy in *A. vinelandii* the resultant strains had to be repeatedly passaged on gentamycin plates to obtain homogenous mutants. A minimum of 10 such passages were done for each clone and homogenetity was confirmed by PCR of chromosomal DNA using primers WJP352 and WJP353.

**2.5 Production of HIS-RpoS Protein and Antiserum.** HIS-RpoS expression, purification and RpoS-antiserum production was executed by Dr. Sonja Kujat-Choi. A brief description of the methods used is presented below.

The *rpoS* gene was amplified by PCR using primer WJP106 (5'-TTGAACATACCatgGGACGACAACG-3') and the M13-20 universal primer. Primer WJP106 contained an *Nco*I site, allowing translational fusion with the start codon of *rpoS*. pRSETa (Invitrogen) and the PCR product were digested with *Hin*dIII and *Nco*I. The ligated product, pRSET-*rpoS*, was transformed into *E. coli* DH5α. HIS-RpoS protein was overexpressed, solubilized and purified by Ni-NTA-agarose (Qiagen) followed by SDS-PAGE and electroelution (Tyler Research Instruments Model EE-04) with a 12-14 kDa cut-off membrane. Electroelution was performed in buffer 1 (0.1% w/v SDS, 0.05

M $NH_4HCO_3$) at 50 V for 12-18 h, followed by elution-dialysis in buffer 2 (0.2% w/v SDS, 0.01 M $NH_4HCO_3$) for an additional 4-6 h at 80 V, at which point the HIS-RpoS protein was recovered at the cathode. The resultant HIS-RpoS protein was used to raise polyclonal antiserum in rabbits.

**2.6 Hydrogen Peroxide Survival Assay.** Bacteria for treatment with hydrogen peroxide were grown in shake flask culture to the desired growth phase, removed from the culture medium by centrifugation at 3,000×g for 5 min at room temperature and resuspended in to a final $OD_{600}$ of 1.0 (ca. $1.8 \times 10^8$ CFU $mL^{-1}$). To maintain the growth state of the cultures, exponentially growing cells were resuspended in sterile BBG whereas stationary phase cells were resuspended in sterile BB. A sample at time zero was immediately plated on BBGN to determine CFU and hydrogen peroxide (final concentration 15, 20, 40 or 60 mM as indicated in the Fig. legend) was added to the remaining culture. Further incubation of the culture continued with shaking. Samples were serially diluted and enumerated on BBGN agar plates.

**2.7 Preparation of Cell Lysates.** *Azotobacter* lysate was prepared based on the method of Benov and Al-Ibraheem (2002). The cells were concentrated by centrifugation at 5000×g for 5 min at 4 °C and immediately frozen at –20 °C. The pellets were resuspended in 1.0 mL 10 mM potassium phosphate buffer (pH 7.7) containing 1 μM Pepstatin A and 1 mM PMSF, then transferred to 13 mm test tubes. One gram of 0.1 mm diameter beads (Cole Parmer) was added and the mixture was agitated on a vortex mixer at full speed for 5 min at 4 °C. The resultant cell lysate was cleared of whole cells by centrifugation for 5 min at 7,100 rpm (model CL, International Equipment), then clarified for 5 min at 16,110×g in a microcentrifuge. *E. coli* lysate was prepared by sonication (two 30 sec bursts at 1 min intervals on ice) using a Braun-Sonic 2000 sonicator. Soluble protein concentrations were determined by the method of Lowry *et al.* (1951).

**2.8 Sphaeroplast Formation.** Liberation of the periplasmic fraction from *A. vinelandii* UWD was accomplished by sphaeroplast formation, based on previously published methods (Hunter *et al.*, 1989; Goodhew *et al.*, 1990) except that the organism was grown

to stationary-phase, allowing for maximal *Av*CCC expression. Bacteria were pelleted at 10,000×g for 10 min, and the spent growth medium was decanted. Pellets were resuspended at 10 × volume/weight (1 mL buffer per 100 mg wet cells) in 0.5 M sucrose, 40 mM Tris HCl, pH 8.0 buffer that had been pre-heated to 30 °C. The cells were exposed to 4 mM EDTA (final concentration) for 2 min, then $CaCl_2$ was added to a final concentration of 10 mM. The cells were incubated a further 30 min with mixing, centrifuged at 15,000×g for 30 min at 4 °C, and the periplasmic fraction was decanted from the firm pellet. The remaining cell debris was ground in a sintered glass homogenizer in the presence of milliQ $H_2O$ to ensure cell lysis. Cytoplasmic fractions were cleared of membrane and PHB components by centrifugation at 15,000×g for 30 min.

**2.9 *In situ* Catalase Zymography.** Aliquots (50 µg) of whole cell lysate were electrophoresed on 7.5% or 10% discontinuous non-denaturing polyacrylamide gels as noted in text (SE500, Hoefer Scientific). Catalases were detected by the activity staining method of Clare *et al.* (1984). Briefly, gels were soaked in 100 mL horse radish peroxidase (10 U $mL^{-1}$ final concentration) for 45 min, followed by incubation in 100 mL 5.0 mM hydrogen peroxide for 15 min. Gels were transferred to 100 mL 0.75 mg $mL^{-1}$ diaminobenzidine, and colour development typically occurred within 5 min. All solutions were made in 50 mM potassium phosphate buffer, pH 7.6.

To assess residual catalase activities following heat treatment, protein was incubated for 30 min in a pre-heated thermocycler (Mastercycler, Eppendorf), then immediately chilled on ice. Samples were added to non-denaturing load dye, immediately loaded to the gel without further heating and electrophoresed and stained as described above.

**2.10 *In situ* Superoxide Dismutase Zymography.** Protein samples (100 µg) were separated on 7.5% non-denaturing gels and stained as in Beauchamp and Fridovich (1971). Gels were soaked for 20 min in 2.45 mM nitroblue tetrazolium. The gels were then soaked for 15 min in 28 mM TEMED, 280 µM riboflavin and finally exposed to a 60 watt light bulb at 10 cm for 5 min. The background developed a solid blue colour

within 10 min, and active bands remained clear. All solutions were made up in 50 mM potassium phosphate buffer, pH 7.6.

**2.11 *In situ* Heme Staining.** Protein was added to loading dye containing β-mercaptoethanol (βME, Kodak) and SDS (Sigma) at the final concentrations of 10% (vol/vol) and 4% (wt/vol). Samples were separated on 10% SDS polyacrylamide gels that had been aged for 24 h to ensure the dissipation of residual peroxides. The gels were then stained for heme groups by the method of Goodhew *et al.* (1990) wherein the gel was soaked in 30:70 methanol:250 mM sodium acetate, pH 5.0, for 30 min, then stained with 1.25 mM 3,3',5,5,' tetramethylbenzidine (TMBZ, Sigma) in the same buffer for 30 min. Colorimetric development (cyan) of the hemic protein bands was initiated by the addition of 20 mM $H_2O_2$.

**2.12 Quantification of RpoS Protein by Western Analysis.** Proteins were separated by 10% SDS-PAGE in a large format Hoefer Scientific SE500 or Dalt*six* (Amersham) gel system. Proteins were transferred to Immun-Blot PVDF membranes (Bio-Rad) by wet-cell electrotransfer (TE Transfor, Hoefer Scientific) in 20% methanol, 27 mM Tris-base, 192 mM glycine, pH 8.1 transfer buffer at 100 mA for 16 h at 4 °C. Membranes were first stained with the primary antibody (HIS-RpoS rabbit polyclonal antibody), then by a fluorescently conjugated secondary antibody (IRDye800-conjugated donkey anti-rabbit antibody, Rockland Immunochemicals). Detection was performed using the Odyssey Infrared Imaging System (Li-Cor) at 84 μm resolution.

Validation of the quantification method was done as described by Jishage and Ishihama (1995). RpoS present in 25 μg of strain UW cell lysate harvested at 48 h was arbitrarily designated as 100 fluorescence units and the same amount of lysate from strain UWS at 48 h as zero fluorescence units. Relative fluorescence (RF) was linear ($R^2$ = 0.996) from 1 to 1500 RF units (RFU).

**2.13 Determination of RpoS Stability.** Culture (35 ml) was removed aseptically from the bioreactor and placed in sterile 125 ml shake flasks. A 5 ml time zero sample was immediately removed and tetracycline (final concentration of 1.7 mM) was added to the

remaining culture which was incubated at 28 °C with shaking at 250 rpm. Samples were pelleted at 1,000×g in a model CL clinical centrifuge (International Equipment Co.) for exactly 2 min at 4 °C, and flash frozen in liquid $N_2$. Samples were lysed and subjected to Western analysis.

**2.14 Northern Analysis.** The hot phenol method (Sandercock and Frost, 1998) was used to purify RNA from *A. vinelandii*, which was subsequently stored in 1 vol isopropyl alcohol, 0.1 vol 3.0 M sodium acetate, pH 5.2, and 0.01 vol RNA guard (GE Healthcare) at -80 °C. RNA (5 μg) was separated on 4% formaldehyde agarose gels and transferred to Hyghbond N+ (GE Healthcare) membranes. *rpoS* template was created by PCR amplification using the primers WJP188 (5'-TGGGAATTCAGGGGACGACAACGAT GGCTG-3') and WJP189 (5'-CTCGAGCTCTTTTCACTGGAACAGCGCATCG-3'). [$\alpha^{32}$P] dCTP-labeled *rpoS* probe was produced from the template by random priming (Roche) and hybridized overnight at 65 °C. Membranes were washed at 68 °C in buffer B (0.5×SSPE, 0.1% SDS), exposed to a phosphor screen (Molecular Dynamics) overnight and imaged using a PhophorImager (Molecular Dynamics). Relative *rpoS* mRNA signal was normalized against a 16s rRNA probe, produced by [$\gamma^{32}$P] dATP-end labeling (Roche) the oligonucleotide (WJP16S 5'-CCGTCAATTCATTTGAGTTT-3'). The 16S probe was hybridized to the same membrane overnight at 51 °C, washed several times with buffer B at 51 °C, and imaged after a 15 min exposure.

**2.15 Identification of Proteins by Trypsin Digestion and LC/MS/MS.** Duplicate samples containing 80 μg of cell lysate were resolved electrophoretically on 7.5% non-denaturing PAG. Half of the gel was stained for catalase activity and the other half was stained using Coomassie blue R-250 (Sigma). The Coomassie stained protein bands corresponding to the activity bands were excised, de-stained, reduced by 37.1 mM DTT, alkylated with 32 mM iodoacetamide and digested with sequencing grade modified trypsin (Promega) overnight. The peptides were extracted from the gel and LC/MS/MS was performed using a CapLC HPLC (Waters) coupled to a Q-TOF-2 mass spectrometer (Waters). Data was analysed using Mascot (Matrix Science). Other than the initial electrophoretic separations and staining methods, all trypsin-LC/MS/MS related methods

were executed by Lorne Burke of the Institute for Biomolecular Design, University of Alberta.

**2.16 Purification of the Stationary Phase Catalase.** Cell lysates were heat-treated for 15 min at 75 °C, chilled on ice, then centrifuged at 5,000×g (SS-34 rotor, Sorvall) to remove heat-labile protein complexes that had precipitated out of solution. The supernatant was then concentrated by ultrafiltration (YM100, Millipore). The retentate was washed with 5 vol ice cold 50 mM Tris-HCl, pH 7.0, to facilitate the removal of low molecular weight proteins and peptides. The resultant solution was fractionated by FPLC (AKTA Explorer 100A, Amersham) using a Q-Sepharose column (Pharmacia). Gradient elution was done using 50 mM Tris-HCl, pH 7.0 containing 1M NaCl. Active fractions were quantitatively assayed for total antioxidant activity (catalatic or peroxidatic) by following the degradation of hydrogen peroxide spectrophotometrically at $A_{240}$ (Beers and Sizer, 1952). Fractions containing 33% of the peak activity were pooled, and concentrated by ultrafiltration (Amicon Ultra-15, Millipore). The solution was pre-filtered through a 0.8 μm filter under vacuum (type AA, Millipore) then separated by size exclusion chromatography (Sephacryl S-300, Pharmacia). The antioxidant activity of each fraction was assessed as detailed above. This degree of purification was sufficient for *de novo* protein identification of the enzyme. However, due to the presence of trace protein contaminants, the enzyme had to be further purified by ultrafiltration (Ultra4, Millipore) for enzyme kinetics studies. This step facilitated both protein concentration and a buffer exchange to 50 mM potassium phosphate, pH 7.0.

Crude protein concentrations were determined by the method of Lowry *et al.* (1951). Due to the relatively low percentage of Trp and Tyr residues in the catalase, protein concentrations of pure fractions were determined by the methods of Waddell (1956) and Wolf (1983).

**2.17 *De novo* LC/MS/MS Identification Following Proteinase K Digestion.** *De novo* identification was achieved by the proteinase K method of Wu et. al. (2003). An equal volume of 200 mM $Na_2CO_3$, pH 11 was added to an aliquot of purified protein. The sample was incubated on ice for 1 h and urea was added to a final concentration of 8 M.

The sample was reduced by 25 mM DTT and alkylated with 30 mM iodoacetamide. Digestion with proteinase K (Sigma) proceeded for either 3 h or 18 h with continuous mixing, and was stopped by the addition of 5% formic acid. LC/MS/MS analysis was performed using a CapLC HPLC (Waters) coupled to a Q-TOF Global Ultima mass spectrometer (Micromass). Data were processed using the automated peptide function of MassLynx 4.0 (Micromass) and analysed using both Mascot (Matrix Science) and PEAKS online (Bioinformatics Solutions). All proteinase-LC/MS/MS related methods were executed by Dr. Amanda Doherty-Kirby of the Siebens-Drake Research Institute, University of Western Ontario.

**2.18 Mass Determination by ESI-MS.** Molecular mass determination of the purified monomer by electrospray ionization (ESI) was performed on a Q-TOF2 mass spectrometer (Micromass) equipped with a Z-spray source and run in positive ion electrospray mode. Proteins were separated by microbore liquid chromatography (1 mm × 15 cm, Phenomenex) on a Waters CapLC system. Data were processed using the MaxEnt1 algorithm which is part of the MassLynx4.0 software suite (Micromass). All experimental methods were executed by Dr. Amanda Doherty-Kirby of the Siebens-Drake Research Institute, University of Western Ontario.

**2.19 Multimer Stability.** The stability of the enzyme was determined in the presence of urea and heat as previously described (Switala *et al.*, 1999). Briefly, urea was added (5.6 M, final concentration) to 200 ng purified enzyme in 50 mM potassium phosphate, pH 7.0. Samples were heated for 10 min (specific temperatures indicated in Results), then cooled to room temperature. Room temperature SDS-urea load dye was added and the samples were separated electrophoretically on 10% denaturing PAG. Protein was visualized using Coomassie R-250.

**2.20 N-terminal Sequence Determination.** Purified catalase was dissociated in denaturing load dye for 3 min at 95 $^{o}$C and separated on 7.5% SDS polyacrylamide gels that had been aged for at least 24 h. Protein was transferred to a PVDF membrane (Immun-Blot, BioRad) in 10 mM 3-cyclohexamino-1-propanesolfonic acid (CAPS), 5%

methanol, pH 11.0 buffer in a wet cell device (TE Transfor, Hoefer Scientific) for 16 h at 100 mA at 4 °C. The membrane was washed twice with 50% methanol and stained using Coomassie blue R-250 (BioRad). The band containing the catalase monomer (51 kDa) was excised and sent to Johns Hopkins University for analysis following Edman degradation on a Procise protein sequencing system (Perkin-Elmer).

**2.21 Spectral Determinations.** The UV-visual absorption profile of purified *Av*CCC was determined using a Unicam UV3-100 (ATI) spectrophotometer with an interval of 0.2 nm. The pH was titrated using 1 M NaOH. An unsuccessful redox titration of *Av*CCC at neutral pH was attempted by the step-wise addition of 30 μM sodium dithionite as previously described (Louie *et al.*, 1997). However reduction could only be achieved by the addition of sodium dithionite under alkaline conditions.

**2.22 Resonance Raman Spectroscopy.** The resonance Raman spectrum of purified protein was determined by measuring light scattered at 135° using a visible resonance Raman spectrometer consisting of a single monochromator (Spex) coupled to a CCD detector (Princeton). The light source was a Ti:Sapphire (Coherent) laser, with the frequency doubled to 412.5 nm to give maximal amplification near the Soret line. Photobleaching was minimized by rapid rotation of the sample in an NMR tube, and the UV-visual spectrum was examined after illumination to ensure that the enzyme was undamaged. All resonance Raman-related experimental methods were executed by Dr. Glen Loppnow, Department of Chemistry, University of Alberta.

**2.23 Electron Paramagnetic Resonance Spectroscopy.** Paramagnetic resonance of sphaeroplast-derived pure *Av*CCC was determined at a constant temperature of 10 °K (ESR900 flowing helium cryostat, Oxford Instruments) with a 10 $G_{pp}$ modulating field (Bruker EleXyo, Bruker Spectrospin) and 2 mW microwave power. Data were recorded and analysed using the software package Xepr (Bruker Spectrospin). The UV-visual spectrum was examined after each experimental stage to ensure that magnetic oversaturation had not occurred. All EPR-related experimental methods were executed by Dr. Richard Rothery, Department of Biochemistry, University of Alberta.

**2.24 Anti-oxidant Enzyme Assays.** Total anti-oxidant activity of partially purified and whole cell fractions was assessed by the method of Beers and Sizer (1952) which follows the degradation of hydrogen peroxide over time at 240 nm. Because the method measures only the loss of peroxide, it cannot discriminate between catalatic and peroxidatic activities.

Catalatic activity was specifically determined using a Gilson oxygraph equipped with a Clarke type electrode by the method of Rørth and Jensen (1967). Phosphate buffer (50 mM, pH 7.0) and substrate (60 mM $H_2O_2$, final concentration) were added to the reaction cell at 37 °C, followed by the addition of enzyme. For inhibition assays, the buffer, enzyme and inhibitor were mixed in the reaction cell at 37 °C for 2 min prior to the addition of substrate.

**2.25 Phylogenetic Analysis.** Protein sequences, with or without protein maturation prediction, were aligned by ClustalW using default settings. Alignments were imported to PAUP (version 4.0) and distance matrices were created using total character distance. Unrooted neighbour joining trees were calculated from total character difference algorithms and the topology of resultant trees were confirmed by the bootstrap method (100 iterations).

# CHAPTER 3

# Regulation of RpoS Expression and Activity in *Azotobacter vinelandii*

## 3.1 INTRODUCTION

The stress and starvation response of many Gram negative organisms is largely regulated by the alternative sigma factor, variably called $\sigma^{38}$, $\sigma^{S}$ or RpoS. Expression of this factor during stationary phase has been shown to regulate (positively or negatively) nearly 360 genes or 10% of the annotated genome of *Escherichia coli* (Patten *et al.*, 2004) while the homologue in *Pseudomonas aeruginosa* regulates 772 genes or 14% of the genome (Schuster *et al.*, 2004). Induction of such a large regulon is of particular interest in the pseudomonads, which can undergo morphological changes such as biofilm formation, allowing for improved colonization of the rhizosphere and other host surfaces Suh *et al.*, 1999; Xu *et al.*, 2001). Although the RpoS of *P. aeruginosa* has not been shown to directly regulate biofilm formation, it is notable that *rpoS* mutants produce modified biofilm structures (Heydorn *et al.*, 2002) and produce significantly less alginate than wild type cultures (Xu *et al.*, 2001). RpoS has also been shown to regulate the production of secondary metabolites and various pathogenicity determinants in pseudomonads (Suh *et al.*, 1999; Heeb *et al.*, 2005). *A. vinelandii*, an organism closely related to *P. aeruginosa* (Kennedy *et al.*, 2001), can differentiate into metabolically dormant cysts during stress conditions (Sadoff, 1975). In *A. vinelandii*, RpoS plays a role in the regulation of alginate (Page *et al.*, 2001) and poly-β-hydroxybutyrate (PHB) synthesis (Peralta-Gil *et al.*, 2002), which are necessary for the formation of mature cysts.

Catalase, peroxidase, and superoxide dismutase are used by aerobes of all kingdoms to degrade reactive oxygen intermediates (ROIs) produced during respiration. It is estimated that 0.1% of oxygen consumed by *E. coli* becomes the superoxide radical $\cdot O_2^-$ due to leakage from the electron transport chain to molecular oxygen (Imlay and Fridovich, 1991). This is considerably more efficient than mitochondria, in which 2-4% of consumed oxygen becomes $\cdot O_2^-$ (Scandalios *et al.*, 1997). However, because of its greater respiration rate, *E. coli* creates more ROIs overall (Imlay and Fridovich, 1991). Superoxide dismutases reduce $\cdot O_2^-$ to $H_2O_2$, so full protection most commonly necessitates the presence of catalase to remove the hydrogen peroxide. Considering that the *A. vinelandii* respiration rate far exceeds that of *E. coli* (Kelly *et al.*, 1990; Soballe and Poole, 1998) catalase expression in this organism is of particular interest. Since this

high respiration rate is widely credited for protecting the nitrogenase enzymes from irreversible inactivation by oxygen (Robson and Postgate, 1980) and because certain cytochrome components are derepressed in *A. vinelandii* during conditions of high oxygen tension *via* the FNR-like regulator, CydR (Wu *et al.*, 2000), there may well be a dedicated catalase isoenzyme that is specifically induced under high oxygen tension or nitrogen-fixing conditions.

During periods of stress or starvation, many organisms supplement their constitutively expressed "house keeping" catalase with an inducible enzyme. For example the RpoS-dependent and RpoS-independent catalases of *Pseudomonas putida* (CatB and CatC, respectively) are produced in stationary phase in addition to the constitutively expressed catalase CatA (Miller *et al.*, 1997; Miller *et al.*, 2001). Alternatively, *P. aeruginosa* modulates the expression of a single "house keeping" catalase (KatA) in response to varying growth conditions (Brown *et al.*, 1995), whereas a second catalase (KatB) is induced under oxygen-stress conditions by the regulator OxyR (Ochsner *et al.*, 2000). In a third scenario, *Pseudomonas syringae* pv. *glycinea*, a rhizosphere associated plant pathogen, expresses as many as 8 different catalases, some in a growth phase dependent manner (Klotz and Hutcheson, 1992).

In addition to stationary phase expression, RpoS induction is also observed when *E. coli* exhausts specific nutrients during growth in minimal medium (Pruteanu and Hengge-Aronis, 2002). The magnitude of RpoS induction is the same (~4-fold) whether the stress is a terminal starvation event such as glucose exhaustion (Pruteanu and Hengge-Aronis, 2002) or a transitional starvation event during diauxic-shift from glucose to lactose (Fischer *et al.*, 1998). Following the stress event, RpoS decreases to a basal level, whether starvation is alleviated or not. Equivalent studies of *Pseudomonas* spp. grown on minimal medium have not been done.

The first objective of this chapter was to examine the role of RpoS in the long-term survival of *A. vinelandii* under starvation conditions and conditions of ROI-stress. The second objective was to measure the stress responses during the unique nitrogen-diauxie experienced by the free-living diazotoph, which was not possible for any of the organisms in which RpoS had been studied to date. Furthermore, given the rapid respiration rates, and subsequently (theoretical) higher ROI loads, observed during

nitrogen fixation it was possible that additional or specific catalase induction might occur during this growth phase. The third objective of this chapter was to determine at what levels RpoS expression, and possibly transcriptional activity, were regulated in *A. vinelandii*. This was to be a preliminary study preceding the identification of specific molecular regulators of RpoS. Given the scope of regulation of the *E. coli* RpoS, involving over 20 individual regulatory moieties, it was not expected that all regulators would be identified within the scope of this thesis. Care was taken to measure the *rpoS* products directly by quantitative Western and Northern analysis, as a careful analysis of previously described transcriptional (Zgurskaya *et al.*, 1997) and translational (Stüdemann *et al.*, 2003) chimeric reporter systems indicated that they exhibit different stabilities than the native *rpoS* products (Mandel and Silhavy, 2005).

## 3.2 RESULTS

**3.2.1 Interruption of the *rpoS* Gene of *A. vinelandii*.** Primers WJP58 and WJP59 were used to amplify a 430 bp PCR product from within the *rpoS* gene of *A. vinelandii*. This PCR product was used as a probe to detect a 2.8 kb *Xho*I fragment of *A. vinelandii* chromosomal DNA containing the *rpoS* gene (pRPOS in Fig. 3.1). This fragment encoded the 3'-end of the *pcm* gene (encoding L-isoaspartate o-methyltransferase, 80% identical to the *P. aeruginosa* protein GenBank P45683), the complete *nlpD* gene (60% identity to the NlpD lipoprotein-precursor of *P. aeruginosa* GenBank P45682) and the 5'-end of *rpoS*. The sequence at the 3'-end of *rpoS* was amplified using primers WJP58 (hybrizing to *rpoS*) and WJP79 (hybridizing to *fdxA*; GenBank 63007). This 2.7 kb DNA fragment (pRPOS3') contained the 3'-end of *rpoS* and the *aerP* and *aerH* genes (Fig. 3.1), which are currently under investigation as aerotaxis transducers. This sequence has been deposited as GeneBank AF421351 and the order of these genes has been confirmed by whole genome sequencing (in progress, http://azotobacter.org).

The *rpoS* gene of *A. vinelandii* is 1005 bp long and encodes a protein of 334 amino acids with a predicted molecular weight of 38,306 Da. The *A. vinelandii* RpoS ORF is 87% identical to that of *P. aeruginosa* (protein accession NP_252312). Sigma factor regions 2.4 and 2.5, responsible for recognition of the -10 and extended -10

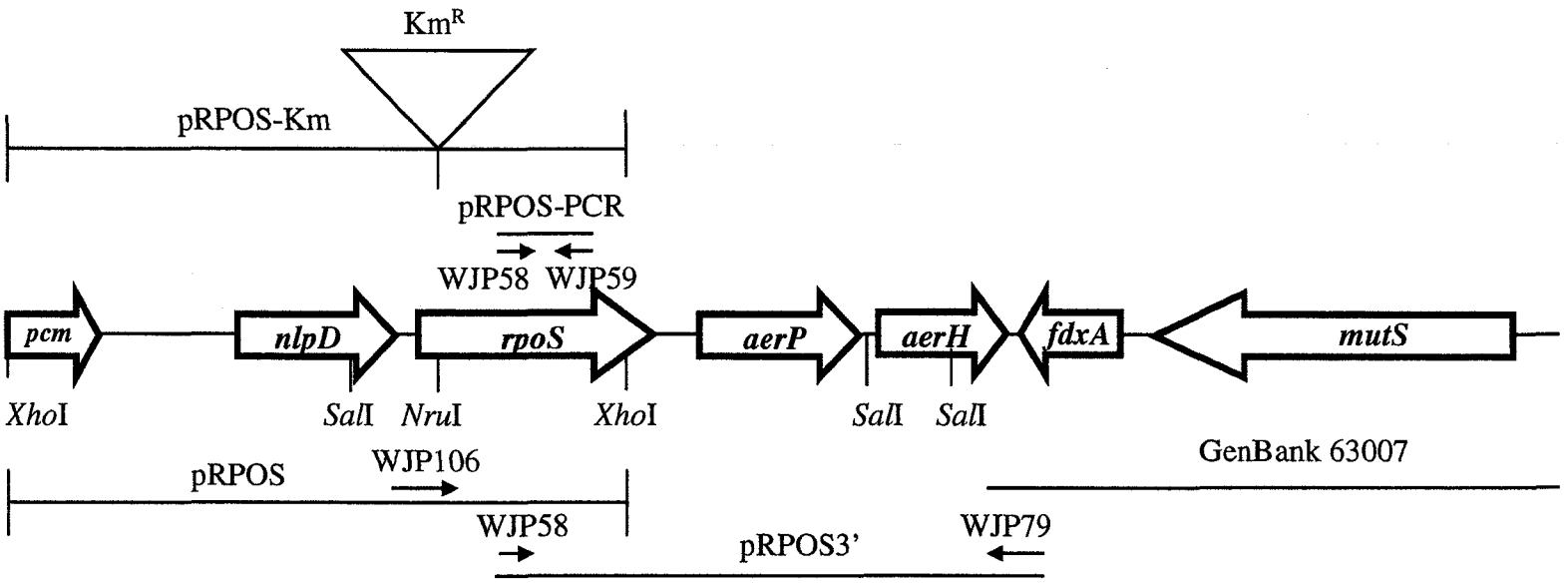



**Figure 3.1 Map of the *rpoS* Gene and the Surrounding Sequences.** Plasmid pRPOS was produced by cloning the *Xho*I fragment into pCR2.1-TOPO. The construct was cut by *Nru*I and the kanamycin resistance cassette ($Km^R$) was inserted, interrupting the *rpoS* open reading frame. Probes for southern analysis were created using the primers WJP58 and WJP59. The 3' end of *rpoS* and the highly variable domain between *mutS* and *rpoS* was cloned and sequenced using primers WJP58 and WJP79.

promoter sequences, were 100% identical, while region 4.2, which is responsible for recognition of the -35 promoter sequences, shared 96% identity. (Lonetto *et al.*, 1998; Fig 3.2).

**3.2.2 RpoS is Required During Starvation Survival.** The *rpoS* gene in pRPOS was disrupted with a $Km^R$ cassette. This was used to produce the *rpoS* strain UWS by homologous gene replacement. Both the wild-type strain UW and the mutant strain UWS had identical growth rates and growth curves in BBGN medium. The long-term survival of strains UW and UWS was tested on slants of solid BBGN medium. Strain UW stored in the dark at room temperature (21 °C) was culturable on BBGN after 16.5 yr, whereas UWS stored under identical conditions remained culturable for only 10 mo. However neither strain was viable if the medium became dry (this work, Page, 1983).

The strains were also grown for 2 d in shake flask culture until glucose became limiting, then cell survival was examined over an additional 8 d (Fig. 3.3). Strain UW cells suffered a 60.5% decrease in viability by the eighth day of starvation. In contrast, the viability of strain UWS decreased 99.5% over the same period of time.

**3.2.3 Expression of an RpoS-dependent Catalase Affects Oxidative Stress Survival.** Strains UW and UWS were grown from early-exponential phase (7 h) through to late-stationary phase (72 h) and treated with 15 mM $H_2O_2$ to determine their sensitivity to ROI stress. Both strains exhibited identical growth curves in shake flask cultures (Fig. 3.4A). Survival of strain UW was worst during exponential growth (at most 3.2% survival; Fig. 3.4B) and best during stationary phase (20 to 93% survival; Fig. 3.4C). In contrast, strain UWS survival was worst during stationary phase (at most 0.001%, Fig. 3.4C) and best during mid- to late-exponential phase (0.007 to 0.024%; Fig. 3.4B). This suggested that catalase activity was being expressed in an RpoS-dependent manner.

To determine if multiple catalases were expressed by *A. vinelandii*, cell lysates were analyzed for catalase activity after non-denaturing PAGE. Strain UW had two bands of catalase activity, designated Kat1 and Kat2, when assayed zymographically for 5 min (Fig. 3.5A). A third faint band of catalase activity (Kat3) was detected after 15 min (Fig 3.5A, right panel). Kat2 was expressed in early-exponential phase but not in 48-

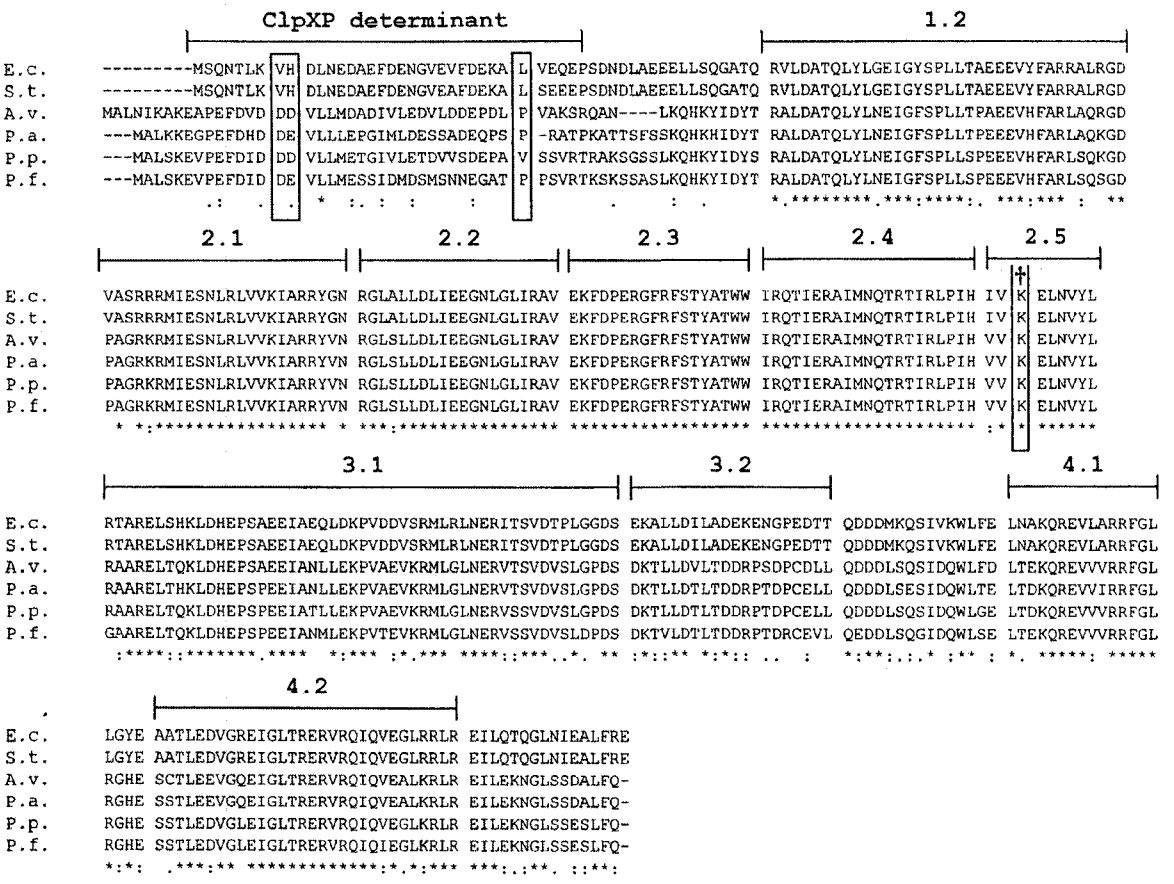



**Figure 3.2 Multiple Alignment of RpoS Sequences.** *E. coli* K-12 (E.c.), *Salmonella typhimurium* LT2 (S.t.), *A. vinelandii* UW (A.v.), *P. aeruginosa* PA01 (P.a.), *P. putida* KT2440 (P.p) and *P. fluorescens* Pf-5 (P.f.) RpoS sequences were aligned performed using ClustalW (Higgins et al.,1996). Accession numbers for the protein sequences used were NP_417221, AAI21804, AAL16091, NP_252312, NP_743780, and AAK30051. The regions of the sigma factors (1.2 through 4.2) are based on landmarks defined in Lonetto et. al. (1992). The essential residue for RssB binding (K173; Becker et al., 1999) is indicated by the cross symbol (†). Residues important for ClpX recognition (V8, H9 and L30; Studemann et al., 2003) are boxed. The "*" symbol indicates complete residue conservation, ":" indicates strong group conservation, "." indicates weak group conservation and blanks indicate no conservation of residues.

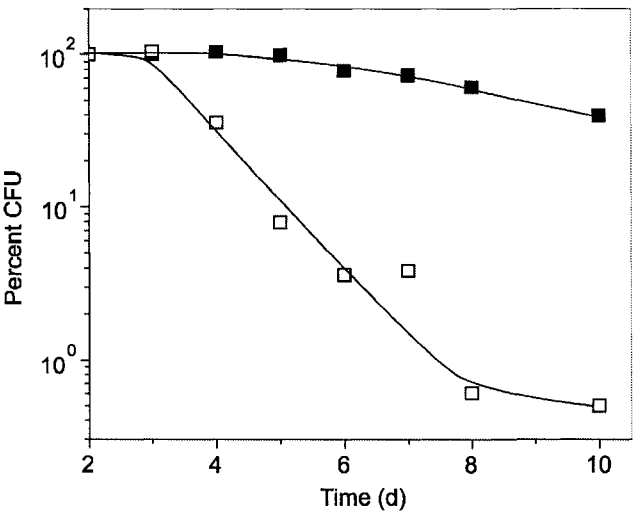



**Figure 3.3 The Effect of RpoS on Viability During Starvation.** Survival of *A. vinelandii* strains UW (■) and UWS (□) when grown in liquid BBGN media over 10 d. Survival is expressed as a percent of the CFU/mL at day 2, when glucose became limiting.

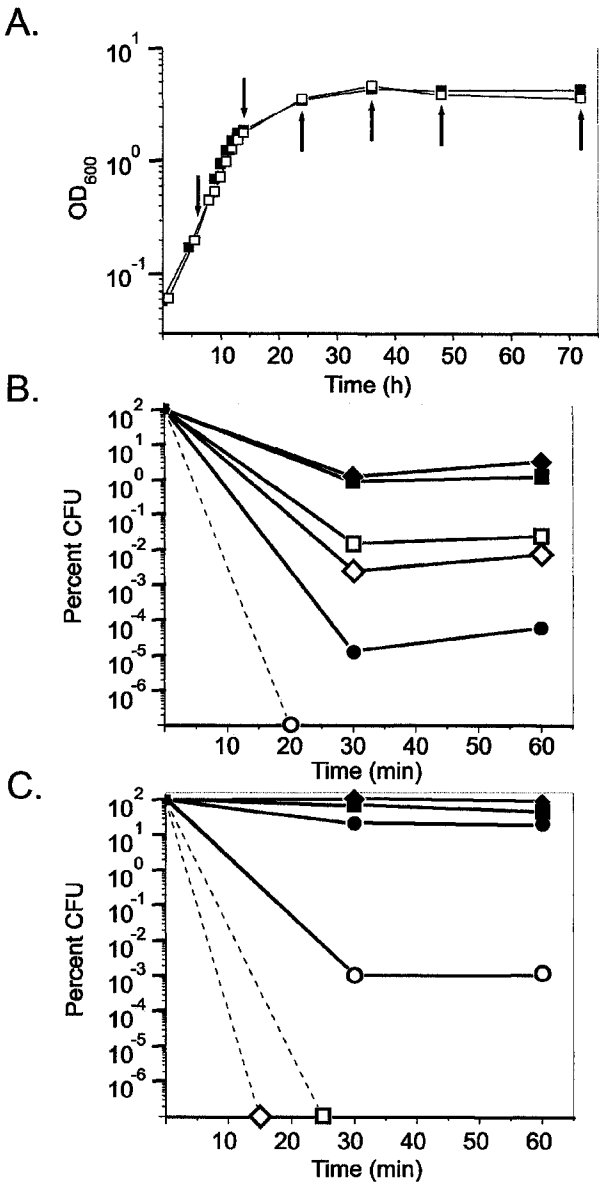



**Figure 3.4 The Effect of RpoS on Viability During Peroxide Challenge.** Percent survival of strains UW (closed symbols) and UWS (open symbols) following treatment with 15 mM $H_2O_2$. Sampling times are indicated by arrows in the growth curve (A). The cultures were normalized to an $OD_{600}$ of 1.0 in BB medium and the percent survival of 7 (●), 14 (◆), and 24 (■) h cultures before and after peroxide treatment are shown in (B). Survival of 36 (●), 48 (◆), and 72 (■) h cultures are shown in (C). Dashed lines indicate that no colonies forming units were recovered following 30 min of peroxide treatment.

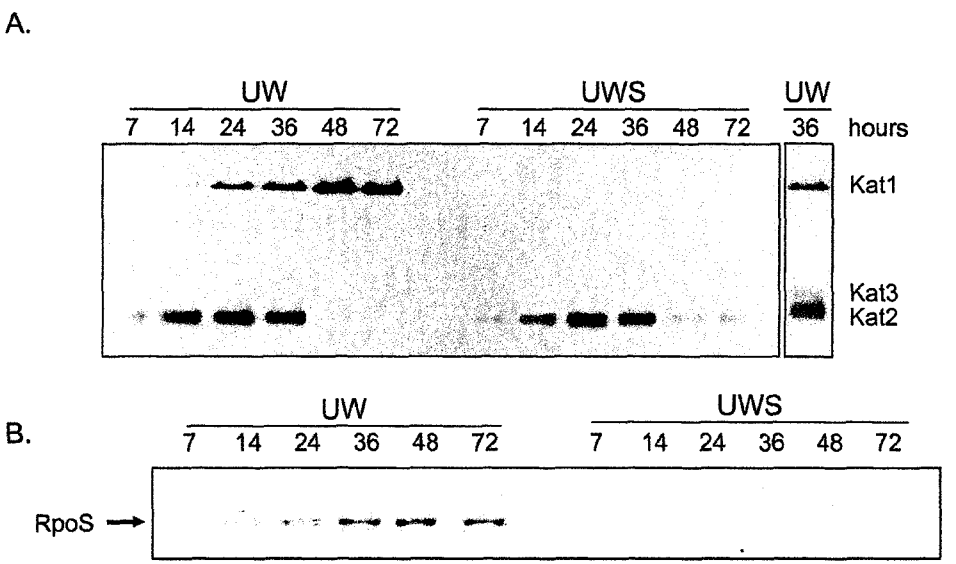



**Figure 3.5 Zymographic Analysis of the Catalases of *A. vinelandii*.** Catalase activity of cell lysates from cultures in Fig. 3.4A. *In situ* catalase zymographs of 50 μg of cell lysate from strains UW and UWS, separated by non-denaturing PAGE are shown in panel A. Two strong bands (Kat1 and Kat2) were detected after a 5 min exposure to diaminobenzidine. Longer exposure times allowed for the detection of an additional catalase band, Kat2. The same lysates (20 μg) were separated by denaturing PAGE and analyzed for RpoS by Western analysis (B). Sample times (in hours) are indicated above each lane.

72 h stationary phase cells. Conversely Kat1 was not expressed in early-exponential phase, but increased in concentration during later phases. The absence of Kat1 in the *rpoS* mutant indicated that this catalase was expressed in an RpoS-dependent manner. The mutant exhibited wild-type expression of Kat2 during exponential growth, but continued to produce low levels of Kat2 in late-stationary phase (Fig. 3.5A).

The apparent anomaly of expressing an RpoS-dependent catalase during exponential growth was explained by Western analysis (Fig. 3.5B). RpoS was detected in strain UW in mid-exponential phase and reached maximum levels by 36-72 h, which coincided with the pattern of Kat1 expression. RpoS was not detected in strain UWS at any time (Fig. 3.5B).

**3.2.4 RpoS Expression During Acetate to Glucose Diauxic Shift.** The expression of an RpoS-dependent catalase and of RpoS itself during exponential growth was unexpected, as a previous study (Castañeda *et al.*, 2001) reported that *A. vinelandii rpoS* mRNA expression occurred only in late-stationary phase. To determine what event might be responsible for RpoS induction during exponential growth, we grew strain UW in batch culture in a bioreactor, then followed RpoS concentration and nutrient availability over 48 h. The use of the bioreactor system allowed for optimal aeration and mixing, thereby avoiding oxygen limitation which occurs when *A. vinelandii* is grown in shake flask culture (Page and Knosp, 1989).

Strain UW used acetate in the BBGN medium preferentially during early-exponential phase with a mean generation time (g) of 2.48 h (specific growth rate of 0.28 $h^{-1}$; Fig. 3.6A). During this time $dO_2$ levels in the bioreactor decreased, indicative of active metabolism. Acetate concentrations decreased rapidly while glucose was consumed at a very slow rate, typical of the carbon source preference of *A. vinelandii* (George *et al.*, 1985). A rapid increase in $dO_2$ concentrations at 8.6 h signaled the cessation of metabolism due to acetate depletion (denoted by an arrow, Fig. 3.6A). Rapid metabolism of the non-preferred carbon source glucose began at 9.5 h and continued into stationary phase.

The Western analysis of cell lysates taken from the fermentor culture can be seen in Fig. 3.7. The RFU trend, plotted in Fig. 3.6A, showed that RpoS was present in very

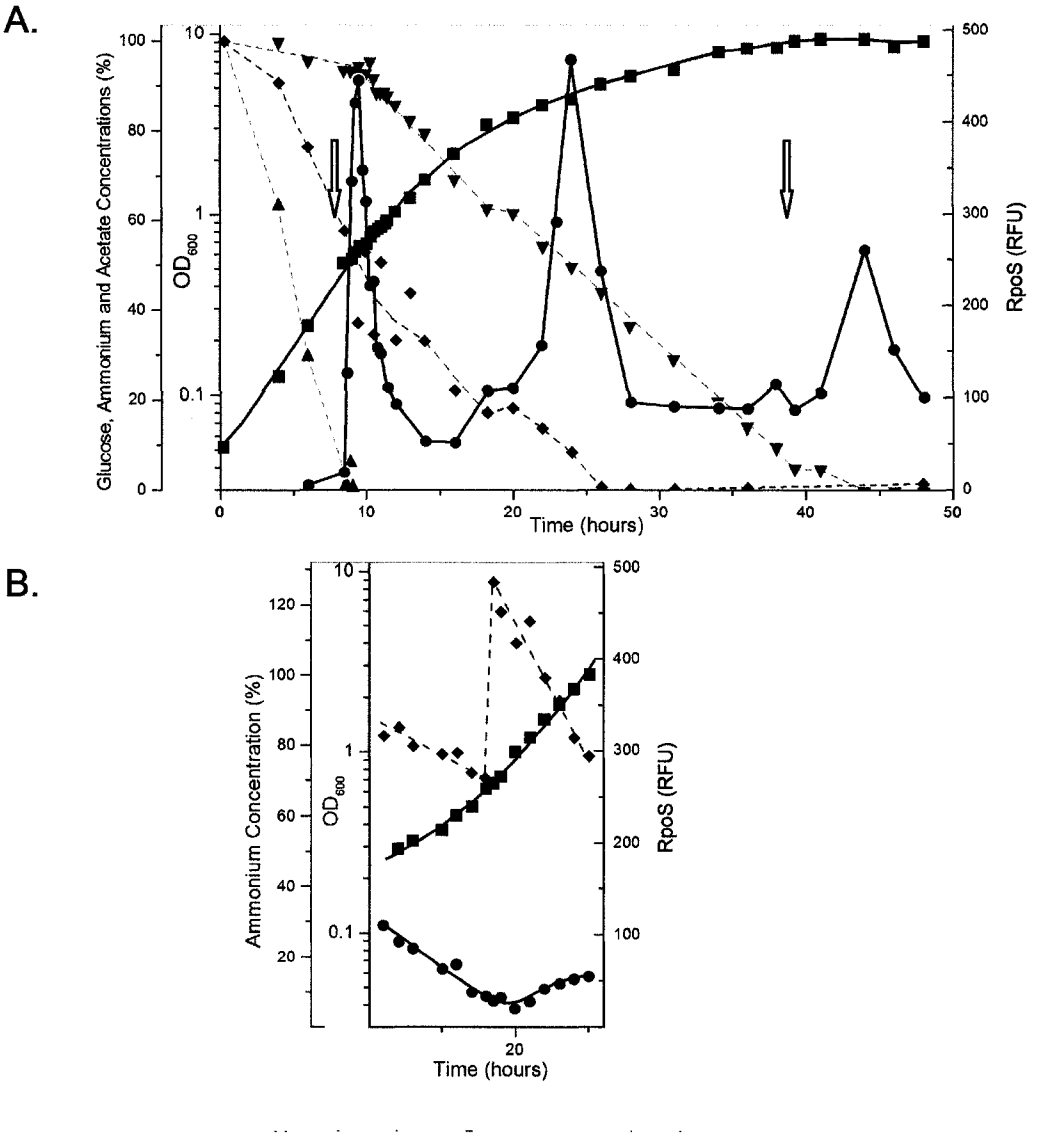



**Figure 3.6 Growth and Stress Responses of Strain UW in BBGN medium in a 2.5 L Bioreactor.** Growth was determined optically ($OD_{600}$; ■). Acetate (▲), ammonium (◆), and glucose (▼) concentrations are expressed as a percentage of their respective initial concentration. Periods of oxygen saturation, indicative of metabolic cessation, are designated by the arrows. The relative amount of RpoS protein was determined by quantitative Western analysis (●). (A) Data from a representative batch culture grown for 48 h. (B) Data from an ammonium-fed culture, with $(NH_4)_2SO_2$ added immediately after 18 h. Panel B is aligned directly below the equivalent times in panel A.

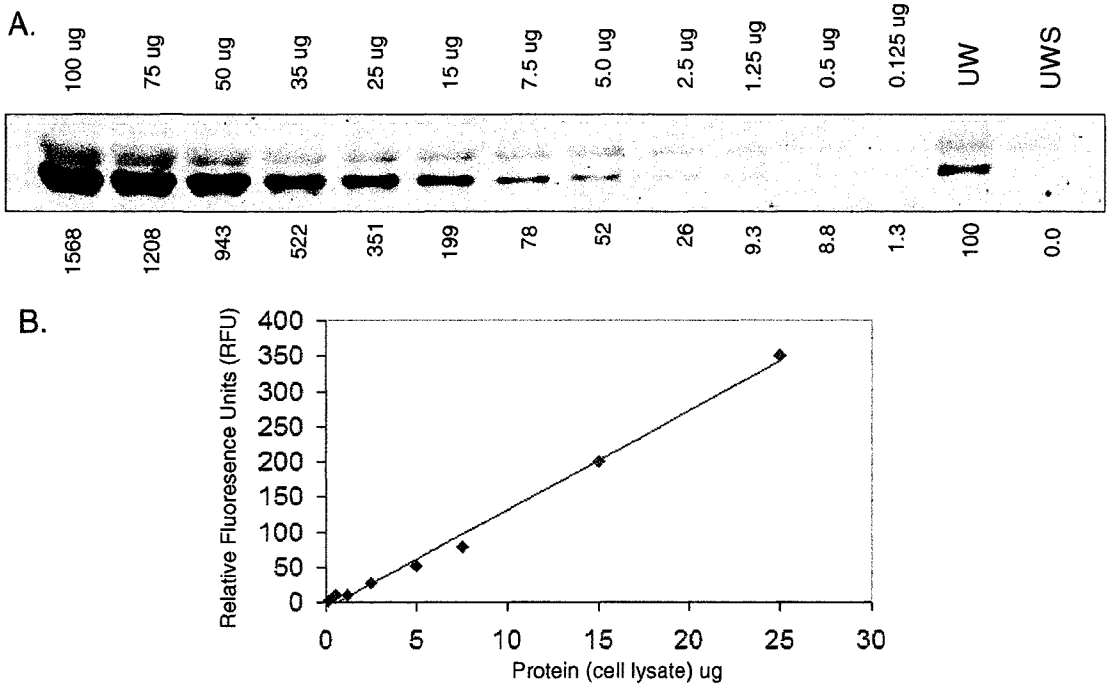



**Figure 3.7 Validation of Quantitative Western Analysis of the RpoS Protein.** (A) Different quantities of whole cell lysate from 24 h of growth were loaded into each lane, and the Relative Fluorescent Units (RFU) of RpoS signal strength were calculated. Strain UWS was used to define zero signal. (B) Graphical representation of the data: the $R^2$ value for RFU per protein mass was calculated to be 0.997, comparable to previously reports for quantitative RpoS Westerns ($R^2 = 0.97$; Jishage and Ishihama, 1995).

low quantities (0 to 17 RFU) during early exponential growth. The RpoS concentration increased within 9 min of acetate depletion and reached 443 RFU within 1 h, a 26-fold induction. RpoS concentration declined to a basal level of 51 RFU by 14 h (Fig. 3.6A), coincident with the recovery of metabolic activity on glucose. *A. vinelandii* resumed growth on glucose with a reduced generation time of 3.8 h.

**3.2.5 RpoS Expression During Ammonium to $N_2$ Diauxic Shift.** At 18 h the culture growth rate decreased (g = 11.3 h) and RpoS levels increased 28-fold (relative to early exponential phase) by 24 h (Fig. 3.6A). Although ammonium levels were not completely depleted by 18 h, it is possible that the cells experienced nitrogen-limitation, since there was a significant decrease in growth rate. RpoS concentration declined rapidly after the 24 h point to a new basal level of 92 RFU.

**3.2.6 RpoS is not Regulated by Oxygen-Limitation.** At the 24 h time point the $dO_2$ of the medium was below 1% saturation, indicating that oxygen was being consumed as fast as it could be added to the culture. To determine if the observed RpoS response was due to nitrogen limitation or oxygen limitation, an ammonium-augmentation experiment was performed (Fig. 3.6B). The culture was incubated for 18 h in BBGN medium, at which time glucose, ammonium, and oxygen were still in excess. Ammonium sulfate was added such that ammonium remained in excess for the duration of the experiment. RpoS levels did not increase, even though the $dO_2$ levels dropped below 1% saturation by 20 h. From this we conclude that the production of RpoS starting at 18 h in Fig. 3.6A was due to ammonium, rather than oxygen, limitation.

**3.2.7 RpoS Expression During Nitrogen Fixation and Glucose Depletion.** After ammonium depletion, the cells were nitrogen-fixing from 27 to 40 h, using glucose as the carbon and energy source. Since a significant amount of energy is required for nitrogen fixation (George *et al.*, 1985) it seemed likely that the organism would undergo a stress response during this period. The growth rate was greatly decreased (g = 31 h) but the RpoS concentration changed very little during this period (92 to 95 RFU).

Glucose was depleted at approximately 42 h with a coincident increase in $dO_2$ indicating cessation of active metabolism (denoted by an arrow, Fig. 3.6A). RpoS levels increased then returned to basal levels within 6 h. Although not measured in these experiments, it is likely that the cells switched to the metabolism of endogenous poly-hydroxybutyrate (PHB) reserves in preparation for encystment (Page, 1983).

**3.2.8 Catalase Expression During Growth on BBGN.** To determine if the RpoS produced by *A. vinelandii* during the metabolic shifts was transcriptionally active we used Katl activity as a reporter of RpoS-dependent transcriptional activity. RpoS and Katl levels were compared from selected time points from the bioreactor culture (Fig. 3.8A and B respectively). The first appearance of Katl coincided with the induction of RpoS during the acetate to glucose diauxic shift at 9.5 h. Katl activity continued to increase until 14 h, although RpoS levels had decreased hours before this. Similarly an increase in Katl activity occurred at 24 h, coincident with the ammonium to $N_2$ diaxic shift, but high levels of Katl activity persisted for an additional 14 h, despite a 4-fold decrease in RpoS in the same time-frame. A comparison of RpoS concentrations and Katl activities at 38 h (nitrogen fixing conditions) and 44 h (following glucose depletion) showed a similar trend (high Katl:low RpoS). Although the catalase staining method is not strictly quantitative, comparison of the relative catalase (Fig. 3.8B) and RpoS (Fig. 3.8A) band strengths at these time points does indicate that there is a non-linear relationship between the RpoS concentration and Katl activity in *A. vinelandii*.

Contrary to what was observed in the shake-flask study (Fig. 3.5A) Kat2 expression occurred throughout the bioreactor experiment and a third catalase (Kat3) became more evident in stationary phase (34 to 48 h; Fig. 3.8B).

**3.2.9 RpoS Stability During Various Growth Conditions.** The stability of the RpoS protein was determined by inhibiting protein synthesis with tetracycline, followed by Western analysis of the sigma factor. The total amount of RpoS produced by the culture during the acetate to glucose diauxie and the specific times at which samples were taken for half-life determinations are shown in Fig. 3.9A. Cultures were incubated in the

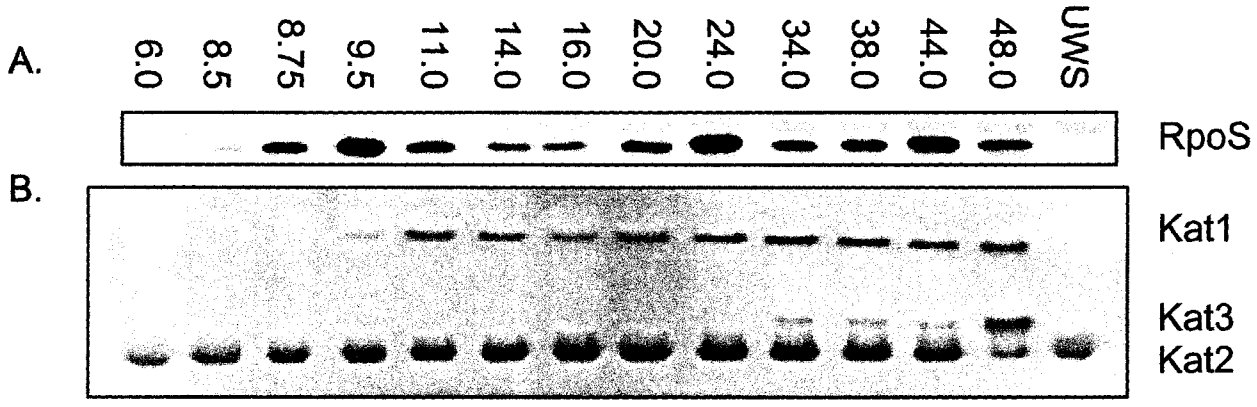



**Figure 3.8 RpoS and Catalase Expression During Stress Responses.** Selected time points from the bioreactor culture in Fig. 3.6A were analysed for RpoS and catalase activity. (A) Western analysis of RpoS from 25 μg cell lysate separated by SDS-PAGE. (B) Catalase zymography of 50 μg cell lysate separated by non-denaturing PAGE. The sample times (in hours) are indicated above each lane.

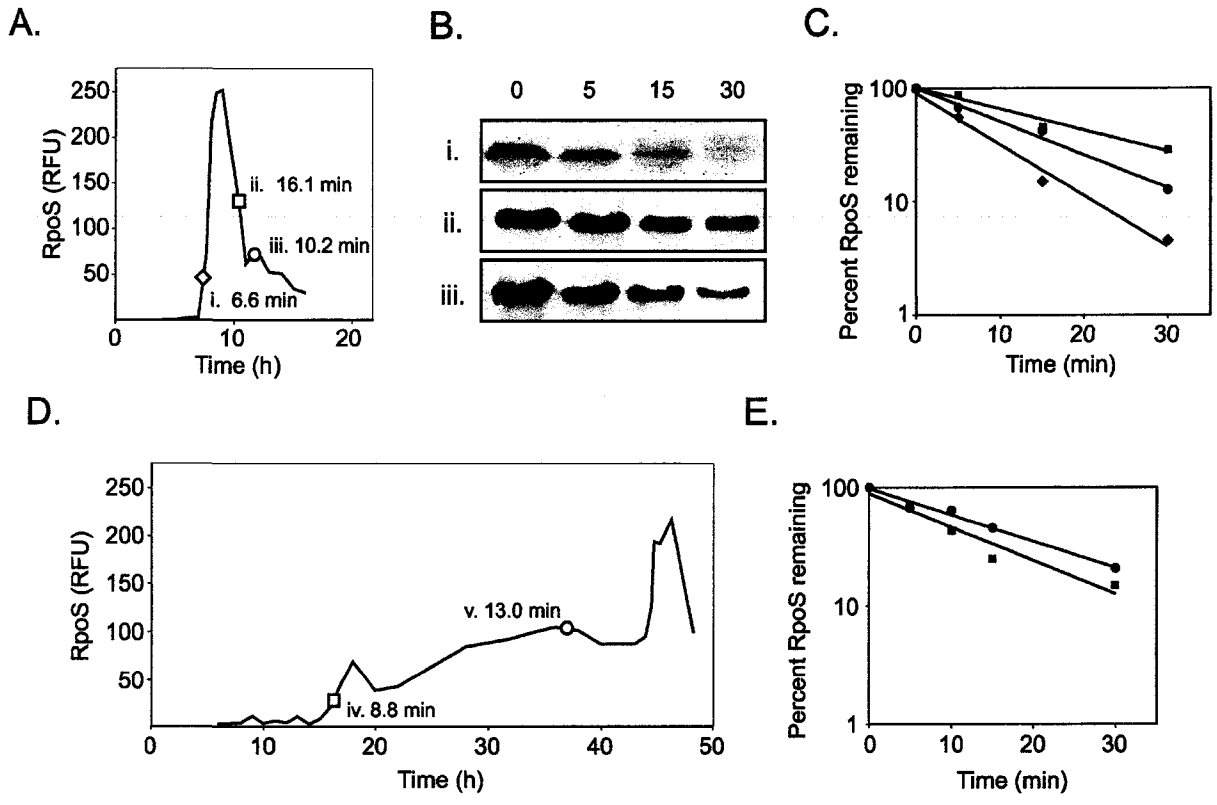



**Figure 3.9 RpoS Stability (Half-Life) Determinations.** Samples were taken during various times throughout bioreactor growth of *A. vinelandii* in BBGN and BBG media. Relative RpoS expression during an acetate to glucose diauxic shift is shown in panel A. Samples were removed during acetate limitation (i), diauxic lag (ii), and growth recovery on glucose (iii), then treated with tetracycline to stop translational initiation. Western analysis of RpoS following tetracycline addition is shown (B). The relative amount of RpoS following treatment is plotted graphically (C) as a percentage of the RpoS present prior to tetracycline addition, and the calculated half-lives are indicated in panel A. RpoS expression during growth on BBG is shown in panel D. Samples were removed at two time points (iv, v) and treated with tetracycline. The relative amount of RpoS following treatment is plotted as a percentage of the RpoS present prior to tetracycline addition (E) and the calculated half-lives are indicated in panel D.

presence of tetracycline for up to 30 min, lysed, and analysed by Western analysis (Fig. 3.9B). A graphical representation of the RpoS stability is shown in Fig. 3.9C. During the acetate to glucose diauxic shift, half-lives of 6.6 to 16.1 min were observed at the time points indicated in Fig. 3.9A.

RpoS expression and stability during growth on BBG medium under identical aeration, mixing, and temperature conditions were also determined (Fig. 3.9D). It is interesting to note that RpoS levels were low for the first 15 h of growth, but then steadily increased over the next 30 h, an observation that directly contradicts the observations by Castaneda et. al. (2001), which suggested that *rpoS* expression occurred strictly during stationary phase. Half-lives of 8.8 to 13.0 min were determined for the time points indicated (Fig. 3.9D) and the RpoS concentrations are plotted in Fig. 3.9E.

**3.2.10 *rpoS* mRNA and RpoS Expression are Asynchronous.** Both the relative RpoS protein and *rpoS* mRNA levels were followed during growth on BBGN in a bioreactor batch culture (Fig. 3.10). Relative *rpoS* mRNA concentrations were calculated using the internal control 16s rRNA (Fig. A, Appendix). During the initial acetate to glucose diauxic shift, the *rpoS* mRNA levels increased just prior to RpoS levels, and likewise decreased slightly in advance of the RpoS decline. Likewise, during ammonium limitation (16 h), *rpoS* mRNA induction preceded RpoS accumulation. However, while RpoS levels subsided during nitrogen fixing growth (see also Fig. 3.8A), the *rpoS* mRNA levels continued to increase, reaching maximal induction (16-fold) by 30 h. The *rpoS* transcript levels remained elevated for an additional 17 h, until glucose became limiting. Following glucose limitation the *rpoS* mRNA levels decreased dramatically, yet RpoS levels increased slightly in the same time-frame.

## 3.3 DISCUSSION

**3.3.1 RpoS and Stationary Phase Survival.** The stationary phase sigma factor RpoS is needed for the long-term survival of *A. vinelandii* and for resistance to reactive oxygen intermediates like hydrogen peroxide. Vegetative growth of wild-type and *rpoS* strains was essentially identical, but the stationary phase survival characteristics were different.

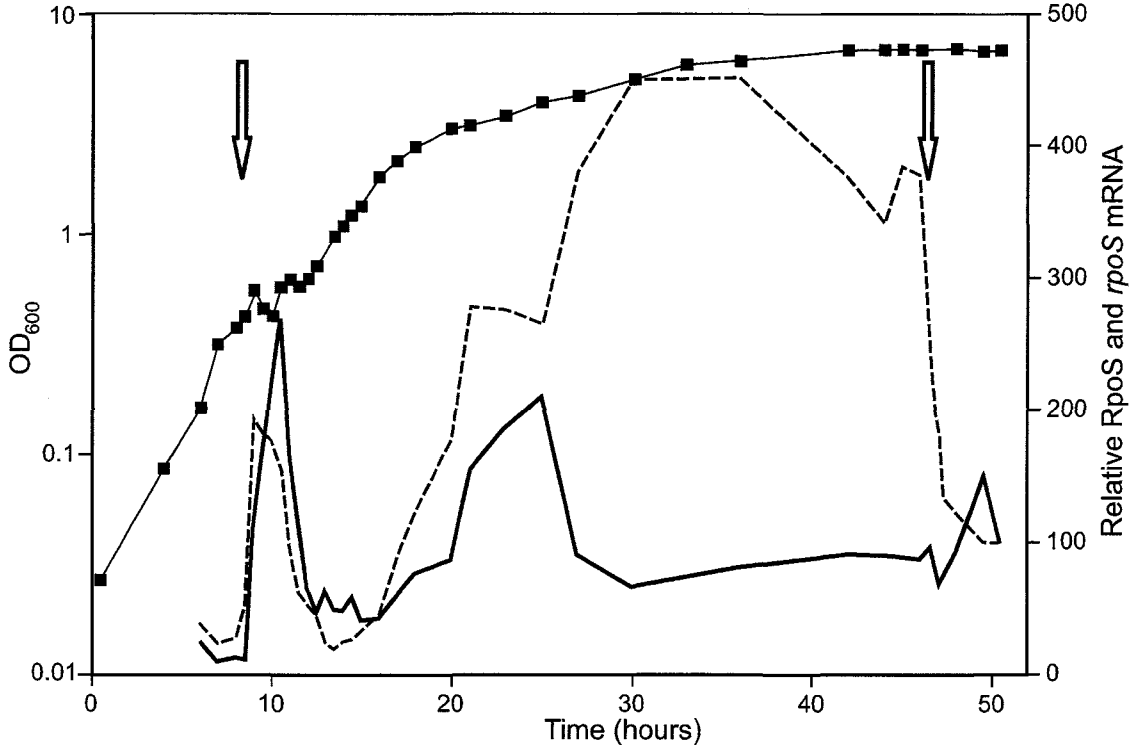



**Figure 3.10 *rpoS* mRNA and RpoS Expression During Growth on BBGN medium in a 2.5 L Bioreactor.** Northern and Western analyses of the products *rpoS* mRNA (dashed line) and RpoS (solid line) from *A. vinelandii* UW cultures grown on BBGN. Bacterial growth was determined spectrophotometrically at 600 nm (■). Periods of oxygen saturation, indicative of metabolic cessation, are designated by vertical arrows.

It is tempting to say that the wild-type strain UW formed cysts while strain UWS did not. However, while old cultures of strain UWS contained no cyst-like structures when viewed under the electron microscope (data not shown), neither does strain UW form fully 'mature' cysts (Page, 1983). The RpoS mutant survived as vegetative cells for about 10 mo, even though the PHB reserves of wild-type cells are expected to last only 2.5 mo (Aladegbami *et al.*, 1979). However, the wild-type was able to survive at least 16.5 yr, suggesting that RpoS is necessary for the development of a dormant cell, independent of cyst formation.

RpoS is required for the survival of stationary phase *A. vinelandii* exposed to hydrogen peroxide. Five putative catalase genes are predicted to be present in the *A. vinelandii* genome, but only three appear to be expressed or functional. It is interesting that despite the high respiratory rate in *A. vinelandii*, it has a relatively low tolerance to exogenous $H_2O_2$ compared to members of the closely related *Pseudomonadaceae* (Jørgensen *et al.*, 1995; Heeb *et al.*, 2005). Mounting evidence suggests that catalases are primarily produced to deal with exogenous peroxide threats, while endogenously created peroxides are removed mainly by peroxidases (Park *et al.*, 2005). Many pathogenic and strongly rhizosphere associated pseudomonads produce large quantities, and in some cases several types, of catalases (Klotz *et al.*, 1992; Kang *et al.*, 2004) for protection from host defenses and inter-species competition. As a free-living organism (Kennedy *et al.*, 2001) *A. vinelandii* may not have developed this degree of ROI defense because it is less likely to encounter exogenous sources of hydrogen peroxide. *A. vinelandii* survival in the presence of hydrogen peroxide was best during stationary phase, which correlated to both RpoS expression and elevated activities of the RpoS-dependent catalase.

In contrast to most bacteria described to date, which either supplement a constitutively expressed catalase with additional catalase species (Miller *et al.*, 1997; Kang *et al.*, 2004) or up-regulate the housekeeping catalase during stationary phase (Brown *et al.*, 1995), *A. vinelandii* appears to undergo complete catalase switching from Kat2 in the exponential phase to Kat1 in the stationary phase. We also did not observe the expression of a phase-specific catalase during or following the transition to nitrogen fixing conditions. However, the increased production of Kat2 and Kat3 at higher aeration does suggest that these enzymes are ROI-inducible, as described for many other catalases

(Ochsner *et al.*, 2000; Chelikani *et al.*, 2004). The reciprocal production of Kat3 at the expense of Kat2 suggests that it may be an isoform of Kat2, as previously observed for other KatG species.

**3.3.2 Regulation of RpoS Expression.** The expression of RpoS in *A. vinelandii* during acetate to glucose diauxie was similar to that described in *E. coli* during the glucose to lactose shift (Fischer *et al.*, 1998), except that the magnitude of induction was much greater in this study. *rpoS* mRNA and protein concentrations rose and fell synchronously, indicating that RpoS accumulation is primarily regulated at the transcriptional level during the acetate to glucose diauxie. Since *A. vinelandii* is able to fix nitrogen, it was used to examine the general stress response during a nitrogen-diauxic shift and subsequent nitrogen-fixing growth. Up to 28-fold induction of RpoS and 16-fold induction of *rpoS* mRNA occurred during the nitrogen diauxic shift. Thus it appears that *A. vinelandii* accumulates similar amounts of RpoS whether it experiences carbon- or nitrogen-limitation. However, during diauxic recovery on $N_2$ the RpoS levels decreased to basal levels even though the mRNA levels continued to increase. Because RpoS stability did not change appreciably in *A. vinelandii* (discussed below), nor has it been observed to do so in other pseudomonads (Venturi, 2003), the most probable interpretation of the RpoS/*rpoS* asynchrony is that a significant degree of post-transcriptional regulation occurs in *A. vinelandii* under certain nutrient-limiting growth conditions.

RpoS levels in *E. coli* are regulated by a dedicated proteolytic circuit involving the recognition factor SprE/RssB (Pratt and Silhavy, 1996). RssB binds to RpoS at the critical residue K173, then transfers it to the protease ClpXP for degradation (Becker *et al.*, 1999). This residue and the surrounding regions 2.4 and 2.5 of *E. coli* RpoS are conserved (96% identity) in *A. vinelandii*, despite an overall protein identity of only 76% (Fig. 3.2). No *rssB* homologue was observed in the *A. vinelandii* genome, despite the presence of an *rssA* homologue (accession ZP_00418183). Whereas RssB-directed regulation in *E. coli* modulates RpoS stability nearly 38-fold (half-lives from 4.1 to 154 min) during carbon-limitation (Mandel and Silhavy, 2005), stability changed a mere 1.4- to 2.4-fold during the nitrogen- and carbon-limiting events tested with *A. vinelandii*.

Therefore it appears that no RssB-like proteolytic regulatory mechanism exists in *A. vinelandii*. Consequently the conservation of RpoS regions 2.4 and 2.5 in these divergent genera is most likely driven by sequence recognition of the -10 promoter sequence.

**3.3.3 Regulation of RpoS Activity.** The efficiency of RNA polymerase:RpoS holoenzyme formation and subsequent transcriptional activity was assessed using the natural reporter, Katl. Considering that RpoS has the lowest affinity for RNA polymerase of all the *E. coli* sigma factors (Maeda *et al.*, 2000) it would be expected that the high magnitude induction of RpoS during the various diauxic shifts would result in increased RNA polymerase:RpoS formation and resultant activity (Malik *et al.*, 1987), and thus greater expression of the reporter Katl. However, while Katl activity was only observed during periods of RpoS expression, its activity was not directly proportional to RpoS levels across the entire growth curve. In particular the Katl levels were similar during periods of peak RpoS expression (24 h, Fig. 3.6) and during periods of basal RpoS expression (34, 38 and 48 h, Fig. 3.6). Similar phenomena have been observed in *E. coli*, where effector molecules such as ppGpp and trehalose increase the formation or activity of the RpoS holoenzyme. From these observations we report that the RpoS of *A. vinelandii* is regulated at the transcriptional, post-transcriptional and formation/activity levels, but not at the post-translational level.

# CHAPTER 4

## Identification of two catalases from *Azotobacter vinelandii*: a KatG homologue and a novel bacterial cytochrome *c* catalase, *Av*CCC.

## 4.1 INTRODUCTION

The obligate aerobe *Azotobacter vinelandii* is able to fix dinitrogen gas during periods of nitrogen-source limitation (Lipman, 1903), or differentiate into quiescent cysts during periods of carbon-source depletion (Winogradsky, 1938). The rapid respiration rate of *A. vinelandii* relative to other bacteria during exponential growth during diazotrophic or "normal" growth has been well documented (Kelly *et al.*, 1990; Soballe and Poole, 1998; Williams and Wilson, 1954) and the elevated rate could be expected to result in the production of large quantities of reactive oxygen species (ROS). In contrast to exponential growth, the metabolism of encysted *A. vinelandii* is very low (Sadoff, 1975). From these observations it can be seen that *A. vinelandii* experiences two metabolic extremes. In one circumstance the ROS load is potentially very high, but active metabolism allows for the rapid production and maintenance of protective enzymes. In the other, the ROS load is low, but the quiescent state of the cell is not amenable to enzyme synthesis. Since the cyst-like structures remain viable for years (Page, 1983; Chapter 3) one can postulate that the quiescent organism either actively produces enzymes to deal with toxic oxygen products or that it synthesizes very stable enzymes which require infrequent replacement.

Several classes of catalases have been described. The monofunctional heme catalases tend to have very high catalatic rates but demonstrate no peroxidatic activity. This group is further divided into the small and large subunit categories. The small subunit catalases tend to be more heat resistant, and in the case of *Escherichia coli* KatE (*Ec*HPII), also protease resistant (Chelikani *et al.*, 2003). The non-heme catalases have a manganese rather than iron cofactor at the active site. These enzymes tend to be heat stable, and with one notable exception (Amo *et al.*, 2002), have lower catalatic activities than the heme based monofunctional catalases (Klotz and Loewen, 2003). Finally, the bifunctional catalases, frequently referred to as HPI or KatG, exhibit both strong catalatic activity and moderate peroxidatic activity.

The bacterial cytochrome *c* peroxidases (BCCPs) constitute another class of antioxidant enzyme. First identified in *Pseudomonas fluorescens* (Lenhoff and Kaplan, 1953), BCCPs reduce hydrogen peroxide to water peroxidatically using small protein

donors such as cytochrome *c*, azurin, or pseudoazurin. Members of this class of proteins contain two type-*c* heme groups with different redox potentials. Each heme group plays a distinct role in peroxide reduction (Ronnberg and Ellfolk, 1979; Pettigrew *et al.*, 2003), requiring the communication of electrons from one heme site to the other. Because heme *c* groups are covalently bound to the protein, they can be distinguished from non-covalently bound heme *b* and *d* groups, such as those found in all the hemic catalases described to date.

It was shown previously that *A. vinelandii* produces two catalase proteins when grown on the minimal medium BBGN (Chapter 3). One of these is expressed during exponential growth, whereas the other catalase is expressed in an RpoS-dependent manner during periods of nutrient limitation and stationary-phase. The objectives of this chapter were to identify each catalase by tandem mass spectroscopy (LC/MS/MS) following protease digestion. The identification of the stationary phase catalase would allow us to use its mRNA as a natural and quantitative metric of $E\sigma^S$ formation/activity. This catalase was found to be a novel heme *c*-type catalase, and was further characterized based on predictions from its primary residue sequence. To prove that the predicted ORF was responsible for the catalase activity, the gene was interrupted in the genome by the insertion of a resistance cassette. Finally, the mutants were assessed for their stress survival when challenged with hydrogen peroxide.

## 4.2 RESULTS

**4.2.1 The Stationary Phase Catalase is Thermostable and Monofunctional; the Exponential Phase Catalase is Heat-Labile and Bifunctional.** The heat stability of the two catalases of *A. vinelandii* was examined. Whole cell extracts of *A. vinelandii* cultures from mid-exponential phase were heated for 30 min at various temperatures, then rapidly cooled on ice. The lysates were separated by 7.5% non-denaturing PAGE and the gel was subsequently stained for catalase activity (Fig. 4.1A). The catalase exhibiting greater electrophoretic mobility, Kat2, was heat-labile above 46 °C, whereas Kat1 retained function following heat treatments as high as 90 °C. Increased temperatures shifted the migration pattern of Kat2 to the previously identified form, Kat3

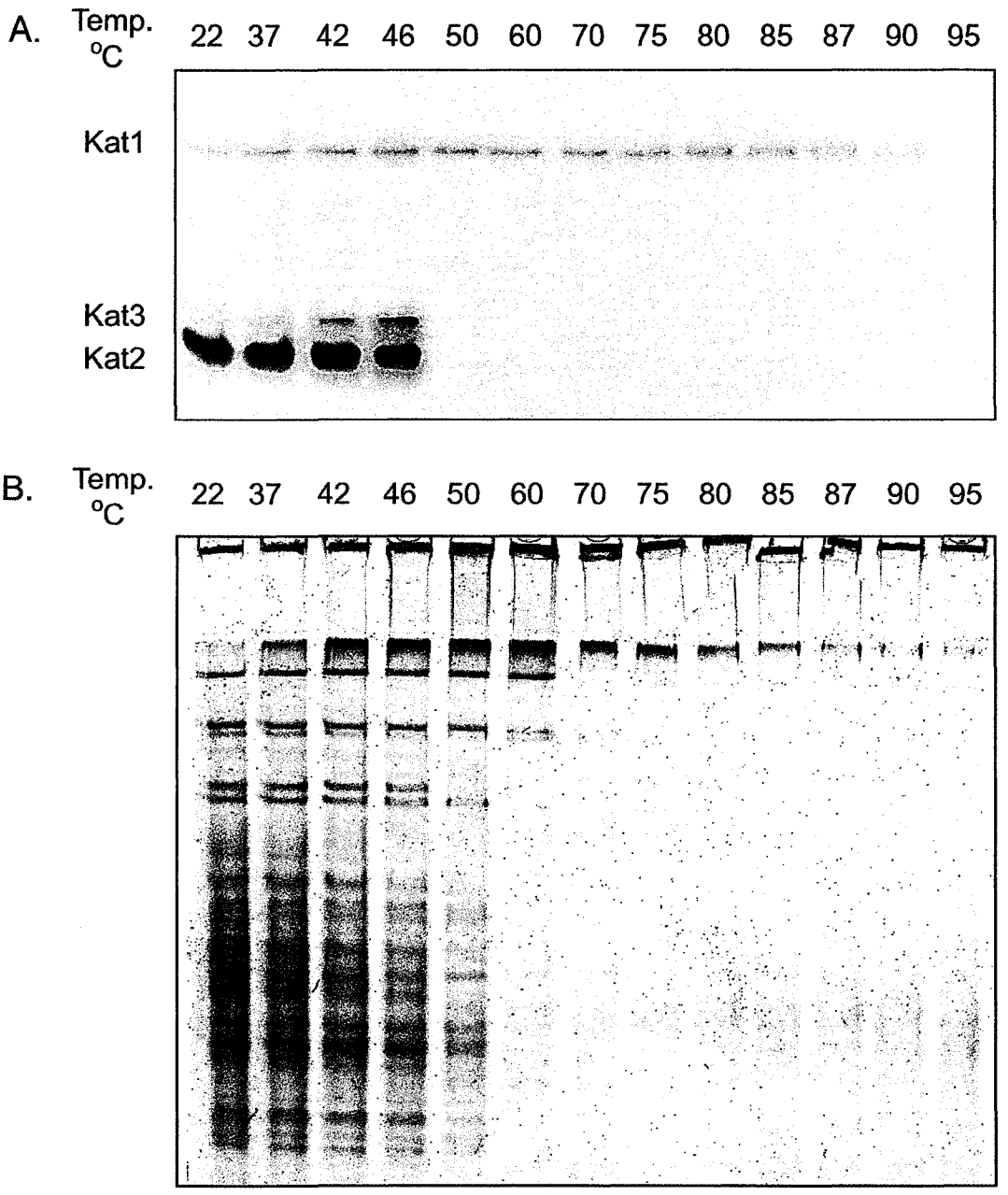



**Figure 4.1 Zymographic Analysis of Residual Catalase Activity.** Cell lysate-derived samples were heat treated for 30 min. Following heat treatment, 50 μL aliquots were separated by 7.5% non-denaturing PAGE and stained for the presence of catalases by the method of Clare et al. (1984) (A). Temperatures used are indicated above each lane. The electrophoretically discernable catalase species Kat1, Kat2 and Kat3 are also indicated. The image was colour inverted (Adobe Photoshop v5.5). Similarly treated aliquots were also electrophoresed and stained for total protein by Coomassie staining following a precipitation step (B).

(Chapter 3). Kat2 and Kat3 were found to have peroxidase activity when stained with the donor 3,3'-diaminobenzidine (DAB) in the absence of horse radish peroxidase (data not shown), indicating that they are capable of bifunctional catalase-peroxidase activities. Kat1 was considered a monofunctional catalase as it demonstrated no peroxidase activity in the presence of DAB.

**4.2.2 Putative Catalase Genes from the *A. vinelandii* Genome.** In an attempt to identify possible catalase gene candidates in the *A. vinelandii* draft genome, protein sequences of known catalases from several classes were compared to the sequence using the BLASTp function of the WAT software package (web annotation tools; Nirav Merchant, unpublished results) found at http://azotobacter.org. Because the genome had undergone an automated annotation process, it was also analysed exhaustively by text search term. Several candidates showing similarity to known catalases were identified (Table 4.1). One protein, identified strictly by search term, shared no similarity to any known catalase, and is presumed to be an erroneous call by the automated system. Another catalase (accession ZP_00417093) shared a moderate degree of identity with the clade 1, 2, and 3 catalases of *P. aeruginosa* (27, 28, and 26% respectively). However this candidate lacked the conserved sequences found in any of these clades (Peter Loewen, personal communication) so was deemed a poor candidate. Of the catalase candidates identified the two best candidate genes encoded a probable bifunctional (accession ZP_00418696) and a probable Mn-catalase (accession ZP_00417015).

**4.2.3 Identification of the Exponential Phase Catalase, a KatG Homologue.** To determine the identities of the enzymes responsible for each of the distinct catalase bands, 80 μg of mid-exponential phase cell lysate was separated in duplicate on a non-denaturing PAG; half of the gel was stained for total protein by Coomassie staining while the other half was stained for catalase activity. Coomassie stained protein bands corresponding to the catalase activity bands were excised, digested by trypsin and subsequently analysed by LC/MS/MS. The *A. vinelandii* KatG protein (accession ZP_00418696) was detected in both the Kat2 and Kat3 bands. Cumulatively, 16.4% peptide coverage of *Av*KatG was obtained (Fig. 4.2). *Av*KatG was found to share 65%

| Catalase type | Genome designation | Accession number | Signal peptide | Mature MW (kDa) | Predicted pI |
|---|---|---|---|---|---|
| Mono[a] | 91.2138 | ZP_00417093 | yes | 33.7[b] | 5.82[b] |
| CPx | 67.160 | ZP_00418696 | no | 81.7 | 5.33 |
| Mn-cat | 91.304 | ZP_00417015 | no | 31.3 | 4.88 |
| KatX | 63.100 | ZP_00417159 | no | 34.8 | 10.2 |

[a] This ORF shared an equivalent degree of similarity with the clade I, II, and III catalases of *P. aeruginosa*, making definitive classification impossible.

[b] The N-terminal end of ORF 91.2138 was truncated by the automated annotation software. Data presented here are based on more probable open reading frame limits (Loewen and Sandercock, unpublished).



**Table 4.1 Identification of Putative Catalase Genes from the *A. vinelandii* Annotation Project.** The genome project was accessed *via* http://azotobacter.org, and the data is based on the shotgun sequence (DOE) and first-round automated annotation (ORNL). Each protein sequence was examined for the presence of a leader peptide sequence (SignalP 3.0). The average theoretical molecular weight and pI value (PeptideMass) was calculated, with an additional 616.5 Da added per predicted heme where applicable. The designation of catalase type was based upon homology comparisons (Blast2seq) with known catalases of *P. aeruginosa* (monofunctional catalases, Mn-catalase) *P. putida* (HPI) or *Bacillus spp* (KatX). CPx type catalases have both catalase and peroxidase activities.


**Figure 4.2 Multiple Alignment of *Av*KatG.** *Av*KatG (accession EAM04829) was aligned with the previously described KatG proteins of *Pseudomonas syringae* DC3000 (*Ps*KatG, accession NP_794283) and *Escherichia coli* (*Ec*KatG, accession AAC76924) by ClustalW (Higgins et. al.,1996). Peptide fragments identified by LC/MS/MS following trypsin digestion are indicated in bold with underlining. Fragments with scores greater than 75, greater than 60, or greater than 50 are underlined with solid, dashed or dotted lines respectively. Scores below 50 were not considered further. The "*" symbol indicates complete residue conservation, ":" indicates strong group conservation, "." indicates weak group conservation and blanks indicate no conservation of residues.

protein identity with *E. coli* KatG (accession AAC76924), and 74% protein identity with *Pseudomonas syringae* KatG (accession NP_794283).

No catalase homologues were identified from the trypsin treated Kat1 band, though a single peptide for a putative antioxidant protein (accession ZP 00418713) similar to the BCCP/MauG family of proteins (COG 1858; pfam 03150) was detected. This identification was suspect since the fragment only gave 3.7% coverage of the protein and because no BCCP has ever been reported to have catalase activity, despite decades of scrutiny. Further analysis demonstrated that the enzyme responsible for the Kat1 band was resistant to the proteases pronase E, trypsin (Fig. 4.3) and AspN (data not shown), but not proteinase K. Such protease resistance and heat resistance are features shared with *Ec*HPII (Chelikani *et al.*, 2003) and KatA of *P. aeruginosa* (Hassett *et al.*, 2000). However no HPII homologue was identified in the genome. Furthermore, while the genetic context around the *P. aeruginosa katA* gene is present in *A. vinelandii* (*rpsD – uvrA*), a two gene lesion has resulted in the loss of the catalase (Fig. B, Appendix). Thus, given the heat resistance of the protein and the genes identified from the genome annotation, it seemed more likely that the stationary phase catalase would be a Mn-type catalase (Kono and Fridovich, 1983; Amo *et al.*, 2002).

**4.2.4 Purification of the Stationary Phase Catalase.** It was possible that the enzyme responsible for catalase activity at band Kat1 was present in the lysate at concentrations undetectable by MS, that it was resistant to trypsin digestion, or that the BCCP homologue was merely a co-migrating contaminant of the actual catalase. Given these and other uncertainties inherent to working with mixed protein samples, further purification of the enzyme was necessary. Because the enzyme was heat-stable, small quantities of cell free extract were heated at various temperatures for 15 min, chilled on ice for 10 min, and clarified by centrifugation. The resultant lysate was separated by non-denaturing PAGE and proteins were visualized by Coomassie staining (Fig. 4.1B). Excellent precipitation of whole cell lysate occurred at 75 °C, with minimal catalatic activity loss, so this parameter was selected for further treatment of the remaining sample of whole lysate. Low molecular weight proteins and peptides were removed using a 100 kDa molecular weight cut-off filter (YM100, Millipore) in a stirred cell device. During

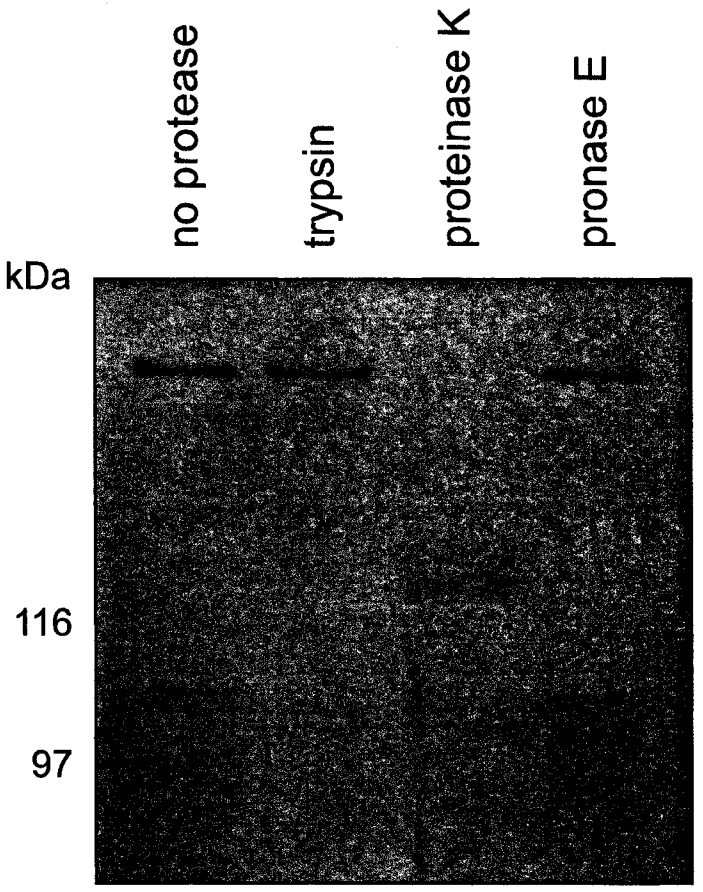



**Figure 4.3 Determination of Kat1 Protease Sensitivity.** Approximately 2 μg of purified Kat1 was digested with 1 unit of protease overnight at 37°C. Reactions were stopped with the addition of SDS-PAGE load dye and heated for 15 min at 37°C. Proteins were separated by 10% denaturing PAGE and visualized using Coomassie Brilliant Blue.

fractionation by FPLC (AKTA Explorer 100A, Amersham) on a Q-Sepharose column (Pharmacia) the approximate protein concentrations of the fractions were followed at 215 nm and the presence of heme was followed at 407 nm. Fractions were subsequently analysed for their relative enzyme activities by the method of Beers and Sizer (1952) (Fig. 4.4A). Three peaks were seen to absorb strongly at 407 nm, one of which corresponded to the fractions demonstrating the strongest enzymatic activity. Fractions containing 33% of the peak activity were pooled, and concentrated by centrifugation at 4,000×g (A-4-62 swinging bucket rotor, Eppendorf) through a 100 kDa molecular weight cut-off filter (Amicon Ultra-15, Millipore). The solution was pre-filtered through a 0.8 μm filter under vacuum (type AA, Millipore) and separated by size exclusion chromatography (Sephacryl S-300, Pharmacia). The protein concentration, heme content and catalatic activity of each fraction was assessed as above. The strongest activity fractions absorbed strongly at 407 nm (Fig. 4.4B). These observations strongly suggested that the enzyme being purified was a heme-catalase, despite the annotation based prediction of a Mn-catalase.

Several fractions were separated by 10% denaturing SDS PAGE and Coomassie stained (Fig. 4.5A) to determine protein purity. Fractions with the highest activities had a dominant protein band at approximately mass of 215 kDa, and three minor contaminants with masses of 58, 38, and 37 kDa. Fractions with at least 20% of the peak activity were pooled. Duplicate samples of the pooled S-300 solution were separated by 10% denaturing SDS PAGE. Half of the gel was stained for total protein by Coomassie staining, and the other half was stained for catalase activity. The zymographically detected catalase band corresponded to the dominant protein band at 215 kDa (Fig. 4.5B).

**4.2.5 Identification of Kat1, a Protein Similar to Snr-1 of *P. aeruginosa*.** Prior to pooling, the S-300 fraction 27 (ca. 57 to 58 mL, Fig. 4.4B) was selected for further analysis. Proteins identified from this fraction by trypsin fragmentation and LC/MS/MS analysis included the 60 kDa GroEL chaperonin (accession AAL25964.1) with 10.99% peptide coverage (Fig. C, Appendix), and the 36.6 kDa acetohydroxy acid isomeroreductase enzyme (accession EAM03644.1) with 9.76% peptide coverage (Fig. D, Appendix). The identity of the 38 kDa protein could not be determined from fraction

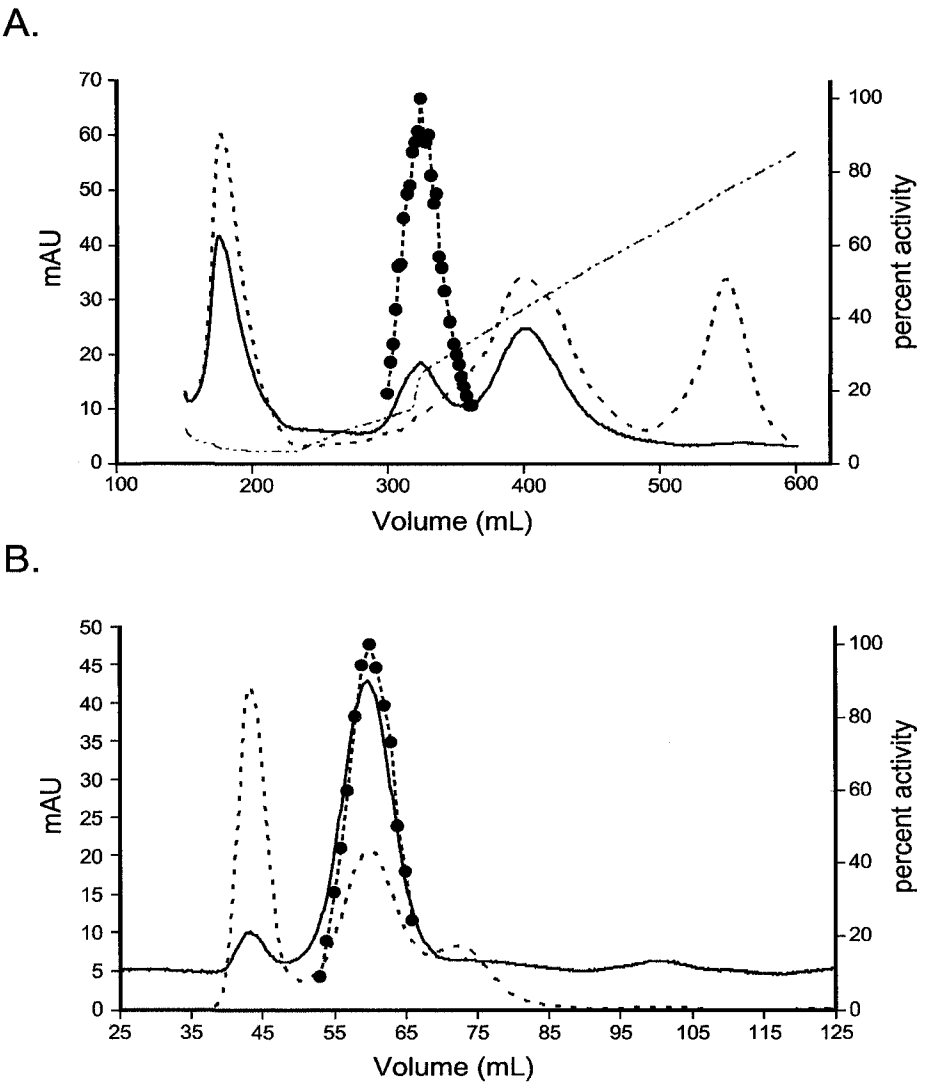



**Figure 4.4 Elution Profiles of *A. vinelandii* Whole Cell Lysates Fractionated by Q-sepharose (panel A) and S-300 (panel B) FPLC.** For each fraction the relative protein concentration was approximated at 215 nm (dashed line), heme content was measured at 407 nm (solid line) and the percent catalase/peroxidase activity was determined (circles). The dashed-dotted line plots the relative NaCl concentration during elution of the Q-sepharose column (A). Protein estimations were scaled down 30 and 12-fold in panels A and B to allow the use of a single axis (mAU). Graphs were prepared using Origin 5.0. Elution profiles of the void volume were not included for the sake of clarity.

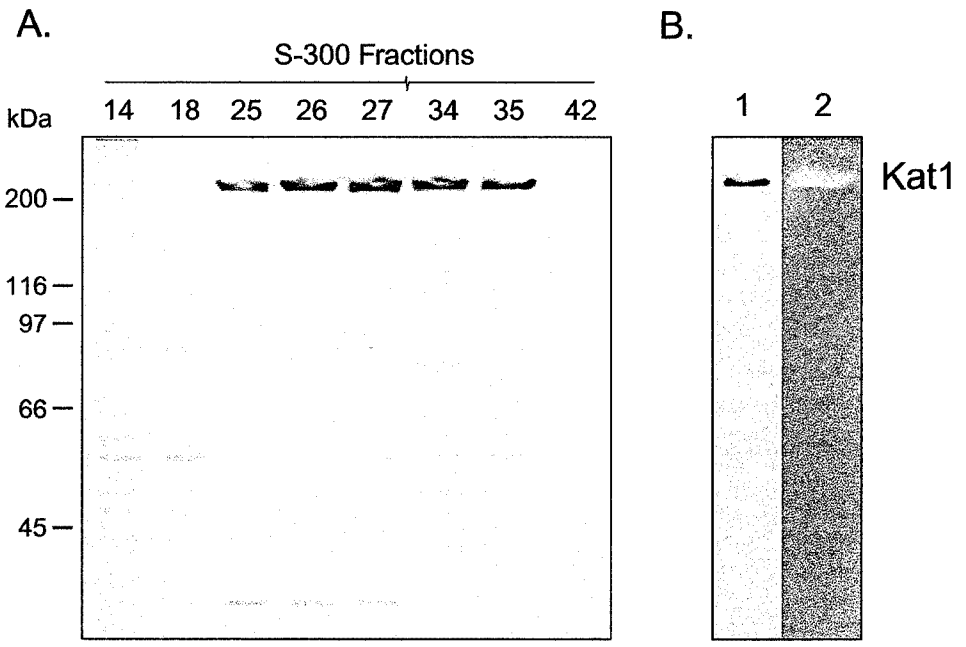



**Figure 4.5 Degree of Catalase Purification following S-300 Fractionation.** Selected fractions that exhibited strong catalase activity following S-300 chromatography were assessed for purity by electrophoretic separation by 10% SDS-PAGE and subsequent staining with Coomassie blue (A). The migration of the molecular weight standards is indicated on the left (BroadRange, BioRad). The most active fractions had a major band at 215 kDa and three minor contaminants with molecular weights of 68, 38, and 37 kDa. The samples were pooled as detailed in the text. Duplicate lanes containing 2.3 μg of the pooled protein were separated by 10% SDS-PAGE (B), and stained for total protein (lane 1) with Coomassie Brilliant blue or for catalase activity (lane 2). The mobility of the catalase activity band, Kat1, corresponded to that of the 215 kDa protein complex.

27. An additional fragment of the previously identified BCCP/MauG homologue (accession ZP_00418713) was also observed.

Analysis of the peptides liberated by proteinase K digestion of fraction 27 gave an additional 21.6% coverage of the theoretical open reading frame of the BCCP-like catalase (Fig. 4.6), and a composite coverage of 27.7% with the trypsin fragments. The degree of similarity between the catalase and the BCCPs was fairly low, as the *A. vinelandii* catalase shared only 26% identity with the cytochrome *c*551 peroxidase of *P. aeruginosa* (*Pa*CCP, accession NP_253277). Rather, the enzyme shared the greatest degree of similarity with two novel diheme proteins; 78% protein identity with the Snr-1 protein of *P. aeruginosa* PA01 (accession NP_251722) and 72% protein identity with the Hsc protein of *Desulfonmonile tiedjei* DCB-1 (accession AAB66558).

**4.2.6 Sequence Analysis of Kat1.** Two *c*-type heme binding sites, conforming to the motif CXXCH, were observed in the primary amino acid sequence at the C-terminal end of the catalase (Fig. 4.6A), in agreement with the diheme proteins *Dt*Hsc and *Pa*Snr-1. This motif distribution is dissimilar to that of the BCCP proteins, in which only one of the heme sites is found in the C-terminal end (Fig. 4.6B). All of the Cys residues found in the protein mapped to the two heme motifs, a surprising observation considering the thermal stability of the protein. Because they are expected to be covalently bound to heme groups, it is unlikely that any of the Cys residues are involved in intrapeptide cross-linkage or that they act as axial ligands to either heme.

Analysis of the amino acid content of the protein indicated a paucity of aromatic amino acids. In particular, only a single Trp residue was observed. This residue was localized near the N-terminus of the protein (W102), but did not map closely to the conserved Trp found in other cytochrome *c* moieties (Fülöp *et al.*, 1995).

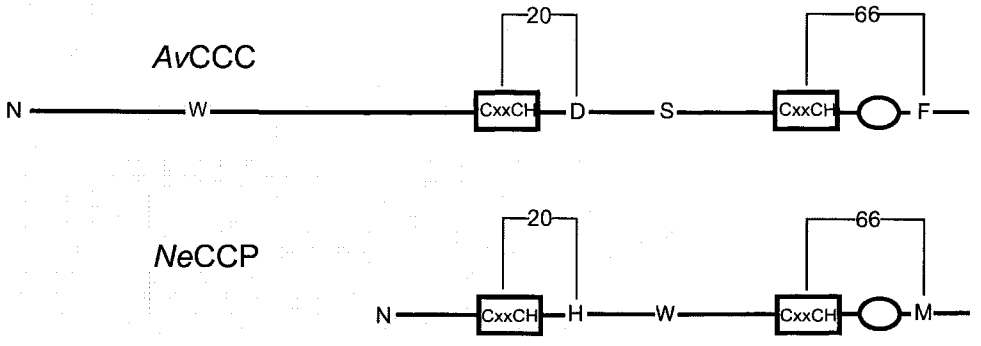
Calcium has been shown to be critical for the activation and multimerization of several BCCPs (Pettigrew *et al.*, 2003; Timóteo *et al.*, 2003; Dias *et al.*, 2004). A sequence bearing some resemblance to the consensus $Ca^{2+}$ binding site of the BCCPs (TXPYXHXG) was observed at the C-terminal end of the protein (423 SPPYLHDG 430, Fig. 4.6).

**Figure 4.6 Multiple Alignment of *Av*CCC.** *Av*CCC (accession ZP_00418713) was aligned with the closest homologues, the inducible cytochrome C of *Desulfonmonile tiedjei* DCB-1 (*Dt*Hsc, accession AAB66558), and the Snr-1 protein of *Pseudomonas aeruginosa* PA01 (*Pa*Snr, accession NP_251722) using ClustalW (Higgins et. al.,1996) as shown in Panel A. Sequences detected by LC/MS/MS are indicated; those identified following trypsin digestion are italicized and underlined, those identified following proteinase K digestion are bold and underlined. The single Trp residue (W) is highlighted by the inverted triangle (▼), two heme binding sites adhering to the CXXCH motif are boxed, and the putative $Ca^{2+}$ site is boxed with a dashed line. Leader peptide sequences predicted by SignalP 3.0 (Bendtsen et al., 2004) are indicated in italics and are underlined with a dashed line. The N-terminal extension unique to *Av*CCC is indicated in lower case, and is discussed in the text. The "*" symbol indicates complete residue conservation, ":" indicates strong group conservation, "." indicates weak group conservation and blanks indicate no conservation of residues. The schema in Panel B highlights the major putative motifs found in *Av*CCC, aligned with the secondary structure of the BCCP of *Nitrosomonas europaea*. Heme *c* motifs are boxed, the $Ca^{2+}$ binding motif is indicated by an oval, and the putative $6^{th}$ axial ligand to each heme (including distances) are indicated by thin lines.

A.

```
AvCCC vppaarrgrrapgqqsrgdrpgacrqppggrarrdragaapasrtvarragrraapgaep


AvCCC vgadrrptgaaaqsrrapatepgalsaagrafrpsfflairplrcarlyllllpldihrr


AvCCC etpMSKATIKAGLAIPIALALSLADSVRA APEEKPTSYSPVAITEDFATTKERLAGEKSS 57
DtHSc ---MKFVRPKLIFVMACCMIVALAGLIYA ESKKVPSSYSPVVITEPFDSIMTRMKAAKPE 57
PaSnr -MKTPAWTRHALWVMPLALGLQSAVVA GDEQPSKTSSYSPVVINEDFATIMKRMTANKPS 59
             :   .:. .: :  *         . .:*****.*.* * :   *: . *..
                                                  ▼
AvCCC VMQTQRRLLDARYDLSDKPVQGVAMSKGKPVQGGVRVKLP-GGTNWDDLGRLNPEEIRGK 116
DtHsc IEKRHTDLLSSRYDLSNKPAQGVTMSRGKAIQEGVRVKLPQGGVTWEQLAAMTPEQIKEK 117
PaSnr IEQAHKTLLEQRYDLSDRPAKGASMTRGKPLQEGIRVKLP-AGTSWEELARLSPEEIRKQ 118
      : : :  **. *****::*.:*.:*::**.:* *:***** .*..*::*. :.**:*: :

AvCCC RLFPAGFMPLPHPKHGEGGMLFPQFMIDEINKQEQRDLTRFDLDYDLPDHFLPEFPAPIF 176
DtHsc NVWPEGFYPLPHPNHPEGGMVFPKTHIEEIKKQEQRDLTRFDLDFDLPDHVLPESPAAIF 177
PaSnr GLFPGGFLPLPHPNHAEGGMVFPKFLIDEIKRQESRDLTRFDLDYDLPDHFLPEFPAPMF 178
       ::* ** *****:* ****:**:  *:**::**.*********:*****.*** **.:*

AvCCC LTTRPDLGDVSQGKQVTLFNYYELFNGILNPKQLEGLRLLLTPFPQQQFNVTEDRRSEHP 236
DtHsc LTTRPDLGDVSKGKLVTIDNYFELFNGILNPKQLEGLRLLVTPFPQQQFNQTDDRRSEKP 237
PaSnr LTTRPDLGDVSKGKLVTIDNYFELFNGILNPKQLEGLRLLLTAFPQQQFNLTDDRRSEHP 238
      ***********:** **: **:******************:*.******* *:*****:*

AvCCC HRGVACLDCHANGHTNAATHLVGDIRPQELRHRIDTPTLRGVNIQRLFGSQRALKSVEDF 296
DtHsc SRGVTCFDCHANGHTNGATHLVGDIRPQEFRHRLDTPTLRGVNIQRLFGSQRALKSVEDF 297
PaSnr SRGVACFDCHANGHTNAATHLAGDVRPQPFRHRIDTPTLRGVNIQRLFGSQRALKTVEDF 298
       ***:*:*********.****.**:*** :***:**********************:****

AvCCC TEFEQRAAYFDGDPVIATKKGVNILERGSQVHFMGEFQALLDFPPAPKLDVEGRLIKGQA 356
DtHsc TEFEQRAAYFDGDPVIATKKGVNVLERGSQVHFMAEFQELLDFPPAPKLDIYGKLDPQKA 357
PaSnr TEFEQRAAYFDGDPVIATKKGVNVLERGSQVHFMGEFQALLDFPPAPKLDVEGRLDPGKA 358
      ***********************:**********.*** **********: *:*   :*

AvCCC SEQELRGEALFHGKAACAGCHAPPYYTDNLMHNLKTERFFEPQTINGRLAKADGPIKTFP 416
DtHsc SELELKGQEVFFGKAKCAVCHPAPYYTDNLMHNLKAERFFKPKMINGRMASADGPIKTFP 417
PaSnr SEQELRGEKLFYGKAACAGCHAPPYFTDNLMHNLKVERFYDPKLVNGVMASADGPIKTFP 418
      ** **:*: :*.*** ** **..**:*********.***:.*: :** :*.*********

AvCCC LRGIKDSPPYLHDGRLLTLEDTVEFFNLVLQRGLSAEEKADLVAFLRTL 465
DtHsc LRGIKESPPYLHDGRLITLEDTVEFFNLIQGLNLNADEKKALVAFMRAL 466
PaSnr LRGIKDSPPYLHDDRLLTLEDTVEFFNLVLERKLSAEEKGDLVAYLRTL 467
      *****:********.**:***********:    *.*:**  ***::*:*
```

B.



Sequence analysis using the publicly available signal peptide prediction software SignalP 3.0 (Bendtsen *et al.*, 2004) indicated, with high probability (0.993), the presence of a signal peptide starting at residue M124 of the *Av*CCC open reading frame. Revision of the N-terminal ORF boundary such that it started at residue M124 rather than V1, and subsequent cleavage of the 26 amino acid leader sequence (Fig. 4.6) gave a theoretical monomeric mass of 50.6 kDa (including two heme groups at 616 Da each) rather than 66.3 kDa (PeptideMass; Wilkins *et al.*, 1997). Furthermore, the adjusted conceptual protein had a theoretical pI of 6.20. It is unlikely that a protein with the original ORF boundaries, and the corresponding theoretical pI of 9.60, would bind Q-sepharose with the buffer used in this study (50 mM Tris-HCl, pH 7.0). A third position codon bias plot (Artemis, Sanger Institute), which can be particularly useful for defining coding regions in G+C rich genomes (Ishikawa and Hotta, 1999), supported the revised N-terminal boundary (Fig. 4.7A and B).

**4.2.7 *A. vinelandii* Kat1 is a Novel Cytochrome *c* Catalase.** In order to determine the validity of the predicted heme binding sites, purified protein was subjected to denaturation by several methods, separated by denaturing PAGE, then stained with TMBZ for the detection of heme groups (Fig. 4.8). Low temperature denaturation (37 °C, 15 min) in the presence of SDS and βME did not cause dissociation of the catalase complex, which migrated at about 215 kDa, and stained poorly (Fig. 4.8, lane 3). Higher temperature treatment (100 °C) in the presence or absence of SDS and βME (Fig. 4.8, lanes 4 and 5) was able to dissociate the active complex to varying degrees. The resolution of a 51 kDa monomer corresponds with the hypothetical ORF mass (including the hemes) of 50.6 kDa. Strong staining of the monomer by TMBZ indicated the presence of covalently bound heme groups. As a result, we have given this enzyme the designation cytochrome *c* catalase (*Av*CCC). Treatment with the chelating agent EGTA resulted in some dissociation, suggesting a role for $Ca^{2+}$ in enzyme stabilization (Fig. 4.8, lane 6). To further elucidate the ORF boundaries of *Av*CCC, N-terminal sequencing by Edman degradation (Procise, Perkin-Elmer) was attempted. However little useful data was obtained (chromatograms shown in Fig. E, Appendix), likely due to natural modifications at the N-terminus.

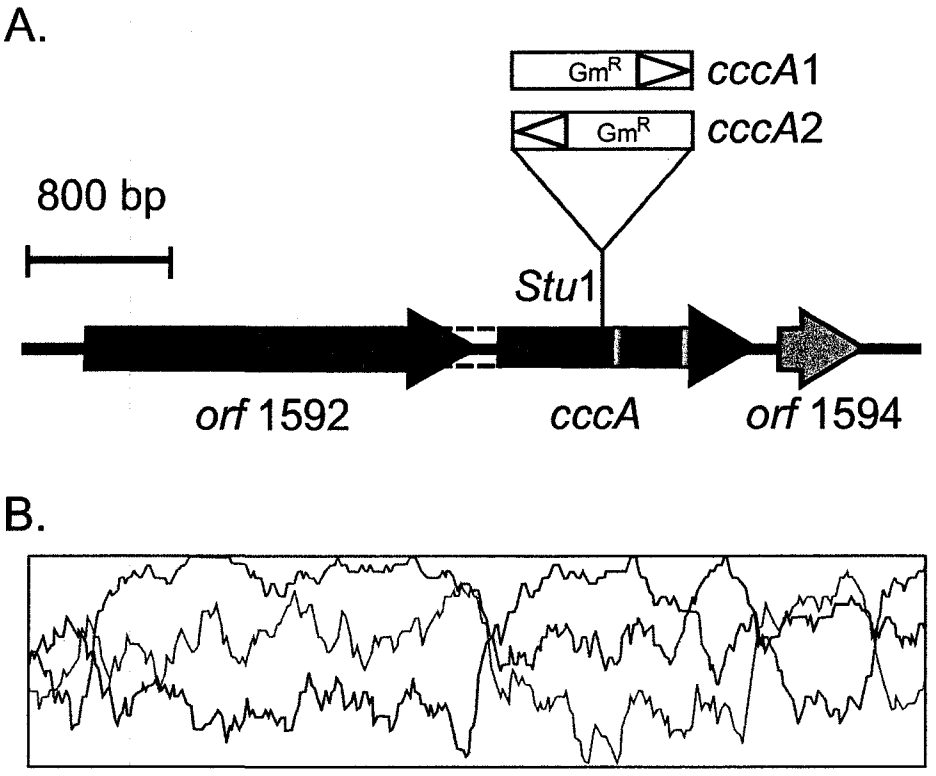



**Figure 4.7 Map of *cccA* and the Surrounding Open Reading Frames.** The map is shown to scale (A). The N-terminal extension predicted by automated annotation is indicated by a dashed box, and the adjusted ORF is indicated by the blue arrow (see text). The position of the putative heme sites within the *cccA* ORF are indicated by shaded boxes. The $Gm^R$ antibiotic cassette was inserted at the *Stu*I site, and the orientation of the cassette in the mutants *cccA*1 and *cccA*2 is indicated. The 3rd position G+C bias plot (Panel B) is aligned and is to scale with panel A. The three different reading frames for the top strand are indicated by the pink, green and black lines. The G+C bias was plotted in Artemis (Sanger Institute).

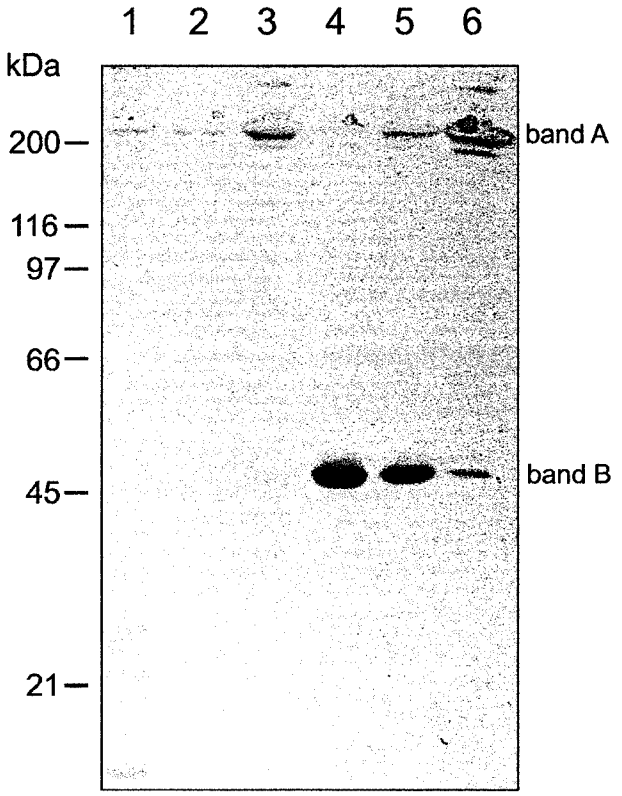



**Figure 4.8 Heme Staining of Proteins Separated by 10% SDS-PAGE.** Lane 1 contains 154 μg whole cell lysate, lane 2 contains 9 μg Q-sepharose pooled protein, and lanes 3 through 6 contain 3.4 μg of S-300 pooled protein. Prior to loading the samples were treated as follows: lanes 1 to 3 were treated with load dye containing BME and SDS for 15 min at 37°C; lane 4 was treated with load dye for 3 min at 100°C; lane 5 was treated for 30 min at 100°C in aqueous buffer then load dye was added at room temperature; lane 6 was treated with 1 mM EGTA and load dye for 15 min at 37°C. The proteins were separated electrophoretically and stained by the method of Goodhew et al. (1990). Band A, which corresponded to the catalase multimer, stained poorly. The monomeric form of the enzyme, band B, stained strongly and had a molecular weight of approximately 51 kDa. The molecular weight standards are indicated on the left (BroadRange, BioRad).

**4.2.8 The Effects of the *cccA*::Gm Interruption on the Expression of Kat1.** The *cccA* gene was amplified by PCR and cloned into pGEM-Teasy (Promega) resulting in the creation of the plasmid pCCC. The plasmid was cleaved by *Stu*I, leaving blunt ends, into which a $Gm^R$ cassette (liberated from p34S-Gm by *Sma*I) was ligated. Plasmids pCCC::Gm1 and pCCC::Gm2 were produced, with the gentamycin cassette in the forward and reverse orientation, respectively (Fig. 4.7A). Each of these plasmids was transformed into strain UWD and the *rpoS* mutant strain, UWDS. Whereas strain UWD produced significant quantities of the catalase *Av*CCC during early stationary phase (Fig. 4.9A), the mutant strains UWD *cccA*1 and UWD *cccA*2 produced none. The *rpoS* strain UWDS did not produce *Av*CCC, consistent with previous observations (Chapter 3), nor did the *rpoS cccA*1 double mutant (Fig. 4.9A). To ensure that the effect of the mutation was direct, and not due the abolition of RpoS expression, Western analysis of RpoS levels was conducted from the same cell lysates (Fig. 4.9B). RpoS expression was not eliminated by the *cccA*::$Gm^R$ mutation, as similar levels were observed in strains UWD, UWD *cccA*1, and UWD *cccA*2. Expression of the sigma factor was completely abolished in the *rpoS* mutant strains.

**4.2.9 Hydrogen Peroxide Survival of the *cccA* Mutant.** The survival of strains UWD and UWD *cccA*1 following oxidative stress as assessed by the addition of $H_2O_2$ to a final concentration of 20, 40, or 60 mM during exponential or stationary phase growth (Table 4.2). Strain UWD was very sensitive to hydrogen peroxide treatment during exponential phase, with only $9.6 \times 10^{-4}$ percent survival following the addition of 20 mM $H_2O_2$, and no detectable CFU following treatment with 40 mM peroxide. Survival of strain UWD was 2700-fold greater during stationary phase, with 2.6% survival following the addition of 20 mM $H_2O_2$ (Table 4.2). The opposite profile was observed for UWD *cccA*1, as survival was 34-fold poorer in stationary phase than in exponential phase after treatment with 20 mM $H_2O_2$. In general the mutant resisted peroxide exposure poorly, and did not demonstrate measurable survival following the addition 40 mM $H_2O_2$.

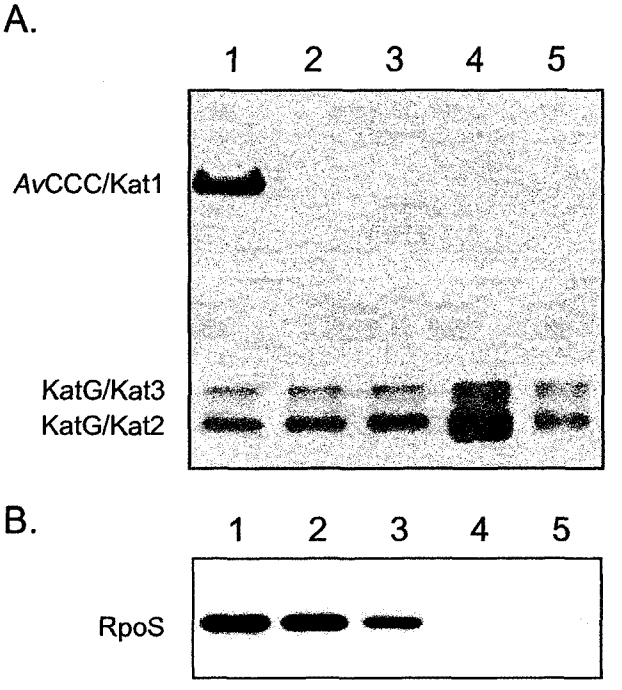



**Figure 4.9 Catalatic and Stress Response of *cccA* Mutants during Stationary Phase.** Cell lysates from strains UWD (lane 1), UWD *cccA*1 (lane 2), UWD *cccA*2 (lane 3), UWD *rpoS* (lane 4), and UWD *cccA*1 *rpoS* (lane 5) were separated by 10% SDS-PAGE. Zymographic analysis for catalase activity was done with 50 μg of cell extract per lane (panel A), whereas Western analysis of RpoS used 25 μg of cell extract (panel B).

| $H_2O_2$ Treatment | Percent Survival | | | |
|---|---|---|---|---|
| | Strain UWD | | Strain UWD *cccA*1 | |
| | Exponential | Stationary | Exponential | Stationary |
| 20 mM | $9.6 \times 10^{-4}$ | 2.6 | $1.0 \times 10^{-2}$ | $3.0 \times 10^{-4}$ |
| 40 mM | 0 | $2.6 \times 10^{-2}$ | 0 | 0 |
| 60 mM | 0 | $1.3 \times 10^{-4}$ | 0 | 0 |

**Table 4.2 Survival of *A. vinelandii* strains during Hydrogen Peroxide Challenge.** Percent survival of strains UWD and UWD *cccA*1 following 20 min exposure to 20, 40 or 60 mM hydrogen peroxide during exponential or stationary phase growth in liquid media.

## 4.3 DISCUSSION

### 4.3.1 *A. vinelandii* Produces a KatG Homologue

One of the dangers inherent to oxygen based metabolism is the generation of toxic co-products such as oxygen radicals, hydroxyl radicals, and hydrogen peroxide, which can damage virtually any bio-molecule in the cell. This is of particular importance to organisms that live in hostile environments (Seib *et al.*, 2006), those that have rapid respiration rates (Imlay and Fridovich, 1991) or bacteria that enter quiescent states measured in years (Nicholson, 2002; Vela, 1974) so are unable to repair damage in a timely manner. Here it is reported that *A. vinelandii*, which has the highest respiration of any known bacterium (Kelly *et al.*, 1990; Soballe and Poole, 1998; Williams and Wilson, 1954), synthesizes two catalases under specific physiological conditions. The first, a bifunctional KatG homologue, shares a high degree of similarity to previously described catalases in this class. Previous work suggested that *Av*KatG may be ROI inducible and demonstrated that the stationary phase form of *Av*KatG exhibits different migratory patterns than the exponential form on polyacrylamide gels (Chapter 3). A similar expression profile during elevated aeration was also seen here (Fig. 4.1). Furthermore, a similar shift in migration rate occurred following heat-treatment of the enzyme. The enzyme was found to be somewhat more heat-labile than *Ec*KatG, losing detectable activity above 46 $^{o}$C. Expression was strong during exponential phase growth, so *Av*KatG can be considered the "house-keeping" catalase.

### 4.3.2 The Stationary Phase Catalase is Thermally- and Chemically-Resistant

The second catalase, a stationary phase inducible enzyme, was found to be both thermostable and protease-resistant. While nearly complete denaturation of the catalase was observed following boiling in the presence of βME and SDS, a small but detectable amount of the enzyme complex remained in the multimeric form (Fig. 4.8, lane 4). The low Cys content observed in *Av*CCC was surprising, considering the thermal stability of the protein. It was determined that only four Cys residues were present in the protein, and that all were confined to the heme binding motifs, so would not be available for the formation of intrapeptide disulfide bonds. The heat tolerance of the enzyme, equivalent

to the catalase of the thermotolerant *Bacillus terminalis* spore (Lawrence and Halvorson, 1954) was surprising, considering that the vegetative form of *A. vinelandii* UWD does not grow above 37 °C (J. H. Robitaille and W. J. Page, unpublished) and that the encysted form of the organism is not considered heat resistant (Socolofsky and Wyss, 1962). A similar degree of heat-resistance has been described for HPII/KatE of the mesophile, *E. coli* (Meir and Yagil, 1985). It has been suggested that HPII either developed protease resistance due to its role as the stationary phase catalase and that heat resistance arose coincidentally (Chelikani *et al.*, 2004), or that *katE* transferred horizontally from a thermophilic organism (Switala *et al.*, 1999). Whether *Av*CCC developed within *A. vinelandii* or the gene arrived by horizontal transfer, the presence of a protease-resistant enzyme is advantageous for the stationary and quiescent phases of *A. vinelandii*, as the frequency with which the enzyme would need to be manufactured would be decreased. Additionally, whereas peroxidatic activity is to some degree metabolically dependent (due to the need for donor regeneration) catalatic protection would have the added advantage of being "metabolically independent" during a period of nutrient limitation.

### 4.3.3 Protein Motifs of *Av*CCC

Unlike the BCCPs, which are in most cases paradoxically expressed under anoxic or microaerophilic conditions (Van Spanning *et al.*, 1997; Lissenden *et al.*, 2000), expression of the stationary phase catalase seemed to be insensitive to the aerobic state of the culture (Chapter 3). Initially recognized by the automated annotation software as a BCCP/MauG homologue (COG1858, pfam 03150), the novel diheme catalase shares limited similarity to other members of that protein family. The only motifs shared between *Av*CCC and the BCCPs appear to be the putative $Ca^{2+}$ binding site at the C-terminus and two CXXCH heme binding sites. Treatment of the protein with EGTA resulted in partial dissociation of the protein, so either $Ca^{2+}$ is less critical for multimer stability than in the BCCPs, or binding at this site is stronger than at the equivalent BCCP site. It should be noted that the heme binding sites had a very unique distribution in *Av*CCC, where both sites were relegated to the C-terminal half of the protein, rather than being equally distributed between the N- and C- termini. It is uncertain what the $6^{th}$ axial

ligands, if any, are for each of the heme groups, as typically spaced methionine and histidine candidates were not observed.

The catalase was also unique in that it contains only one tryptophan residue (W102). From the crystal structure of *Pa*CCP, it is known that a highly conserved N-terminal Trp residue is localized in three dimensional space between the two heme groups, and is believed to mediate the transfer of electrons from the high spin accepting heme to the low spin peroxidatic heme (Fülöp *et al.*, 1995). The primary sequence context of this residue is highly conserved in the BCCPs (92 QFWDGRA 98, *Pa*CCP numbering) and is similar in the MauG family. However, this conserved sequence is not found around the Trp of *Av*CCC (100 TNWDDLG 106), which could indicate that W102 is not involved in electron transfer. One hypothesis is that the absence of the conserved Trp would preclude intra-heme communication. Such an insulated state could give rise to two independent catalytic centers.

*Av*CCC bears the greatest degree of similarity to the less studied diheme cytochrome *c* proteins, *Pa*Snr-1 and *Dt*Hsc. It is uncertain if these proteins have significant catalase activity, since definitive enzymatic and physiological functions have yet to be elucidated (J.J. Rowe, personal communication; W.W. Mohn, personal communication). However *Av*CCC and *Dt*Hsc share many physiochemical properties, including a similar monomer mass around 51 kDa, a multimer mass just above 200 kDa (and so apparent tetrameric organization) and the presence of covalently bound heme groups (Louie *et al.*, 1997).

### 4.3.4 Physiological Importance of *Av*CCC.

As seen in Chapter 3, *A. vinelandii* resistance to hydrogen peroxide stress during exponential phase, when *Av*KatG was the dominant catalase, was quite low. Although survival during stationary phase was 2,700-fold higher than in exponential phase, the resistance was much poorer than that of other soil pseudomonads (Jørgensen *et al.*, 1995; Heeb *et al.*, 2005). This may seem paradoxical, considering the high metabolic rate of *A. vinelandii*, but mounting evidence indicates that endogenously produced $H_2O_2$ constitutes a minor fraction of the peroxides encountered by a bacterium, and that it is largely removed by peroxidases (Park *et al.*, 2005). In the case of *A. vinelandii*, which is a free-

living soil organism, exposure to exogenous peroxides by plants and other bacteria may be a relatively rare event necessitating less antioxidant protection. Interruption of the *cccA* gene had little effect on survival during exponential phase, but resulted in a nearly 8,700-fold lower survival rate than the wild-type strain in stationary phase. So unlike many of the BCCPs, which have only assumed physiological roles at this time (Seib *et al.*, 2006), a clear role has been determined for *Av*CCC.

### 4.3.5 Summary

Despite their energetic potential to function in a catalatic manner, neither the yeast CCP (Santoni *et al.*, 2004) nor the BCCP exhibit measurable catalase activities. It could be argued that *Av*CCC does not normally have catalase activities, but became damaged during the heating stage of the purification process, resulting in the release of new enzymatic capabilities. Such a "functional conversion" is observed when certain cytochrome *c* proteins are treated with proteases, resulting in the activation of hitherto unobserved peroxidase and catalase functions (Jeng *et al.*, 2004). However, *Av*CCC also exhibited strong catalatic activities when liberated by gentle cell lysis methods in the absence of heat (Fig. 4.1; see also Chapter 3). As a result, we report here the first BCCP-like protein, and possibly the first naturally occurring cytochrome *c* type protein, to demonstrate significant catalase activity. Should the closely related proteins *Dt*Hsc and *Pa*Snr-1 demonstrate significant catalatic function, this would constitute the first new class of catalase since the description of the manganese (non-heme) catalases (Kono and Fridovich, 1983).

# CHAPTER 5

## Characterization of the Di-Heme Cytochrome *c* Catalase (*Av*CCC) of *Azotobacter vinelandii.*

## 5.1 INTRODUCTION

Aerobic metabolism allows for efficient capture of energy during nutrient oxidation (due to the large difference in potential between the $NAD^+ + 2H^+/NADH + H^+$ and $2H^+ + \frac{1}{2}O_2/H_2O$ half-reactions) but also results in the production of a large number of damaging reactive oxygen intermediates. As a result, all aerobic life-forms utilize a host of protective enzymes; SODs dismutate superoxide radicals into $H_2O_2$, peroxidases reduce $H_2O_2$ to $H_2O$, and catalases disproportionate peroxide to $H_2O$ and $O_2$ in sequential oxidative and reductive reactions. Many of these protective enzymes are induced during periods of ROI stress, particularly by the regulators SoxRS and OxyR, which are in turn activated by superoxide and $H_2O_2$, and $H_2O_2$ respectively (reviewed in Storz and Imlay, 1999). Other anti-oxidant proteins are induced during the general stress response (reviewed in Chelikani *et al.*, 2004) following the transition to non-optimal growth conditions or varied (i.e. not specifically ROI-induced) stresses.

Several thermotolerant bacteria produce heat resistant catalases. This is particularly true of *Bacillus* spp. following differentiation to their quiescent spore state (Lawrence and Halvorson, 1954). Surprisingly, thermo-stable catalases have also been documented for mesophilic organisms, such as the HPII/KatE protein of *E. coli* (Meir and Yagil, 1985; Switala and Loewen, 2002). The RpoS-dependent catalase of *A. vinelandii*, *Av*CCC, exhibits heat-stability equivalent to those of *Ec*HPII and the *Bacillus terminalis* catalase, though it shares little homology with any previously described catalase. Rather *Av*CCC shares the greatest degree of similarity with the bacterial cytochrome *c* peroxidases (BCCPs). Though studied in considerable depth biochemically, few efforts have been made to identify the molecular regulators of BCCP expression (Seib *et al.*, 2006). Paradoxically the expression of most BCCPs occurs under anaerobic or microaerobic conditions (Van Spanning *et al.*, 1997; Lissenden *et al.*, 2000), when peroxide stress is likely to be lowest. As a result it is unlikely that any of these belong to ROI-inducible regulons. Since these peroxidases are typically expressed under non-optimal growth conditions, it remains a tantalizing possibility that they could fall within the RpoS regulons of their respective organisms.

Heme *c*-type proteins differ from other hemic proteins in that their hemes are covalently bound to the peptide chain. Thiol bridging of the heme to two cysteines is

termed maturation. This process is mediated by the periplasmic proteins expressed from the *ccmABCDEFGH* operon (Thöny-Meyer, 1997). The stability of heme *c* enzymes should, theoretically, be greater than other hemic enzymes since the heme group cannot easily dissociate. Nevertheless, the BCCPs tend to be unstable during purification from their natural hosts, and particularly low yields following transgenic expression in *E. coli* have been reported in all but one case.

Following purification, BCCPs are typically found in an inactive oxidized state, but can be reactivated by reduction in the presence of $Ca^{2+}$ (Dias et. al., 2004; Echalier et. al., 2006). Each of the hemes has a specific role; in the favoured model a high-potential C-terminal heme is reduced and can accept electrons from a donor molecule, whereas the low-potential N-terminal heme remains oxidized. Electrons flow to the N-terminal heme *via* a conserved Trp residue, allowing the heme to act as the peroxidatic core (DeSmet et. al., 2006). Complete reduction of the BCCPs by saturating quantities of dithionite results in peroxidase inactivation. Each heme also has a distinct spin state, which is constrained by the axial ligands surrounding the heme group. Though these can be distinguished by UV/visual spectroscopy, the preferred methods are resonance Raman and electron paramagnetic resonance. Crystal structures, and so residue positions, can also be used to determine the identities of the axial ligands.

Within the BCCP/MauG family, those most closely related proteins to *Av*CCC are *P. aeruginosa* Snr-1 and *D. tiedjei* Hsc. Expressed during periods of nitrate- or halo-respiration, these di-hemic proteins were originally believed to be components of the electron-transport chain (Louie *et al.*, 1997; Kerschen *et al.*, 2001). However the redox midpoint of *Dt*Hsc is too low to be useful during respiration on 3-chlorobenzoate or other halo-organic compounds (Louie *et al.*, 1997). The redox midpoint of *Pa*Snr-1 has not been determined, but neither has the importance of Snr-1 for nitrate reduction been unequivocally demonstrated. Indeed, definitive activities have been described for neither enzyme, so claims regarding their catalatic or peroxidatic natures cannot be made at this time (J.J. Rowe, personal communication; T.M. Louie, personal communication).

The objectives of this chapter were to determine the cellular localization of the novel heme *c* catalase of *A. vinelandii*, further substantiating the amended ORF boundaries discussed in Chapter 3. The heat resistance, enzymatic parameters, and

sensitivity of the enzyme to common catalatic inhibitors were examined. Various methods were employed to probe the heme structures of the dihemic catalase, including UV-visual, resonance Raman and electron paramagnetic resonance spectroscopy, as well as structure prediction. Finally, comparative methods were used to examine the validity of classifying *Av*CCC as a BCCP/MauG family member (COG 1858, pfam 03150).

## 5.2 RESULTS

### 5.2.1 Monomeric Mass Determination.

Purification of the stationary-phase catalase of *A. vinelandii* strain UWD was done as in Chapter 3, but the final purity was increased (Fig. 5.1, Table 5.1) by the addition of a final ultrafiltration step across a 100 MWCO filter. This step completely removed the contaminating protein, GroEL, leaving only trace amounts of the contaminant acetohydroxy acid isomeroreductase (band B, Fig. 5.1). The molecular weight of the catalase subunit was determined to be 50,606.81 ± 3.90 Da by ESI-MS (Fig. 5.2), confirming the previous SDS-PAGE derived estimate of 51 kDa (Fig. 3.8). The observed molecular weight is much closer to the predicted protein mass for the adjusted ORF (50.6 kDa with 2 hemes) than the genome-annotated ORF (66.3 kDa with hemes) lending additional support to our suggestion that the original annotation is incorrect.

### 5.2.2 Catalytic Properties.

*Av*CCC was purified 183-fold to a specific activity of 38,654 U/mg as detailed in Table 5.1. The enzyme rate and substrate affinity of the catalase were determined using a Clarke-type electrode, which specifically measures catalase but not peroxidase activity. From the Michaelis Menton plot it can be seen that inactivation of the enzyme occurred at high peroxide concentrations (Fig. 5.3), which is common for many small subunit catalases. Surprisingly two distinct enzymatic rates were observed, which is particularly

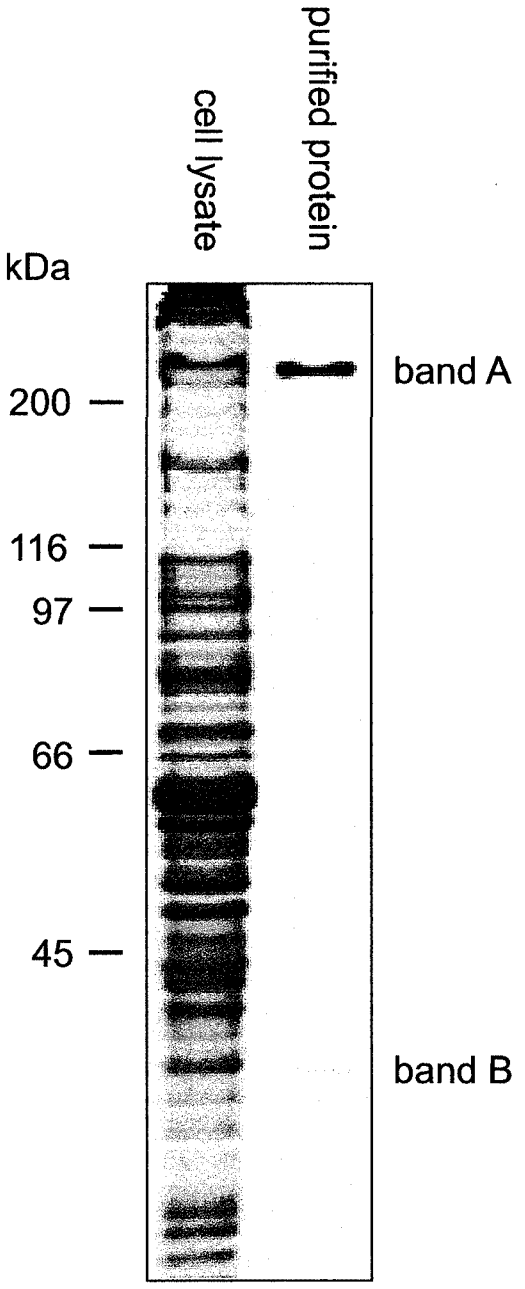



**Figure 5.1 Final Purification of AvCCC by Ultrafiltration.** Pooled samples after size exclusion chromatography (S-300) were filtered across a 100 MWCO filter. Whole cell lysate (lane 1) and purified *Av*CCC (lane 2) were separated by 10% SDS PAGE after being heated in load dye for 15 min at 37°C. Band A migrated with the catalase activity band (data not shown) whereas band B demonstrated no catalase activity.

| Sample | Activity (U/mL) | Protein (μg/mL) | Specific Activity (U/mg) | -Fold Purification |
|---|---|---|---|---|
| Cell lysate | 543.3 | 2570.0 † | 211.4 | 1.0 |
| Heat treated lysate | 481.7 | 817.8 † | 589.0 | 2.8 |
| Q-sepharose pool | 223.3 | 146.2 ‡ | 1528.0 | 7.2 |
| S-300 pool | 1050.0 | 57.0 ‡ | 18421.0 | 87.1 |
| Buffer exchange | 434.2 | 11.2 ‡ | 38654.4 | 182.8 |

**Table 5.1 Purification of Cell-Lysate Derived Cytochrome *c* Catalase.** Catalatic activity was determined by measuring the amount of oxygen generation in the presence of varying concentrations of $H_2O_2$. Protein concentrations were determined by the method of Lowry et al (1951) for crude fractions (†), and Waddell (1956) and Wolf (1983) for purified fractions (‡).

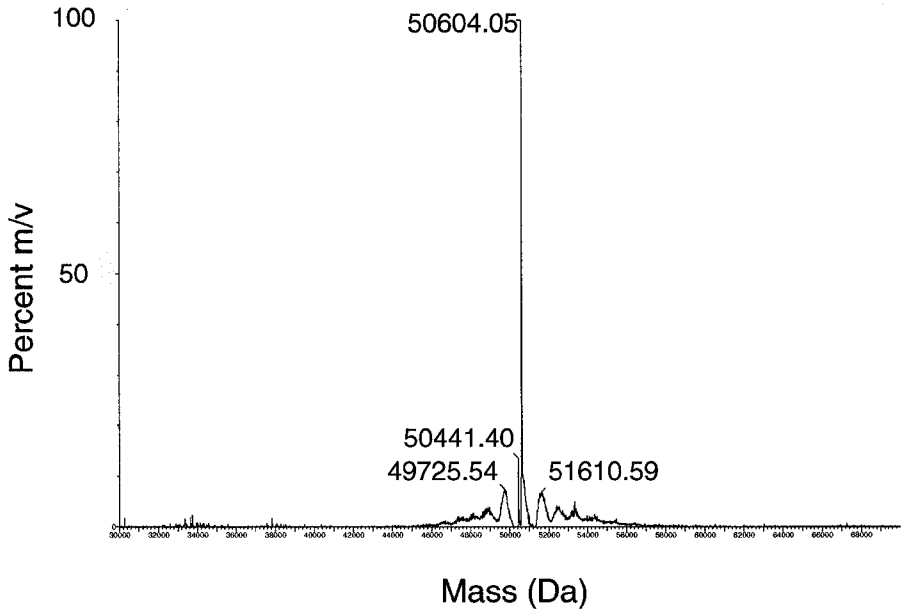



**Figure 5.2 Molecular Weight Determination of the *Av*CCC Monomer (Da) by ESI-MS.** Determinations were performed on a Q-TOF2 spectrometer with masses calculated using MaxEnt1 (Micromass).

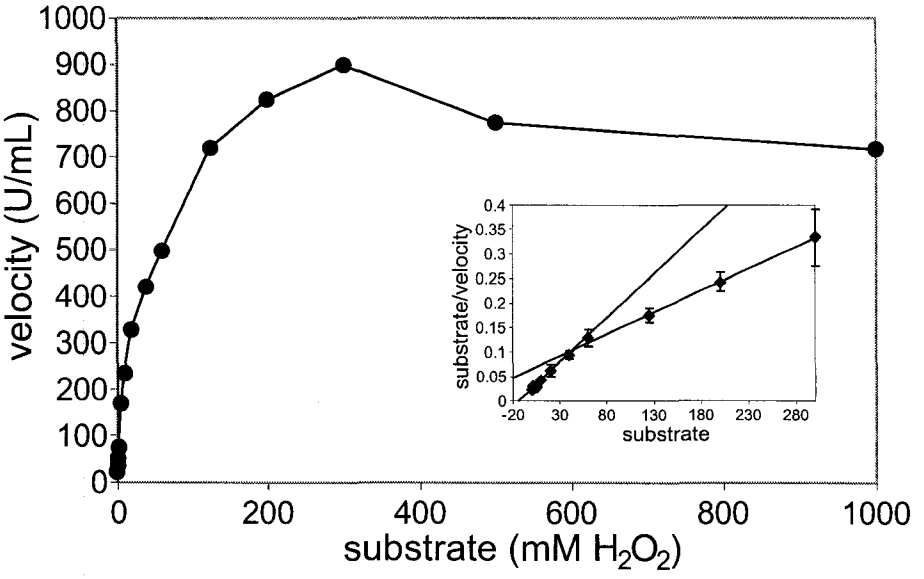



**Figure 5.3 Catalatic Enzyme Kinetics of Purified *Av*CCC.** Michaelis-Menton plot of catalase activity (U/mL) determined using a Clarke-type electrode at various hydrogen peroxide concentrations by the method of Rørth and Jensen (1967). Two apparent $K_M$ for the substrate $H_2O_2$ can be seen in the Hanes plot (inset).

apparent in the Hanes plot (inset, Fig. 5.3), with $K_M$ values of 8.4 and 67.3 mM, and $V_{max}$ values of 370 and 1080 U $mg^{-1}$.

The effect of various inhibitors on enzyme activity was also determined. The enzyme was insensitive to 3-amino-1,2,4-triazole at concentrations up to 1000 mM (data not shown) but showed various degrees of sensitivity to NaCN (Fig. 5.4A) and $NaN_3$ (Fig. 5.4B) with 50% activity at 311 and 83 μM respectively. Treatment with $NH_2OH$ inhibited *Av*CCC activity by 50% at 107 nM (Fig. 5.4C).

### 5.2.3 Heat Stability of *Av*CCC.

It was shown previously that the di-hemic catalase of *A. vinelandii* in crude lysates was heat resistant, retaining residual activity after heat treatments as high as 90 °C (Fig. 3.1). To better define the heat stability of *Av*CCC, purified protein was heated across a range of temperatures for 30 min and the residual activities were assessed (Fig. 5.5); the enzyme retained 18% activity after heat treatment at 92 °C, but catalase activity was not detected after identical treatments at 95 °C. In contrast to this, *Av*KatG in crude lysates retained no detectable activity following heat treatments above 46 °C. The purified *Av*CCC was also heated and held at 65 °C over a 24 h period; the activity half-life (50% residual activity remaining) under these conditions was determined to be 11.3 h.

The effect of chaotropic agents upon multimer stability was also examined across a range of temperatures in the presence of 5.6 M urea (Fig. 5.6). *Av*CCC was resistant to denaturation at temperatures up to 70 °C. Equivalent studies have been done with some catalases, but not with members of the BCCP/MauG family.

### 5.2.4 Cellular Localization of the Catalases of *A. vinelandii.*

*A. vinelandii* culture was grown to stationary phase, and sphaeroplasted to specifically liberate the periplasmic fraction of the cells. *Av*CCC co-purified with the periplasmic fraction (lane 3, Fig. 5.7B), confirming the theorized periplasmic localization of the catalase. It can be seen that complete sphaeroplasting of *A. vinelandii* was not achieved as in Hunter *et al.* (1989), since post-sphaeroplast cell lysates still contained a significant amount of *Av*CCC (lane 2, Fig. 5.7B). The poor sphaeroplasting efficiency

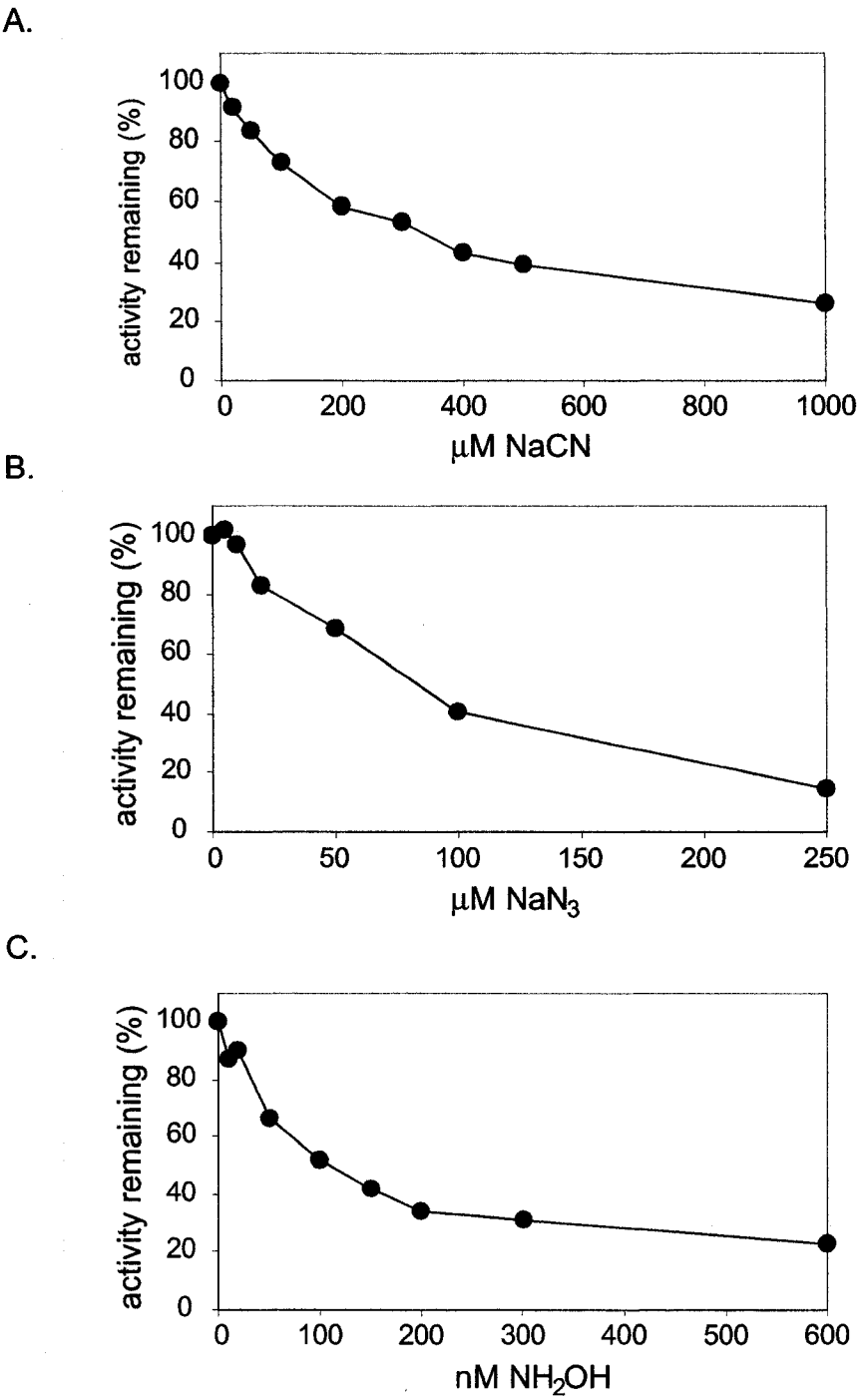



**Figure 5.4 Catalase Inhibitor Studies.** Assays using the inhibitors NaCN, $NaN_3$ and $NH_2OH$ are shown in panels A, B, and C, respectively. Phosphate buffer, enzyme and inhibitor were pre-incubated for 2 min prior to the addition of the substrate, $H_2O_2$ (60 mM final concentration). Activity is expressed as a percentage of the reaction when no inhibitor was added to the pre-incubation mix.

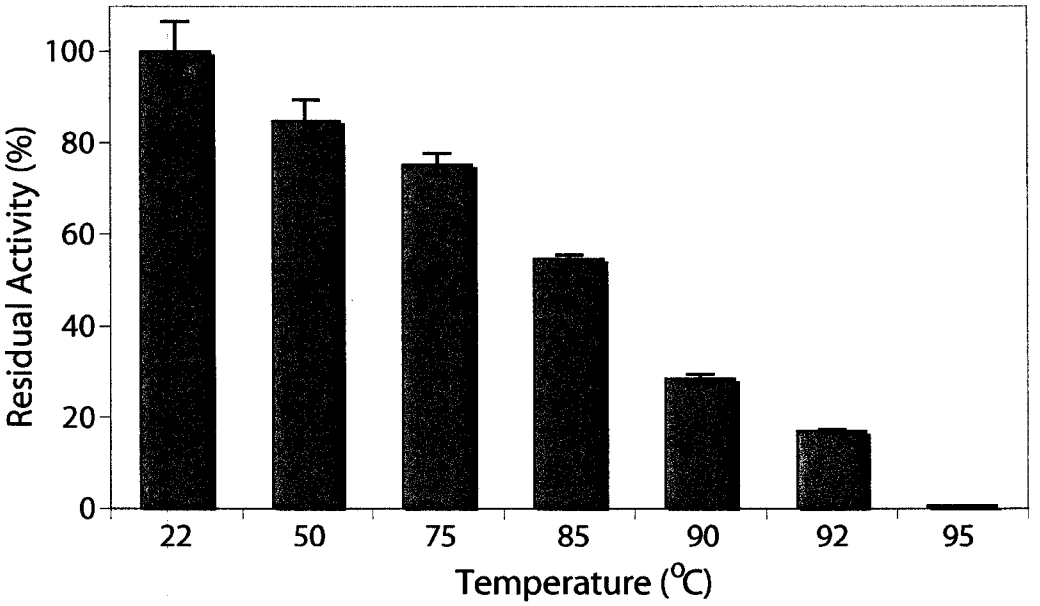



**Figure 5.5 Residual Activity Determinations of *Av*CCC Following Heat Treatment.** Enzyme was heated in phosphate buffer for 30 min, cooled to room temperature, then assayed for catalase activity by the standard method (Rørth and Jensen,1967). Activity is expressed as a percentage of the reaction at 22°C rather than 37°C.

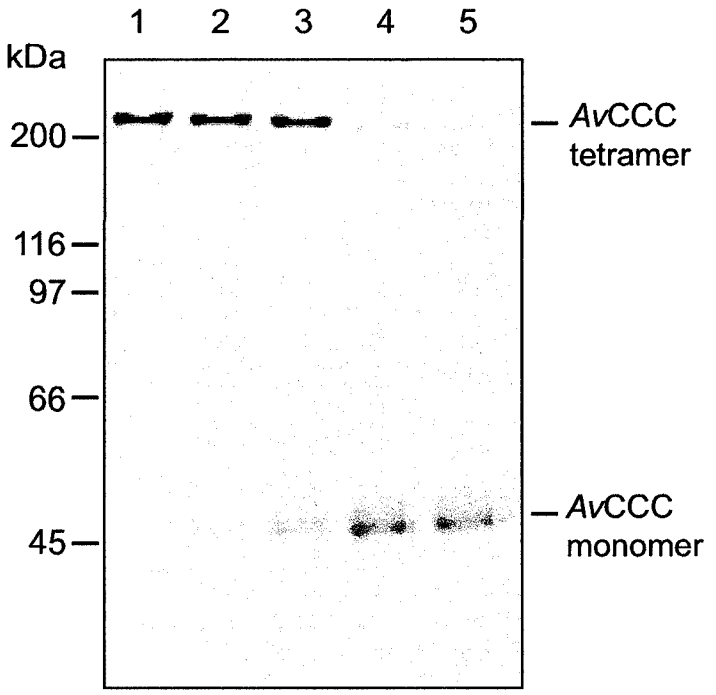



**Figure 5.6 Denaturation of Purified *Av*CCC in the Presence of Urea.** The chaotropic agents urea and SDS were added to final concentrations of 5.6 M and 10% to 6.5 μg purified *Av*CCC. The mixture was heat treated at 50°C (lane 1), 60°C (lane 2), 70°C (lane 3), 80°C (lane 4) or 85°C (lane 5) for 15 min then cooled to room temperature. Protein was separated by 10% SDS-PAGE and stained with Coomassie Brilliant Blue.

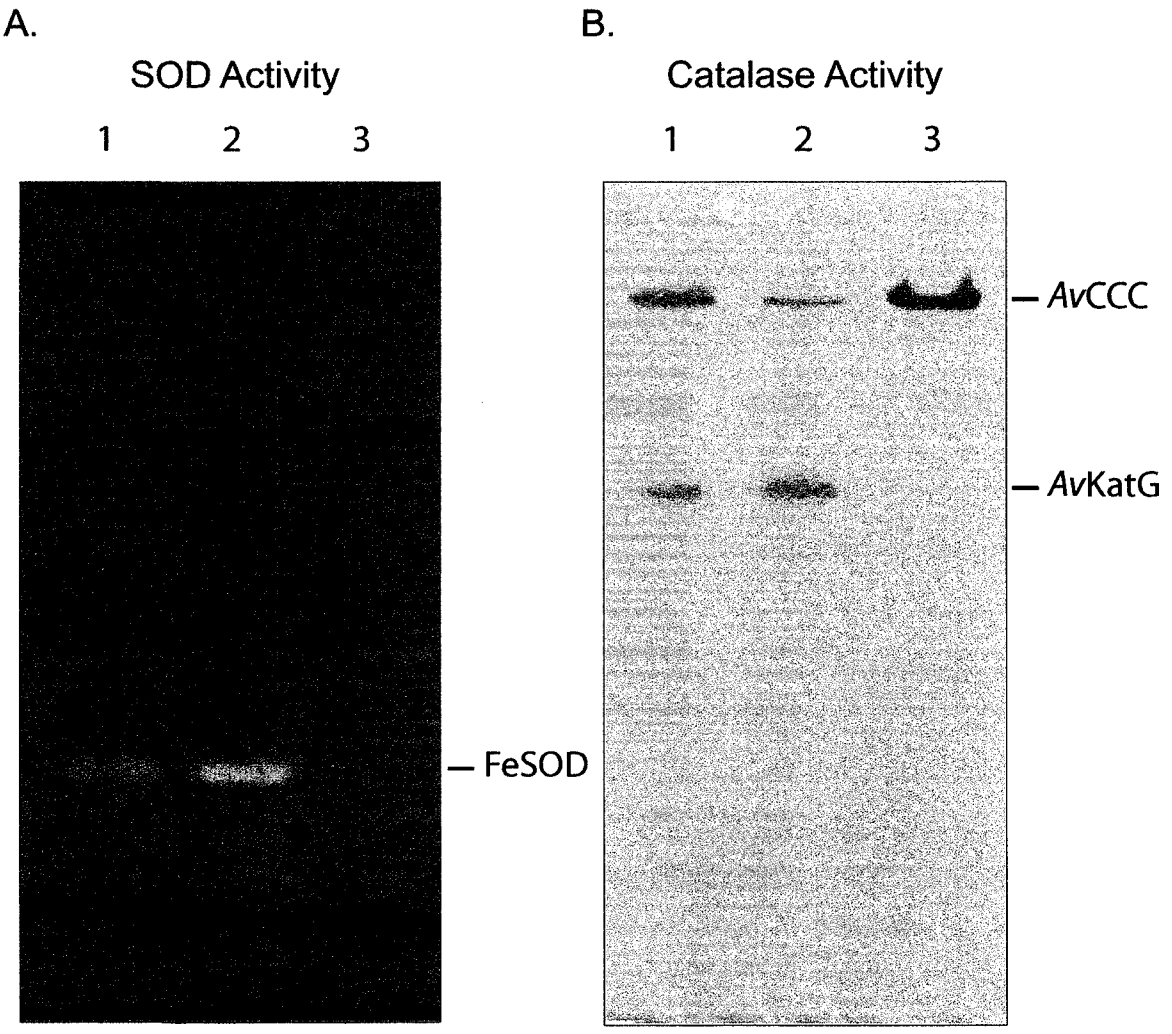



**Figure 5.7 Sphaeroplast Formation of Stationary Phase *Azotobacter vinelandii.*** Whole cell lysates (lane 1), post-sphaeroplast lysates (i.e. cytosolic fraction, lane 2) and the periplasmic fraction (lane 3) were separated by non-denaturing 7.5% PAGE. In panel A the gel was stained for superoxide dismutase activity, whereas the gel in panel B was stained for catalase activity. The image in panel B was inverted (Adobe Photoshop) for enhanced clarity.

reported here is likely due to the unavoidable use of stationary phase cells. Although more aggressive sphaeroplasting regimes could have been pursued, significant cytoplasmic leakage may have resulted (Schnell and Steinman, 1995). Contamination of the periplasmic fraction with the known cytosolic marker FeSOD (Qurollo *et al.*, 2001) was not observed (compare lanes 2 and 3, Fig. 5.7A), indicating that cytoplasmic leakage had not occurred, thereby ruling out the possibility that *Av*CCC is cytoplasmically localized. Surprisingly the known cytosolic marker CuZnSOD, previously described by Qurollo and coworkers (2001), was not observed in this study. Chelation inactivation of this SOD by EDTA during sphaeroplasting can be ruled out as other studies using the French press method were also unable to detect the enzyme (Cornish and Page, unpublished).

The bifunctional catalase-peroxidase, *Av*KatG, was found in the cytosolic fraction and was not present to any degree in the periplasmic fraction, confirming the prediction that it is strictly cytoplasmic in nature.

### 5.2.5 Optical Characteristics of *Av*CCC.

Purification of *Av*CCC (Fig. 5.8) was performed as previously described, except that the initial material was the periplasmic fraction rather than whole cell lysate and that this material did not undergo an elevated heat treatment. Following ion exchange, size exclusion chromatography and concentration across a 100 MWCO filter, a minor contaminating band with high molecular weight persisted under non-denaturing conditions (not shown). The contaminating protein complex resolved to two distinct species (ca. 45 and 41 kDa) following separation by SDS -PAGE (Fig. 5.8). Bands separated by SDS-PAGE were identified by tandem mass spectroscopy following trypsin digestion. The lower molecular weight contaminant was identified as a thiolase (accession ZP_00415203) with 61% coverage of the open reading frame (Fig. F, Appendix). The higher-mass contaminant was identified as citrate synthase I (Genebank 67156800). However only one trypsin fragment was detected for the higher mass protein, giving a mere 4% coverage of the open reading frame (Fig. G, Appendix).

The absorption spectrum of *Av*CCC, purified from sphaeroplasted samples, revealed a very broad Soret band at 398 nm with a CT1 band at 627 nm and two

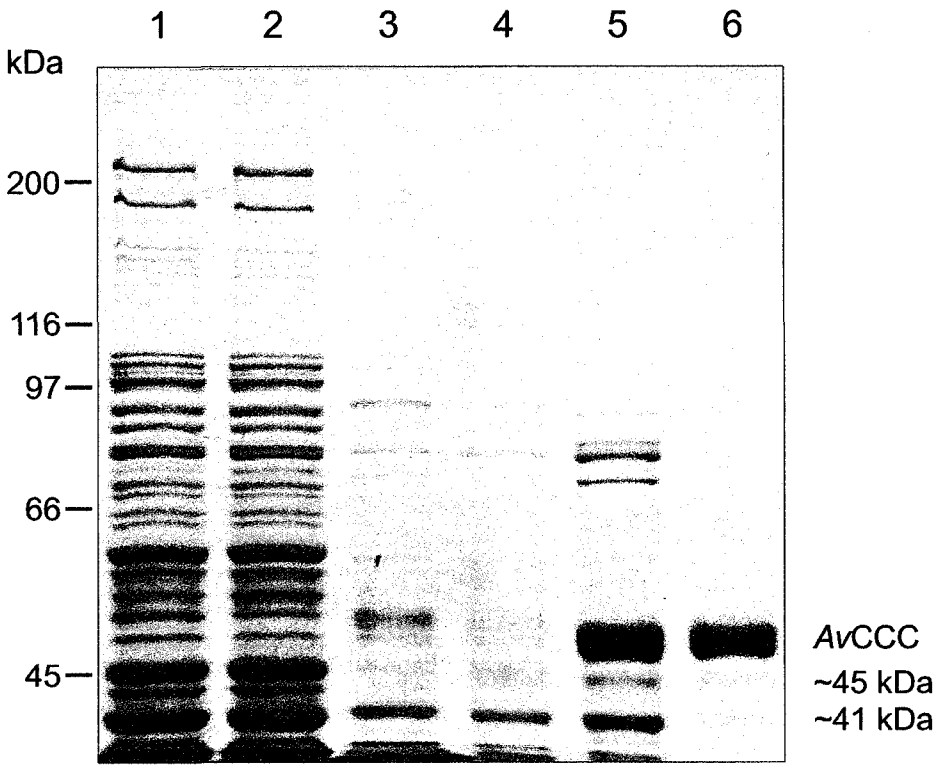



**Figure 5.8 Purification of *Av*CCC Protein from Sphaeroplasted Cultures.** Whole cell lysate (50 μg, lane 1), post-sphaeroplast lysate (i.e. cytosolic fraction; 50 μg, lane 2), periplasmic (50 μg, lane 3), filtered pre-FPLC periplasmic (50 μg, lane 4), post Q-sepharose pooled (5 μg, lane 5) and post Sephadex-300 pooled (2.5 μg, lane 6) fractions were separated by 7.5% SDS PAGE and stained with Coomassie Brilliant Blue. The molecular weight markers are shown in the far left lane.

shoulders at 500 and 540 nm (Fig. 5.9A), suggestive of the presence of a heme iron in a high (S = 5/2) – intermediate (S = 3/2) admixed spin state (Maltempo, 1976, Indiani *et al.*, 2000). The Soret band was symmetrical, which suggests little heme dissociation had occurred. Alkaline titration resulted in some red shifting of the major peaks and shoulders, but nothing as extreme as reported for the closely related *Dt*Hsc. A similar spectrum was observed for the oxidized form of whole cell lysate-derived pure *Av*CCC (Soret 403, shoulders at 500 and 550 nm; Fig. 5.9B and Table 5.2). However the whole cell lysate-derived catalase exhibited significantly more red shifting upon pH titration to 13.2, with a Soret band at 413 nm and the weakening of the band at 500 nm. These observations suggest that while little structural damage had been caused by the heating step, the protein had become more susceptible to alkaline treatments when prepared from whole cell lysates. Reduction of the catalase using sodium dithionite was also attempted; under alkaline conditions reduction resulted in a shift of the Soret band to 418 nm and allowed for the resolution of the β and α bands at 522 and 550 nm respectively (Fig. 5.9B). Like many catalases, and distinct from most peroxidases, *Av*CCC was resistant to reduction by sodium dithionite at neutral pH. The absorption maxima differed markedly from those of other hemic catalases, sharing more similarity with cytochrome *c/c'* proteins (Table 5.3).

### 5.2.6 Resonance Raman Spectroscopy.

*Av*CCC was examined by resonance Raman in an attempt to identify the axial ligands to the heme(s). Because strongly absorbing chromaphores are not associated with either of the minor contaminating proteins found in the purified solution, further purification was deemed unnecessary for analysis (Glen Loppnow, personal communication). The protein mixture was assessed by reiterative resonance Raman at room temperature at 412.5 nm. The high frequency scan of *Av*CCC is shown in Fig. 5.10. Because the protein solution was relatively dilute (~2.5 mg/mL protein) the background signal was high, obscuring the lower frequency bands (data not shown). The higher frequency bands $\nu_2$ and $\nu_3$ typically indicate the coordination state, and by extension an approximation of spin-state, of the hemic iron moieties. The $\nu_2$ band of *Av*CCC was strongest at 1579 $cm^{-1}$, a frequency typically observed in 5c HS proteins.

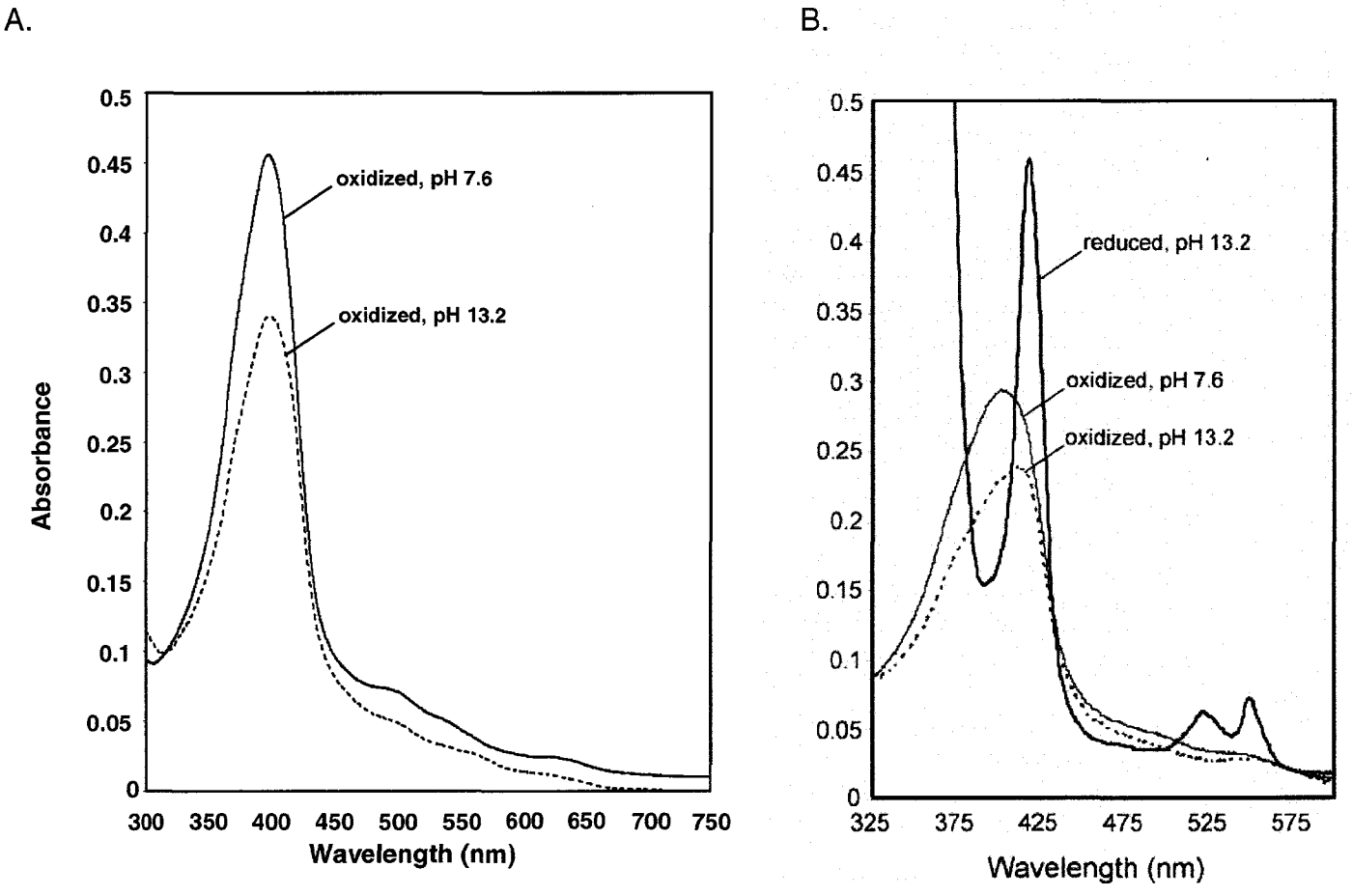



**Figure 5.9 The Optical Absorption Spectra of *Av*CCC Protein.** The spectra of sphaeroplast-derived pure *Av*CCC are shown in panel A. Absorption of oxidized enzyme at pH 7.6 (light line) and pH 13.2 (dashed line) were measured at room temperature. Spectra of the oxidized form of the cell lysate-derived pure enzyme at pH 7.6 (light line) and pH 13.2 (dotted line), and the reduced form of the enzyme at pH 13.2 (solid line) are indicated in panel B.

| | Sphaeroplast Purified *Av*CCC | | | Heat Purified *Av*CCC | | | Purified D*t*Hsc[1] | | |
|---|---|---|---|---|---|---|---|---|---|
| | Oxidized | Alkylated | Reduced | Oxidized | Alkylated | Reduced | Oxidized | Alkylated | Reduced |
| Soret | 398 | 402 | n/d | 403 | 413 | 418 | 399 | 413 | 416 |
| β-band | - | - | n/d | - | - | 522 | - | - | 521 |
| α-band | - | - | n/d | - | 540-550 | 550 | - | 540 | 549 |
| CT1 | 627 | 627 | n/d | 630 | n/d | n/d | 630† | - | - |
| Shoulders | 500, 540 | 505, 540 | n/d | 500,~550 | - | - | 480†, 540† | - | - |

[1] Louie et al., 1997

**Table 5.2 Summary of the Major Absorption Bands of *Av*CCC.** Oxidized, alkylated and reduced spectra of the catalase were determined for the protein purified from sphaeroplast or whole-cell lysate preparations. Major bands of the closely related dihemic protein *Dt*Hsc are also listed. Bands not detected are indicated by the symbol "-" and incidents where the spectra were not scanned are indicated by "n/d". Values estimated from published figures are indicated by the "†" symbol

| Protein | Soret band | | β band | α band | Heme | |
|---|---|---|---|---|---|---|
| | Oxidized | Reduced | Reduced | Reduced | Type | Reference |
| *Av*CCC | 398 | 418 | 522 | 550 | c | this work |
| *Dt*Hsc | 399 | 416 | 521 | 549 | c | Louie et al., 1997 |
| *Rp*Cyto c' | 395‡ | 418‡ | 521‡ | 550‡ | c | Strekas and Spiro, 1974 |
| *Pf*CCP | - | 414 | 520 | 550 | c | Ellfolk and Soininen, 1970 |
| *Pd*MauG | 406 | 416 | 524 | 552 | c | Wang et al., 2003 |
| *Ec*HPI | - | 408 | 500 | 639 | d | Hillar and Loewen, 1995 |
| HRP | - | 404 | 510 | 639 | b | Hillar and Loewen, 1995 |
| *Ec*HPII | 407 | - | 590† | 636† | d | Sevinc et al., 1995 |

† *Ec*HPII could not be reduced, so the reported α and β band values are for the oxidized enzyme

**Table 5.3 Comparison of UV-visible Absorption Bands of Representative Heme Proteins.** Values estimated from published figures are indicated by the "‡" symbol, and instances where a band is not observed are indicated by the dash.

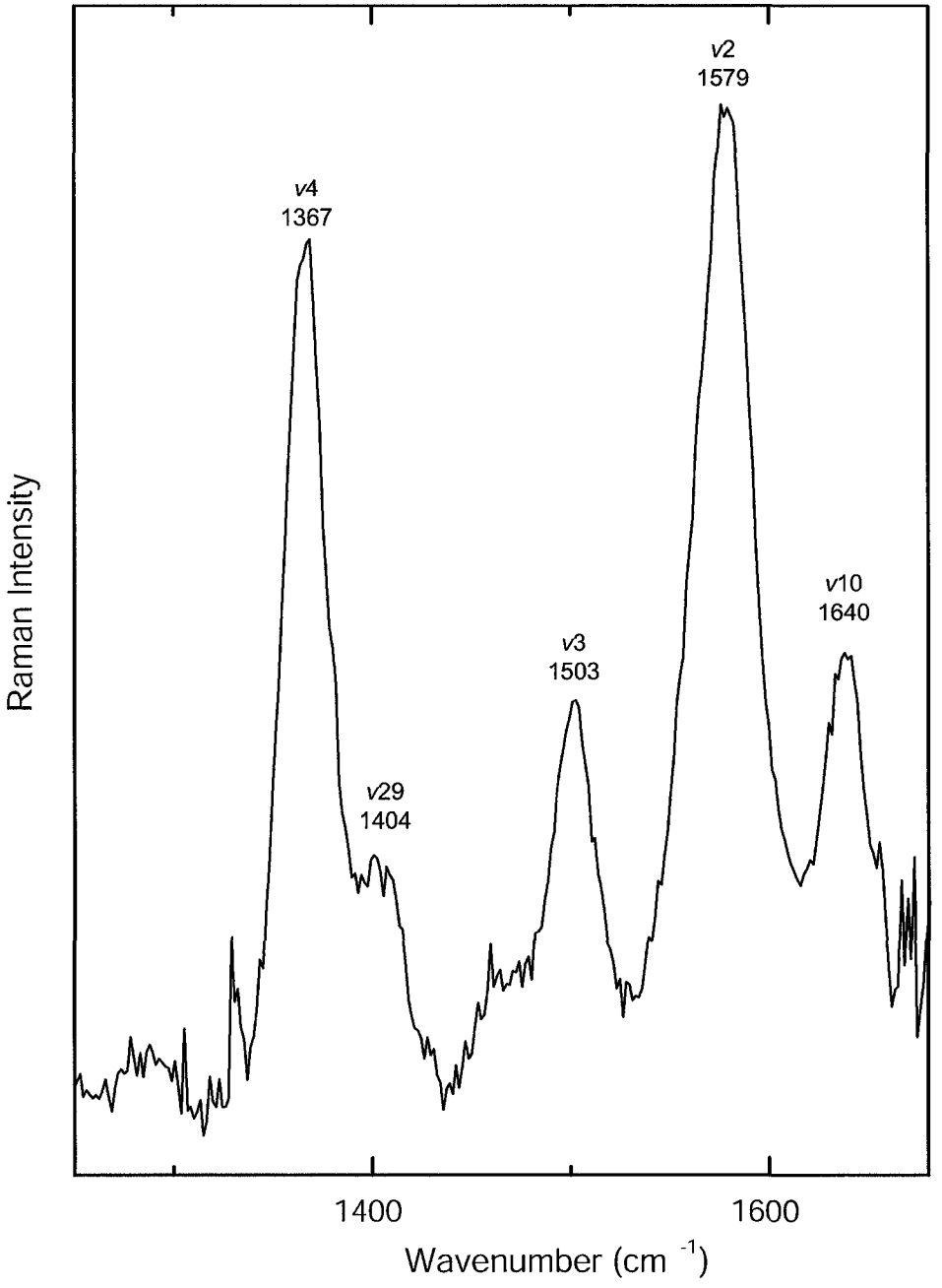



**Figure 5.10 Resonance Raman Spectrum of Ferric *Av*CCC.** The high frequency resonance Raman spectrum of ferric (sphaeroplast-derived) pure *Av*CCC at pH 7.0 was measured at room temperature. Excitation of the protein sample was achieved using a Sapphire laser at 412.5 nm.

The $\nu_3$ band was strongest at 1503 $cm^{-1}$ which has been observed in 6c LS and occasionally 5c QS moieties. The presence of a LS line under the broad $\nu_2$ band was ruled out because LS heme signals tend to be significantly stronger than those of HS hemes (Pauleta *et al.*, 2001). Thus we prefer the 5c QS interpretation over the LS interpretation for the $\nu_3$ line at 1503 $cm^{-1}$, particularly considering that the Soret band is not significantly red shifted (Lukat-Rodgers *et al.*, 2000). A similar interpretation has been made for cytochrome *c'* of *Rhodopseudomonas palustris* based on the presence of anomalous and intermediate $\nu_2$ (1578/1581 $cm^{-1}$ split signal) and $\nu_3$ (1500 $cm^{-1}$) bands (Strekas and Spiro, 1974; Maltempo and Moss, 1976).

A very broad peak was observed, centered around 1640 $cm^{-1}$, for the $\nu_{10}$ band. Higher $\nu_{10}$ frequencies tend to correlate to hemes with planar topologies which are common for QS and LS hemes, whereas values down to 1610 $cm^{-1}$ tend to be associated with HS heme moieties under higher torsional strain (Strekas and Spiro, 1974; Haddad *et al.*, 2003). Further elucidation of the hemic topologies requires accurate measurements in the lower frequency ranges which were not possible within the scope of this work.

### 5.2.7 Electron Paramagnetic Resonance Spectroscopy.

The EPR spectra of the sphaeroplast-purified catalase (2.5 mg/mL protein) was determined at 10 $^{o}K$. Two distinct hemic spin signatures were observed; a typical low-spin signal with values of $g_x$ 2.64, $g_y$ 2.46 and $g_z$ 1.66 , and a non-typical $g_{xy}$ band at 5.48 (Fig. 5.11). The latter likely corresponds to a heme with an admixture of intermediate and high spin, though the associated $g_x$ band, if actually present, was obscured by the copper signature at 2.08. A high spin interpretation for the non-typical band is unlikely since HS hemes have $g_{xy}$ values in excess of 6. Likewise a pure intermediate spin is unlikely, as spin = 3/2 hemes exhibit $g_{xy}$ value below 4 (Maltempo and Moss, 1976). The copper signature was also observed in the buffer blank, indicating that the cupric presence was due to Tris reagent or possibly Milli Q water impurities, rather than the presence of a docked protein such as azurin or pseudoazurin. A very weak signal due to chelated iron was observed at g = 4.3.

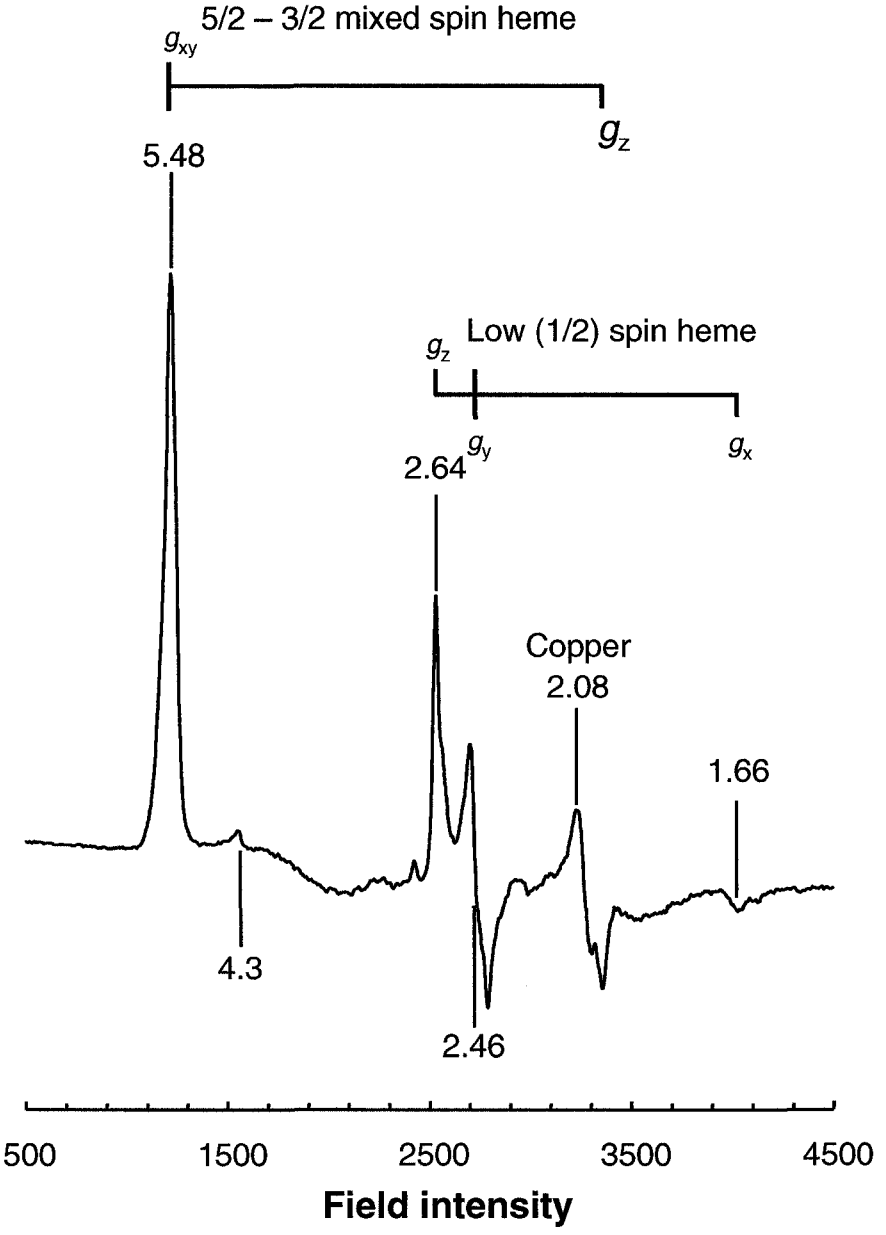



**Figure 5.11 Ferric *Av*CCC X-band EPR Spectrum.** Sphaeroplast-derived pure *Av*CCC, pH 7.0, was measured at 10 °K. Field modulation was $10G_{PP}$, with 2 mW microwave power.

### 5.2.8 Phylogenetic Designation of the *A. vinelandii* Heme *c* Catalase.

*Av*CCC was predicted to be a member of the MauG/BCCP protein family (COG 1858, pfam 03150) by the programs utilized in the *A. vinelandii* annotation project. The enzyme secondary structure was predicted using PROFsec (PredictProtein server; Rost *et al.*, 2003) to be approximately 30% α-helical in nature, with very little β-sheet structure, as is true of most MauG/BCCP proteins. However significant differences were observed in the placement of several motifs, particularly the relegation of both heme sites to the C-terminal half of *Av*CCC (Fig. 3.6). To address these differences, several phylogenetic trees were generated (PAUP, version 4.0). Complete sequences were collected from GenBank (accession numbers given in Table B, Appendix) and aligned (Clustal W; Higgins *et al.*, 1996) with the adjusted protein sequence of *Av*CCC. When neighbour joining trees were calculated from total character differences using the whole ORF sequences or the predicted mature protein sequences (cleavage sites and probabilities indicated in Table B), the same groups were observed, though some minor changes in branch length were observed. A representative tree based on the alignment of full length ORFs is shown in Fig. 5.12. Four unique protein groups were apparent as there was no overlap between cross-group (minimal distances between members of different families) and intra-group (maximum distances between members of the same family) distances (Table 5.4).

### 5.2.9 Predictive Structure Modeling.

The hemic catalase protein structure was modeled using publicly available structure prediction programs that performed well in Critical Assessment of Techniques for Protein Structure Prediction (CASP) competitions. ESyPred3D is an automated program that identifies and then models protein structures based on their homology to previously determined structures (Lambert *et al.*, 2002). Robetta is an *ab initio* protein structure prediction software package which models the structure based on a series of general protein folding rules (Chivian *et al.*, 2005), although it will preferentially return homology-determined models if appropriate pre-defined structures exist. The programs were run without imposing human-intervention on any of the parameters.

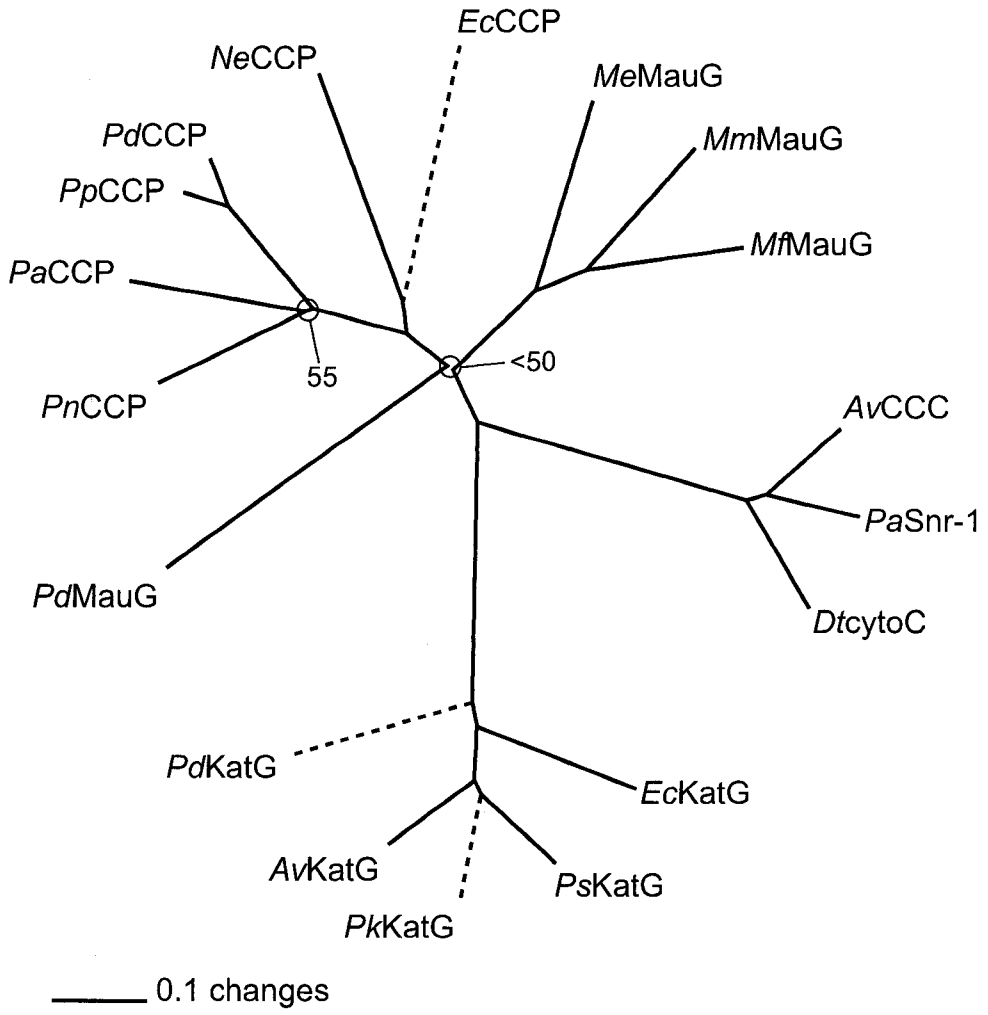



**Figure 5.12 Phylogenetic Analysis of the Di-Hemic Protein Families, BCCP, MauG and *Av*CCC.** Neighbour joining tree calculated based on total character differences between whole protein sequences of members of the BCCP, MauG, KatG and novel dihemic catalase (*Av*CCC) families. Reiterative Bootstrap analysis (100 iterations) gave > 95% certainty for all but two nodes, which are indicated by circles. Dashed lines indicate hypothetical enzymes, solid lines indicate experimentally demonstrated enzymes. Accession source files and organism names from which the sequences were derived are found in Table A (Appendix).

| Family/Group | Representative Used | Distance (differences) |
|---|---|---|
| *Av*CCC[a, b] | *Av*CCC – *Dt*Hsc | 0.2533 |
| MauG[b] | *Me*MauG – *Mm*MauG | 0.4316 |
| BCCP[b] | *Pa*CCP – *Ne*CCP | 0.5765 |
| KatG[b] | *Ec*KatG – *Ps*KatG | 0.3482 |
| | | |
| BCCP to KatG[c] | *Pn*CCP – *Av*KatG | 0.8820 |
| MauG to KatG[c] | *Me*MauG – *Av*KatG | 0.8669 |
| BCCP to MauG[c] | *Pn*CCP – *Me*MauG | 0.6604 |
| *Av*CCC to MauG[c] | *Av*CCC – *Me*MauG | 0.7998 |
| *Av*CCC to BCCP[c] | *Av*CCC – *Pn*CCP | 0.8149 |
| *Av*CCC to KatG[c] | *Av*CCC – *Av*KatG | 0.8926 |

[a] The *Av*CCC family includes the proteins *Av*CCC, *Pa*Snr-1 and *Dt*CytC

[b] Breadth of the protein family was determined using the most distantly related (experimentally confirmed) members of the family.

[c] Distance between groups was determined using the most closely related (experimentally confirmed) members of each protein family.

**Table 5.4 Distance Matrix of Selected Proteins Based on their Published Open Reading Frame Sequences.** Distances of protein sequences (including signal peptides when applicable) were determined following alignment (ClustalW) and subsequently calculated using Paup. The intra-group "breadth" of each family, defined as the maximum difference between family members, is indicated in the top portion of the table. The cross-group distance, defined as the minimum distance between members of different families, is indicated in the bottom portion of the table.

The closest structural homologue of *Av*CCC was found to be the BCCP of *Nitrosomonas europaea* (PDB 1IQC, Fig. 5.13) by PSI-BLAST (as executed by ESyPred3D). Use of the metaserver 3D-Jury (Ginalski *et al.*, 2003) confirmed that homology folding of *Av*CCC (439 residues submitted) using this PDB structure as a template was reasonable, as good Jscores (161.43 with the INUB algorithm; 154.43 with PDB-PSI) were obtained. This was surprising considering that the protein sequences only shared 32% protein sequence identity across 134 residues at the C-terminal end. Interestingly Robetta also returned a homology-based structure that mapped well to *Ne*CCP (discussed below). The Robetta derived *ab initio* structure was not considered further as it diverged more significantly than did the homology modeled structure.

Superimposition of the Robetta-predicted *Av*CCC structure on the known *Ne*CCP crystal structure (Fig. 5.14B) gave a fairly high RMSD of 4.339 over 167 aligned α-carbon atoms (of 263). However the low overall structure similarity was primarily due to differences within the N-terminal portion of the *Av*CCC structure. Alignments constrained to only the C-terminal portion (K344-L465) of the Robetta structure (Fig. 5.14D) yielded a favorable RMSD of 2.54 (over 116 aligned residues) which was refined to 0.695 (over 95 aligned residues). The favoured model was predicted by the ESyPred3D server (Fig. 5.14A), with an overall RMSD of 3.851 (over 167 aligned residues). Constraint of the structure at the C-terminal end (Fig. 5.14C) gave an RMSD of 2.71 (over 116 aligned residues) with refinement to 0.529 (over 93 residues).

Examination of the constrained ESyPred3D structure at the predicted N-terminal heme (Fig. 5.15A and C) shows considerable overlap between helix 2 of *Av*CCC and *Ne*CCP. As defined by the heme binding motif (CXXCH) the proximal side of the heme is coordinated by a histidine residue for both proteins. The distal pocket of *Ne*CCP contains three residues, Q92 under ring C and P96 under ring A (both map to helix 2) and E102 which is located next to Q92. E102 has been invoked as the acid/base catalyst for peroxidatic activity in *Ne*CCP (Shimizu *et al.*, 2001). The corresponding helix 2 residues in the *Av*CCC ESyPred3D model are F296 and E300 (Fig. 5.15B). In the Robetta model helix 2 projects away from the heme group at about 180° (Fig. 5.16A), likely highlighting the deficiencies of the model.


**Figure 5.13 Cartoon View of the Known Crystal Structure of NeCCP** (source file PDB 1IQC). The N-terminal portion of the protein (top) is clearly defined from the C-terminal portion (bottom) by a large envaginated cleft. These domains are bridged by the residue Trp82 (orange). The low (red) and high (dark red) potential hemes are separated by 21 Å in free space, but only 9.32 and 9.07 Å via Trp82. All structures were viewed using PyMol v.0.99rc6.

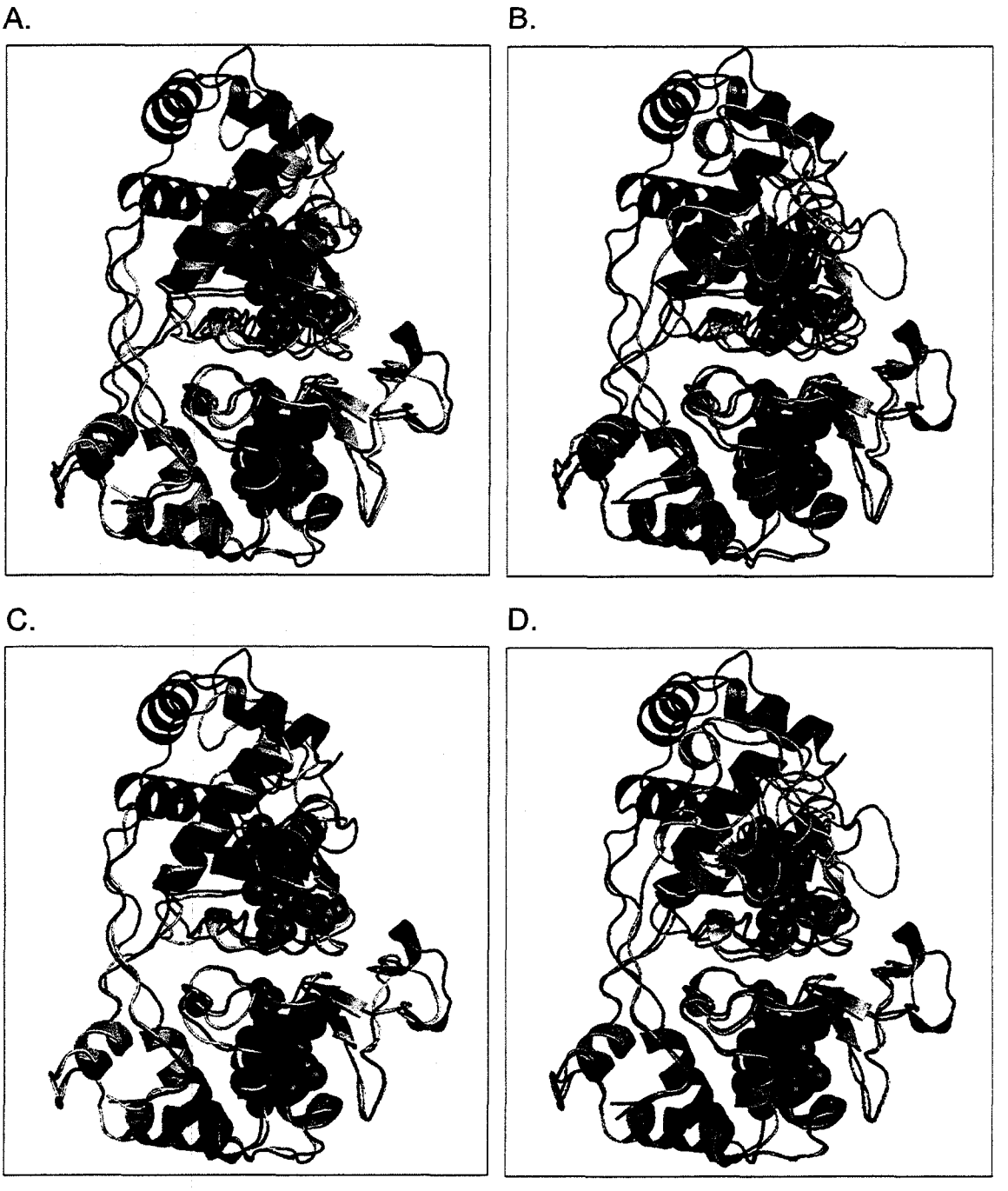



**Figure 5.14 Composite Alignments of Predicted *Av*CCC Structures with that of NeCCP.** ESyPred3D (yellow) and ROBETTA (green) structures aligned to *Ne*CCP (purple) across the entire protein sequence are shown in panels A and B. Equivalent structures aligned with constraint at the C-terminus gave lower RSMD values and are shown in panels C and D.

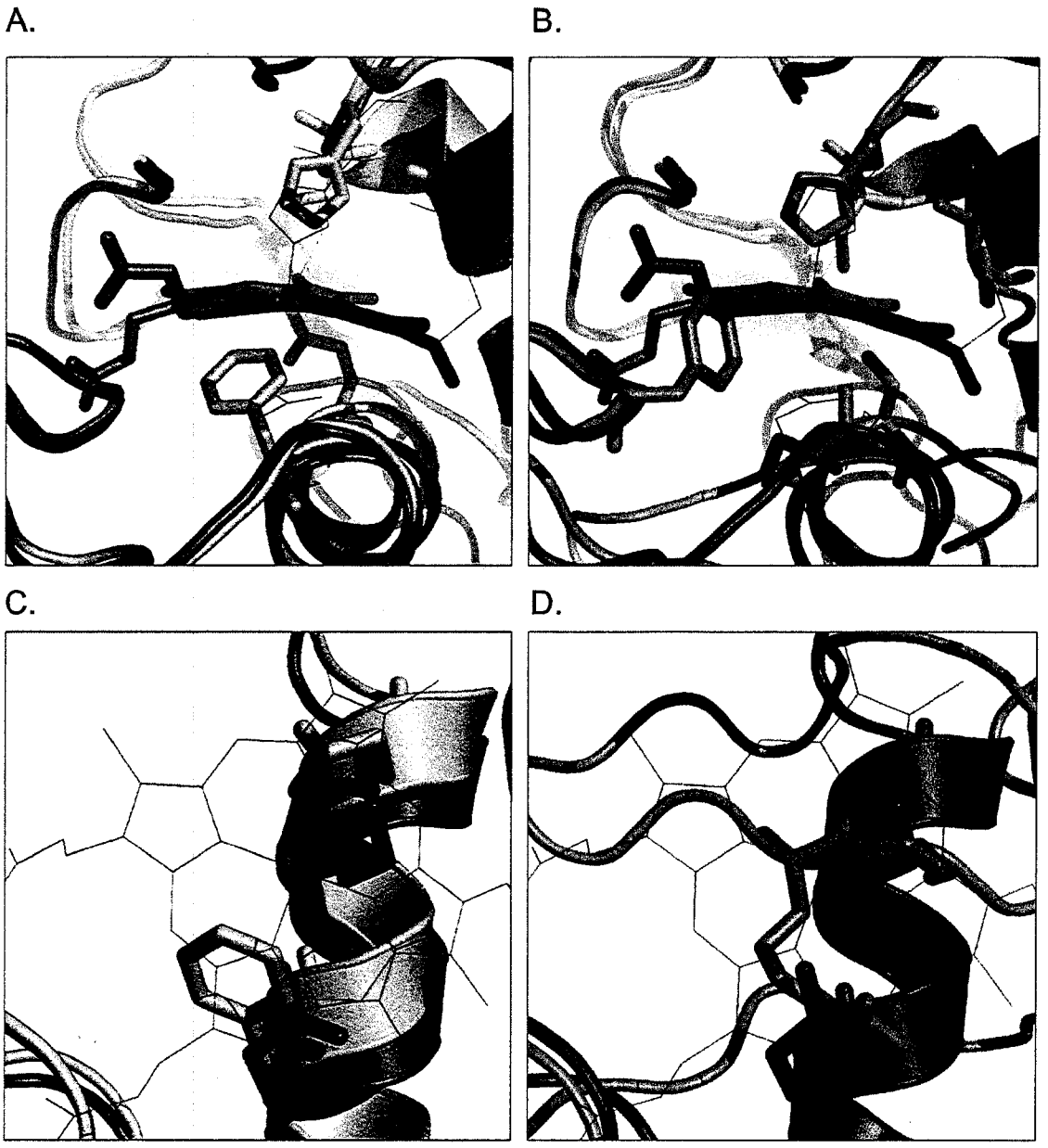
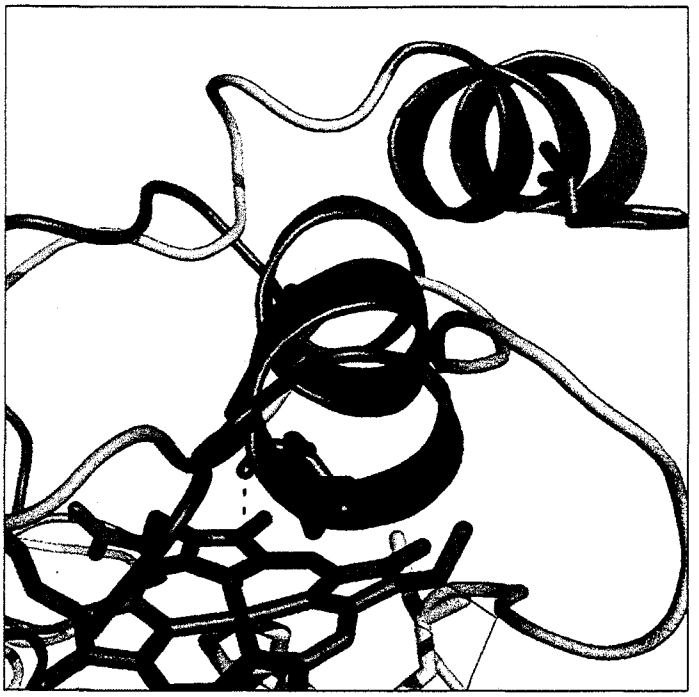



**Figure 5.15. Views of the Predicted *Av*CCC Axial Ligands at the N-Terminal Heme.** The predicted structures were superimposed upon the known structure of *Ne*CCP (purple). The *Ne*CCP proximal ligand His43, the thioester bonded residues Cys39 and Cys42, and the distal ligand Met258 are indicated by thin lines, whereas predicted *Av*CCC axial ligands are indicated by sticks. Side views of the N-terminus of the catalase

A.



B.

**Figure 5.16. Structural Alignment of the Predicted *Av*CCC Helix 2 to that of *Ne*CCP.** Panel A shows that the *Av*CCC helix 2, as predicted by ROBETTA (cyan and blue) is oriented away from the N-terminal heme, whereas *Ne*CCP helix 2 (purple and blue) is located directly above the heme. Panel B shows that *Av*CCC helix 2, as predicted by ESyPred3D (yellow), aligns closely to that of *Ne*CCP.

Both the ESyPred3D and Robetta models are similar at the C-terminus, sharing considerable structural homology with that of *Ne*CCP (Fig. 5.17). Again, all hemes are predicted to have a proximal His residue due to the constraints of the CXXCH pentapeptide. The distal axial ligand in *Ne*CCP is the sulfur group of residue M258, which provides a strong ligand field to the ferric ion. The equivalent residue at the C-terminal heme of *Av*CCC is F442, which sits directly below the heme iron in the ESyPred3D model (Fig. 5.17 C). Based on studies of class 1 cytochrome *c'* proteins, which also have distal Phe residues, it seems likely that this structure would not provide any significant degree of heme coordination.

The interheme region which is critical for BCCP function was also considered, particularly since Trp102 is the only tryptophan in the *Av*CCC primary sequence. Superposition of the favoured ESyPred3D model to that of the *Ne*CCP structure indicated that Trp102 residue does not correspond to that of *Ne*CCP, and that a large void is predicted in the *Av*CCC model (Fig. 5.18). It is unlikely that such a large void would actually persist in the catalase protein however, as small movements of neighbouring residues would likely occur, filling some or all the space (J. N. Glover). This issue can only be resolved by determining the crystal structure.

## 5.3 DISCUSSION

### 5.3.1 Cellular Localization of the Catalases of *A. vinelandii.*

As reported for many cytochrome *c* proteins, the dihemic catalase of *A. vinelandii* is expected to translocate to the periplasmic space. Sphaeroplasting data presented here indicate that this stationary-phase catalase is indeed periplasmically localized (Qurollo *et al.*, 2001) whereas the exponential-phase catalase, *Av*KatG, localizes to the cytoplasm with the exponential-phase FeSOD (Qurollo *et al.*, 2001). Presumably co-localization of the exponentially expressed antioxidant proteins to the cytoplasm gives the greatest protection from endogenously produced reactive oxygen intermediates (ROI). Periplasmic localization of the stationary-phase antioxidant proteins is believed to confer optimal protection from exogenously produced reactive oxygen species (ROS) since they can be degraded before they enter the cytoplasm (Korshunov and Imlay, 2002). As a

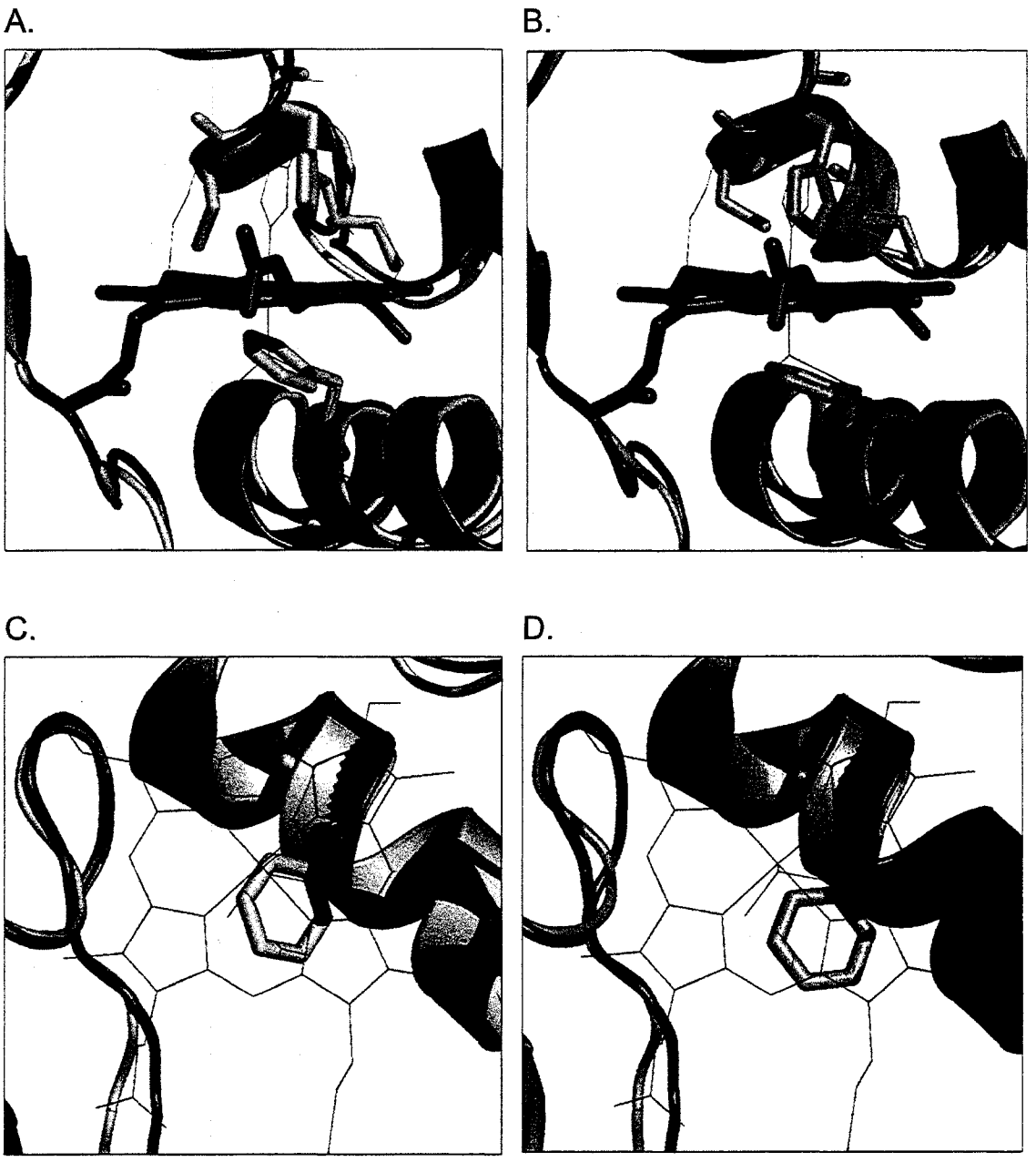



**Figure 5.17 Views of the Predicted *Av*CCC Axial Ligands at the C-Terminal Heme.** The predicted structures were superimposed upon the known structure of *Ne*CCP (purple). The *Ne*CCP proximal ligand His187, and the thioester bonded residues Cys183 and Cys186, are indicated by thin lines, whereas predicted *Av*CCC axial ligands are indicated by sticks. Side views of the catalase C-terminus predicted by ESyPred3D (yellow) and ROBETTA (green) are shown in panels A and B. Top views in which the proximal ligands are hidden are shown in panels C and D.

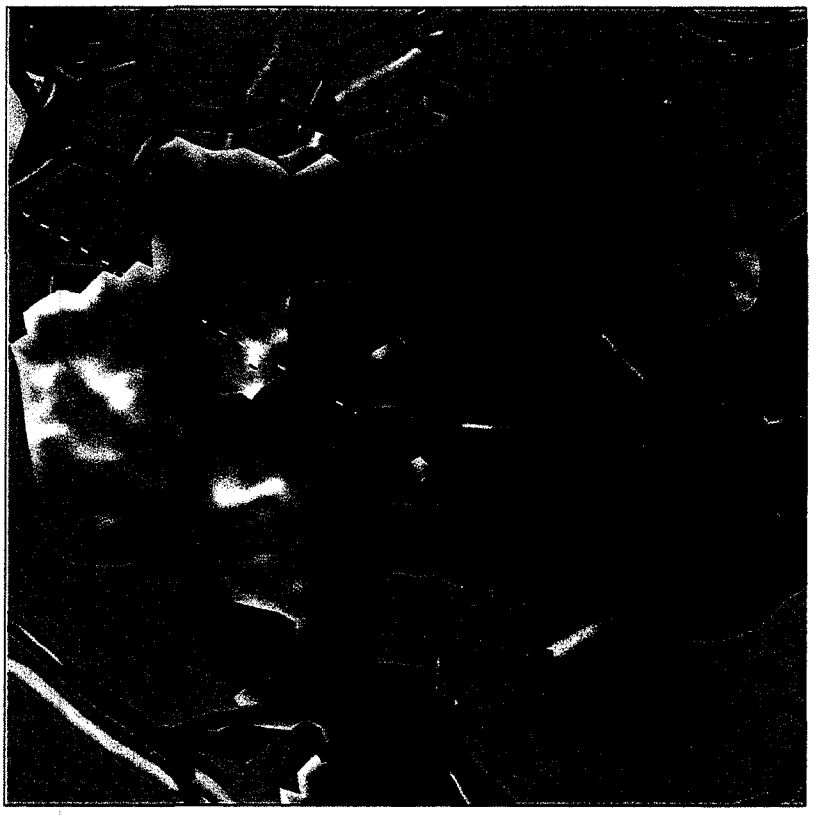


**Figure 5.18 Comparison of the Intra-Heme Cleft Region of *Ne*CCP to that of *Av*CCC.** The conserved BCCP Trp is shown in purple, with distance lines (yellow) indicating the least distance path to the heme irons. A void in the space-filling model of *Av*CCC (grey surface, cyan cartoon) is predicted due to a Trp/Ser replacement in the catalase.

result this may be the preferred method of protection for quiescent cells, as the need for cellular repair can be completely avoided (Loewen *et al.*, 2000). Aspects of these assumptions have been studied in *E. coi* but have not been satisfactorily answered at this time (Messner and Imlay, 1999; Korshunov and Imlay, 2006). It was surprising to find that the periplasmic CuZnSOD, which had previously been reported as being essential for *A. vinelandii* survival (Qurollo *et al.*, 2001), was not detected in this or other studies (Cornish and Page, unpublished).

### 5.3.2 Spectroscopic Analysis of Hemic Coordination and Spin State.

The optical absorption of purified *Av*CCC is virtually identical to the spectra of *Chromatium* ferricytochrome *c*' (Maltempo, 1976) and soy bean peroxidase (Indiani *et al.*, 2000) under neutral oxidizing conditions, for which 5-coordinate quantum-mixed intermediate-high spin heme states have previously been determined. This interpretation is further supported by resonance Raman and EPR spectra for the oxidized protein. As a result, despite the near identical nature of the *Av*CCC spectrum to that of *Dt*Hsc, we prefer the 5c intermediate-high mixed spin interpretation over the HS interpretation (Louie *et al.*, 1997). Whether the mixed state is quantum or thermal in nature has not yet been elucidated, though the presence of intermediate-high spin mixed signals at both room temperature (resonance Raman) and 10 $^{o}$K (EPR) make a compelling argument for a quantum admixture.

Unlike the resonance Raman and UV-visual spectra, a distinctly low spin signal was observed in the EPR spectrum. Based on this observation it seems likely that the two hemes of *Av*CCC have discernable structures and electronic characteristics. It remains unclear if the distinct LS signal observed by EPR is a cryogenic artifact.

Upon alkylation, the appearance of a broad band between 540 and 550 nm with a coincident red-shift of the Soret band to 413 nm indicates that one or more hemes have undergone a spin-coordination switch to 6c-LS, with the sixth axial ligand position being filled with a hydroxyl ion (Smulevich *et al.*, 2006). However this shift is not as extreme for the sphaeroplast-derived enzyme, which highlights the possibility that the 75 $^{o}$C step (during purification from whole cell lysate) might cause damage to the catalase. The

protein is resistant to direct reduction as is true for many catalases and cytochrome *c'* moieties. A coincident alkaline treatment is required to resolve α- and β-bands.

### 5.3.3 Structure Prediction of the Heme Sites of *Av*CCC.

Both the best and second best structure models support a 5c structure at the C-terminal heme, as the predicted distal residue F442 is unlikely to provide coordination to the heme iron. However steric hindrance by the phenyl ring could be sufficient to reduce solvent access, thereby eliminating the formation of aquo-6c ligand structures and related signals. Whereas the N-terminal ligand structure of *Av*CCC conforms poorly to those of known cytochrome *c* crystal structures, which have invariant Met-His coordination, group 1 cytochrome *c'* proteins are known to have Phe-His axial coordination (Tahirov *et al.*, 1996). Coincidentally these Phe-His structures frequently exhibit 5c QS resonance Raman and EPR spectra (Strekas and Spiro, 1974; Maltempo and Moss, 1976; Tsan *et al.*, 2001) which were also observed for *Av*CCC. Taken together these observations bolster the structure prediction at the C-terminal heme.

The N-terminal heme is also predicted to have a nearby distal phenylalanine, F296 (Fig. 5.15C). The extremely tentative nature of structure prediction aside, the nearby residue E300 has the potential of being involved in acid/base reactions with hydrogen peroxide in a manner reminiscent of the active site Glu of chloroperoxidases (Shimizu *et al.*, 2001). *Ne*CCP contains a nearby Glu at the N-terminal heme which has been invoked, though not proven, to coordinate the equivalent reactions (Shimizu *et al.*, 2001).

### 5.3.4 Physiochemical and Kinetic Properties of the Di-Hemic Catalase.

Two distinct $K_M$ were observed for *Av*CCC, a unique observation considering that this phenomenon has not been reported for any other catalase. Previous work has suggested multiple substrate binding sites for some ascorbic peroxidases and their electron donor molecules (Celik *et al.*, 2000). It has also been observed that bifunctional catalases have two distinct catalytic rates; one for catalatic and another for peroxidatic activities. However, to our knowledge no antioxidant enzyme has been shown to exhibit multiple catalatic kinetic rates. Two explanations of this phenomenon are favoured.

First, the enzyme may have two discrete active sites with different enzymatic properties. In the BCCPs the hemes have distinct roles, with the high potential heme accepting electrons from donor molecules, which are then passed to the low potential peroxidatic heme *via* the bridging amino acid, Trp (DeSmet *et al.*, 2006). Given the paucity of Trp and other potential "bridging residues" it may be that the heme groups of *Av*CCC are insulated from each other, each constituting a separate enzymatic centre. Furthermore, catalatic dismutation of hydrogen peroxide does not require exogenously supplied electrons, so the acquisition (or maintenance) of such an inter-hemic electron-flow pathway, would be superfluous. It is interesting to note that a significant unfilled pocket exists in the *Av*CCC model where it overlaps the conserved Trp of the BCCPs (Fig. 5.18). In the BCCPs this bridging residue reduces the electron transit distance to only 9.3 Å (heme iron to Trp), rather than the full inter-hemic distance. If the *Av*CCC model is accurate, it would mean there is an approximately 20 Å gap between the hemes, which would reduce, if not block, heme-heme communication. Indeed, W97F and W97A mutations result in drastically reduced peroxidatic rates in *Rc*CCP (DeSmet *et al.*, 2006). Second, it is possible that there are two or more distinct peroxide binding sites, and that the observed kinetics are due to (antagonistic) cooperative binding at higher substrate concentrations. Isolation of the function of each heme by mutagenesis or crystallography may be required to determine which of these hypotheses, if either, is correct.

The enzyme was sensitive to the inhibitors NaCN, $NaN_3$, and $NH_2OH$ but not to 3-amino-1,2,4-triazole. Cyanide typically acts as an exogenous ligand at the active heme site, essentially blocking $H_2O_2$ access to the active site. Azide, which acts much as cyanide does, likely acts as an exogenous ligand in an equivalent manner but binds at a different angle. Taken together, the sensitivity of the enzyme to these inhibitors suggests that the active sites are either penta-coordinated or hexa-coordinated with solvent, so are sensitive to exogenous ligands. The insensitivity of *Av*CCC to triazole may indicate that a suitable amino acid target (i.e. histidine) is not present in the distal heme pocket, or that the inhibitor cannot get to the active centre due to the physiochemical nature of the access channel or channels. The successful inhibition by hydroxylamine, which acts by a similar mechanism (Wariishi *et al.*, 2000), suggests that the triazole resistance is primarily due to the inability of the compound to navigate the access channel.

As reported for *Av*CCC in whole-cell lysates (Chapters 3 & 4), the purified catalase was also heat resistant. We found that the residual activity of the purified enzyme remained high following heat treatment at 75 °C, and that the enzyme retained 18% activity following heat treatments as high as 92 °C. The half-life during treatment at 65 °C was comparable to that of *Ec*HPII (12 h; Switala and Loewen, 2002). Multimer stability of *Av*CCC was slightly greater than that of *Ec*HPII in the presence of urea. Furthermore the *Av*CCC complex only partially dissociated when heated to 100 °C in aqueous solution (prior to the addition of load dye) whereas *Ec*HPII dissociated completely at this temperature (Switala *et al.*, 1999). Thus, as was seen with *Ec*HPII, *Av*CCC is particularly resistant to heat and chemical denaturation, which is somewhat surprising for enzymes purified from mesophilic organisms; doubly so considering the observed instability of many dihemic BCCP and MauG proteins during purification (Foote *et al.*, 1983; Timoteo *et al.*, 2003; Wang *et al.*, 2003).

### 5.3.5 Phylogenetic Analysis of *Av*CCC.

Considering that the number of differences between *Av*CCC and the nearest MauG (*Me*MauG, 0.7998 changes) or BCCP (*Pn*CCP, 0.8149 changes) family members are nearly as great as those to the outgroup, KatG (*Av*KatG, 0.8962), we suggest that *Av*CCC and its closest homologues do not share a (recent) phylogenetic heritage with any of these previously identified protein families. That is to say, on the basis of amino acid sequence, *Av*CCC and the proteins *Pa*Snr-1 and *Dt*Hsc constitute a unique and distinct protein family, despite their co-grouping with the BCCP/MauG families by COG and pfam. Indeed, even the validity of the grouping of the MauG proteins with the closely related BCCPs has been challenged (Wang et. al., 2003). Considering the low sequence homology, it is interesting to see the high degree of structure homology with the dihemic peroxidases. These similarities likely reflect the functional constraints for antioxidant functions around a heme *c* moiety.

### 5.3.6 Summary

The stationary phase catalase of *A. vinelandii* exhibits considerable resistance to heat and chemical damage, properties previously observed for *E. coli* HPII and the Mn

catalases, yet the enzyme is distinctly heme *c*. The enzyme is also unique in that it exhibits two distinct affinities for the substrate, hydrogen peroxide. This may be indicative of two active sites, or of cooperative (antagonistic) substrate binding.

Phylogenetic comparisons of *Av*CCC to the bacterial CCP, MauG, and the outgroup KatG suggest that the dihemic catalase shares little if any phylogenetic relationship with other dihemic proteins. However structure modeling suggests that it is structurally similar to the dihemic peroxidase of *N. europaea*. While the significance of this dichotomy remains unresolved in this work, the structure models are largely supported by the UV-visible, resonance Raman and EPR spectral interpretations. These methods strongly suggest that at least the C-terminal heme is found in an intermediate-high spin mixed state and that, based upon previous cytochrome *c'* studies, the admixture is quantum rather than thermal in nature. Some support for the observed LS EPR signal may be found in the ESyPred3D structure at the N-terminal heme.

# CHAPTER 6

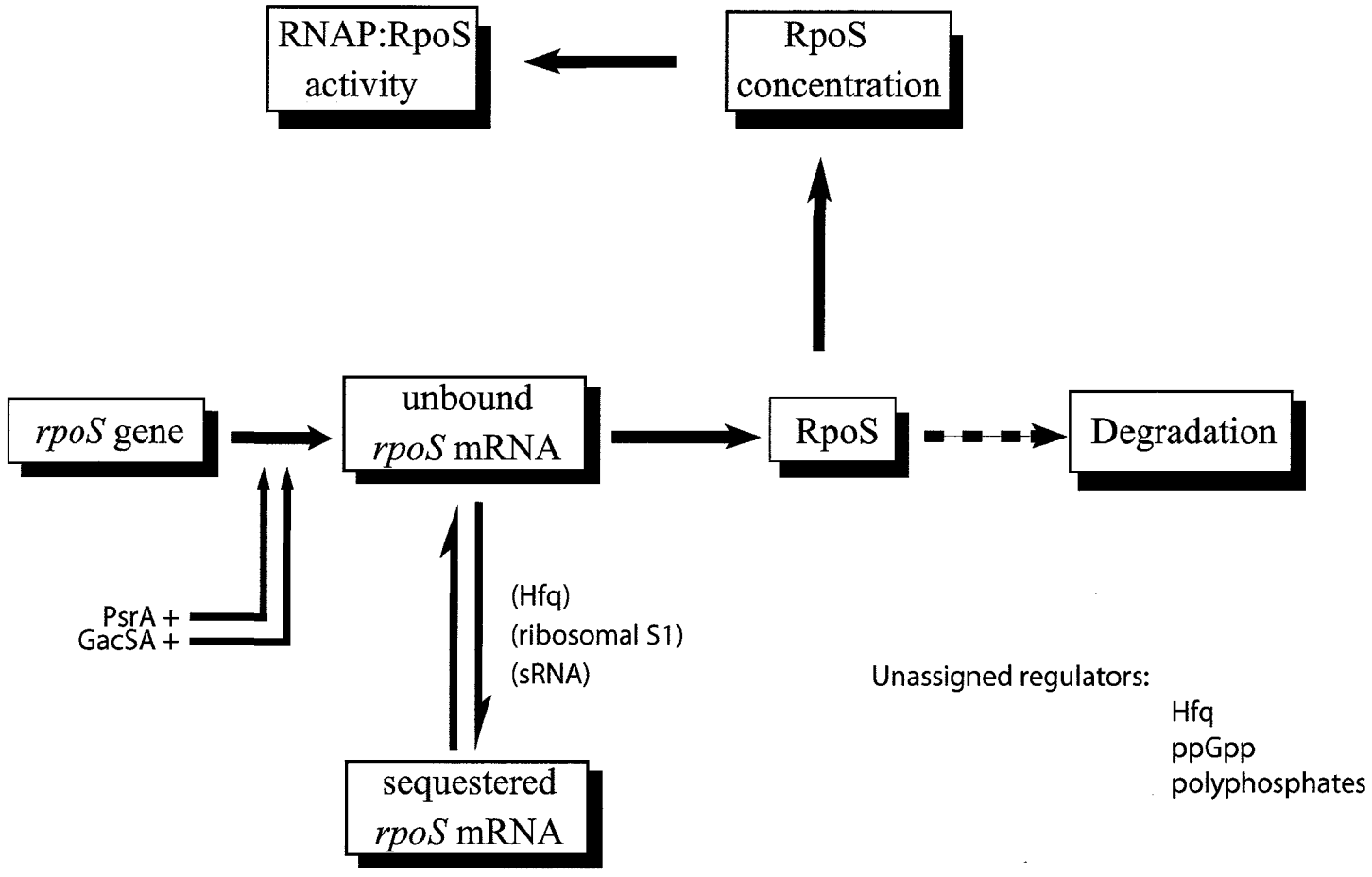
# General Discussion & Future Studies

## 6.1 The Expression of RpoS in *Azotobacter vinelandii*

The diazotrophic bacterium *Azotobacter vinelandii* exhibited a very high degree of RpoS induction during carbon-source and nitrogen-source diauxic shifts. This was followed by nearly as rapid a decrease in RpoS levels upon recovery of growth. During the acetate to glucose diauxic shift the *rpoS* mRNA levels followed a similar pattern, just slightly in advance of the RpoS concentrations. Thus we propose a model in which the general stress response of *A. vinelandii* is strongly regulated at the level of transcription (Fig. 6.1) during carbon-diauxie. It is unclear if the two component regulator GacS/A, already studied in *A. vinelandii* (Castañeda *et al.*, 2000), accounts for all the transcriptional control at the *rpoS* promoter(s) or if other regulators such as PsrA might also have a role. Judging from the complexity of regulation at the *E. coli rpoS* promoters, it seems likely that additional factors will be found to regulate *rpoS* transcription in *A. vinelandii*.

No particularly unique stress response was observed at the protein level during nitrogen diauxie, except that basal RpoS concentrations were slightly elevated following growth recovery. However significantly elevated levels of *rpoS* mRNA were found to persist during nitrogen-fixing growth, despite a simultaneous decline in RpoS protein levels. It is possible that the cells are transcribing large quantities of the mRNA during this period, explaining the high transcript levels. An elevated transcription rate seems counter-intuitive however, considering that the organisms are attempting to grow under the considerable energy burden of nitrogen-fixation. From an energy-conservation perspective we instead propose that *A. vinelandii* sequesters the *rpoS* mRNA in a protected but inactive form, such that the cells are found in a "stress-poised" state (Fig. 6.1). The observation following glucose exhaustion, that the RpoS concentration increases despite a rapid decrease in *rpoS* mRNA levels, could be explained by the release of the transcript from the sequestered state; the mRNA would be translationally re-activated but would also become sensitive to degradation. Both proteinaceous and anti-sense RNA moieties have been shown to bind and protect transcripts from RNase degradation (Jerome and Frost, 1999; Tomizawa, 1986). However the largely irreversible nature of RNA:RNA complex formation likely precludes anti-sense or sRNA candidates, as it would be difficult to accommodate the reversibility of the model. At this time the

**Figure 6.1** A Model of RpoS Regulation in *Azotobacter vinelandii*. A schema indicating the levels at which expression or function are controlled for the sigma factor, RpoS, in *A. vinelandii*. The levels at which *A. vinelandii* RpoS regulation has been clearly demonstrated are shown by the solid arrows, whereas levels at which there is no specific regulation are shown by the dashed arrows. Molecular regulators that are known to act in *P. aeruginosa* RpoS regulation are indicated by small arrows or are indicated under the category "unassigned" if the level(s) at which they act are unknown. The newly proposed *rpoS* mRNA sequestration branch is shown by reverse arrows since the pathway is believed to be reversible. The most likely regulatory candidates are indicated in brackets.


RNAP:RpoS activity
RpoS concentration
rpoS gene
unbound rpoS mRNA
RpoS
Degradation
PsrA +
GacSA +
(Hfq)
(ribosomal S1)
(sRNA)
sequestered rpoS mRNA
Unassigned regulators:
Hfq
ppGpp
polyphosphates

most likely protein candidates include Hfq, which plays a central regulatory role in the translation efficiency of *E. coli rpoS* mRNA and is known to regulate RpoS expression in *P. aeruginosa* (Sonnleitner *et al.*, 2003). In addition to helping pair sRNAs to their target RNA species, Hfq also offers protection from RNase E degradation (Moll *et al.*, 2003; Gottesman 2004), which could explain the "excess *rpoS* mRNA" phenomenon observed during $N_2$-fixation. Another prime candidate is the ribosomal accessory protein, S1, which was found to bind the 5' UTR of *P. aeruginosa rpoS* mRNA during exponential growth but not at all during stationary phase (Sevo *et al.*, 2004). Though shown to be required for the "activation" of a very significant portion of mRNAs prior to translation (Sorensen *et al.*, 1998), there are instances when S1 is believed to be in excess, so is not associated with the ribosome. Under these conditions it is possible that S1 sequesters rather than "activates" mRNA species. Furthermore S1:RNA complexes are proposed to occlude RNase E activity on these transcripts (Koleva *et al.*, 2006) possibly offering some degree of protection from degradation. Unfortunately studies have been hampered by the essential nature of the protein. However the larger hypothesis, that the *rpoS* mRNA is being sequestered in a stable form, could be tested by mRNA turnover studies.

Despite the strong conservation of the *E coli* RpoS RssB binding domain centered on K173, no evidence for regulation at the protein stability level was observed in the protein half-life studies of *A. vinelandii* RpoS. We propose that the high degree of sequence conservation throughout regions 2.4 and 2.5 of these sigma factors has more to do with promoter sequence recognition than proteolytic regulation.

It was also observed in these studies that the expression of the RpoS-dependent reporter (catalase *Av*CCC) was not directly proportional to the RpoS levels throughout much of the growth on BBGN. This observation suggests that regulation of RpoS also occurs at the $E\sigma^S$ formation-activity level, as has been frequently cited for the enterics, but not the pseudomonads. Now that the gene responsible for the RpoS-dependent catalase has been identified, future work should examine the regulation of RpoS activity by comparing RpoS:*cccA* mRNA expression ratios.

In summary, the general stress response of *A. vinelandii* is regulated at the transcriptional level, the translational level involving a possibly unique sequestration

mechanism, and at the level of $E\sigma^S$ formation/activity. At this time the only confirmed regulator of the general stress response in *A. vinelandii* is the response-regulator pair, GacA/S (Castañeda *et al.*, 2000). In agreement with earlier studies on *P. aeruginosa*, there appears to be no proteolytic regulation mechanism as described for the enteric bacteria.

## 6.2 Further Investigations into the Roles and Regulation of *A. vinelandii* Sigma S

### 6.2.1 The Lowest Hanging Fruit: Potential Molecular Regulators of RpoS

Despite the considerable advances in genetic manipulation techniques in *E. coli* and to a lesser degree the pseudomonads, *A. vinelandii* techniques have lagged behind for a number of technical and economic reasons. While transformation methods have been employed for years (Page and von Tigerstrom, 1978) the creation of mutants remains a non-trivial issue for *A. vinelandii* and a near impossibility for the closely related species *Azotobacter salinestris* and *Azotobacter chroococcum*. None-the-less several high priority regulatory moieties could be the focus of immediate genetic study.

A homologue of the pseudomonas regulator, PsrA, has been found in an analysis of the *A. vinelandii* genome, and shares 67% identity with that of *P. aeruginosa*. Comparison of *rpoS* mRNA levels in "wild-type" and *psrA* strains during nitrogen diauxie would be of interest, since the regulator has been shown to be important for *rpoS* transcription. Combined with mRNA half-life studies it should be possible to determine if the *rpoS* mRNA excess observed during this period of growth/growth recovery is due to increased transcriptional activity or transcript stabilization.

If mRNA stabilization is found to be significant during nitrogen diauxie, then two prime candidates would include the RNA chaperones Hfq and ribosomal protein S1. While *hfq* interruption should be relatively simple to achieve, it would be impossible to execute equivalent studies with S1 because the protein is essential. Instead *rpoS* mRNA:S1 *in vitro* binding studies could be pursued, as could *in vivo* cross-linking studies (Sevo *et al.*, 2004; Koleva *et al.*, 2006).

sRNA have only recently been identified in *P. aeruginosa* by homology searches (Livny *et al.*, 2006). If the frequency of *rpoS*-regulating sRNAs observed in *E. coli* holds

true, then at least five sRNAs should regulate RpoS translation in *P. aeruginosa.* Though it would seem easiest to identify sRNA species in *A. vinelandii* by homology methods, it should be remembered that the phylogenetic differences between pseudomonas species often exceeds those between enteric genera. Indeed, preliminary sRNA searching using the automated homology searching program, IGCompare v0.99 (Stothard and Sandercock, unpublished), was able to replicate previous studies between *E. coli* and *S. enterica* (Wassarman *et al.*, 2001), but performed poorly when comparing *A. vinelandii* and *P. aeruginosa.* With the recent release of more annotated pseudomonas genomes such comparative work will continue to be executed on a variety of clinical and environmental *P. aeruginosa* isolates (Livny *et al.*, 2006). However alternative methods will have to be employed for the *Azotobacteracaea* since these species are unlikely to garner the financial support necessary for the completion of additional genome projects.

*E. coli* expresses two HU subunits, HUα and HUβ, allowing the cell to produce either a homo- or hetero-dimer during different phases of growth (Claret and Rouviere-Yaniv, 1997; Balandina *et al.*, 2001). While it is known that HU regulates RpoS translation, it is not yet clear whether the transition from the homodimer to heterodimer form is important for this regulation, particularly as only double mutants have been studied in *E. coli*. Intriguingly *A. vinelandii* has three HU homologues. Studies employing both single and multiple mutations could yield interesting information regarding *rpoS* regulation in the pseudomonads, while simultaneously catapulting this area of study beyond that of *E. coli*.

Tetraphosphate is known to regulate RpoS at both the expression and activity levels in the enterics. While it is known that ppGpp actively regulates RpoS expression in *P. aeruginosa* (Erickson *et al.*, 2004) it is unclear at what level this regulation occurs. Furthermore it is unclear if ppGpp affects $E\sigma^S$ formation or activity rates. Finally, the type of nutrient limitation has a significant effect on the degree to which RpoS is regulated by ppGpp; C-source limitation results in a much more significant induction of RpoS under ppGpp control than does N-limitation. Unfortunately for the purposes of this thesis most ppGpp-related studies in *P. aeruginosa* have focused on its role in quorum sensing rather than the general stress response, and have never considered single nutrient limiting events. Both *relA* and *spoT* genes have been identified in the *A. vinelandii*

genome and should be amenable to mutation. Studies employing radioactively tagged exogenous ppGpp may not be as successful in *A. vinelandii* however, as the organism has proven recalcitrant to the uptake of many nutrients (W. J. Page, personal communication).

### 6.2.2 Development of an Effective Reporter of $E\sigma^S$ Activity

With the identification of an absolutely RpoS-dependent gene it is now possible to measure the degree of activity of RpoS during various phases of growth or stress conditions in *A. vinelandii*. It will be important to first delineate the promoter region of *cccA* to ensure that it is independent of other regulators, which would vastly improve its value as a biological read-out of $E\sigma^S$ activity. If few additional regulators act at the promoter then Northern analysis of the *cccA* transcripts would constitute the most immediate reporter of RpoS activity. Measurement of catalase activity, an example of a "proxy reporter method" analogous to the frequently utilized *lacZ* fusion method, could also be used. While infinitely easier to perform than Northern analysis, such proxy methods can result in anomalous results, especially when the stability of the protein product differs significantly from that of its transcript.

## 6.3 Significance of the Newly Described Heme *c*-Type Catalase

Cytochrome *c* type proteins differ from other hemic proteins in that the protoheme IX moiety is covalently bound to the peptide *via* one or two thioether bonds. The question of how essentially identical hemes can catalyse such divergent functions as oxygen transport, electron transfer, photosynthetic reaction centers, hydrocarbon activation, peroxidatic or catalatic dissociation of peroxides, and oxygenation of tryptophan residues has been posited innumerable times (e.g. Perutz 1992; Wang *et al.*, 2003). The fact that no heme *c* catalase had been reported and the consideration of the extreme sensitivity of KatG catalatic but not peroxidatic functions to perturbations (Hillar *et al.*, 2000; Chelikani *et al.*, 2004; Santoni *et al.*, 2004) suggested (negatively) that there might be more restrictive structural limitations on the catalatic function. With the discovery of a dihemic *c*-type enzyme capable of the "missing" catalase function, we may

now have a subject from which the minimal context of each function can be determined for the heme *c* moieties. We report here the identification and initial characterization of a cytochrome *c* enzyme capable of the catalatic degradation of hydrogen peroxide.

The catalase exhibits a high degree of heat, protease and chemical resistance, which is somewhat surprising for a mesophilic organism such as *A. vinelandii*. However the cyst form of the organism, being quiescent yet viable over decades (Vela, 1974), likely requires an anti-oxidizing agent that remains active over at least a portion of this time. As such, it is likely that the observed heat and chemical resistances are simply coincidental to the selected protease resistance (Chelikani *et al.*, 2004). Additionally a catalatic, rather than peroxidatic, enzyme should be theoretically preferable for dormant cells since a catalase would be free of the peroxidatic requirement for active metabolism (to supply reducing power). The physiological importance of the catalase to stationary phase *A. vinelandii* was clearly demonstrated during hydrogen peroxide stress survival studies (Table 4.2). In contrast, many of the most heavily studied cytochrome *c*/*c'* proteins still do not have known functions. Equally surprising, while the biochemistry of the BCCPs have been thoroughly delineated, the most basic information regarding their regulation and physiological purposes have, with one notable exception (*Ng*CCP; Lissenden *et al.*, 2000; Turner *et al.*, 2003; Stohl *et al.*, 2005) not been determined.

Studies examining the importance of *Av*CCC to dormant cell survival are ongoing and will require testing over a period of years. As a result findings could not be reported here.

## 6.4 Further Study of the Heme *c* Catalase

### 6.4.1. Biochemical Tests

Several interesting preliminary biochemical observations have been made for the dihemic catalase, which highlight the novelty of the enzyme. Amongst these are the apparent dual kinetic rates of activity, the resistance of the protein to reduction, the degree of resistance to typical catalase inhibitors, and the unusual 5c mixed high-intermediate spin properties observed by three different spectrometric methods. The favoured hypothesis regarding the apparent dual $K_M$ for $H_2O_2$, that there are two distinct

catalatic sites within the enzyme, could be best tested by disrupting the heme regions by site directed mutagenesis or perhaps deletion analysis (discussed below). The possibility that the phenomenon is due to negative cooperative binding of the substrate could be tested by more classical enzymatic means employing different inhibitory compounds.

The spectroscopic response of the catalase to various inhibitors can also provide valuable insights into the nature of the ligands at each of the hemes. Studies in which substrate analogues such as $CN^-$, $N_3^-$, CO, NO or $F^-$ are introduced could yield a wealth of information regarding enzyme function, structure and classification.

It will also be important to determine if the observed high-intermediate mixed spin state of the enzyme is thermal or quantum in nature. The fact that strong mixed signals were observed during both room temperature and cryogenic experiments suggests that the phenomenon is not thermally derived. However a more robust magnetic susceptibility study across a wide range of temperatures (e.g. 6-250 $^{o}$K; Tsan *et al.*, 2001) could provide not only a more definitive answer, but also discern the degree of contribution from the 5/2 and 3/2 forms.

While the resonance Raman and EPR data are in agreement regarding the presence of a mixed-spin heme, the paramagenetic experiment also indicated the presence of a low spin signal. It is possible that the LS product is a cryogenic artifact. Alternatively the excitation frequency used in the Raman study (412.5 nm), which was quite distant from the maximum absorption of the catalase (ca. 400 nm), might not have excited the putative LS heme sufficiently for detection. Other research groups found that probing at multiple wavelengths can allow for the detection of weak or hidden coordination spectra (Othman *et al.*, 1996). Such a system could employ a $Kr^+$ laser with a wavelength of 406.7 nm and a He-Cd laser with a wavelength of 441.6 nm to probe at lower and higher frequencies, respectively.

### 6.4.2 Structure Determinations

The structure prediction models proposed in Chapter 5 are supported in part by the resonance Raman and EPR data, particularly for the structure at the C-terminal heme. If accurate, the prediction of a distal phenylalanine, combined with the observed 5c high-intermediate admixed spin signals, suggest that the catalase should be classified as a

Group 1 cytochrome *c'* protein (Tahirov *et al.*, 1996). It should be noted however that this would be a strictly phenomenological classification, and that phylogenetic classifications for the bacterial heme *c/c'* proteins is in considerable disarray due to historically derived eukaryotic heme protein classification schema.

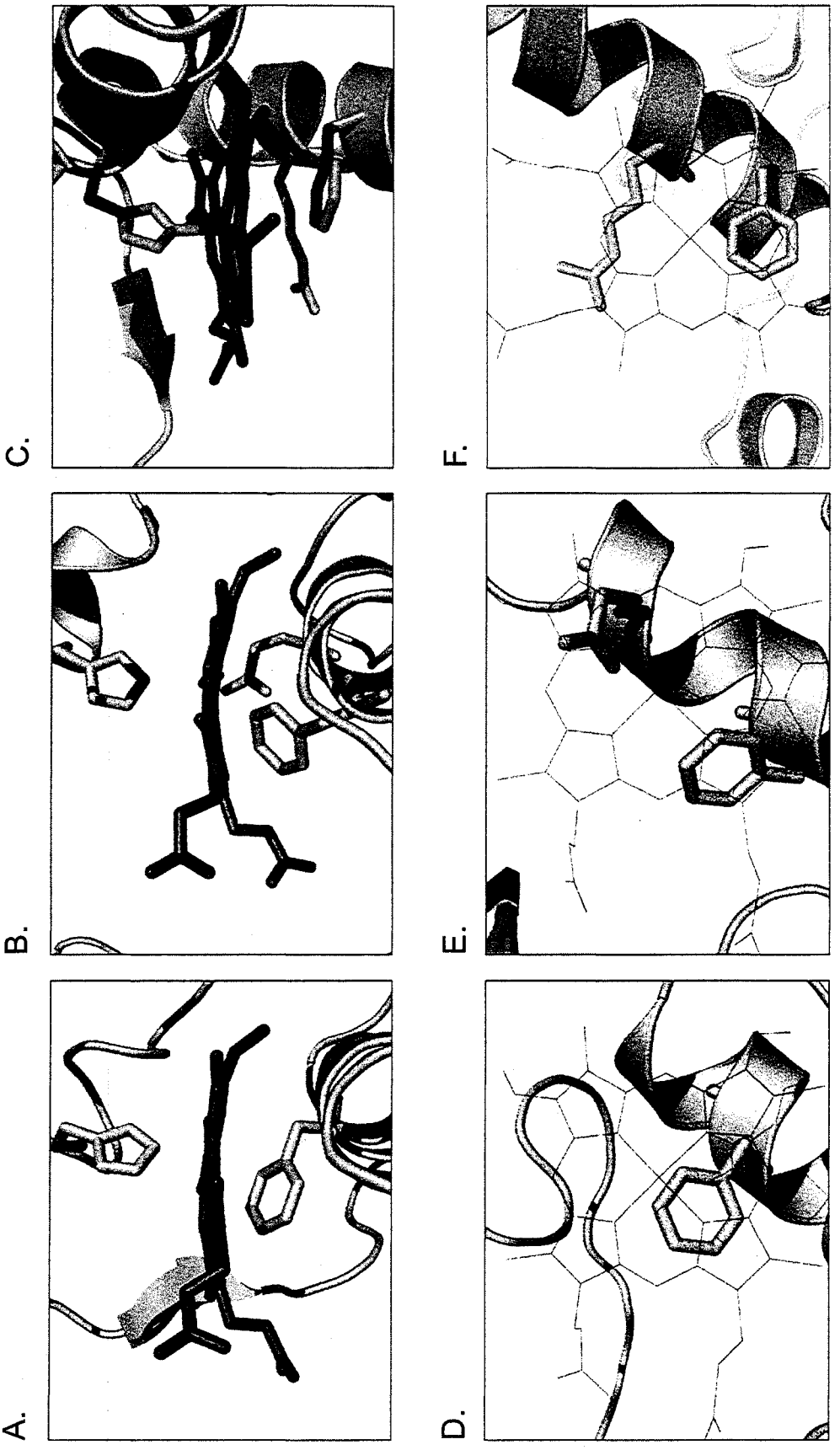
Information regarding the porphyrin topologies could be discerned by low frequency resonance Raman as several bands have been shown to be indicative of the degree and types of heme torsion (Haddad *et al.*, 2003). Preliminary data from the $\nu_{10}$ band (1640 $cm^{-1}$; Fig. 5.10) suggest that the *Av*CCC hemes are in an extremely planar conformation, with values similar to those seen for other mixed high-intermediate spin hemes (BP-1, Howes *et al.*, 1999; cytochrome *c'*, Hobbs *et al.*, 1990). Furthermore, the predicted heme-pocket structures of *Av*CCC are remarkably similar to that of the unrelated plant peroxidase BP-1 (Fig. 6.2), except that the phenylalanine side chains of *Av*CCC appear to be in closer proximity to the porphyrin ring.

Crystallographically derived structures could probably resolve most of the questions regarding the ligand fields at each of the *Av*CCC hemes. Additionally, the non-hemic structures of the protein that are not amenable to spectroscopic study, such as the multimerization region, the unique N-terminal extension, the inter-hemic region and the putative $Ca^{2+}$ binding site, could be resolved. However crystal formation often occurs only under exotic solvent or non-physiological conditions. Of particular note is the extreme acid (pH 4.0 – 5.3) requirement for *Pp*CCP crystal formation (Echalier *et al.*, 2006). It is questionable as to whether this structure is reflective of the protein under normative conditions. So while crystal structures may well be considered the gold-standard, particular attention should be given to spectroscopic determinations such as resonance Raman and EPR to ensure that the derived structure is meaningful.

### 6.4.3 Functional-Ligand Probing

Site directed mutagenesis may be an effective method for determining what it is about the *Av*CCC context that allows its *c*-type heme(s) to exhibit catalatic rather than peroxidatic functions. Additionally, purification could be greatly aided by the introduction of a $His_6$ or other epitope tag. However several caveats should be high-

**Figure 6.2** Comparison of the predicted *Av*CCC heme ligands to those of the known 5c QS heme of BP-1. The front view of the predicted C-terminal and N-terminal heme regions of *Av*CCC are shown in panels A and B, respectively. Potential axial ligands are shown as sticks against the cartoon background. The known structure of the BP-1 heme pocket is shown for comparison in panel C. In both cases the distal phenylalanine is believed to disrupt coordination at the heme, resulting in hemes exhibiting 5-coordinate quantum admixed spin signals. The N-terminal structure (panel B) may have enough room to accommodate a water molecule between the Phe and Glu residues, accounting for the [aquo-His 6c] LS signal observed by EPR. Top down views are shown for the *Av*CCC C-terminal, N-terminal and BP-1 hemes in panels D through F, respectively.


A.
B.
C.
D.
E.
F.

lighted. First, *A. vinelandii* contains multiple genomes during certain phases of growth, making non-selective mutagenesis unrealistic. Furthermore the lack of a *relA* mutant makes plasmid maintenance equally difficult, meaning that any expression advantages gained by cloning the gene into a commercial vector would be lost upon crossover to the genome. An obvious solution is to express the enzyme transgenically in a host such as *E. coli*. However the successful expression of dihemic *c*-type peroxidases has rarely been reported, and even more critically only one publication has reported success in the mutagenesis and transgenic expression of a BCCP (De Smet *et al.*, 2006). It is unclear why transgenic expression of this protein family is so poor, but it may be due to attempts to exclusively use *E. coli* as the host. Expression in *P. aeruginosa* may yield better results, at least for an *A. vinelandii* derived protein. Though not utilized for BCCP expression a heme *c* maturation helper plasmid, previously reported to improve heme *c* maturation rates during over-expression studies, may prove helpful with *Av*CCC expression (Thöny-Meyer, 1997).

Should these expression issues be overcome, some immediate targets for mutagenesis should include the heme axial ligands. At the C-terminus F442 replacement with other typical heme ligands, such as methionine, histidine and cysteine, could reveal much about the coordination state of the heme. An equivalent residue exchange at the N-terminus would also be valuable, though both of the putative ligands (F296 and E300) should be examined. Mutagenesis of the heme *c* binding pentamer to AXXCH or CXXAH rarely results in the production of apo-proteins since the surrounding amino acid context frequently allows for correct heme positioning (Allen *et al.*, 2005). More radical changes in the pentamer sequence could potentially produce the desired heme-vacant site, but are more likely to result in undesirable structure changes. A better method would be to replace the $6^{th}$ axial ligand with a non-coordinating residue such as alanine. Such a minimal interruption should allow for the determination of which heme constitutes the active site(s). Mutagenesis of the heme sites individually would allow us to test the "two active site" hypothesis put forward in Chapter 6.

Mutagenesis of the inter-hemic region is less likely to yield interesting results. The introduction of the BCCP/MauG conserved Trp residue at S276 is unlikely to result in the "recovery" of peroxidatic function, as other residues needed for such functions

likely degenerated long ago, assuming that the protein ever had peroxidatic capacities in the first place. However it remains possible that S276P or S276W mutants could exhibit some novel activities due to the introduction of interheme communication.

Regardless of the mutagenic method employed mutant products would need to be examined by at least Circular Dichroism to ensure that the overall structure had not been changed to an unacceptable degree.

### 6.4.4 Characterization of the Phylogenetically Related Di-Heme Proteins

Despite mutagenic and biophysical studies of the closely related proteins, *Pa*Snr-1 and *Dt*Hsc, no definitive functions have been ascribed to either. In the case of *Dt*Hsc, the UV-visual spectra are nearly identical to those of *Av*CCC, suggesting that the heme axial-ligations are similar. Additionally, considerable conservation at the primary sequence level can be seen around the putative axial ligands (F296, E300 and F442), the putative calcium binding domain (S423-G430), and at the inter-hemic region surrounding S276 (Fig 4.6). *Dt*Hsc also has an exceptionally low midpoint potential of –342 mV (Louie *et al.*, 1997), making it a reasonable candidate for both peroxidatic and catalatic functions. The description of a new catalase in *P. aeruginosa* would be rewarding, considering the degree to which the organism has been studied. The fact that the genome also encodes a highly probable Mn catalase, which has never been observed experimentally, highlights the complexity of the organism and the limited capacity of laboratory culturing methods to fully describe the capacity of these organisms. Should either of these homologues exhibit catalatic activities they, together with *Av*CCC, would constitute the first new class of catalase since the discovery of the Mn catalases (Kono and Fridovich, 1983).

# REFERENCES


Aladegbami, S.L., Tsai J.C., and Vela, G.R. (1979) Adenilate energy charge of *Azotobacter vinelandii* during encystment. *Curr Microbiol* **2:** 327–329.

Allgood, G.S., and Perry, J.T. (1986) Characterization of a manganese-containing catalase from the obligate thermophile, *Thermoleophilum album*. *J Bact* **168:** 563-567.

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., 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.

Altuvia, S., Zhang, A., Argaman, L., Tiwari, A., and Storz, G. (1998). The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO J* **17:** 6069-6075.

Amo, T., Atomi, H., and Imanak, T. (2002) Unique presence of a manganese catalase in a hyperthermophilic archaeon, *Pyrobaculum calidifontis* VA1. *J Bacteriol* **184:** 3305-3312.

Andersson, R.A., Palva, E.T., and Pirhonen, M. (1999) The response regulator *expM* is essential for the virulence of *Erwinia carotovora* subsp. *carotovora* and acts negatively on the sigma factor RpoS (sigma S). *Mol Plant Microbe Interact* **12:** 575-584.

Andreoletti, P., Field, M.J., Gouet, P., and Jouvre, H.M. (1995) Simulations of electron transfer in the NADPH-bound catalase from *Proteus mirabilis* PR. *Biochim Biophys Acta* **1252:** 172-176.

Andreoletti, P., Pernoud, Sainz, G., Gouet, P., and Jouve, H.M. (2003) Structural studies of *Proteus mirabilis* catalase in its ground state, oxidized state and in complex with formic acid. *Acta Cryst* **59:** 2163-2168.

Archibald, F.S., and Fridovich, I. (1983) Oxygen radicals, oxygen toxicity and the life of microorganisms. *Acta Med Port* **4:** 101-112.

Arcierio, D.M., and Hooper, A.B. (1994) A di-heme cytochrome *c* peroxidase from *Nitrosomonas europaea* catalytically active in both the oxidized and half-reduced states. *J Biol Chem* **269:** 11878-11886.

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H., and Altuvia, S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli. Curr Biol* **11:** 941-950.

Asada, K. Yoshikawa, K., Takahashi, M., Maeda, Y., and Emanji, K. (1975) Superoxide dismutase from a blue green algae, *Plectonema boryanum. J Biol Chem* **250:** 2801-2807.

Atlung, T., and Ingmer, H. (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* **24:** 7-17.

Azam, T.A., Hiraga, S., and Ishihama, A. (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* **5:** 613-626.

Azam, T.A., and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli.* Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274:** 33105-33113.

Balandina, A., Claret, L., Hengge-Aronis, R., and Rouviere-Yaniv, J. (2001) The *Escherichia coli* histone-like protein HU regulates *rpoS* translation. *Mol Microbiol* **39:** 1069-1079.

Ballesteros, M., Kusano, S., Ishihama, A., and Vicente, M. (1998) The ftsQ1p gearbox promoter of Escherichia coli is a major sigma S-dependent promoter in the *ddlB-ftsA* region. *Mol Microbiol* **30:** 419-430.

Barker, M.M., Gaal, T., and Gourse, R.L. (2001) 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. (2001) 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.

Barth, M., Marschall, C., Muffler, A., Fischer, D., and Hengge-Aronis, R. (1995) Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of sigma S and many sigma S-dependent genes in *Escherichia coli*. *J Bacteriol* **177:** 3455-3464.

Barynin, V.V., Hempstead, P.D., Vagin, A.A., Antonyuk, S.V., Melik-Adamyan, W.R., Lamzin, V.S., Harrison, P.M., and Artymiuk, P.J. (1997) The three-dimensional structure of the di-Mn catalase and the environment of the di-Mn sites in different redox states. *J Inorg Biochem* **67:** 196

Barynin, V.V., Whittaker, M.M., Antonyuk, S., Lamzin, V.S., Harrison, P.M., Artymiuk, P.J., and Whittaker, J.W. (2001) Crystal structure of manganese catalase from *Lactobacillus plantarum*. *Structure* **9:** 725-738.

Battistoni, A. 2003. Role of prokaryotic Cu,Zn superoxide dismutase in pathogenesis. *Biochem Soc Trans* **31:** 1326-1329.

Bearson, S.M., Benjamin, W.H. Jr., Swords, W.E., and Foster, J.W. (1996). Acid shock induction of RpoS is mediated by the mouse virulence gene *mviA* of *Salmonella*

*typhimurium. J Bacteriol* **178:** 2572-2579.

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

Becker, G., and Hengge-Aronis, R. (2001) What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). *Mol Microbiol* **39:** 1153-1165.

Becker, G., Klauck, E., and Hengge-Aronis, R. (1999) Regulation of RpoS proteolysis in *Escherichia coli*: the response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. *Proc Natl Acad Sci USA* **96:** 6439-6444.

Becking, J.H. 1974. Nitrogen-fixing bacteria of the genus *Beijerinckia. Soil Sci* **118:** 196-212.

Beers R.F. Jr., and Sizer, I.W. (1952) A spectrophotometric method for measuring the breakdown of hydrogen peroxide by catalase. *J Biol Chem* **195:** 133-140.

Bellei, M., Jakopitsch, C., Battistuzzi, G., Sola, M., and Obinger, C. (2006) Redox thermodynamics of the ferric-ferrous couple of wild-type *Synechocystis* KatG and KatG(Y249F). *Biochemistry* **45:** 4768-4774.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved Prediction of Signal Peptides: SignalP 3.0. *J Mol Biol* **340:** 783-795.

Benov, L., and Al-Ibraheem, J. (2002) Disrupting *Escherichia coli*: a comparison of methods. *J Biochem Mol Biol* **35:**428-431.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra,

G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., and Hopwood, D.A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417:** 141-147.

Bernardo, L., Johansson, L., Solera, D., Skärfstad, E., and Shingler, V. (2006) The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of $\sigma^{54}$-dependent transcription. *Mol Microbiol* **60:** 749-764.

Bertani, I., Ševo, M., Kojic, M., and Venturi, V. (2003) Role of GacA, LasI, RhlI, Ppk, PsrA, Vfr and ClpXP in the regulation of the stationary-phase sigma factor *rpoS*/RpoS in *Pseudomonas*. *Arch Microbiol* **180:** 264-271.

Bertsova, Y.V., Bogachev, A.V., and Skulachev, V.P. (1997) Generation of protonic potential by the *bd*-type quinol oxidase of *Azotobacter vinelandii*. *FEBS Lett* **414:** 369-372.

Bingle, W.H., Doran, J.L., and Page, W.J. (1984) Regular surface layer of *Azotobacter vinelandii*. *J Bacteriol* **159:** 251-259.

Bishop, P. E. (1993) Three genetically distinct nitrogenase systems in *Azotobacter vinelandii*. In Iron Chelation in Plants and Soil Microorganisms. Barton, L.L., and Hemming B.C. (eds) California: Academic Press.

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

Boelrijk, A.E., and Dismukes, G.C. (2000) Mechanism of hydrogen peroxide dismutation

by a dimanganese catalase mimic: dominant role of an intramolecular base on substrate binding affinity and rate acceleration. *Inorg Chem* **39:** 3020-3028.

Bouche, S., Klauck, E., Fischer, D., Lucassen, M., Jung, K., and Hengge-Aronis, R. (1998) Regulation of RssB-dependent proteolysis in *Escherichia coli*: a role for acetyl phosphate in a response regulator-controlled process. *Mol Microbiol* **27:** 787-795.

Bougdour, A, Wickner, S, and Gottesman, S (2006) Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in *Escherichia coli*. *Genes Dev* **20:** 884-897.

Brescia, C.C., Kaw, M.K., and Sledjeski, D.D. (2004) The DNA binding protein H-NS binds to and alters the stability of RNA *in vitro* and *in vivo*. *J Mol Biol* **339:** 505-514.

Brescia, C.C., Mikulecky, P.J., Feig, A.L., and Sledjeski, D.D. (2003) Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA* **9:** 33-43.

Briat, J.F. (1992). Iron assimilation and storage in prokaryotes. *J Gen Microbiol* **138:** 2475-2483.

Brown, L., and Elliott, T. (1996) Efficient translation of the RpoS sigma factor in *Salmonella typhimurium* requires host factor I, an RNA-binding protein encoded by the *hfq* gene. *J Bacteriol* **178:** 3763-3770.

Brown, L., and Elliott, T. (1997) Mutations that increase expression of the *rpoS* gene and decrease its dependence on *hfq* function in *Salmonella typhimurium*. *J Bacteriol* **179:** 656-662.

Brown, S.M., Howell, M.L., Vasil, M.L., Anderson, A.J., and Hassett, D.J. (1995) Cloning and characterization of the *katB* gene of *Pseudomonas aeruginosa* encoding a hydrogen peroxide-inducible catalase: purification of KatB, cellular localization, and demonstration

that it is essential for optimal resistance to hydrogen peroxide. *J Bacteriol* **177:** 6536-6544.

Burr, T., Mitchell, J., Kolb, A., Minchin, S., and Busby, S. (2000) DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Res* **28:** 1864-1870.

Carpena, X., Loprasert, S., Mongkolsuk, S., Switala, J., Loewen, P.C., and Fita, I. (2003) Catalase-peroxidase KatG of *Burkholderia pseudomallei* at 1.7Å resolution. *J Mol Biol* **327:** 475-489.

Castañeda, M., Guzman, J., Moreno, S., and Espín, G. (2000) The GacS sensor kinase regulates alginate and poly-beta-hydroxybutyrate production in *Azotobacter vinelandii*. *J Bacteriol* **182:** 2624-2628.

Castañeda, M., Sánchez, J., Moreno, S., Núñez, C., and Espín, G. (2001) The global regulators GacA and $\sigma^S$ form part of a cascade that controls alginate production in *Azotobacter vinelandii*. *J Bacteriol* **183:** 6787-6793.

Cashel, M., Hsu, L.M., and Hernandez, V.J. (2003) Changes in conserved region 3 of *Escherichia coli* sigma 70 reduce abortive transcription and enhance promoter escape. *J Biol Chem* **278:** 5539-5547.

Cayley, S., Lewis, B.A., and Record, M.T., Jr. (1992) Origins of the osmoprotective properties of betaine and proline in *Escherichia coli* K-12. *J Bacteriol* **174:** 1586-1595.

Celik, A., Cullis, P.M., and Lloyd Raven, E. 2000. Catalytic oxidation of p-cresol by ascorbate peroxidase. *Arch Biochem Biophys* **373:** 175-181.

Chelikani, P., Donald, L.J., Duckworth, H.W., and Loewen, P.C. (2003) Hydroperoxidase II of *Escherichia coli* exhibits enhanced resistance to proteolytic cleavage compared to other catalases. *Biochemistry* **42:** 5729-5735.

Chelikani, P., Fita, I., and Loewen, P.C. (2004) Diversity of structures and properties among catalases. *Cell Mol Life Sci* **61:** 192-208.

Chivian, D., Kim, D.E., Malmstrom, L., Schonbrun, J., Rohl, C.A., and Baker, D. (2005) Prediction of CASP6 structures using automated Robetta protocols. *Proteins* **61:** 157-166.

Clairborne A. and Fridovich I. (1979) Purification of the o-di-anisidine peroxidase from *Escherichia coli* B. *J Biol Chem* **254:** 4245-4252.

Clare, D. A., Duong, M.N., Darr, D., Archibald, F., and Fridovich, I. (1984) Effects of molecular oxygen on detection of superoxide radical with nitroblue tetrazolium and on activity stains for catalase. *Anal Biochem* **140:** 532-537.

Claret, L., and Rouviere-Yaniv, J. (1997) Variation in HU composition during growth of *Escherichia coli*: the heterodimer is required for long term survival. *J Mol Biol* **273:** 93-104.

Claus, D., and Hempel, W. (1970) Specific substrates for isolation and differentiation of *Azotobacter vinelandii*. *Arch Mikrobiol* **73:** 90-96.

Colland, F., Barth, M., Hengge-Aronis, R., and Kolb, A. (2000) sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and lrp transcription factors. *EMBO J* **19:** 3028-3037.

Costanzo, A., and Ades, S.E. (2006) Growth phase-dependent regulation of the extracytoplasmic stress factor, $\sigma^E$, by guanosine 3',5'-bispyrophosphate (ppGpp). *J Bacteriol* **188:** 4627-4634.

Coulson, A.F.W., Erman, J.E. and Yonetani, T. (1971) XVII. Stoichiometry and mechanism of the reaction of compound ES with donors. *J Biol Chem* **246:** 917-924.

Cunning, C., Brown, L., and Elliott, T. (1998) Promoter substitution and deletion analysis of upstream region required for *rpoS* translational regulation. *J Bacteriol* **180:** 4564-4570.

Curratti, L., Brown, C.S., Ludden, P.W., and Rubio, L.M. (2005) Genes required for rapid expression of nitrogenase activity in *Azotobacter vinelandii*. *Proc Natl Acad Sci USA* **102:** 6291-6296.

Dennis, J.J., and Zylstra, G.J. (1998) Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl Environ Microbiol* **64:** 2710-2715.

Desideri, A., and Falconi, M. (2003) Prokaryotic Cu,Zn superoxide dismutases. *Biochem Soc Trans* **31:** 1322-1325.

DeSmet, L., Savvides, S.N., Van Horen, E., Pettigrew, G., and Van Beeumen, J.J. (2006) Structural and mutagenesis studies of the cytochrome *c* peroxidase from *Rhodobacter capsulatus* provide new insights into structure-function relationships of bacterial di-heme peroxidases. *J Biol Chem* **281:** 4371-4379

Deutscher, J., Reizer, J., Fischer, C., Galinier, A., Saier, M.H. Jr., and Steinmetz, M. (1994) Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the *ptsH* gene confers catabolite repression resistance to several catabolic genes of *Bacillus subtilis*. *J Bacteriol* **176:** 3336-3344.

Dias, J., Alves, T., Bonifácio, C., Pereira, A.S., Trincão, J., Bourgeois, D., Moura, I., and Romão, M.J. (2004) Structural basis for the mechanism of $Ca^{2+}$ activation of the di-heme cytochrome *c* peroxidase from *Pseudomonas nautica* 617. *Structure* **14:** 961-973.

Ding, Q., Kusano, S., Villarejo, M., and Ishihama, A. (1995) Promoter selectivity control of *Escherichia coli* RNA polymerase by ionic strength: differential recognition of osmoregulated promoters by E sigma D and E sigma S holoenzymes. *Mol Microbiol* **16:** 649-656.

Dingler, C. and Oelze, J. (1987) Superoxide dismutase and catalase in *Azotobacter vinelandii* grown in continuous culture at different dissolved oxygen concentrations. *Arch Microbiol* **147:** 291-294.

Dombroski, A.J., Walter, W.A., Record, M.T. Jr., Siegele, D.A., and Gross, C.A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70:** 501-512.

Echalier A., Goodhew, C.F., Pettigrew, G.W., and Fülöp, V. (2006) Activation and catalysis of the di-heme cytochrome *c* peroxidase from *Paracoccus pantotrophus*. *Structure*. **14:** 107-117.

Ellfolk, N., Rönnberg, M., Aasa, R., Andreasson, L.-E., and Vanngard,T. (1983) Properties and function of the two hemes in *Pseudomonas* cytochrome *c* peroxidase. *Biochim Biophys Acta* **743:** 23-30.

Elliott, S.J., Bradley, A.L., Arciero, D.M., and Hooper, A.B. (2007) Protonation and inhibition of *Nitrosomonas europaea* cytochrome *c* peroxidase observed with protein film voltammetry. *J Inorg Biochem* **101:** 173-179.

Erickson, D.L., Lines, J.L., Pesci, E.C., Venturi, V., and Storey, D.G. (2004) *Pseudomonas aeruginosa relA* contributes to virulence in *Drosophila melanogaster*. *Infect Immun* **72:** 5638-5645.

Espinosa-Urgel, M., Chamizo, C., and Tormo, A. (1996) A consensus structure for sigma S-dependent promoters. *Mol Microbiol* **21:** 657-659.

Espinosa-Urgel, M., Chamizo, C., and Tormo, A. (1996) Sigma s regulates pLS1 maintenance in stationary-phase *Escherichia coli*. *FEMS Microbiol Lett* **135:** 45-50.

Everse, J. (1998) The structure of heme proteins Compounds I and II: some misconceptions. *Free Radic Biol Med* **24:** 1338-1346.

Fang, F.C., DeGroote, M.A., Foster, J.W., Baumler, A.J., Ochsner, U., Testerman, T., Bearson, S., Giard, J.C., Xu, Y., Campbell, G., and Laessig, T. (1999) Virulent *Salmonella typhimurium* has two periplasmic Cu,Zn-superoxide dismutases. *Proc Natl Acad Sci USA* **96:** 7502-7507.

Fang, M., Majumder, A., Tsai, K.J., and Wu, H.Y. (2000) ppGpp-dependent *leuO* expression in bacteria under stress. *Biochem Biophys Res Commun* **276:** 64-70.

Farewell, A., Kvint, K., and Nystrom, T. (1998) Negative regulation by RpoS: a case of sigma factor competition. *Mol Microbiol* **29:** 1039-1051.

Fischer, D., Teich, A., Neubauer, P., and Hengge-Aronis, R. (1998) The general stress sigma factor $\sigma^S$ of *Escherichia coli* is induced during diauxic shift from glucose to lactose. *J Bacteriol* **180:** 6203-6206.

Foote, N., Thompson, A.C., Barber, D., and Greenwood, C. (1983) *Pseudomonas* cytochrome *C*-551 peroxidase. A purification procedure and study of CO-binding kinetics. *Biochem J* **209:** 701-707.

Foote, N., Turner, R., Brittain, T., and Greenwood, C. (1992) A quantitative model for the mechanism of action of the cytochrome *c* peroxidase of *Pseudomonas aeruginosa*. *Biochem J* **283:** 839-843.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M. (1995) The minimal gene complement of *Mycoplasma genitalium*. *Science* **270:** 397-403.

Fülöp, V., Ridout, C.J., Greenwood, C., and Hajdu, J. (1995) Crystal structure of the di-haem cytochrome *c* peroxidase from *Pseudomonas aeruginosa*. *Structure* **15:** 1225-1233.

Gabbianelli, R., D'Orazio, M., Pacello, F., O'Neill, P., Nicolini, L., Rotilio, G., and Battistoni, A. (2004) Distinctive functional features in prokaryotic and eukaryotic Cu,Zn superoxide dismutases. *Biol Chem* **385:** 749-754.

Gajhede, M., Schuller, D.J., Henriksen, A., Smith, A.T., and Poulos, T.L. (1997). Crystal structure of horseradish peroxidase C at 2.15 A resolution. *Nat Struct Biol* **4:** 1032-1038.

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

George, S.E., Costenbader, C.J., and Melton, T. (1985) Diauxic growth in *Azotobacter vinelandii*. *J Bacteriol* **164:** 866-871.

Ginalski, K., Elofsson, A., Fischer, D., and Rychlewski, L. (2003) 3S-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* **22:**1015-1018.

Goodhew, C.F, Brown, K.R., and Pettigrew, G.W. (1986) Haem staining in gels, a useful tool in the study of bacterial *c*-type cytochromes. *Biochim Biophys Acta* **852:** 288-294.

Goodhew, C.F., Wilson, I.B., Hunter, D.J., and Pettigrew, G.W. (1990) The cellular location and specificity of bacterial cytochrome *c* peroxidases. *Biochem J* **271:**707-712.

Gottesman, S. (2004) The small RNA regulators of *Escherichia coli*: roles and

mechanisms. *Annu Rev Microbiol* **58:**303-328.

Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb Symp Quant Biol* **63:** 141-155.

Guerinot, M.L. (1994). Microbial iron transport. *Annu Rev Microbiol* **48:** 743-772.

Haddad, R.E., Gazeau, S., Pecaut, J., Marchon, J-C., Medforth, C.J., and Shelnutt, J.A. (2003) Origin of the red shifts in the optical absorption bands of nonplanar tetraalkylporphyrins. *J Am Chem Soc* **125:** 1253-1268.

Halbleib, C.M., and Ludden, P.W. (2000) Regulation of biological nitrogen fixation. *J Nutr* **130:** 1081-1084.

Harbury, H.A. 1957. Oxidation-reduction potentials of horseradish peroxidase. *J Biol Chem* **225:** 1009-1024.

Harley, C.B., and R.P. Reynolds. (1987) Analysis of *Escherichia coli* promoter sequences. *Nucleic Acids Res* **15:** 2343-2361.

Hassan, H.M., and Fridovich, I. (1979) Paraquat and *Escherichia coli*: mechanism of production of extracellular superoxide radicals. *J Biol Chem* **254:** 10846-10852.

Hassett, D. (1996) Anaerobic production of alginate by *Pseudomonas aeruginosa*: alginate restricts diffusion of oxygen. *J Bacteriol* **178:** 7322-7325.

Hassett, D.J., Alsabbagh, E., Parvatiyar, K., Howell, M.L., Wilmott, R.W., and Ochsner, U.A. (2000) A protease-resistant catalase, KatA, released upon cell lysis during stationary phase is essential for aerobic survival of a *Pseudomonas aeruginosa oxyR* mutant at low cell densities. *J Bacteriol* **182:** 4557-4563.

Hassett, D.J., Woodruff, W.A., Wozniak, D.J., Vasil, M.L., Cohen, M.S., and Ohman, D.E. (1993) Cloning and characterization of the *Pseudomonas aeruginosa sodA* and *sodB* genes encoding manganese- and iron-cofactored superoxide dismutase: demonstration of increased manganese superoxide dismutase activity in alginate-producing bacteria. *J Bacteriol* **175:** 7658-7665.

Heeb, S., Valverde, C., Gigot-Bonnefoy, C., and Haas, D. (2005) Role of the stress sigma factor RpoS in GacA/RsmA-controlled secondary metabolism and resistance to oxidative stress in *Pseudomonas fluorescens* CHA0. *FEMS Microbiol Lett* **243:** 251-258.

Hengge-Aronis, R. (2002). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev* **66:** 373-395.

Hengge, R., and Bukau, B. (2003) Proteolysis in prokaryotes: protein quality control and regulatory principles. *Mol Microbiol* **49:** 1451-1462.

Henriksen, A., Welinder, K.G., and Gajhede, M. (1998) Structure of barley grain peroxidase refined at 1.9 Å resolution. A plant peroxidase reversibly inactivated at neutral pH. *J Biol Chem* **273:** 2241-2248.

Hershberger, C.D., Ye, R.W., Parsek, M.R., Xie, Z.D., and Chakrabarty, A.M. (1995) The *algT* (*algU*) gene of *Pseudomonas aeruginosa*, a key regulator involved in alginate biosynthesis, encodes an alternative sigma factor (sigma E). *Proc Natl Acad Sci USA* **92:** 7941-7945.

Hersleth, H.P., Ryde, U., Rydberg, P., Gorbitz, C.H., and Andersson, K.K. (2006) Structures of the high-valent metal-ion haem-oxygen intermediates in peroxidases, oxygenases and catalases. *J Inorg Biochem* **100:** 460-476.

Heydorn, A., Ersbøll, B., Kato, J., Hentzer, M., Parsek, M.R., Tolker-Nielsen, T., Givskov, M., and Molin, S. (2002) Statistical analysis of *Pseudomonas aeruginosa* biofilm development: impact of mutations in genes involved in twitching motility, cell-to-cell signaling, and stationary-phase sigma factor expression. *Appl Environ Microbiol* **68:** 2008-2017.

Higgins, D.G, Thompson, J.D., and Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymology* **266:** 383-402.

Hillar, A., and Loewen, P.C. (1995) Comparison of isoniazid oxidation catalyzed by bacterial catalase-peroxidases and horseradish peroxidase. *Arch Biochem Biophys* **323:** 438-446.

Hillar, A., Peters, B., Pauls, R., Loboda, A., Zhang, H., Mauk, A.G., and Loewen, P.C. (2000) Modulation of the activities of catalase-peroxidase HPI of *Escherichia coli* by site-directed mutagenesis. *Biochemistry* **39:** 5868-5875.

Hirsch, M., and Elliott, T. (2005) Stationary-phase regulation of RpoS translation in *Escherichia coli. J Bacteriol* **187:** 7204-7213.

Hobbs, J.D., Larsen, R.W., Meyer, T.E., Hazzard, J.H., Cusanovich, M.A., and Ondrias, M.R. (1990) Resonance Raman characterization of *Chromatium vinosum* cytochrome c'. Effect of pH and comparison of equilibrium and photolyzed carbon monoxide species. *Biochemistry* **29:** 4166-4174.

Howes, B.D., Schiodt CB, Welinder KG, Marzocchi MP, Ma JG, Zhang J, Shelnutt JA, Smulevich G. (1990) The quantum mixed-heme state of barley peroxidase: A paradigm for class III peroxidases. *Biophys J* **77:** 478-492.

Hughes, K.T., and Mathee, K. (1998) The anti-sigma factors. *Annu Rev Microbiol* **52:** 231-286

Hughes, K.T., Gillen, K.L., Semon, M.J., and Karlinsey, J.E. (1993) Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science* **262:**1277-1280.

Hunter, D.J., Brown, K.R., and Pettigrew, G.W. (1989) The role of cytochrome *c*4 in bacterial respiration. Cellular location and selective removal from membranes. *Biochem J* **262:** 233-240.

Imlay, J. A., and Fridovich, I. (1991) Assay of metabolic superoxide production in *Escherichia coli*. *J Biol Chem* **266:** 6957-6965.

Indiani C., Feis, A., Howes, B.D., Marzocchi, M.P., and Smulevich, G. (2000) Effect of low temperature on soybean peroxidase: spectroscopic characterization of the quantum-mechanically admixed spin state. *J Inorg Biochem* **79:** 269-274.

Ishihama, A. (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol* **54:** 499-518.

Ishikawa J., and Hotta, K. (1999) FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol Lett* **174:** 251-253.

Jeng, W.Y., Tsai, Y.H., and Chuang, W.J. (2004) The catalase activity of Nalpha-acetyl-microperoxidase-8. *J Pept Res* **64:**104-109.

Jerome, L.J., and Frost, L.S. (1999) Degradation of FinP antisense RNA from F-like plasmids: the RNA-binding protein, FinO, protects FinP from ribonuclease E. *J Mol Biol* **285:** 1457-1473.

Jishage, M., and Ishihama, A. (1995) Regulation of RNA polymerase sigma subunit

synthesis in *Escherichia coli*: intracellular levels of $\sigma^{70}$ and $\sigma^{38}$. *J Bacteriol* **177:** 6832-6835.

Jishage, M., and Ishihama, A. (1998) A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc Natl Acad Sci USA* **95:** 4953-4958.

Jishage, M., and Ishihama, A. (1999) Transcriptional organization and *in vivo* role of the *Escherichia coli rsd* gene, encoding the regulator of RNA polymerase sigma D. *J Bacteriol* **181:** 3768-3776.

Jishage, M., Iwata, A., Ueda, S., and Ishihama, A. (1996). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol* **178:** 5447-5451.

Jishage, M., Kvint, K., Shingler, V., and Nystrom, T. (2002) Regulation of sigma factor competition by the alarmone ppGpp. *Genes Dev* **16:** 1260-1270.

Jørgensen, F., Bally, M., Chapon-Herve, V., Michel, G., Lazdunski, A., Williams, P., and Stewart, G.S.A.B. (1995) RpoS-dependent stress tolerance in *Pseudomonas aeruginosa*. *Microbiology* **145:** 835-844.

Jurtshuk, P., Liu, J-K., and Moore, E.R.B. (1984) Comparative cytochrome oxidase and superoxide dismutase analyses on strains of *Azotobacter vinelandii* and other related free-living nitrogen-fixing bacteria. *Appl Environ Microbiol* **47:** 1185-1187.

Jurtshuck, P., and Yang, T. (1980) Oxygen reactive hemoprotein components in bacterial respiratory systems. In Diversity of Bacterial Respiratory Systems. Knowles, C.J (ed) Boca Raton: CRC Press.

Kaberdin, V.R., Walsh, A.P., Jakobsen, T., McDowall, K.J., and von Gabain, A. (2000)

Enhanced cleavage of RNA mediated by an interaction between substrates and the arginine-rich domain of *E. coli* ribonuclease E. *J Mol Biol* **301:** 257-264.

Kang, B.R., Cho, B.H., Anderson, A.J., and Kim, Y.C. (2004) The global regulator GacS of a biocontrol bacterium *Pseudomonas chlororaphis* O6 regulates transcription from the *rpoS* gene encoding a stationary-phase sigma factor and affects survival in oxidative stress. *Gene* **325:** 137-143.

Kelly, M.J., Poole, R.K., Yates, M.G., and Kennedy, C. (1990) Cloning and mutagenesis of genes encoding the cytochrome *bd* terminal oxidase complex in *Azotobacter vinelandii*: mutants deficient in the cytochrome *d* complex are unable to fix nitrogen in air. *J Bacteriol* **172:** 6010-6019.

Kelley, W.L., and Georgopoulos, C. (1997) Positive control of the two-component RcsC/B signal transduction network by DnaJ: a member of the DnaJ family of molecular chaperones in *Escherichia coli*. *Mol Microbiol* **25:** 913-931.

Kennedy, C., Rudnick, P., MacDonald, M.L., and Melton, T. (2001) Genus III, *Azotobacter Beijerinck* 1901. In Bergey's manual of systematic bacteriology, 2nd ed., vol. 1. Boone, D.R., and Castenholz, R.W. (eds), Springer-Verlag, New York, N.Y.

Kennedy, C., and Toukdarian, A. (1987) Genetics of azotobacters: applications to nitrogen fixation and related aspects of metabolism. *Annu Rev Microbiol* **41:** 227-258.

Kerschen E.J., Irani, V.R., Hassett, D.J., and Rowe, J.J. (2001) *snr*-1 gene is required for nitrate reduction in *Pseudomonas aeruginosa* PA01. *J Bacteriol* **183:** 2125-2131.

Klauck, E., Bohringer, J., and Hengge-Aronis, R. (1997) The LysR-like regulator LeuO in *Escherichia coli* is involved in the translational regulation of *rpoS* by affecting the expression of the small regulatory DsrA-RNA. *Mol Microbiol* **25:** 559-569.

Klotz, M.G., and Hutcheson, S.W. (1992) Multiple periplasmic catalases in phytopathogenic strains of Pseudomonas syringae. *Appl Environ Microbiol* **58:** 2468-2473.

Klotz, M.G., Klassen, G.R., and Loewen, P.C. (1997) Phylogenetic relationships among prokaryotic and eukaryotic catalases. *Mol Biol Evol* **14:** 951-958.

Klotz, M.G., and Loewen, P.C. (2003) The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota. *Mol Biol Evol* **20:**1098-1112.

Kojic, M., Aguilar, C., and Venturi, V. (2002) TetR family member PsrA directly binds the *Pseudomonas rpoS* and *psrA* promoters. *J Bacteriol* **184:** 2324-2330.

Kojic, M., and Venturi, V. (2001) Regulation of *rpoS* gene expression in *Pseudomonas*: involvement of a TetR family regulator. *J Bacteriol* **183:** 3712-3720.

Koleva, R.I., Ausin, C.A., Kowaleski, J.M., Neems, D.S., Wang, L., Vary, C.P.H., and Schlax, P.J. (2006) Interactions of ribosomal protein S1 with DsrA and *rpoS* mRNA. *Biochem Biophys Res Comm* **348:** 662-668.

Kono, Y., and Fridovich, I. (1983) Isolation and characterization of the pseudocatalase of *Lactobacillus plantarum*. *J Biol Chem* **258:** 6015-6019.

Kornberg, A., Rao, N.N., and Ault-Riche, D. (1999) Inorganic polyphosphate: a molecule of many functions. *Annu Rev Biochem* **68:** 89-125.

Korshunov S.S., and Imlay, J.A. (2002) A potential role for periplasmic superoxide dismutase in blocking the penetration of external superoxide into the cytosol of Gram-negative bacteria. *Mol Microbiol* **43:** 95-106.

Korshunov S.S., and Imlay, J.A. (2006) Detection and quantification of superoxide formed

within the periplasm of *Escherichia coli*. *J Bacteriol* **188:** 6326-6334.

Kroll, J.S., Langford, P.R., Wilks, K.E., and Keil, A.D. (1995) Bacterial [Cu,Zn]-superoxide dismutase: phylogenetically distinct from the eukaryotic enzyme, and not so rare after all! *Microbiology* **141:** 2271-2279.

Kruger, S., Stulke, J., and Hecker, M. (1993) Catabolite repression of beta-glucanase synthesis in *Bacillus subtilis*. *J Gen Microbiol* **139:** 2047-2054.

Kusano, S., and Ishihama, A. (1997) Functional interaction of *Escherichia coli* RNA polymerase with inorganic polyphosphate. *Genes Cells* **2:** 433-441.

Kusano, S., and Ishihama, A. (1997) Stimulatory effect of trehalose on formation and activity of *Escherichia coli* RNA polymerase E sigma38 holoenzyme. *J Bacteriol* **179:** 3649-3654.

Kutsukake, K., and Iino, T. (1994) Role of the FliA-FlgM regulatory system on the transcriptional control of the flagellar regulon and flagellar formation in *Salmonella typhimurium*. *J Bacteriol* **176:** 3598-3605.

Kuzma, M.M., Hunt, S., and Layzell, D.B. (1993) Role of Oxygen in the Limitation and Inhibition of Nitrogenase Activity and Respiration Rate in Individual Soybean Nodules. *Plant Physiol* **101:** 161-169.

Kvint, K., Farewell, A., and Nystrom, T. (2000) RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). *J Biol Chem* **275:** 14795-14798.

Lambert, C., Leonard, N., De Bolle, X., and Depiereux, E. (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18:** 1250-1256.

Lange, R., Fischer, D., and Hengge-Aronis, R. (1995) Identification of transcriptional start sites and the role of ppGpp in the expression of *rpoS*, the structural gene for the sigma S subunit of RNA polymerase in *Escherichia coli*. *J Bacteriol* **177:** 4676-4680.

Lange, R., and Hengge-Aronis, R. (1991) Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Mol Microbiol* **5:** 49-59.

Lange, R., and Hengge-Aronis, R. (1994) The cellular concentration of the sigma S subunit of RNA polymerase in *Escherichia coli* is controlled at the levels of transcription, translation, and protein stability. *Genes Dev* **8:** 1600-1612.

Lange, R., and Hengge-Aronis, R. (1994) The *nlpD* gene is located in an operon with *rpoS* on the *Escherichia coli* chromosome and encodes a novel lipoprotein with a potential function in cell wall formation. *Mol Microbiol* **13:** 733-743.

Lawrence N.L., and Halvorson, H.O. (1954) Studies on the spores of aerobic bacteria. IV. A heat resistant catalase from spores of *Bacillus terminalis*. *J Bacteriol* **68:** 334-337.

Lease, R.A., Cusick, M.E., and Belfort, M. (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci USA* **95:** 12456-12461.

Lenhoff, H. M., and Kaplan, N.O. (1953) A cytochrome peroxidase from *Pseudomonas fluorescens*. *Nature* **172:** 730-731.

Leoni, L., Orsi, N., de Lorenzo, V., and Visca, P. (2000) Functional analysis of PvdS, an iron starvation sigma factor of *Pseudomonas aeruginosa*. *J Bacteriol* **182:** 1481-1491.

Lepock, J.R., Arnold, L.D., Torrie, B.H., Andrews, B, and Kruuv, J. (1985) Structural analysis of various $Cu^{2+}$, $Zn^{2+}$-superoxide dismutases by differential scanning calorimetry and Raman spectroscopy. *Arch Biochem Biophys* **241:** 243-251.

Lin, L.P., and Sadoff, H.L. (1968) Encystment and polymer production by *Azotobacter vinelandii* in the presence of β-hydroxybutyrate. *J Bacteriol* **95:** 2336-2343.

Lin, L.P., and Sadoff, H.L. (1969) Chemical composition of *Azotobacter vinelandii* cysts. *J Bacteriol* **100:** 480-486.

Lipman, J.G. (1903) Experiments on the transformation and fixation of nitrogen by bacteria. *Rep. New Jers. St. Agric. Exp. Stn.* **24$^{th}$**(1902/1903)**:** 217-285.

Lissenden, S., Mohan, S., Overton, T., Regan, T., Crooke, H., Cardinale, J.A., Householder, T.C., Adams, P., O'Connor, C.D., Clark, V.L., Smith, H., and Cole, J.A. (2000) Identification of transcription activators that regulate gonococcal adaptation from aerobic to anaerobic or oxygen-limited growth. *Mol Microbiol* **37:** 839-855.

Livesay, D.R., Jambeck, P., Rojnuckarin, A., and Subramaniam, S. (2003) Conservation of electrostatic properties witin enzyme families and subfamilies. *Biochem* **42:** 3464-3473.

Livny, J., Brencic, A., Lory, S., and Waldor, M.K. (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatics tool sRNAPredict2. *Nucleic Acids Res* **34:** 3484-3493.

Loew, O. (1900) Physiological studies of Conneticut leaf tobacco. *US Dept of Agri Repts* **56:** 5-57.

Loewen, P.C., and Hengge-Aronis, R. (1994) The role of the sigma factor sigma S (KatF) in bacterial global regulation. *Annu Rev Microbiol* **48:** 53-80.

Loewen, P.C., Klotz, M.G., and Hassett, D.J. (2000) Catalase – an "old" enzyme that continues to surprise us. *ASM News* **66:** 76-82.

Loewen, P.C., and Stauffer, G.V. (1990) Nucleotide sequence of *katG* of *Salmonella typhimurium* LT2 and characterization of its product, hydroperoxidase I. *Mol Gen Genet* **224:** 147-151.

Lonetto, M.A., Brown, K.L. Rudd, K.E.,and Buttner, M.J. (1994) Analysis of the *Streptomyces coelicolor* sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *Proc Natl Acad Sci USA* **91:** 7573-7577.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* **174:** 3843-3849.

Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase sigma$^{70}$ subunit. *J Mol Biol* **284:** 1353–1365.

Losick, R., and Pero, J. (1981) Cascades of Sigma factors. *Cell* **25:** 582-584.

Louie, T.M., Ni, S., Xun, L., and Mohn, W.W. (1997) Purification, characterization and gene sequence analysis of a novel cytochrome *c* co-induced with reductive dechlorination activity in *Desulfomonile tiedjei* DCB-1. *Arch Microbiol* **168:** 520-527.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* **193:** 265-275.

Lukat-Rodgers G.S., Wengenack, N.L., Rusnak, F., and Rodgers, K.R. (2000) Spectroscopic comparison of the heme active sites in WT KatG and its S315T mutant. *Biochemistry* **39:** 9984-9993.

Maeda, H., Fujita, N., and Ishihama, A. (2000) Competition among seven *Escherichia coli* σ subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res* **28:**

3497-3503.

Maj, M., Loewen, P., and Nicholls, P. (1998) *E. coli* HPII catalase interaction with high spin ligands: formate and fluoride as active site probes. *Biochimica Biophys Acta* **1384:** 209-222.

Majdalani, N., Chen, S., Murrow, J., St John, K., and Gottesman, S. (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol* **39:** 1382-1394.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., and Gottesman, S. (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* **95:** 12462-12467.

Majdalani, N., Hernandez, D., and Gottesman, S. (2002) Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46:** 813-826.

Malik, S., Zalenskaya, K., and Goldfarb, A. (1987) Competition between sigma factors for core RNA polymerase. *Nucleic Acids Res* **15:** 8521-8530.

Maltempo, M.M. (1976) Visible absorption spectra of quantum mixd-spin ferric heme proteins. *Biochim Biophys Acta* **434:** 513-518.

Maltempo, M.M., and Moss, T.H. (1976) The spin 3/2 state and quantum spin mixtures in haem proteins. *Quart Rev Biophys* **9:** 181-215.

Mandel, M.J., and Silhavy, T.J. (2005) Starvation for different nutrients in *Escherichia coli* results in differential modulation of RpoS levels and stability. *J Bacteriol* **187:** 434-442.

Martínez-Bueno, M.A., Tobes, R., Rey, M., and Ramos, J.L. (2002) Detection of multiple extracytoplasmic function (ECF) sigma factors in the genome of Pseudomonas putida

KT2440 and their counterparts in Pseudomonas aeruginosa PA01. *Environ Microbiol* **4:** 842-855.

Martínez-Salazar, J.M., Moreno, S., Nájera, R., Boucher, J.C., Espín, G., Soberón-Chávez, G., and Deretic, V. (1996) Characterization of the genes coding for the putative sigma factor AlgU and its regulators MucA, MucB, MucC, and MucD in *Azotobacter vinelandii* and evaluation of their roles in alginate biosynthesis. *J Bacteriol* **178:** 1800-1808.

Matsui, T., Ozaki, S., Liong, E., Phillips, G.N. Jr., and Watanabe, Y. (1999) Effects of the localization of distal histidine in the reaction of myoglobin with hydrogen peroxide. *J Biol Chem* **274:** 2838-2844.

Meir, E., and Yagil, E. (1985) Further characterization of the two catalases in *Escherichia coli. Curr Microbiol* **12:** 315-320.

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.

Mika F, and Hengge R. (2005) A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in *E. coli. Genes Dev* **19:** 2770-2781.

Miller, A-F. (2004) Superoxide dismutases: active sites that save, but a protein that kills. *Curr Opin Chem Biol* **8:** 162-168.

Miller, C. D., Kim, Y.C., and Anderson, A.J. (1997) Cloning and mutational analysis of the gene for the stationary-phase inducible catalase (*catC*) from *Pseudomonas putida. J Bacteriol* **179:** 5241-5245.

Miller, C.D., Kim, Y.C., and Anderson, A.J. (2001) Competitiveness in root colonization

by *Pseudomonas putida* requires the *rpoS* gene. *Can J Microbiol* **47:** 41-48.

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V.R., and Blasi, U. (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **11:** 1308-1314.

Moreno, J, Gonzalez-Lopez, J., and Vela, G.R. (1986) Survival of *Azotobacter* spp. in dry soils. *Appl Environ Microbiol* **51:** 123-125.

Moshiri, F., Kim, J.W., Fu, C., and Maier, R.J. (1994) The FeSII protein of *Azotobacter vinelandii* is not essential for aerobic nitrogen fixation, but confers significant protection to oxygen- mediated inactivation of nitrogenase *in vitro* and *in vivo*. *Mol Microbiol* **14:** 101-114.

Muffler, A., Barth, M., Marschall, C., and Hengge-Aronis, R. (1997) Heat shock regulation of sigmaS turnover: a role for DnaK and relationship between stress responses mediated by sigmaS and sigma32 in *Escherichia coli*. *J Bacteriol* **179:** 445-452.

Muffler, A., Fischer, D., Altuvia, S., Storz, G., and Hengge-Aronis, R. (1996) The response regulator RssB controls stability of the sigma(S) subunit of RNA polymerase in *Escherichia coli*. *EMBO J* **15:** 1333-1339.

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.

Mukhopadhyay, S., Audia, J.P., Roy, R.N., and Schellhorn, H.E. (2000) Transcriptional induction of the conserved alternative sigma factor RpoS in *Escherichia coli* is dependent on BarA, a probable two-component regulator. *Mol Microbiol* **37:** 371-381.

Mukhopadhyay, S., and Schellhorn, H.E. (1997) Identification and characterization of

hydrogen peroxide-sensitive mutants of *Escherichia coli*: genes that require OxyR for expression. *J Bacteriol* **179:** 330-338.

Mulvey, M.R., Switala, J., Borys, A., and Loewen, P.C. (1990) Regulation of transcription of *katE* and *katF* in *Escherichia coli*. *J Bacteriol* **172:** 6713-6720.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296:** 1285-1290.

Nagai, H., and Shimamoto, N. (1997) Regions of the *Escherichia coli* primary sigma factor sigma70 that are involved in interaction with RNA polymerase core enzyme. *Genes Cells* **2:** 725-734.

Neilands, J.B. (1981) Microbial iron compounds. *Annu Rev Biochem* **50**: 715–731.

Nettleton, C.J., Bull, C., Baldwin, T.O., and Fee, J.A. (1984) Isolation of the *Escherichia coli* iron superoxide dismutase gene: evidence that intracellular superoxide concentration does not regulate oxygen-dependent synthesis of the manganese superoxide dismutase. *Proc Natl Acad Sci USA* **81:** 4970-4973.

Nicholson, W.L. (2002) Roles of *Bacillus* endospores in the environment. *Cell Mol Life Sci* **59:** 410-416.

Ochsner, U.A., Vasil, M.L., Alsabbagh, E., Parvatiyar, K., and Hassett, D.J. (2000) Role of the *Pseudomonas aeruginosa oxyR-recG* operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of *katB-ankB, ahpB*, and *ahpC-ahpF*. *J Bacteriol* **182:** 4533-4544.

Ohnuma, M., Fujita, N., Ishihama, A., Tanaka, K., and Takahashi, H. (2000) A carboxy-terminal 16-amino-acid region of sigma(38) of *Escherichia coli* is important for

transcription under high-salt conditions and sigma activities *in vivo*. *J Bacteriol* **182:** 4628-4631.

Othman, S., Richaud, P., Verméglio, A., and Desbois, A. (1996) Evidence for a Proximal Histidine Interaction in the Structure of Cytochrome *c'* in Solution: A Resonance Raman Study. *Biochemistry* **35:** 9224-9234.

Pagani, S., Colnaghi, R., Palagi, A., and Negri, A. (1995) Purification and characterization of an iron superoxide dismutase from the nitrogen-fixing *Azotobacter vinelandii*. *FEBS Lett* **357:** 79-82.

Page, W.J. (1983) Formation of cystlike structures by iron-limited *Azotobacter vinelandii* strain UW during prolonged storage. *Can J Microbiol* **29:** 1110-1118.

Page, W.J., and Knosp, O. (1989) Hyperproduction of poly-β-hydroxybutyrate during exponential growth of *Azotobacter vinelandii* UWD. *Appl Environ Microbiol* **55:** 1334–1339.

Page, W.J., and Sadoff, H.L. (1975) Relationship between calcium and uroinic acids in the encystment of *Azotobacter vinelandii*. *J Bacteriol* **122:** 145-151.

Page, W.J., Tindale, A., Chandra, M., and Kwon, E. (2001) Alginate formation in *Azotobacter vinelandii* UWD during stationary phase and the turnover of poly-beta-hydroxybutyrate. *Microbiology* **147:** 483-490.

Page, W.J, and von Tigerstrom, M. (1978) Induction of transformation competence in Azotobacter vinelandii iron-limited cultures. *Can J Microbiol* **24:** 1590-1594.

Park, S., You, and X., Imlay, J.A. (2005) Substantial DNA damage from submicromolar intracellular hydrogen peroxide detected in Hpx- mutants of *Escherichia coli*. *Proc Natl Acad Sci USA* **102:** 9317-9322.

Parker, C.T., Kloser, A.W., Schnaitman, C.A., Stein, M.A., Gottesman, S., and Gibson, B.W. (1992) Role of the *rfaG* and *rfaP* genes in determining the lipopolysaccharide core structure and cell surface properties of *Escherichia coli* K-12. *J Bacteriol* **174:** 2525-2538.

Patten, C.L., Kirchhof, M.G., Schertzberg, M.R., Morton, R.A., and Schellhorn, H.E. (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Mol Genet Genomics 272:* 580-591.

Pau, R.N., Eldridge, M.E., Lowe, D.J., Mitchenall, L.A., and Eady, R.R. (1993) Molybdenum-independent nitrogenases of *Azotobacter vinelandii*: a functional species of alternative nitrogenase-3 isolated from a molybdenum-tolerant strain contains an iron-molybdenum cofactor. *Biochem J* **293:** 101-107.

Pauleta S.R., Lu, Y., Goodhew, C.F., Moura, I., Pettigrew, G.W., and Shelnutt, J.A. (2001) Calcium-dependent conformation of a heme and fingerprint peptide of the diheme cytochrome *c* peroxidase from *Paracoccus pantotrophus*. *Biochemistry* **40:** 6570-6579.

Peralta-Gil, M., Segura, D., Guzmán, J., Servín-González, L., and Espín, G. (2002) Expression of the *Azotobacter vinelandii* poly-beta-hydroxybutyrate biosynthetic *phbBAC* operon is driven by two overlapping promoters and is dependent on the transcriptional activator PhbR. *J Bacteriol* **184:** 5672-5677.

Perederina, A., Svetlov, V., Vassylyeva, M.N., Artsimovitch, I., Yokoyama, S., and Vassylyev, D.G. (2004) Regulation through the secondary channel – structural framework for ppGpp-DksA syngergism during transcription. *Cell* **118:** 297-309.

Pernestig, A.K., Melefors, O., and Georgellis, D. (2001) Identification of UvrY as the cognate response regulator for the BarA sensor kinase in *Escherichia coli*. *J Biol Chem* **276:** 225-231.

Perron, K., Comte, R., and van Delden, C. (2005) DksA represses ribosomal gene transcription in *Pseudomonas aeruginosa* by interacting with RNA polymerase on ribosomal promoters. *Mol Microbiol* **56:** 1087-1102.

Perutz, M.F. (1992) Introductory Lecture: What are enzyme structures telling us? *Faraday Discuss* **93:** 1-11.

Peterson, C.N., Ruiz, N., and Silhavy, T.J. (2004) RpoS proteolysis is regulated by a mechanism that does not require the SprE (RssB) response regulator phosphorylation site. *J Bacteriol* **186:** 7403-7410.

Pettigrew, G.W., and Moore, G.R. (1987) Cytochromes *c*: Biological aspects. Springer-Verlag, Berlin.

Pettigrew, G.W., Goodhew, C.F., Cooper, A., Nutley, M., Jumel, K., and Harding, S.E. (2003) The electron transfer complexes of cytochrome *c* peroxidase from *Paracoccus denitrificans*. *Biochemistry* **42:** 2046-2055.

Phillips, G.N., Arduini, R.M., Springer, B.A., and Sligar, S.G. (1990) Crystal structure of myoglobin from a synthetic gene. *Proteins* **7:** 358-365.

Ponnambalam, S., Webster, C., Bingham, A., and Busby, S. (1986) Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal -35 region sequences. *J Biol Chem* **261:** 16043-16048.

Poulos, T.L., and Kraut, J. (1980) The stereochemistry of peroxidase catalysis. *J Biol Chem* **255:** 8199-8205

Pratt, L.A., and Silhavy, T.J. (1996) The response regulator SprE controls the stability of RpoS. *Proc Natl Acad Sci USA* **93:** 2488-2492.

Pruteanu, M., and Hengge-Aronis, R. (2002) The cellular level of the recognition factor

RssB is rate-limiting for $\sigma^S$ proteolysis: Implications for RssB regulation and signal transduction in $\sigma^S$ turnover in *Escherichia coli*. *Mol Microbiol* **45**: 1701-1714.

Qurollo, B.A., Bishop, P.E., and Hassan, H.M. (2001) Characterization of the iron superoxide dismutase gene of *Azotobacter vinelandii*: *sodB* may be essential for viability. *Can J Microbiol* **47:** 63-71.

Reddy, P.S., Raghavan, A., and Chatterji, D. (1995) Evidence for a ppGpp-binding site on *Escherichia coli* RNA polymerase: proximity relationship with the rifampicin-binding domain. *Mol Microbiol* **15:** 255-265.

Repoila, F., and Gottesman, S. (2001) Signal transduction cascade for regulation of RpoS: temperature regulation of DsrA. *J Bacteriol* **183:** 4012-4023.

Reusch, R.N., and Sadoff, H.L. (1979) 5-n-Alkylresorcinols from encysting *Azotobacter vinelandii*: isolation and characterization. *J Bacteriol* **139:** 448-453.

Robson, R.L., and Postgate, J.R. (1980) Oxygen and hydrogen in biological nitrogen fixation. *Annu Rev Microbiol* **34:** 183-207.

Rockabrand, D., Arthur, T., Korinek, G., Livers, K., and Blum, P. (1995) An essential role for the *Escherichia coli* DnaK protein in starvation- induced thermotolerance, $H_2O_2$ resistance, and reductive division. *J Bacteriol* **177:** 3695-3703.

Ronnberg, M., and Ellfolk, N. (1979) Heme-linked properties of *Pseudomonas* cytochrome *c* peroxidase. Evidence for non-equivalence of the hemes. *Biochim Biophys Acta* **581:** 325-333.

Rørth, M., and Jensen, P.K. (1967) Determination of catalase activity by means of a Clark oxygen electrode. *Biochim Biopys Acta* **139:** 171-173.

Rost B. and Liu, J. (2003) The PredictProtein server. *Nucleic Acids Res* **31:** 3300-3304.

Rowen, D.W., and Deretic, V. (2000) Membrane-to-cytosol redistribution of ECF sigma factor AlgU and conversion to mucoidy in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients. *Mol Microbiol* **36:** 314-327.

Ruiz, N., Peterson, C.N., and Silhavy, T.J. (2001) RpoS-dependent transcriptional control of *sprE*: regulatory feedback loop. *J Bacteriol* **183:** 5974-5981.

Sabra, W., Zeng, A.P., Lunsdorf, H., and Deckwer, W.D. (2000) Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. *Appl Environ Microbiol* **66:** 4037-4044.

Sadoff, H.L. (1975) Encystment and germination in *Azotobacter vinelandii*. *Bacteriol Rev* **39:** 516-539.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sandercock, J.R., and Frost, L.S. (1998) Analysis of the major domains of the F fertility inhibition protein, FinO. *Mol Gen Genet* **259:** 622-629.

Santoni, E., Jakopitsch, C., Obinger, C., and Smulevich, G. (2004) Comparison between catalase-peroxidase and cytochrome *c* peroxidase. The role of the hydrogen-bond networks for protein stability and catalysis. *Biochemistry* **43:** 5792-5802.

Sarniguet, A., Kraus, J., Henkels, M.D., Muehlchen, A.M., and Loper, J.E. (1995) The sigma factor sigma s affects antibiotic production and biological control activity of *Pseudomonas fluorescens* Pf-5. *Proc Natl Acad Sci USA* **92:** 12255-12259.

Scandalios, J.G., Guan, L., and Polidoros, A.N. (1997) Catalase in plants: gene structure,

properties, regulation, and expression. In Oxidative stress and the molecular biology of antioxidant defenses. Scandalios, J.G. (ed.), Cold Spring Harbor Laboratory Press, New York, N.Y.

Schnell, S., and Steinman, H.M. (1995) Function and stationary-phase induction of periplasmic copper-zinc superoxide dismutase and catalase/peroxidase in *Caulobacter crescentus*. *J Bacteriol* **177:** 5924-5929.

Schonbaum, G.R., and Lo, S. (1972) Interaction of peroxidases with aromatic peracids and alkyl peroxidases. Product analysis. *J Biol Chem* **247:** 3353-3360.

Schuster, M., Hawkins, A.C., Harwood, C.S., and Greenberg, E.P. (2004) The *Pseudomonas aeruginosa* RpoS regulon and its relationship to quorum sensing. *Mol Microbiol* **51:** 973-985.

Schweder, T., Lee, K.H., Lomovskaya, O., and Matin, A. (1996) Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease. *J Bacteriol* **178:** 470-476.

Seib, K.L., Wu, H-J., Kidd, S.P., Apicella, M.A., Jennings, M.P., and McEwan, A.G. (2006) Defenses against oxidative stress in *Neisseria gonorrhoeae*: a system tailored for a challenging environment. *Microbiol Mol Biol Rev* **70:** 344-361.

Severinov, K., and Muir, T.W. (1998) Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J Biol Chem* **273:** 16205-16209.

Sevo, M., Buratti, E., and Venturi, V. (2004) Ribosomal Protein S1 specifically binds to the 5' untranslated region of the *Pseudomonas aeruginosa* stationary phase sigma factor *rpoS* mRNA in the logarithmic phase of growth. *J Bacteriol* **186:** 4903-4909.

Shimizu H., Schuller, D.J., Lanzilotta, W.N., Sundaramoorthy, M., Arciero, D.M., Hooper, A.B., and Poulos, T.L. (2001) Crystal structure of *Nitrosomonas europaea* cytochrome *c*

peroxidase and the structural basis for ligand switching in bacterial di-heme peroxidases. *Biochemistry* **40:** 13483-13490.

Siegele, D.A., Hu, J.C., Walter, W.A., and Gross, C.A. (1989) Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *J Mol Biol* **206:** 591-603.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. (2000) Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. *Proc Natl Acad Sci USA* **97:** 8898-8903.

Sledjeski, D., and Gottesman, S. (1995) A small RNA acts as an antisilencer of the H-NS-silenced *rcsA* gene of *Escherichia coli*. *Proc Natl Acad Sci USA* **92:** 2003-2007.

Sledjeski, D.D., and Gottesman, S. (1996) Osmotic shock induction of capsule synthesis in *Escherichia coli* K-12. *J Bacteriol* **178:** 1204-1206.

Sledjeski, D.D., Gupta, A., and Gottesman, S. (1996) The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J* **15:** 3993-4000.

Smulevich G., Jakopitsch, C., Droghetti, E., and Obinger, C. (2006) Probing the structure and bifunctionality of catalase-peroxidase (KatG). *J Inorg Biochem* **100:** 568-585.

Soballe, B., and Poole, R.K. (1998) Requirement for ubiquinone downstream of cytochrome(s) *b* in the oxygen-terminated respiratory chains of *Escherichia coli* K-12 revealed using a null mutant allele of *ubiCA*. *Microbiology* **144:** 361-373.

Socolofsky, M.D., and Wyss, O. (1962) Resistance of the *Azotobacter* cyst. *J Bacteriol* **84:** 119-124.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jager, K-E., and Blasi, U. (2003) Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. *Microb Pathog* **35:** 217-228.

Sonnleitner, E., Moll, I., and Blasi, U. (2002) Functional replacement of the *Escherichia coli hfq* gene by the homologue of *Pseudomonas aeruginosa*. *Microbiology* **148:** 883-891.

Sorensen, M.A., Fricke, J., Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* in vivo. *J Mol Biol* **208:** 561-569.

Sorger-Domenigg, T., Sonnleitner, E., Kaberdin, V.R., and Blasi, U. (2007) Distinct and overlapping binding sites of *Pseudomonas aeruginosa* Hfq and RsmA proteins on the non-coding RNA RsmY. *Biochem Biophys Res Commun* **352:** 769-773.

Steinman, H.M. (1993) Function of periplasmic copper-zinc superoxide dismutase in *Caulobacter crescentus*. *J Bacteriol* **175:** 1198-1202.

Stevenson, L.H., and Socolofsky, M.D. (1972) Encystment of *Azotobacter vinelandii* in liquid culture. *Antonie van Leeuwenhoek J Microbiol Serol* **38:** 605-616.

Stohl, E.A., Criss, A.K., and Seifert, H.S. (2005) The transcriptome response of *Neisseria gonorrhoeae* to hydrogen peroxide reveals genes with previously uncharacterized roles in oxidative damage protection. *Mol Microbiol* **58:** 520-532.

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

Strekas, T.C., and Spiro, T.G. (1974) Resonance-Raman evidence for heme structures in cytochrome *c'* from *Rhodospeudomonas palustris*. *Bioch Biophys Acta* **351:** 237-245.

Strickland, J.D.H., and Parsons, T.R. (1972) A practical handbook of sea water analysis.

*Fish Res Bd Can Bull* **167:**71-89.

Stüdemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D., and Hengge, R. (2003) Sequential recognition of two distinct sites in $\sigma^S$ by the proteolytic targeting factor RssB and ClpX. *EMBO J* **22:** 4111-4120.

Suh, S.J., Runyen-Janecky, L.J., Maleniak, T.C., Hager, P., MacGregor, C.H., Zielinski-Mozny, N.A., Phibbs, P.V. Jr., and West, S.E. (2002) Effect of *vfr* mutation on global gene expression and catabolite repression control of *Pseudomonas aeruginosa*. *Microbiology* **148:** 1561-1569.

Suh, S-J., Silo-Suh, L., Woods, D.E., Hassett, D.J., West, S.E.H., and Ohman, D.E. (1999) Effect of *rpoS* mutation on the stress response and expression of virulence factors in *Pseudomonas aeruginosa*. *J Bacteriol* **181:** 3890-3897.

Switala, J., and Loewen, P.C. (2002) Diversity of properties among catalases. *Arch Biochem Biophys* **401:** 145-154.

Switala, J., O'Neil, J.O., and Loewen, P.C. (1999) Catalase HPII from *Escherichia coli* exhibits enhanced resistance to denaturation. *Biochemistry* **38:** 3895-3901.

Tahirov, T.H., Misaki, S., Meyer, T.E., Cusanovich, M.A., Higuchi, Y., and Yasuoka, N. (1996) High-resolution crystal structures of two polymorphs of cytochrome *c'* from the purple phototrophic bacterium *Rhodobacter capsulatus*. *J Mol Biol* **259:** 467-479.

Takayanagi, Y., Tanaka, K., and Takahashi, H. (1994) Structure of the 5' upstream region and the regulation of the *rpoS* gene of *Escherichia coli*. *Mol Gen Genet* **243:** 525-531.

Timóteo, C.G., Tavares, P., Goodhew, C.F., Duarte, L.C., Jumel, K., Gírio, F.M., Harding, S., Pettigrew, G.W., and Moura, I. (2003) $Ca^{2+}$ and the bacterial peroxidases: the cytochrome *c* peroxidase from *Pseudomonas stutzeri*. *J Biol Inorg Chem* **8:** 29-37.

Thompson, J.P., and Skerman,V.B.D. (1979) *Azotobacteraceae*: The Taxonomy of the Aerobic Nitrogen-Fixing Bacteria (New York: Academic Press).

Thöny-Meyer, L. (1997) Biogenesis of respiratory cytochromes in bacteria. *Microbiol Mol Biol Rev* **61:** 337-376.

Thöny-Meyer, L. (2002) Cytochrome *c* maturation: a complex pathway for a simple task? *Biochem Soc Trans* **30:** 633-638.

Tomizawa, J. (1986) Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell* **47:** 89-97.

Tsan, P., Caffrey, M., Daku, M.L., Cusanovich, M., Marion, D., and Gans, P. (2001). Magnetic susceptibility tensor and heme contact shifts determinations in the *Rhodobacter capsulatus* ferricytochrome *c'*: NMR and magnetic susceptibility studies. *J Am Chem Soc* **123:** 2231-2242.

Turner, S.M., Reid, E.G., Smith, H., and Cole, J.A. (2003) A novel cytochrome *c* peroxidase from *Neisseria gonorrhoeae*, a lipoprotein from a Gram-negative bacterium. *Biochem J* **373:** 865-873.

Van Delden C., Comte, R., and Bally, A.M. (2001) Stringent response activates quorum sensing and modulates cell density-dependent gene expression in *Pseudomonas aeruginosa*. *J Bacteriol* **183:** 5376-5384.

Van Spanning, R.J.M., De Boer, A.P.N., Reijnders, W.N.M., Westerhoff, H.V., Stouthamer, A.H., and Van Der Oost, J. (1997) FnrP and NNR of *Paracoccus denitrificans* are both members of the FNR family of transcriptional activators but have distinct roles in respiratory adaptation in response to oxygen limitation. *Mol Microbiol* **23:** 893-907.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417:** 712-719.

Vela, G.R. (1974) Survival of *Azotobacter* in dry soil. *Appl Microbiol* **28:** 77-79.

Venturi, V. (2003) Control of rpoS transcription in *Escherichia coli* and *Pseudomonas*: why so different? *Mol Microbiol* **49:** 1-9

Viducic, D., Ono, T., Murakami, K., Susilowati, H., Kayama, S., Hirota, K., and Miyake, Y. (2006) Functional analysis of *spoT*, *relA* and *dksA* genes on Quinolone tolerance in *Pseudomonas aeruginosa* under nongrowing condition. *Microbiol Immunol* **50:** 349-357.

Visca, P., L. Leoni, L., Wilson, M.J., and Lamont, I.L. (2002) Iron transport and regulation, cell signaling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Mol Microbiol* **45:** 1177-1190.

Waddell, W.J. (1956) A simple ultraviolet spectrophotometric method for the determination of protein. *J Lab Clin Med* **48:** 311-314.

Wai, S.N., Nakayama, K., Umene, K., Moriya, T., and Amako, K. (1996) Construction of a ferritin-deficient mutant of *Campylobacter jejuni*: contribution of ferritin to iron storage and protection against oxidative stress. *Mol Microbiol* **20:** 1127-1134.

Waldo, G.S., and Penner-Hahn, J.E. (1995) Mechanism of Manganese catalase peroxide disproportionation: determination of manganese oxidation states during turnover. *Biochemistry* **34:** 1507-1512.

Wang Y., Graichen, M.E., Liu, A., Pearson, A.R., Wilmot, C.M., and Davidson, V.L. (2003) MauG, a novel diheme protein required for tryptophan tryptophylquinone biogenesis. *Biochemistry* **42:** 7318-7325.

Wang, Lu, Yao, Xie and Huang (2003) The distal heme coordination environments and heme-binding stabilities of His39Ser and His39Cys mutants of cytochrome $b_5$. Protein *Engineering* **16:** 1047-1054.

Wang, J., Mauro, J.M., Edwards, S.L., Oatley, S.L., Fishel, L.A., Ashford, V.A., Xuong, N.-H., and Kraut, J. (1990) X-ray structures of recombinant yeast cytochrome *c* peroxidase and three heme-cleft mutants prepared by site-directed mutagenesis. *Biochemistry* **29:** 7160–7173.

Wariishi, H., Nonaka, D., Johjima, T., Nakamura, N., Naruta, Y., Kuboi, S., and Fukuyama, K.. (2000) Direct binding of hydroxylamine to the heme iron of *Arthromyces ramosus* peroxidase. Substrate analogue that inhibits compound I formation in a competitive manner. *J Biol Chem* **275:** 32919–32924.

Wassarman, K.M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15:** 1637-1651.

Welinder, K.G. (1991) Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochim Biophys Acta* **1080:** 215-220.

Welsh, D.T., Reed, R.H., and Herbert, R.A. (1991) The role of trehalose in the osmoadaptation of *Escherichia coli* NCIB 9484: interaction of trehalose, K+ and glutamate during osmoadaptation in continuous culture. *J Gen Microbiol* **137:** 745-750.

Whistler, C.A., Corbell, N.A., Sarniguet, A., Ream, W., and Loper, J.E. (1998) The two-component regulators GacS and GacA influence accumulation of the stationary-phase sigma factor sigmaS and the stress response in *Pseudomonas fluorescens* Pf-5. *J Bacteriol* **180:** 6635-6641.

Wilderman, P.J., Sowa, N.A., FitzGerald, D.J., FitzGerald, P.C., Gottesman, S., Ochsner U.A., and Vasil M.L. (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc Natl Acad Sci USA* **101:** 9792-9797.

Wilkins, M.R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J.C., Hochstrasser, D.F., and Appel, R.D. (1997) Detailed peptide characterization using PEPTIDEMASS--a World-Wide-Web-accessible tool. *Electrophoresis* **18:** 403-408.

Williams, A.M., and Wilson, P.W. (1954) Adaptation of *Azotobacter* cells with tricaboxylic acid substrates. *J Bacteriol* **67:** 353-360.

Winogradsky, S. (1938) Études sur la microbiologie du sol et des eaux. Sur la morphologie et l'oecologie des *Azotobacter*. *Annls Inst Pasteur, Paris* **60:** 351-400.

Wolf, P. (1983) A critical reappraisal of Waddell's technique for ultraviolet spectrophotometric protein estimation. *Anal Biochem* **129:** 145-155.

Worhunsky, D.J., Godek, K., Litsch, S., and Schlax, P.J. (2003) Interactions of the non-coding RNA DsrA and RpoS mRNA with the 30S ribosomal subunit. *J Biol Chem* **24:** 15815-15824.

Wray, L.V. Jr., Pettengill, F.K., and Fisher, S.H. (1994) Catabolite repression of the *Bacillus subtilis hut* operon requires a *cis*-acting site located downstream of the transcription initiation site. *J Bacteriol* **176:** 1894-1902.

Wu, G., Cruz-Ramos, H., Hill, S., Green, J., Sawers, G., and Poole, R.K. (2000) Regulation of cytochrome *bd* expression in the obligate aerobe *Azotobacter vinelandii* by CydR (Fnr). Sensitivity to oxygen, reactive oxygen species, and nitric oxide. *J Biol Chem* **275:** 4679-4686.

Wu, C.C., MacCoss, M.J., Howell, K.E., and Yates III, J.R. (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotech* **21:** 532-537.

Xiao, H., Kalman, M., Ikehara, K., Zemel, S., Glaser, G., and Cashel, M. (1991) Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. *J Biol Chem* **266:** 5980-5990.

Xu, K. D., Franklin, M.J., Park, C-H., McFeters, G.A., and Stewart, P.S. (2001) Gene expression and protein levels of the stationary phase sigma factor, RpoS, in continuously-fed *Pseudomonas aeruginosa* biofilms. *FEMS Microbiol Lett* **15:** 67-71.

Yamashino T., Ueguchi C., and Mizuno T. (1995) Quantitative control of the stationary phase-specific sigma factor, $\sigma^S$, in *Esherichia coli*: involvement of the nucleoid protein H-NS. *EMBO J* **14:** 594-602

Yonetani, T. (1976) The Enzymes. vol. 13. Boyer, P.D. (ed) Academic Press, New York.

Yonetani, T., Schleyer, H., and Chance, B. (1966) Crystaline cytochrome *c* peroxidase and its enzyme-substrate complex. *Science* **152:** 678.

Zahn, J.A., Arciero, D.M., Hooper, A.B., Coats, J.R., and DiSpirito, A.A. (1997) Cytochrome *c* peroxidase form *Methylococcus capsulatus* Bath. *Arch Microbiol* **168:** 362-372.

Zamocky, M. (2004) Pylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. *Eur J Biochem* **271:** 3297-3309.

Zgurskaya, H.I., Keyhan, M., and Matin, A. (1997) The $\sigma^S$ level in starving *Escherichia coli* cells increases solely as a result of its increased stability, despite decreased synthesis. *Mol Microbiol* **24:** 643-651.

Zhang, A., Altuvia, S., Tiwari, S., 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.

Zhang, A., Wassarman, K.M., Ortega, J., Steven, A.C., and Storz, G. (2002) The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9:** 11-22.

Zhang, J.P., and Normark, S. (1996) Induction of gene expression in *Escherichia coli* after pilus-mediated adherence. *Science* **273:** 1234-1236.

Zhou, Y., and Gottesman, S. (1998) Regulation of proteolysis of the stationary-phase sigma factor RpoS. *J Bacteriol* **180:** 1154-1158.

Zhou Y-N, and Gottesman, S. (2006) Modes of regulation of RpoS by H-NS. *J Bacteriol* **188:** 7022-7025.

Zhou, Y.N., and Jin, D.J. (1998) The *rpoB* mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in *Escherichia coli*. *Proc Natl Acad Sci USA* **95:** 2908-2913.

# APPENDIX

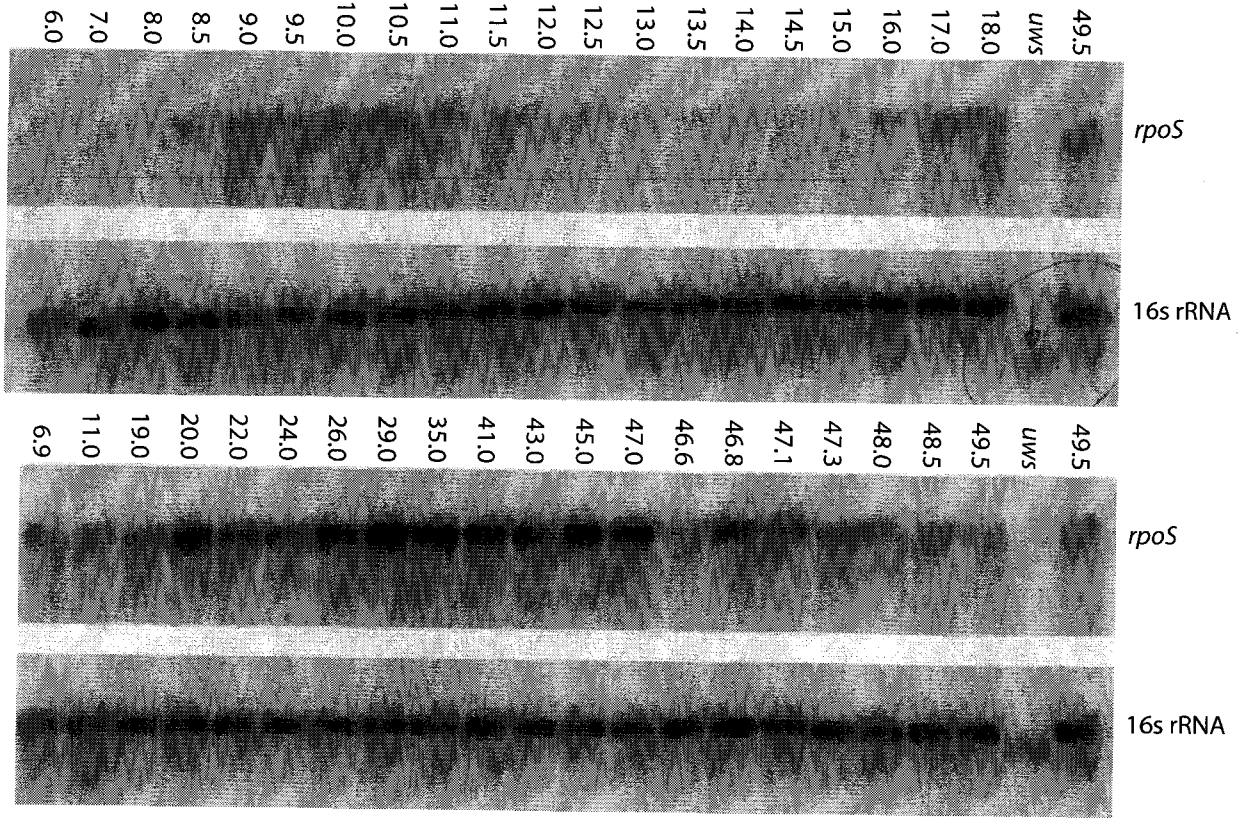



**Figure A.** Northern analysis of *rpoS* mRNA and the internal standard, 16s rRNA. RNA was harvested from flash-frozen cells by the hot phenol method, precipitated and quantified spectrometrically. Equal quantities of total RNA were loaded to, and separated on denaturing formaldehyde agarose gels, transfered to Hibond N+ membranes, probed with *rpoS* mRNA specific probe and analysed using a phosphor-imager. The membranes were stripped and re-probed with 16s rRNA specific probe as an internal standard. The culture time-points (h) are indicated above each lane. Electrophoretic migration of the UWS RNA was higher (indicated by the arrow) due to higher salt concentrations following precipitation.

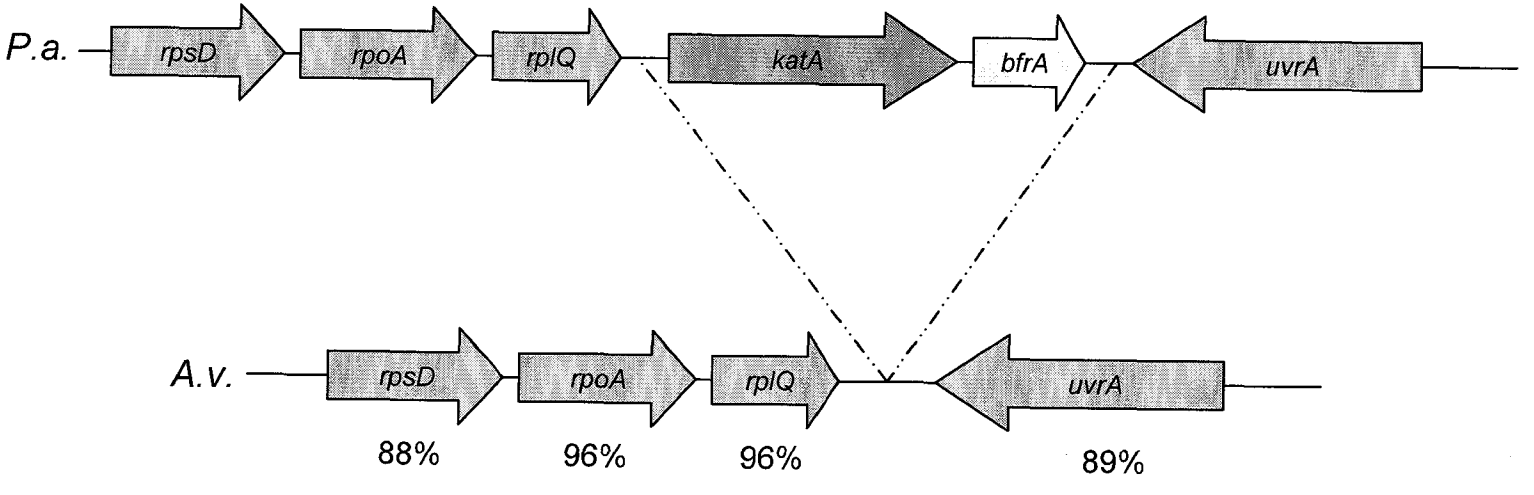



**Figure B. Genetic map of the region around *katA* of *P. aeruginosa* PA01 and the equivalent region in *A. vinelandii*.** The degree of identity between the protein sequences are given below the ORFs in the lower panel. Both *katA* and *bfrA* are absent in *A. vinelandii,* possibly due to a translocation event since a *bfrA* homologue (73% identical) is found elsewhere in the genome.

1 MAAKE VKFGD SARKK MLVGV NVLAD AVKAT LGPKG RNVVL DKSFG APTIT
51 KDGVS VAKEI ELKDK FENMG AQLVK DVASK ANDEA GDGTT TATVL AQAIV
101 NEGLK AVAAG MNPMD LKRGI DKATI AIVAE LKSLA KPCSD SKAIA QVGTI
151 SANSD ESIGN IIAEA MNKVG KEGVI TVEEG SGLEN ELSVV EGMQF DRGYL
201 SPYFI NKPDT MVAEL DNPLL LLVDK KISNI RELLP VLEAV AKSGR PLLIV
251 AEDVE GEALA TLVVN NMRGI VKVAA VKAPG FNDRR KAMLQ VIAIL TGATV
301 ISEKV GLSLE SATLE HLGKP KALVL NKENT TIMHG AGAQA DIEAA VAQIR
351 KQIEE TSSDY DREKL QERLA KLAGG VAVIK VGAAT EVEMK EKKAR VEAAL
401 HATRA AVEEG VVPGG GVALV RALQA IEGLK GDNED QNVGI ALLRR AVEAP
451 LRQIV ANAGD EPSVV VDKVK QGSGN FGFNA ASGVY GDMIE MGILD PAKVT
501 RSALQ AASSI GGLMI TTEAM VADIV EDKAA PAMPD MGGMG GMGGM M

**Figure C. Mass Spectroscopic Identification of GroEL.** The protein sequence of GroEL (accession AAL25964.1) of *A. vinelandii* and the peptide fragments identified by LC/MS/MS following trypsin digestion, indicated in bold with underlining. A total coverage of 10.99% of the open reading frame was obtained.

```
1       MKVYY DKDCD LSIIQ SKKVA IIGYG SQGHA HACNL KDSGV DVYVG LRAGS
51      ASVAK AEAHG LTVKS VKDAV AAADV VMILT PDEFQ GRLYK DEIEP NLKKG
101     ATLAF AHGFS IHYNQ VVPRA DLDVI MIAPK APGHT VRSEF VRGGG IPDLI
151     AVYQD ASGNA KNLAL SYACG VGGGR TGIIE TTFKD ETETD LFGEQ AVLCG
201     GCVEL VKAGF ETLVE AGYAP EMAYF ECLHE LKLIV DLMFE GGIAN MNYSI
251     SNNAE YGEYV TGPEV INEQS RQAMR NALKR IQDGE YAKMF ITEGA ANYPS
301     MTAYR RNNAA HQIEV VGEKL RTMMP WIAAN KIVDK TKN
```



**Figure D. Mass Spectroscopic Identification of Acetohydroxy Acid Isomoreductase.** The protein sequence of acetohydroxy acid isomeroreductase (accession EAM03644.1) of *A. vinelandii* and the peptide fragments identified by LC/MS/MS following trypsin digestion, indicated in bold with underlining. A total coverage of 9.76% of the open reading frame was obtained.

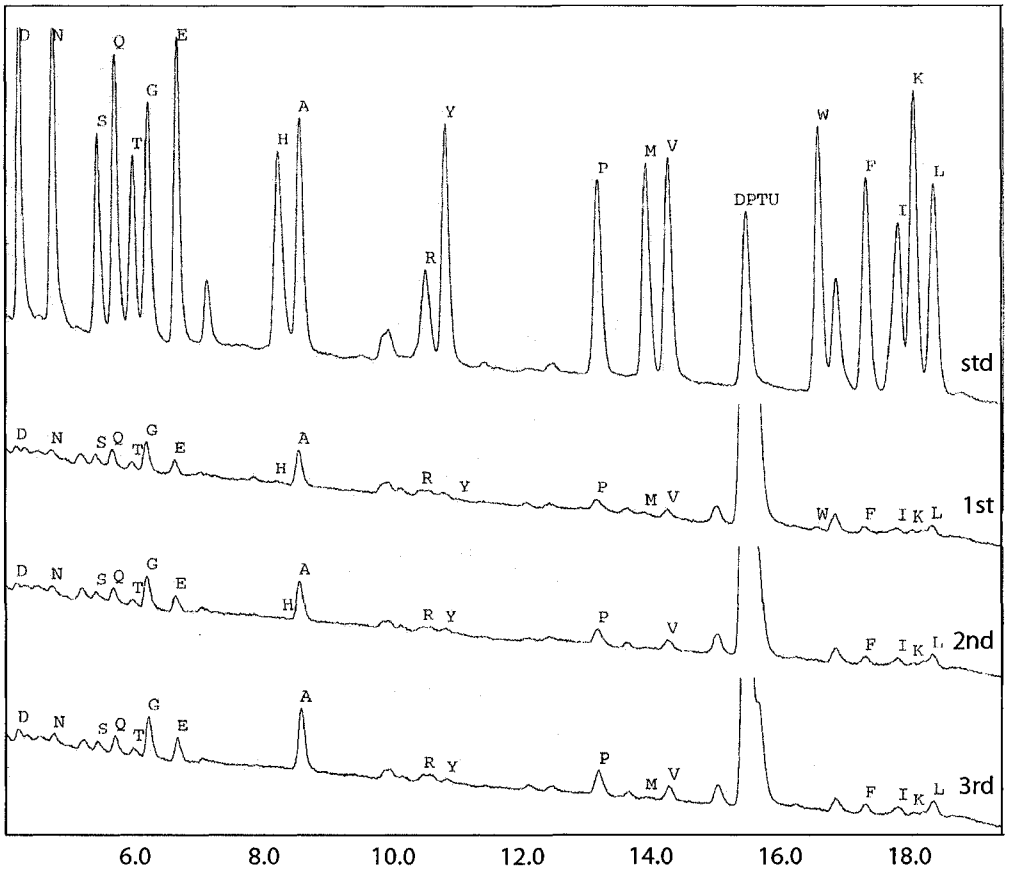



**Figure E. N-terminal sequencing chromatograms of purified *Av*CCC.** The standard curve is the top trace, and the identities of 19 individual amino acids are indicated using the single letter code. The first, second, and third N-terminal amino acids of *Av*CCC are labeled below the standard curve.

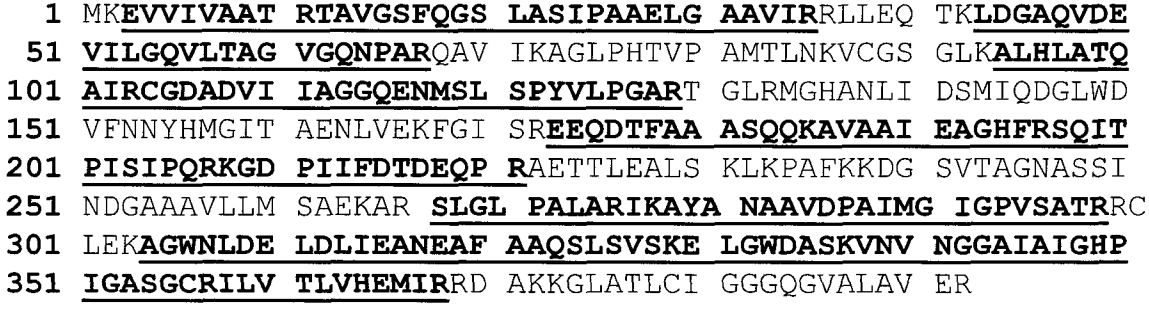



**Figure F. Mass Spectroscopic Identification of Thiolase.** The protein sequence of thiolase (accession ZP_00415203) of *A. vinelandii* and the peptide fragments identified by LC/MS/MS following trypsin digestion, indicated in bold with underlining. A total coverage of 61% of the open reading frame was obtained.

```
  1 MADKKAQLII EGNAPVELPM LTGTVGPDVI DVRGLTSTGH FTFDPGFMST
 51 ASCESKITYI DGDKGILLHR GYPIEQLAEK SDYLETCYLL LNGELPNADQ
101 KAKFVNNIKN HTMVHEQLKN FLNGFRRDAH PMAVMCGVVG ALSAFYHDSL
151 DINDPQHREI SAIRLVAKMP TLAAMVYKYS LGQPLMYPRN DLDYAENFLH
201 MMFNTPCEVK PISPVLAKAM DRIFVLHADH EQNASTSTVR LAGSSGANPF
251 ACIAAGIAAL WGPAHGGANE AVLTMLDEIG DVSNIEKFVA KAKDKNDPFK
301 LMGFGHRVYK NFDPRAKVMK QTCDEVLSEL GINDPQLELA MKLEEIARND
351 PYFKERNLYP NVDFYSGIIL KAIGIPTSMF TVIFALARTV GWISHWKEML
401 SGPYKIGRPR QLYTGYEKRD YISVEKR
```

**Fig. G. Mass Spectroscopic Identification of Citrate Synthase I.** The protein sequence of citrate synthase I (genbank 67156800) of *A. vinelandii* and the peptide fragments identified by LC/MS/MS following trypsin digestion, indicated in bold with underlining. A total coverage of 4% of the open reading frame was obtained.

| Designation | Closest Homologue[a] | Identity to FecI |
|---|---|---|
| RpoD | 85% to RpoD of *P. fluorescens* | None |
| RpoS | 95% to RpoS of *P. aeruginosa* | None |
| RpoN | 81% to RpoN of *P. putida* | None |
| RpoN*[b] | N/D | - |
| RpoH | 89% to RpoH of *P. aeruginosa* | None |
| RpoF | 67% to RpoF of *C. salexigens* | None |
| SigX | 88% to SigX of *P. aeruginosa* | 26% |
| PvdS1 | 55% to FecI of *P. fluorescens* | 27% |
| PvdS2 | 72% to PvdS of *P. aeruginosa* | 30% |
| AlgU | 94% to AlgT of *P. putida* | 21% |
| ECF1 | 52% to probable ECF of *P. aeruginosa* | 26% |
| ECF2 | 57% to probable ECF of *P. aeruginosa* | 45% |
| ECF3 | 62% to FecI of *P. fluorescens* | 51% |
| ECF4 | 40% to probable ECF of *P. aeruginosa* | 28% |
| ECF5 | 45% to ECF of *B. parapertussis* | 31% |
| ECF6 | 63% to ECF *P. putida* | 44% |
| ECF7 | N/D | - |
| ECF8 | 70% to probable ECF of *P. aeruginosa* | 29% |
| ECF9 | 55% to probable ECF of *P. aeruginosa* | 52% |
| ECF10 | 55% to FecI of *P. fluorescens* | 50% |
| ECF11 | 65% to ECF of *N. europaea* | 50% |
| ECF12 | 46% to ECF of *R. solanacearum* | 22% |

[a] Closest homologue defined as the ORF giving the highest % amino acid identity by Blast search.

[b] The gene encoding RpoN* is interrupted by unrelated DNA sequence, putting the ORF out of frame. The product is unlikely to be functional.

**Table A. The Sigma Factors of *A. vinelandii.*** Putative sigma factors identified by the automated programs run during the analysis of the *A. vinelandii* raw genome sequence (http://azotobacter.org). Designations are based on the labeling methods used in the genome project. Where no gene product name was given the term "ECF" and an arbitrary number are used. The closest homologue is defined as the open reading frame giving the highest raw score for protein identity following an unrestricted BLAST search. The degree of similarity to the ECF FecI of *E. coli* is also given.

**Table B. Leader Peptide Cleavage-Site Predictions of Proteins used in Fig. 5.12, Calculated by the Program SignalP 3.0.** Putative leader peptides are indicated in bold and primary cleavage sites by a space. Probabilities are given where appropriate, and account for all potential cleavage sites. Instances where the leader peptide sequence was already removed in the accession file are indicated by "N/A".

| Protein | Organism | Leader Peptide Sequence and Cleavage Site [a] | Probability [b] | Accession # |
|---|---|---|---|---|
| *Av*CCC | *Azotobacter vinelandii* UW | **MSKATIKAGLAIPIALALSLADSVRA** APEEKPTSY [c] | 0.993 | ZP_00418713 [e] |
| *Dt*cytoC | *Desulfomonile tiedjei* DCB-1 | **MKFVRPKLIFVMACCMIVALAGLIYA** ESKKVPSSY [c] | 0.999 | AAB66558 |
| *Pa*Snr-1 | *Pseudomonas aeruginosa* PA01 | **MKTPAWTRHALWVMPLALGLQSAVVA** GDEQPSKTS [c,d] | 0.993 | NP_251722 |
| *Ec*CCP [e] | *Escherichia coli* K-12 | **MKMVSRITAIGLAGVAICYLGLSGYVWY** HDNKRSK | 0.228 | AAC76543 |
| *Ne*CCP | *Nitrosomonas europaea* | **MQIIFLLLLLMASSYGYA** QNYTQSFEPLPVSIIDE [c] | 1.000 | ZP_00670794 |
| *Pa*CCP | *Pseudomonas aeruginosa* PA01 | **MQSSQLLPLGSLLLSFATPLAQA** DALHDQASALFK [c] | 1.000 | PA4587 |
| *Pd*CCP | *Paracoccus denitrificans* PD1222 | **MVQ...KLVPFLTATSLALSALPVMA** QTEAIDDSA [c] | 0.995 | ZP_00628927 [e] |
| *Pn*CCP | *Pseudomonas nautica* 617 | DNLMERANSMFEPIPKYPPVIDGNELTQAKVELGK [c] | N/A | 1NMLA |
| *Pp*CCP | *Paracoccus pantotrophus* | ETEAIDNGALREEAKGVFEAIPEKMTAIKQTEDNP [c] | N/A | 2C1V_A |
| *Me*MauG | *Methylobacterium extorquens* AM1 | **MRAILPIPVLIAWAMVVCGGAYA** VTTCSGAATATA [c] | 0.994 | Q49128 |
| *Mf*MauG | *Methylobacillus flagellatus* | **MRFCSVFLVFGLAAAMGVSHA** ADLPPRESYQRPAD | 1.000 | AAF03761 |
| *Mm*MauG | *Methylophilus methylotrophus* | **MLFRHLVLIISTLMVANTAWS** ANLPPREKFKRPDS [c] | 1.000 | Q50233 |
| *Pd*MauG | *Paracoccus denitrificans* PD1222 | **MLRLACLAPLAILIPAAGTA** EQARPADDALAALGA [c] | 0.998 | Q51658 |
| *Av*KatG | *Azotobacter vinelandii* UW | MSSEKSETKCPFNHAVGGGTSNRDWWPNQLRLDLL | None | EAM04829 |
| *Ec*KatG | *Escherichia coli* K-12 | MSTSDDIHNTTATGKCPFHQGGHDQSAGAGTTTRD [d] | None | AAC76924 |
| *Pd*KatG [f] | *Paracoccus denitrificans* PD1222 | MDGNDIRSTGKCPVMHGGNTAMGSSVTAWWPNALN | None | ZP_00632112 |
| *Pk*KatG [f] | *Pseudomonas putida* KT2440 | MSNESKCPFHQTAGGGTTNRDWWPDQLNLRILHQH | None | NP_745804 |
| *Ps*KatG | *Pseudomonas syringae* DC3000 | MSTESKCPFNHAAGGGTTNRDWWPKQLNLKILHQH | None | NP_794283 |

[a] cleavage sites predicted by SignalP3.0

[b] probabilities given are total values, including probabilities of secondary and tertiary cleavage sites

[c] cleavage demonstrated experimentally

[d] previous prediction methods suggested no leader peptide

[e] truncation from genome derived sequence required due to incorrect ORF boundaries

[f] ORF predicted to encode protein of class, no experimental proof.