THE REGULATION OF PRODUCTION OF THE RHODOBACTER CAPSULATUS GENE TRANSFER AGENT, RcGTA

RYAN MERCER

The Regulation of Production of the *Rhodobacter capsulatus* Gene Transfer Agent, RcGTA

by

© Ryan Mercer

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biology

Memorial University of Newfoundland

September 2012

St. John's, Newfoundland

ABSTRACT

Horizontal gene exchange is the transfer of genetic material between cells and it plays an important role in the evolution of bacterial genomes. One mechanism contributing to high levels of transfer is virus-mediated transduction by bacteriophage (phage). Most bacterial species encode prophages within their genome and play host to actively transducing phage, which can lead to acquisition of new genes. An unusual form of genetic exchange is carried out by a phage-like particle called a gene transfer agent (GTA). The purple, non-sulfur, α-proteobacterium *Rhodobacter capsulatus* produces RcGTA that transfers ~4 kb of random genomic DNA in a process similar to generalized transduction. Expression of the ~15 kb gene cluster encoding the RcGTA particle is under the control of a cellular response regulator protein, CtrA, a homologue of the well-studied cell cycle regulator in *Caulobacter crescentus*. Despite the requirement for CtrA, the complete regulatory system controlling RcGTA has previously been poorly understood. The research in this thesis has shown that loss of CtrA has pleiotropic effects, as the protein is required for proper expression of ~6% of genes in the *R. capsulatus* genome. This includes all flagellar motility genes and genes encoding a number of other putative regulatory proteins. Mutations of the *cckA* and *chpT* genes in *R. capsulatus* revealed that the proteins encoded by these genes, along with CtrA~P, are required for motility and release of RcGTA from cells. It was also determined that both phosphorylated and unphosphorylated CtrA are capable of stimulating RcGTA gene expression. Proper expression of a set of genes encoding putative σ factor regulatory proteins was also found to require CtrA and mutations in these genes affected RcGTA expression, motility,

colony morphology and stationary phase cell viability. The results indicate that CtrA regulates RcGTA, at least in part, by controlling the activity of an important σ factor in *R. capsulatus*. The results of this thesis work have provided an excellent framework for further deciphering the regulation of this intriguing genetic exchange element.

ACKNOWLEDGEMENTS

Many people have helped and supported my progress through my thesis research over the years. I would like to thank my graduate committee members Dr. Ross McGowan and Dr. Matt Rise for their suggestions, guidance and use of materials in designing and analyzing my research. As well, a big thank you is required for Dr. Sherri Christian for her continuous advice and help with data analysis and statistics. I had tremendous help from current and former members of the Lang laboratory in timing experiments, support and exchange of ideas. Thanks to Alex Hynes, Dr. Daniel Jones and Alexandra Rose for their input on my thesis and the entire Memorial University Biology Department for providing an enjoyable graduate school experience. Finally, a huge thank you goes to my supervisor Dr. Andrew Lang for his continual guidance, support and suggestions in adjusting to graduate school and completing my doctoral thesis. Dr. Lang provided me with the tools and help I needed to be successful in academic research and I greatly appreciate his help throughout my degree.

Research for my doctoral thesis has been supported by graduate student fellowships from the Memorial University School of Graduate Studies and the Natural Science and Engineering Research Council (NSERC). Dr. Andrew Lang's laboratory is supported by grants from the Natural Sciences and Engineering Research Council (NSERC), the Newfoundland and Labrador Research and Development Corporation, and the Candian Foundation for Innovation. The data presented in Chapter 2 was published as a research article in the Journal of Bacteriology in 2010 [192(11)]. The research in Chapter 3 was published as a research article in FEMS Microbiology Letters in 2012

[331(1)]. Therefore, the foundations of these chapters benefited from the editorial input of the manuscript co-authors and the suggestions of the peer reviewers.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

ANOVA - Analysis of variance

ATP - Adenosine triphosphate

BACTH - Bacterial adenylate cyclase two-hybrid

BLAST - Basic local alignment search tool

BPROM - Bacterial promoters prediction

CHAPS - 3-(3-cholamidopropyl)dimethylammonio-1-propane sulphonate

CHASE4 - Cyclases/histidine kinases associated sensory extracellular type 4 (protein domain)

COG - Cluster of orthologous groups

CRISPR - Clustered regularly interspaced short palindromic repeats

DNA - Deoxyribonucleic acid

DTT - Dithiothreitol

DUF - Domain of unknown function (protein domain)

ECF - Extra-cytoplasmic function

EDTA - Ethylenediaminetetraacetic acid

ESI - ElectroSpray Ionization

FDG - Fluorescein di-β-D-galactopyranoside

FTICR - Fourier transform ion cyclotron resonance

GEO - Gene expression omnibus

GTA - Gene transfer agent

HA/HATPase - Histidine kinase-like ATPase (protein domain)

HBB - Hook-basal body

HGT - Horizontal gene transfer

HPLC - High performance liquid chromatography

Hpt - Histidine phosphotransferase (protein domain)

HTH - Helix-turn-helix (protein domain)

IMG - Intergrated microbial genomes

IPTG - Isopropyl-β-D-thiogalactopyranoside

kb - Kilobase pair

LB - Luria-Bertani

LC-MS/MS - Liquid chromatography tandem mass spectophotometry

Mb - Megabase pair

n/a - Not applicable

NCBI - National Center for Biotechnology Information

ND - Not detected

Ni-NTA - Nickel-nitrilotriacetic acid

OD - Optical density

ONPG - o-nitrophenol-β-D-galactoside

ORF - Open reading frame

PAS - PER-ARNT-SIM (protein domain)

PCR - Polymerase chain reaction

PP2C - Protein phosphatase 2C-like (protein domain)

PVDF - Polyvinyl difluoride

RcGTA - *Rhodobacter capsulatus* gene transfer agent

RCV - *Rhodobacter capsulatus* medium V

REC - Receiver (protein domain)

RMA - Robust multi-array average

RNA - Ribonucleic acid

RNAP - Ribonucleic acid polymerase

RR - Response regulator

Rsb - Regulators of sigma B

Rsd - Regulator of sigma D

SCX - Strong cation exchange

SDS-PAGE - Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SHK - Sensor histidine kinase

SPE - Solid phase extraction

STAS - Sulfate Transporter and Anti-Sigma factor antagonist (protein domain)

TBST - Tris-buffered saline + tween

TE - Tris-EDTA

TIGR - The Institute for Genomic Research

YPS - Yeast extract/peptone/salts

σ - Sigma

List of Appendices

Hynes, A. P., **R. G. Mercer**, D. E. Watton, C. B. Buckley and A. S. Lang (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol. Microbiol.* **85**(2): 314-325.

Chapter 1 – Introduction and Overview

1.1. Horizontal gene transfer

Horizontal gene transfer (HGT) is the lateral transmission of genetic material between closely or distantly related species. The transfer of DNA between cells plays an important role in the evolution of prokaryotic genomes and can lead to acquisition of new genetic traits (Olendzenski & Gogarten, 2009). This can enhance the adaptation of a recipient cell to a variety of environmental conditions, such as through the ability to metabolize alternative nutrients. HGT is a common occurrence in bacteria with high frequencies of exchange occurring between genetically related species occupying the same environmental niche (Andam *et al.*, 2010; Beiko *et al.*, 2005; Kloesges *et al.*, 2011; Lang *et al.*, 2012). Analyses of diverse bacterial genomes have estimated an average proportion of horizontally transferred genes per genome at ~12%, with other studies suggesting >60% of genes in prokaryotic genomes have been affected by HGT (Nakamura *et al.*, 2004; Zhaxybayeva *et al.*, 2006). Regardless of the specific value, bacterial genomes are impacted by HGT, and the three most common mechanisms by which genes are exchanged are conjugation, transformation, and virus-mediated transduction. Conjugation is the transfer of genetic material, mainly plasmids, through direct cell-cell contact (Chen *et al.*, 2005; Lederberg & Tatum, 1946) via a complex extracellular structure called a pilus (Thomas & Nielsen, 2005). Transformation involves the uptake and stable maintenance of free DNA from the environment by a bacterial cell (Chen & Dubnau, 2004). Cells expressing proteins required for DNA uptake enter a

physiological condition known as 'competence', a prerequisite for bacterial transformation (Chen *et al.*, 2005). Transduction is the result of host cell genetic material being packaged by viral particles and subsequently transferred to a recipient cell (Ochman *et al.*, 2000; Zinder & Lederberg, 1952). Examples of important consequences of these processes in human society include the development of bacterial virulence and antibiotic resistance (Weinbauer, 2004).

1.1.1. Bacteriophages and prophages

Bacteriophages (phages) are viruses that infect bacteria and represent the most abundant biological entity on the planet (Ashelford *et al.*, 2003; Bergh *et al.*, 1989). Studies have estimated that ~10^{31} phage virions are present in the biosphere (Suttle, 2007). Phages are widely distributed and highly diverse, with variations in genome composition, structure, and complexity (Krupovic *et al.*, 2011). However, up to 95% of identified phages belong to the dsDNA tailed-phage order, *Caudovirales* (Ackermann, 2007; Suttle, 2007). The abundance of phages can influence microbial communities, as the overall phage population can determine host cell numbers and/or physiological state (Hennes *et al.*, 1995; Winter *et al.*, 2004). Most phages have a narrow host-range and are restricted to specific species whereas other phages have broader host ranges and are capable of infecting multiple species. One example of this broad host range is enterobacteriophage Mu, which is capable of infecting bacteria from members of the *Enterobacteriacae* family including the genera *Citrobacter*, *Escherichia*, *Erwinia*, *Salmonella* and *Shigella* (Paolozzi & Ghelardini, 2006). In natural habitats of bacteria, phages play a substantial role in mediating HGT. This is supported by the abundance of

phage-derived genetic information found in bacterial genomes (Canchaya *et al.*, 2003; Casjens, 2003) and the amount of bacterial genes found in viral metagenome studies (Rosario & Breitbart, 2011). Overall, it is believed that viruses play pivotal roles in the evolution of prokaryotic genomes and microbial diversity (Filee *et al.*, 2003; Pal *et al.*, 2007; Rodriguez-Valera *et al.*, 2009).

Differences in phage regulatory mechanisms and replication strategies can result in various host cell physiological effects. Lytic phage, such as enterobacteriophage T4 that infects *E. coli*, undergo lytic replication characterized by replication of the virus followed by lysis of the host cell to release progeny phage. The progeny can then infect neighbouring host cells, repeating the processes of penetration, replication, maturation and lysis. Some phage, such as enterobacteriophage λ that also infects *E. coli*, may also undergo lysogeny, where the phage integrates into the host cell genome becoming a prophage, and expression of most viral genes is essentially shut off (Dodd *et al.*, 2005). Phage replication is typically regulated by phage-encoded transcription factors, such as the λ CI repressor, which functions in appropriate timing of phage gene expression. The λ repressor binds the operator sequences and prevents expression of genes required for lytic development (Maurer *et al.*, 1980; Meyer *et al.*, 1980). Activation of host DNA repair systems can lead to proteolytic cleavage of CI, and expression of lytic genes (Little, 1984). Lytic phages also specifically regulate their gene expression. For example, phage T4 relies heavily on multiple phage-encoded factors (e.g. MotA, AsiA) for the temporal regulation of early, middle and late phage genes (Hinton, 2010).

Most bacterial genomes contain a number of prophages that can be functional or non-functional phage-like elements (Stanton, 2007), and these can account for up to 20% of a bacterial genome (Casjens, 2003). Functional prophages can be stimulated to enter the lytic cycle by a variety of stressors, such as chemicals (e.g. Mitomycin C) and UV light. A non-functional prophage is incapable of self-replication, but may still provide some biological function and serve as an important contributor to genome evolution (Stanton, 2007). The lysogenic cycle is believed to have an important influence on evolution of both the phage (i.e. through recombination with other phages that enter the cell) and the host (i.e. through provision of novel genes) (Krupovic *et al.*, 2011).

Genetic exchange mediated by phages is called transduction. There are two classes of transduction depending upon the genetic material being packaged within the bacteriophage capsid: specialized and generalized. Specialized transduction involves the aberrant excision and packaging of host DNA located adjacent to the phage integration site, such as the *gal* gene packaged occasionally by *E. coli* phage λ. In generalized transduction, any segment of the bacterial genome can be packaged, resulting in particles occasionally containing host genetic material, approximately the same size as the phage genome (Lang *et al.*, 2012 and references therein).

1.1.2. Phage-like gene transfer agents (GTAs)

Gene transfer agents (GTAs) are unusual phage-like particles that transfer random segments of host cell DNA in a process similar to generalized transduction (Lang *et al.*, 2012; Stanton, 2007). GTAs were first identified in the purple, non-sulfur α-proteobacterium, *Rhodobacter capsulatus* in the 1970s (Marrs, 1974). In this species,

DNase resistant genetic exchange of antibiotic resistance markers was not dependent on cell-cell contact, suggesting a transfer process analogous to transduction (Marrs, 1974). The phage-like particle responsible for this genetic transfer was smaller in size relative to other phages; therefore, it was called a 'gene transfer agent'. All known GTAs are structurally similar to tailed phage and are predicted to be released via cell lysis (Lang *et al.*, 2012). GTAs are encoded in (and replicate with) the host cell genome, likely representing descendants of defective prophages. Unlike a traditional generalized transducing phage, the DNA packaged in GTAs is less than what is required to encode the particle itself. A GTA cannot transfer all the required structural genes and give the ability to produce GTA to a recipient cell (Lang *et al.*, 2012; Stanton, 2007). Production of the different GTAs seems to be stimulated by different factors and their evolutionary significance for the producing species is not entirely understood.

1.1.2.1. Distribution of known gene transfer agents

Since the initial discovery of the *R. capsulatus* GTA, or RcGTA, functional GTAs have been identified in diverse prokaryotic species: *Desulfovibrio desulfuricans* (Rapp & Wall, 1987), *Methanococcus voltae* (Eiserling *et al.*, 1999) and *Brachyspira hyodysenteriae* (Humphrey *et al.*, 1997). The discovery of genetically unrelated GTAs from these α-proteobacteria, δ-proteobacteria, spirochetes and archaea suggest a possible beneficial role of GTAs arising independently in unrelated lineages. The increase in sequenced bacterial genomes over the last decade has allowed for a more comprehensive search for the genetic potential to produce GTAs. Phylogenetic studies using RcGTA as a model (Lang & Beatty, 2007) described the evolutionary relationships between species

identified to possess complete and incomplete GTA gene clusters. The most highly scoring nucleotide BLAST (Basic Local Alignment Search Tool) hits to the RcGTA sequences are from the α-proteobacteria, where most species contain at least some of the genes required to encode a GTA (Lang & Beatty, 2007). The genetic potential for GTA production, as well as the expression of GTA proteins, is particularly widespread in marine bacteria that are classified in the order Rhodobacterales, such as *Ruegeria pomeroyi* (Biers, 2008; Fu *et al.*, 2010; Lang & Beatty, 2007). It is believed that RcGTA-like genes have been vertically transmitted from a GTA-containing ancestor, with the entire RcGTA gene cluster being conserved in the Rhodobacterales order (Lang & Beatty, 2007). This evolutionary conservation supports the hypothesis that GTAs provide a selective advantage that benefits the host species (Lang *et al.*, 2012).

It is estimated that there are up to a billion virus particles in every millilitre of seawater (Suttle, 2005), and a portion of these could be represented by GTAs (Kristensen *et al.*, 2010; Lang & Beatty, 2007; Rohwer & Vega Thurber, 2009). The Roseobacter lineage in the Rhodobacterales may account for >25% of marine prokaryotic communities and numerous isolates of these bacteria have been found to express GTA capsid proteins (Fu *et al.*, 2010). Some research has suggested that cross-species gene transfer via GTAs occurs in the natural environment, with up to 47% of laboratory culturable recipient strains able to receive genetic material by this mechanism (McDaniel *et al.*, 2010). The large number of bacterial species possessing the necessary genes required to encode GTAs, combined with the potential that many of these species are actually producing

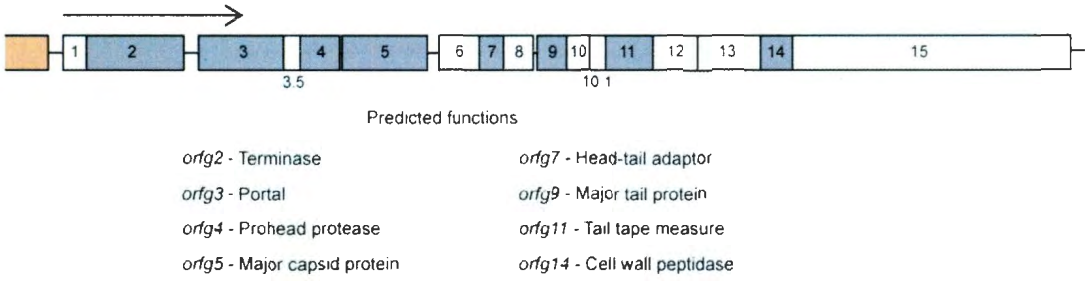
functional particles, suggests that GTAs may play a pivotal role in shaping microbial communities and the overall evolution of bacterial genomes.

1.1.3. *Rhodobacter capsulatus* and RcGTA

The first discovered and one of the best characterized of the currently known GTAs is RcGTA, which serves as a model for studying these elements. The gram-negative, rod-shaped bacterium *R. capsulatus* (formally *Rhodopseudomonas capsulata*) is classified within the Rhodobacterales order in the class α-proteobacteria and is found in a variety of environments (Madigan & Jung, 2009). This species has the ability to grow under aerobic or anaerobic conditions using various sources of organic carbon and terminal electron acceptors (Imhoff & Bias-Imhoff, 1995). Purple, non-sulfur bacteria, like *R. capsulatus*, are also photoheterotrophs that can carry out photosynthesis using light energy generated from photon capture (Pemberton *et al.*, 1998). Due to its metabolic versatility, *R. capsulatus* has previously been utilized for studying photosynthesis, nitrogen fixation and cellular energetics (Haselkorn *et al.*, 2001; Imhoff & Bias-Imhoff, 1995; Madigan, 1995; Pemberton *et al.*, 1998). The genome of *R. capsulatus* consists of a single 3.7-Mb chromosome and a 134-kb circular plasmid, with a relatively high GC content of 66.6% (Haselkorn *et al.*, 2001; Strnad *et al.*, 2010). The genome contains 3685 open reading frames (ORFs) and a coding density of 91% (Strnad *et al.*, 2010). Large regions of functionally similar genes are found throughout the genome, including a ~45-kb photosynthesis gene cluster, 8 clustered regularly interspaced short palindromic repeats (CRISPR) regions and 237 phage-like genes, including the RcGTA gene cluster (Strnad *et al.*, 2010).

The structural genes encoding RcGTA are organized in a ~15-kb cluster (*rcc01682 – rcc01698*) in the *R. capsulatus* genome (Lang & Beatty, 2000). The cluster is comprised of 17 ORFs in an operon (*orfg1-orfg15* with additional *orfg3.5* and *orfg10.1*) arranged in a typical bacteriophage head-to-tail organization (Casjens *et al.*, 1992; Lang & Beatty, 2000) (Fig 1.1A). Many of the RcGTA ORFs have no functional annotation, but some of the genes share homology with known phage genes. RcGTA *orfg2* and *orfg4* share homology with *E. coli* T4 terminase and HK97 prohead protease, respectively. *Streptomyces* phage φC31 portal and capsid proteins are similar to *orfg3* and *orfg5*, respectively, and *orfg5* has been confirmed to encode the major capsid protein (Lang & Beatty, 2000). There are no lysin or holin genes encoded in or near the cluster which initially complicated the understanding of the mechanism of particle release, but a putative endolysin/holin pair required for RcGTA release has now been identified (Hynes *et al.*, 2012; Appendix 6) Furthermore, there are no phage-like regulatory genes nearby (Lang & Beatty, 2000). The gene cluster is flanked by cellular genes coding for fatty acid biosynthesis/conserved hypothetical proteins (*fabF/chp*; 5' of *orfg1*) and cysteine biosynthesis proteins (*cysE*; 3' of *orfg15*). Proteomics on RcGTA particles has revealed the presence of other gene products not encoded in the 15-kb gene cluster (Chen, 2008).

A



B

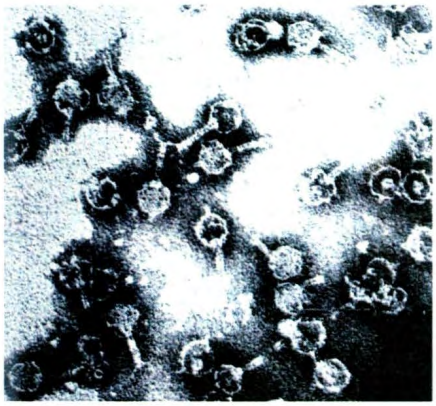


Figure 1.1. Genetics and structure of the *R. capsulatus* gene transfer agent (RcGTA). A. Genetic arrangement of the RcGTA gene cluster. The direction of transcription is indicated by the arrow and predicted or known functions are indicated below for some of the gene products. The greyed ORFs indicate genes with known function based on biochemical data (*orfg5*) or predicted function based on amino acid sequence similarity with known phages. The coloured ORF indicates a cellular gene that precedes the RcGTA gene cluster. B. Electron micrograph of RcGTA particles (with permission; Yen *et al.*, 1979), which resemble tailed-phages with an approximate head diameter of 30 nm and tail length of 50 nm.

Resembling a small tailed phage, the RcGTA particle has a head diameter of ~30 nm and a tail length of ~50 nm (Yen *et al.*, 1979) (Fig. 1.1B). The particle incorporates ~4 kb of random genomic segments into the capsid, with the packaging frequency not

affected by genome position (Solioz & Marrs, 1977; Yen *et al.*, 1979). RcGTA production is greatest in the stationary phase of culture growth with the highest transfer frequency determined to be 4×10^{-4} per recipient cell (Solioz *et al.*, 1975). A culture-dependent quorum-sensing system has been shown to modulate RcGTA production, agreeing with the observed increase in gene transfer in the later growth phases (Schaefer *et al.*, 2002; Solioz *et al.*, 1975). There is no observable lysis in RcGTA producing cultures (Marrs, 1974; Solioz *et al.*, 1975; Yen *et al.*, 1979), however, it has been hypothesized that the particle is only released (by cell lysis) from a sub-population (~3%) of cells in stationary phase cultures (Hynes *et al.*, 2012). RcGTA production is widespread amongst wild type strains of *R. capsulatus* but it is specific in its host range as inter-species transfer has not been demonstrated, despite some comprehensive screening attempts (Wall *et al.*, 1975).

The RcGTA genes were found to be positively regulated by two cellular genes, *cckA* and *ctrA* (Lang & Beatty, 2000). These genes were predicted to encode a sensor histidine kinase and response regulator protein, respectively, that participated in a two-component signalling system controlling RcGTA (Lang & Beatty, 2000). However, the exact regulatory function of these putative proteins on RcGTA expression and how the pathway is stimulated is poorly understood. The sequences of these genes indicated that they are homologues of a well studied phosphorelay system involving cell cycle regulation in the model organism *Caulobacter crescentus* (Lang & Beatty, 2000)

1.2. Bacterial regulation of gene expression

The control of gene expression is critically important for bacteria in order to efficiently respond to different environmental conditions and stimuli. The first tier of gene regulation is at transcription initiation, where signalling pathways and transcriptional regulators control when genes are expressed and to what level. Post-transcriptional regulation can also occur through effects on transcript half-life (i.e. mRNA stability) and regulation of translation of the mRNA. Post-translational modifications can control proper protein folding or target proteins for degradation by proteolysis. Functional protein activity is also modulated by regulatory pathways, activating or inactivating proteins by covalent modification or direct protein-protein interactions. The research presented in this thesis focuses on regulation of expression of the RcGTA genes by signalling systems and the control of transcriptional regulatory protein activity.

1.2.1. Two-component and phosphorelay signalling

Two-component and phosphorelay signalling systems are some of the most dominant signal transduction pathways of Bacteria (Galperin & Gomelsky, 2005) and mediate rapid alterations in gene expression in response to environmental and intracellular changes. One of the first discovered and most well-characterized two-component signalling systems is the EnvZ-OmpR pathway of *E. coli*, which modifies outer membrane porin protein gene expression in response to osmotic conditions (Forst & Roberts, 1994). A two-component signalling system is typically characterized by a histidine protein kinase and a response regulator protein. The histidine kinase contains a highly diverse 'input' sensory domain, which recognizes a particular signal (Krell *et al.*,

2010). Upon stimulus, the histidine kinase undergoes autophosphorylation on a conserved histidine residue within the phosphotransferase sub-domain (Hoch, 2000). This results in a conformational change, such as dimerization, switching the protein to an active signalling state (Casino *et al.*, 2010). The histidine kinase then participates in a histidyl-aspartyl phosphotransfer to a conserved aspartate residue on a cognate response regulator protein. A response regulator is composed of an Asp-containing receiver (REC) domain and an output domain such as a helix-turn-helix (HTH) DNA-binding domain. The phosphorylation-induced conformational change of a response regulator typically results in activation of the protein, enabling it to carry out a specific function (Stock *et al.*, 2000). Early work showed that some of these proteins, such as OmpR, were transcriptional regulators that directly activate or repress target genes (Kenney, 2002). However, response regulators can vary in their output domains and function, whereby activation can alter protein-protein interactions and enzymatic cycling of secondary signalling molecules, in addition to transcriptional regulation (Galperin, 2006). The number of different histidine kinase input and response regulator output domains reflects the diversity in stimuli and responses of two-component signalling systems. As well, some histidine kinases have multiple phosphotransfer targets and response regulators can be phosphorylated by multiple kinases. This many-to-one or one-to-many relationship of some two-component systems differs from signal cross-talk where response regulators are inadvertently phosphorylated by a non-cognate histidine kinases (Laub *et al.* 2007; Laub & Goulian, 2007). However, signalling between histidine kinases and response regulators is tightly buffered, and while cross-talk between non-cognate partners has been identified,

this event is rare and signalling partners are usually very specific (Laub & Goulian, 2007).

Two component signal transduction systems have evolved into more complicated pathways where the components can vary in structure, function and mechanism of signal transfer (Mascher *et al.*, 2006). Histidine kinases can be membrane bound, or not, and many are bifunctional, possessing phosphatase activity for dephosphorylation of a cognate response regulator (Laub & Goulian, 2007). Also, there may be additional phosphorelay proteins that act to bridge the signal between the histidine kinases and response regulators (Hoch, 2000). When these systems are composed of three or more proteins that participate in the phosphotransfer, they are typically referred to as phosphorelay signal transduction systems (Cheung & Hendrickson, 2010). A phosphorelay system can involve a hybrid histidine kinase that possesses an additional receiver domain with a conserved aspartate residue on the protein. In this instance, stimulus-induced autophosphorylation of the conserved histidine residue is followed by intramolecular transfer of the phosphoryl group to the aspartate residue (Stock *et al.*, 2000). This is followed by phosphotransfer to another intermediate histidine-containing phosphotransferase protein, which shuttles the phosphoryl group to the aspartate residue of the cognate response regulator to elicit the output response (Laub *et al.*, 2007). A schematic representation of two-component and phosphorelay signal transduction proteins is shown in Figure 1.2.

The complex circuits of two-component signalling proteins and phosphorelay pathways provide bacteria with an efficient response network for adaption to stressors,

growth conditions and environments (Laub & Goulian, 2007). The average bacterial genome (~3.5 Mbp) contains 50-100 two-component signalling protein-encoding genes (Whitworth & Cock, 2009). Some species, such as some members of the Cyanobacteria and Myxobacteria, encode >240 two-component signalling proteins (Whitworth & Cock, 2008). The large number of two-component signalling proteins identified has required the use of large comprehensive databases such as the Prokaryotic 2-Component Systems (P2CS) (Barakat *et al.*, 2009). This database is a compilation of the >105000 two-component/phosphorelay proteins, including more than 45000 histidine kinases and 52000 response regulators identified through genomic or metagenomic sequencing (as of August 2012 - http://www.p2cs.org/ ; (Barakat *et al.*, 2009).

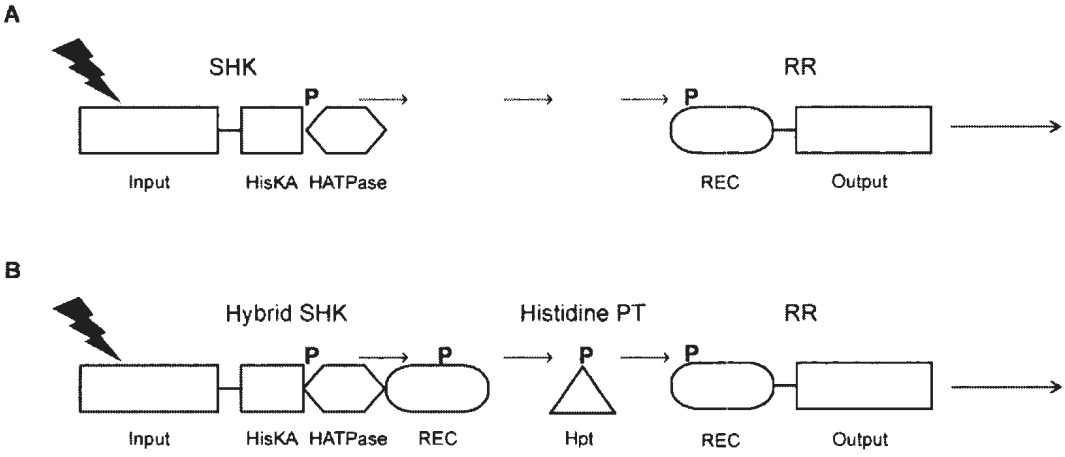


Figure 1.2. Two-component signal transduction and phosphorelay pathways. A. Two-component signal transduction between a sensor histidine kinase (SHK) and a response regulator (RR). A stimulus (lightning bolt) initiates the transfer of a phosphoryl group (P) from a histidine containing domain (HisKA) in the SHK to an aspartate residue located in the REC domain of a RR. B. A phosphorelay system involving a hybrid SHK and an additional histidine phosphotransferase (PT) protein. After autophosphorylation of the

hybrid SHK, the phosphoryl group is transferred to a REC domain on the same protein. The phosphotransfer to a histidine-containing Hpt protein and then to the REC domain of the cognate RR produces an output response. The conserved domains presented along with generic inputs and outputs include: HisKA, histidine kinase A; HATPase, histidine kinase-like ATPase; REC, receiver; and Hpt, histidine phosphotransferase.

1.2.1.1. *Caulobacter crescentus* and CtrA

One of the best studied phosphorelay signalling systems regulates the cell cycle of the gram-negative α-proteobacterium *Caulobacter crescentus*. This system involves the hybrid histidine kinase, CckA (Jacobs *et al.*, 1999), and response regulator, CtrA (Quon *et al.*, 1996), both of which are essential for the bacterium's viability. *C. crescentus* is a model system for studying the bacterial cell cycle and cell division, which yields two cell types, a motile swarmer cell and sessile stalked cell (England & Gober, 2001; Jensen *et al.*, 2002). CtrA is known as the 'master' regulator of the cell cycle and is responsible for the direct control of 95 genes involved in cell division, polar morphogenesis and DNA replication/methylation (Laub *et al.*, 2002). It also influences the initiation of chromosome replication by interacting with sequences at the origin of replication (Quon *et al.*, 1998). The activity of CtrA is tightly controlled by both phosphorylation/dephosphorylation and proteolysis (Domian *et al.*, 1997). Phosphorylation of CtrA is controlled by the histidine kinase CckA and histidine phosphotransferase ChpT (Jacobs *et al.*, 2003). Additional phosphorelay proteins such as DivK, DivJ, PleC and DivL are involved in regulating the activity of this pathway (Jacobs *et al.*, 2001; Sciochetti *et al.*, 2005). CtrA~P levels are also controlled by CckA phosphatase activity (Chen *et al.*, 2009) and proteolysis, stimulated by an additional

response regulator, CpdR (Iniesta *et al.*, 2006). Feedback inhibition of *ctrA* expression and CtrA-regulated genes by the transcription factor SciP provides another layer of control of CtrA~P activity (Gora *et al.*, 2010). The accumulation and activity of CckA at opposite cell poles results in a CtrA~P gradient that influences differentiation and asymmetric cell division (Angelastro *et al.*, 2010). The complex, multi-level control in this pathway reflects the importance of tightly regulating cell cycle gene expression, from transcription to protein activity. This regulatory pathway is illustrated in Figure 1.3 and discussed further in Chapter 3 of this thesis.

Homologues of CtrA have been identified and studied in other α-proteobacterial species including *Sinorhizobium meliloti* (Barnett *et al.*, 2001), *Rhodopseudomonas palustris* (Lang & Beatty, 2001), *Brucella abortus* (Bellefontaine *et al.*, 2002) and *R. capsulatus* (Lang & Beatty, 2000). Indeed, CtrA appears to be conserved in every genome sequence from species in the α-proteobacteria with the exception of the *Pelagibacter* lineage (Brilli *et al.*, 2010; Lang & Beatty, 2007). The phylogenetic relationships and divergence in amino acid sequence and function of CtrA amongst these species are elaborated in Chapter 2. The *R. capsulatus* CtrA protein shares 71% amino acid sequence identity with the *C. crescentus* protein, contains the conserved Asp51 phosphorylation site and a virtually identical DNA-binding region (Lang & Beatty, 2000). Unlike the *C. crescentus* protein, the *R. capsulatus* CtrA and CckA homologues are not essential but have been found to be involved in regulating RcGTA production and motility (Lang & Beatty, 2000, 2002).

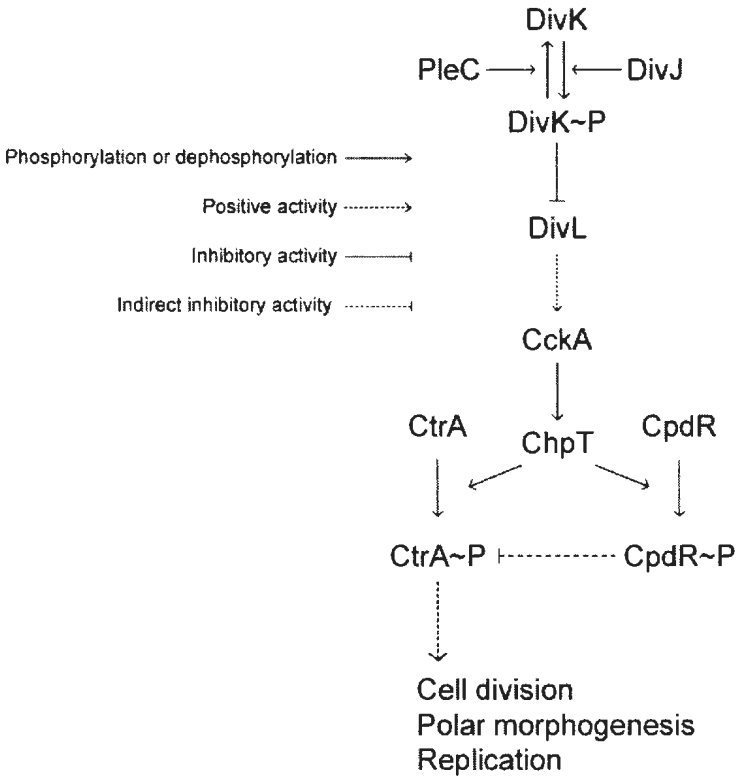


Figure 1.3. The CtrA signalling pathway of *C. crescentus*. A simplified version of this complex signalling pathway is shown, involving phosphorylation of several proteins that can influence the phosphotransfer between other members of the system.

1.2.2. Sigma factors and regulation of gene expression

Transcription in Bacteria is dependent on an essential subunit of the RNA polymerase (RNAP) holoenzyme called the sigma (σ) factor, which determines promoter specificity (Helmann & Chamberlin, 1988). The σ factor interacts with the RNAP core complex transcriptional machinery and recognizes specific sequences in target gene promoters to initiate transcription (Paget & Helmann, 2003). All bacteria have at least 1 essential σ factor (frequently referred to as the major vegetative sigma factor and denoted

σ^D or σ^{70}, where '70' refers to its molecular weight in kDa in *E. coli*) for the transcription of 'housekeeping' genes (Campbell *et al.*, 2008) in addition to several alternative σ factors for the expression of genes/operons in response to specific stimuli or stresses (Helmann, 1999; Mooney *et al.*, 2005). Alternative σ factors substitute for the major vegetative σ factor and redirect RNAP to initiate transcription from alternative promoters (Helmann & Chamberlin, 1988). The first identified alternative σ factor was the stress response factor, σ^B, of *Bacillus subtilis* (Binnie *et al.*, 1986). Most alternative σ factors (σ^{24}, σ^{28}, σ^{32}, σ^{38}, etc.) are structurally similar to the σ^{70}-family, and share high amino acid sequence identity and structural similarities in RNAP and promoter contact sites (Campbell *et al.*, 2008; Helmann & Chamberlin, 1988). The σ^{70}-family of σ factors has two highly conserved domains, σ_2 and σ_4. The σ_2 domain binds the -10 promoter sequence upstream of the transcription initiation point and interacts with the β' subunit of RNAP. The -35 promoter sequence is bound by the σ4 domain, which interacts with the RNAP β subunit (Campbell *et al.*, 2008). Not all alternative σ factors are similar and the σ^{54} family members, frequently involved in nitrogen metabolism, are structurally and functionally distinct (Merrick, 1993).

The σ^{70} family of σ factors can be divided into 4 groups based on their gene structure and function (Lonetto *et al.*, 1992). Group 1 consists of the essential primary σ factors, including the major vegetative σ^D or σ^{70}, while groups 2 and 3 are closely related but non-essential proteins, with roles in stress responses, development and secondary metabolism (Paget & Helmann, 2003). Group 4 represents the largest and most diverse group of extracytoplasmic function (or ECF) σ factors, which frequently respond to

extracellular signals (Helmann, 2002; Staroń *et al.*, 2009). Most bacteria have a single group 1, or primary σ factor, but the number of group 2-4 proteins is reflective of the species' physiological complexity (Osterberg *et al.*, 2011). For example, *E. coli* has 7 σ factors (1 primary and 6 alternative) while the antibiotic-producing bacterium, *Streptomyces coelicolor*, encodes up to 63 σ factors (49 group 4) in its genome (Bentley *et al.*, 2002).

The activity of a σ factor can be controlled by highly diverse proteins known as anti-σ factors (Campbell *et al.*, 2008). Most alternative σ factors remain inactive in the absence of stimuli, a state maintained by direct protein-protein interaction with a cognate anti-σ factor (Brown & Hughes, 1995; Helmann, 1999). Some anti-σ factors, such as the *E. coli* bacteriophage T4-encoded AsiA protein, interact with the major vegetative σ factor (Severinova *et al.*, 1998). Anti-σ factors themselves can be negatively regulated directly by anti-anti-σ factor proteins, which thereby promote σ factor activity and positively regulate target genes (Sharma *et al.*, 2011). These σ-regulatory factors provide another layer of control of transcription initiation and have been extensively studied in model species such as *E. coli* and *B. subtilis* (Campbell *et al.*, 2008). The 'partner switching' of anti-σ factors with cognate σ factors and anti-anti-σ factors allows for robust control of σ factor activity and, consequently, expression of target genes/operons (Fig 1.4). The mechanism of 'partner switching' has been observed in many systems in *Bacillus*, *Streptomyces* and *Sinorhizobium* spp. (Homerova *et al.*, 2011; Lee *et al.*, 2004; Staroń & Mascher, 2010; Sevcikova *et al.*, 2011).

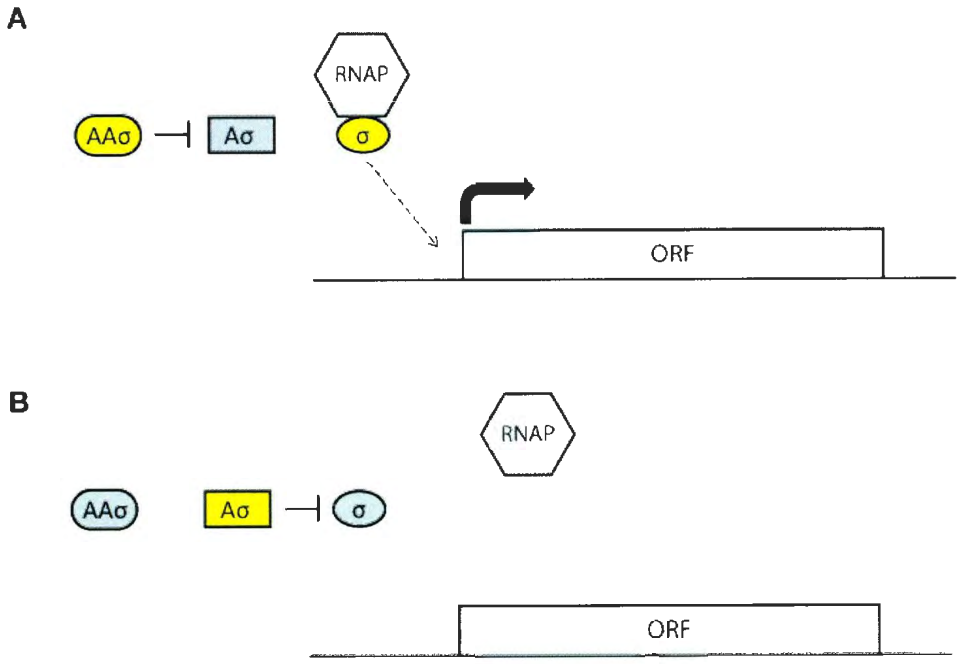


Figure 1.4. Partner switching in regulation of σ factor activity. A. Transcriptional activation by a σ factor. The anti-σ factor (Aσ) is maintained in an inactive (grey) state by an anti-anti-σ (AAσ) factor, allowing the cognate σ factor to associate with the RNAP to initiate transcription (indicated by the thick black arrow) from the promoter of an ORF. B. Repression of a σ factor by an anti-σ-factor. When the anti-anti-σ factor is inactive, the anti-σ factor can sequester the cognate σ factor from complexing with RNAP, preventing transcription of target ORFs. Active proteins are indicated by colour (yellow).

The activity of anti-σ factors and anti-anti-σ factors can be influenced by signalling systems, and may involve phospho-dependent activation and inactivation of these regulators. Phosphorelay systems can affect anti-anti-σ factors by activating or inactivating the antagonist via phosphorylation (Price, 2010; Staroń & Mascher, 2010). In the general stress response of many α-proteobacteria, putative sensor histidine kinase(s) initiate a phosphotransfer to a unique anti-anti-σ factor, PhyR, upon stimulus detection.

The phosphorylation of PhyR allows it to interact with and inhibit the antagonistic activity of the anti-σ factor, NepR (Staroń & Mascher, 2010). This system controls the activity of EcfG-like σ factors and has been found to play an important role in general stress response in *Bradyrhizobium japonicum* (Gourion *et al.*, 2009), *C. crescentus* (Alvarez-Martinez *et al.*, 2007) and *Sinorhizobium meliloti* (Bastiat *et al.*, 2010). In contrast to this system, the general stress response of some gram positive species involves inactivation of an anti-anti-σ by phosphorylation (Hecker *et al.*, 2007). Some anti-σ factors can act as kinases and possess the ability to phosphorylate and inactivate their cognate anti-anti-σ factor. Alleviating the anti-anti-σ factor inhibition allows the anti-σ factor to interact with the cognate σ factor and repress its activity (Alper *et al.*, 1994). Phosphorelay systems can dephosphorylate anti-anti-σ factors, allowing them to once again bind and inhibit anti-σ factors, and promote expression of the σ factor target genes (Price, 2010). One well studied σ factor regulatory system involving phosphorelay is the Rsb pathway in *Bacillus* spp., which will be discussed in more detail in Chapter 4.

1.2.3. Regulation of RcGTA gene expression and production

As previously mentioned, CtrA is required for transcription of the RcGTA gene cluster. The genes are believed to be transcribed as a single, polycistronic operon with subsequent mRNA processing and differing transcript half-lives (Lang & Beatty, 2000). Typically, phage gene expression is controlled by phage-encoded transcriptional regulators, so it is highly unusual to find a cellular response regulator acting to control expression and production of a phage-like particle. Other factors have been identified to also be involved in regulating production of RcGTA, although the complete loss of

RcGTA expression has only been observed in a *ctrA* mutant (Lang & Beatty, 2000). The presence of an identical DNA binding region in the CtrA proteins in *R. capsulatus* and *C. crescentus* suggests that the *R. capsulatus* protein binds the same full (TTAA-N7-TTAA) and half (TTAACCAT) consensus sequences characterized in *C. crescentus* (Quon *et al.*, 1996). However, no predicted CtrA binding sites have been identified upstream of the RcGTA gene cluster. Production of RcGTA is also controlled in part by quorum sensing involving an N-acyl-homoserine lactone (acyl-HSL) synthase called GtaI (Schaefer *et al.*, 2002). Work by Leung *et al.* (2012) identified a LuxR-like quorum sensing receptor, GtaR, encoded with GtaI in a two-gene operon. The identification and characterization of the GtaR and GtaI proteins, alongside the observation that other acyl-HSL compounds can induce RcGTA production, support the model of increased stationary phase production of RcGTA in response to high cell density (Leung *et al.*, 2012). It has also been found that certain nutrient limitations can affect intra- and extracellular RcGTA capsid protein levels (Taylor, 2004). These combined findings suggest that the regulation of RcGTA production is quite complex.

1.3. Research goals

Although expression of the RcGTA genes is dependent on CtrA in *R. capsulatus*, the absence of putative CtrA binding sites in the promoter region of the RcGTA gene cluster suggests that CtrA may not act directly to positively regulate RcGTA. Lang and Beatty (2000) proposed that perhaps CtrA was regulating a σ factor that then controls expression of RcGTA genes, however it was unknown which *R. capsulatus* σ factor was required for targeting the RNAP to the RcGTA promoter. This led to the initial

hypothesis for my thesis research that CtrA controls the expression of other regulators that directly control RcGTA gene expression. The goal of my research was to identify these additional regulators of RcGTA expression and any other components of the regulatory network involving CtrA. The hypothesis was approached first by utilizing a whole-genome transcriptome analysis of a *ctrA* mutant to determine genes that are affected by the absence of CtrA (Chapter 2). I also investigated the involvement of putative CtrA regulators on production of RcGTA to try and determine the conservation of the CtrA signalling pathway in this species (Chapter 3). Throughout this research, important regulatory sequences in the RcGTA gene cluster have been identified and mutations of CtrA-regulated genes have led to the discovery of putative σ factor regulators that affect RcGTA expression and production in this species (Chapter 4). The results of the work presented in this thesis further elucidate the complex regulatory systems controlling RcGTA production.

1.4. References

Ackermann, H.W. (2007). 5500 phages examined in the electron microscope. *Arch. Microbiol.* **152**: 227-243.

Alper, S., L. Duncan and R. Losick (1994). An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in *B. subtilis*. *Cell* **77**(2): 195-205.

Alvarez-Martinez, C. E., R. F. Lourenço, R. L. Baldini, M. T. Laub and S. L. Gomes (2007). The ECF sigma factor σ^T is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. *Mol. Microbiol.* **66**(5): 1240-1255.

Andam, C. P., D. Williams and J. P. Gogarten (2010). Biased gene transfer mimics patterns created through shared ancestry. *Proc. Natl Acad. Sci. USA* **107**(23): 10679-10684.

Angelastro, P. S., O. Sliusarenko and C. Jacobs-Wagner (2010). Polar localization of the CckA histidine kinase and cell cycle periodicity of the essential master regulator CtrA in *Caulobacter crescentus*. *J. Bacteriol.* **192**(2): 539-552.

Ashelford, K. E., M. J. Day and J. C. Fry (2003). Elevated abundance of bacteriophage infecting bacteria in soil. *Appl. Environ. Microbiol.* **69**(1): 285-289.

Barakat, M., P. Ortet, C. Jourlin-Castelli, M. Ansaldi, V. Mejean and D. Whitworth (2009). P2CS: a two-component system resource for prokaryotic signal transduction research. *BMC Genomics* **10**(1): 315.

Barnett, M. J., D. Y. Hung, A. Reisenauer, L. Shapiro and S. R. Long (2001). A homolog of the CtrA cell cycle regulator is present and essential in *Sinorhizobium meliloti*. *J. Bacteriol.* **183**: 3204-3210.

Bastiat, B., L. Sauviac and C. Bruand (2010). Dual control of *Sinorhizobium meliloti* RpoE2 sigma factor activity by two PhyR-type two-component response regulators. *J. Bacteriol.* **192**(8): 2255-2265.

Beiko, R. G., T. J. Harlow and M. A. Ragan (2005). Highways of gene sharing in prokaryotes. *Proc. Natl. Acad. Sci. USA* **102**: 14332-14337.

Bellefontaine, A.-F., C. E. Pierreux, P. Mertens, J. Vandenhaute, J.-J. Letesson and X. D. Bolle (2002). Plasticity of a transcriptional regulation network among alpha-proteobacteria is supported by the identification of CtrA targets in *Brucella abortus*. *Mol. Microbiol.* **43**: 945-960.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**(6885): 141-147.

Bergh, O., K. Y. Borsheim, G. Bratbak and M. Heldal (1989). High abundance of viruses found in aquatic environments. *Nature* **340**(6233): 467-468.

Biers, E. J. (2008). Occurrence and expression of gene transfer agent genes in marine bacterioplankton. *Appl. Environ. Microbiol.* **74**: 2933-2939.

Binnie, C., M. Lampe and R. Losick (1986). Gene encoding the sigma 37 species of RNA polymerase sigma factor from *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* **83**(16): 5943-5947.

Biondi, E. G., S. J. Reisinger, J. M. Skerker, M. Arif, B. S. Perchuk, K. R. Ryan and M. T. Laub (2006). Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**(7121): 899-904.

Brilli, M., M. Fondi, R. Fani, A. Mengoni, L. Ferri, M. Bazzicalupo and E. Biondi (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Syst. Biol.* **4**(1): 52.

Brown, K. L. and K. T. Hughes (1995). The role of anti-sigma factors in gene regulation. *Mol. Microbiol.* **16**(3): 397-404.

Campbell, E. A., L. F. Westblade and S. A. Darst (2008). Regulation of bacterial RNA polymerase σ factor activity: a structural perspective. *Curr. Opin. Microbiol.* **11**(2): 121-127.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin and H. Brussow (2003). Prophage genomics. *Microbiol. Mol. Biol. Rev.* **67**: 238-276.

Casino, P., V. Rubio and A. Marina (2010). The mechanism of signal transduction by two-component systems. *Curr. Opin. Struct. Biol.* **20**(6): 763-771.

Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* **49**(2): 277-300.

Casjens, S., G. Hatfull and R. Hendrix (1992). Evolution of dsDNA tailed-bacteriophage genomes. *Sem. Virol.* **3**: 383-397.

Chen, F. (2008). Proteomic analysis and identification of the structural and regulatory proteins of the *Rhodobacter capsulatus* gene transfer agent. *J. Proteome Res.* **8**: 967-973.

Chen, I., P. J. Christie and D. Dubnau (2005). The ins and outs of DNA transfer in bacteria. *Science* **310**: 1456-1460.

Chen, I. and D. Dubnau (2004). DNA uptake during bacterial transformation. *Nature Rev. Microbiol.* **2**: 241-249.

Chen, Y. E., C. G. Tsokos, E. G. Biondi, B. S. Perchuk and M. T. Laub (2009). Dynamics of two phosphorelays controlling cell cycle progression in *Caulobacter crescentus*. *J. Bacteriol.* **191**(24): 7417-7429.

Cheung, J. and W. A. Hendrickson (2010). Sensor domains of two-component regulatory systems. *Curr. Opin. Microbiol.* **13**(2): 116-123.

Dodd, I. B., K. E. Shearwin and J. B. Egan (2005). Revisited gene regulation in bacteriophage λ. *Curr. Opin. Genet. Dev.* **15**(2): 145-152.

Domian, I. J., K. C. Quon and L. Shapiro (1997). Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell* **90**(3): 415-424.

Eiserling, F., A. Pushkin, M. Gingery and G. Bertani (1999). Bacteriophage-like particles associated with the gene transfer agent of *Methanococcus voltae* PS. *J. Gen. Virol.* **80**: 3305-3308.

England, J. C. and J. W. Gober (2001). Cell cycle control of cell morphogenesis in *Caulobacter*. *Curr. Opin. Microbiol.* **4**(6): 674-680.

Filee, J., P. Forterre and J. Laurent (2003). The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies. *Res. Microbiol.* **154**(4): 237-243.

Forst, S. A. and D. L. Roberts (1994). Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *Res. Microbiol.* **145**(5-6): 363-73.

Fu, Y., D. M. MacLeod, R. B. Rivkin, F. Chen, A. Buchan and A. S. Lang (2010). High diversity of Rhodobacterales in the subarctic North Atlantic Ocean and gene transfer agent protein expression in isolated strains. *Aquat. Microb. Ecol.* **59**: 283-293.

Galperin, M. Y. (2006). Structural classification of bacterial response regulators: Diversity of output domains and domain combinations. *J. Bacteriol.* **188**(12): 4169-4182.

Galperin, M. Y. and M. Gomelsky (2005). Bacterial signal transduction modules: from genomics to biology. *ASM News.* **71:** 326-333.

Gora, K. G., C. G. Tsokos, Y. E. Chen, B. S. Srinivasan, B. S. Perchuk and M. T. Laub (2010). A cell-type-specific protein-protein interaction modulates transcriptional activity of a master regulator in *Caulobacter crescentus*. *Mol. Cell* **39**(3): 455-467.

Gourion, B., S. Sulser, J. Frunzke, A. Francez-Charlot, P. Stiefel, G. Pessi, J. A. Vorholt and H.-M. Fischer (2009). The PhyR-σEcfG signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **73**(2): 291-305.

Haselkorn, R., A. Lapidus, Y. Kogan, C. Vlcek, J. Paces, V. Paces, P. Ulbrich, T. Pecenkova, D. Rebrekov, A. Milgram, M. Mazur, R. Cox, N. Kyrpides, N. Ivanova, V. Kapatral, T. Los, A. Lykidis, N. Mikhailova, G. Reznik, O. Vasieva and M. Fonstein (2001). The *Rhodobacter capsulatus* genome. *Photosynth. Res.* **70**(1): 43-52.

Hecker, M., J. Pané-Farré and V. Uwe (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu. Rev. Microbiol.* **61**(1): 215-236.

Helmann, J. D. (1999). Anti-sigma factors. *Curr. Opin. Microbiol.* **2**(2): 135-141.

Helmann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* **46:** 47-110.

Helmann, J. D. and M. J. Chamberlin (1988). Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**: 839-872.

Hennes, K. P., C. A. Suttle and A. M. Chan (1995). Fluorescently labeled virus probes show that natural virus populations can control the structure of marine microbial communities. *Appl. Environ. Microbiol.* **61**(10): 3623-7.

Hinton, D. (2010). Transcriptional control in the pre-replicative phase of T4 development. *Virol. J.* **7**(1): 289.

Hoch, J. A. (2000). Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* **3**(2): 165-170.

Homerova, D., B. Sevcikova, B. Rezuchova and J. Kormanec (2011). Regulation of an alternative sigma factor σ^I by a partner switching mechanism with an anti-sigma factor PrsI and an anti-anti-sigma factor ArsI in *Streptomyces coelicolor* A3(2). *Gene* **492**(1): 71-80.

Humphrey, S. B., T. B. Stanton, N. S. Jensen and R. L. Zuerner (1997). Purification and characterization of VSH-1, a generalized transducing bacteriophage of *Serpulina hyodysenteriae*. *J. Bacteriol.* **179**: 323-329.

Hynes, A. P., R. G. Mercer, D. E. Watton, C. B. Buckley and A. S. Lang (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol. Microbiol.* **85**(2): 314-325.

Imhoff, J. F. and U. Bias-Imhoff (1995). Lipids, quinones and fatty acids of anoxygenic phototrophic bacteria. In *Anoxygenic Photosynthetic Bacteria*. R. E. Blankenship, M. T. Madigan and C. E. Bauer (Eds.). Dordrecht, Kluwer Academic Publishers: pp. 179-205.

Iniesta, A. A., P. T. McGrath, A. Reisenauer, H. H. McAdams and L. Shapiro (2006). A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc. Natl. Acad. Sci. USA* **103**(29): 10935-10940.

Jacobs, C., N. Ausmees, S. J. Cordwell, L. Shapiro and M. T. Laub (2003). Functions of the CckA histidine kinase in *Caulobacter* cell cycle control. *Mol. Microbiol.* **47**(5): 1279-1290.

Jacobs, C., I. J. Domian, J. R. Maddock and L. Shapiro (1999). Cell cycle-dependent polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division. *Cell* **97**: 111-120.

Jacobs, C., D. Hung and L. Shapiro (2001). Dynamic localization of a cytoplasmic signal transduction response regulator controls morphogenesis during the *Caulobacter* cell cycle. *Proc. Natl. Acad. Sci. USA* **98**(7): 4095-4100.

Jensen, R. B., S. C. Wang and L. Shapiro (2002). Dynamic localization of proteins and DNA during a bacterial cell cycle. *Nat. Rev. Mol. Cell. Biol.* **3**(3): 167-176.

Kenney, L. J. (2002). Structure/function relationships in OmpR and other winged-helix transcription factors. *Curr. Opin. Microbiol.* **5**(2): 135-141.

Kloesges, T., O. Popa, W. Martin and T. Dagan (2011). Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Mol. Biol. Evol.* **28**(2): 1057-1074.

Krell, T., J. Lacal, A. Busch, H. Silva-Jimenez, M.-E. Guazzaroni and J. L. Ramos (2010). Bacterial sensor kinases: Diversity in the recognition of environmental signals. *Annu. Rev. Microbiol.* **64**(1): 539-559.

Kristensen, D. M., A. R. Mushegian, V. V. Dolja and E. V. Koonin (2010). New dimensions of the virus world discovered through metagenomics. *Trends Microbiol.* **18**: 11-19.

Krupovic, M., D. Prangishvili, R. W. Hendrix and D. H. Bamford (2011). Genomics of Bacterial and Archaeal viruses: Dynamics within the prokaryotic virosphere. *Microbiol. Mol. Biol. Rev.* **75**(4): 610-635.

Lang, A. S. and J. T. Beatty (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. USA* **97**: 859-864.

Lang, A. S. and J. T. Beatty (2001). The gene transfer agent of *Rhodobacter capsulatus* and "constitutive transduction" in prokaryotes. *Arch. Microbiol.* **175**: 241-249.

Lang, A. S. and J. T. Beatty (2002). A bacterial signal transduction system controls genetic exchange and motility. *J. Bacteriol.* **184**: 913-918.

Lang, A. S. and J. T. Beatty (2007). Importance of widespread gene transfer agent genes in α-proteobacteria. *Trends Microbiol.* **15**(2): 54-62.

Lang, A. S., O. Zhaxybayeva and J. T. Beatty (2012). Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Micro.* **10**(7): 472-482.

Laub, M. T., E. G. Biondi and J. M. Skerker (2007). Phosphotransfer profiling: Systematic mapping of two-component signal transduction pathways and phosphorelays. *Methods in Enzymology- Two-Component Signaling Systems Part B.* M. I. Simon, B. R. Crane and A. Crane (Eds.). Pasadena, California, Academic Press. **Volume 423:** 531-548.

Laub, M. T., S. L. Chen, L. Shapiro and H. H. McAdams (2002). Genes directly controlled by CtrA, a master regulator of the *Caulobacter* cell cycle. *Proc. Natl. Acad. Sci. USA* **99**(7): 4632-4637.

Laub, M. T. and M. Goulian (2007). Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.* **41**(1): 121-145.

Lederberg, J. and E. L. Tatum (1946). Gene recombination in *Escherichia coli*. *Nature* **158**: 558.

Lee, E.-J., Y.-H. Cho, H.-S. Kim, B.-E. Ahn and J.-H. Roe (2004). Regulation of σ^B by an anti- and an anti-anti-sigma factor in *Streptomyces coelicolor* in response to osmotic stress. *J. Bacteriol.* **186**(24): 8490-8498.

Leung, M. M., C. A. Brimacombe, G. B. Spiegelman and J. T. Beatty (2012). The GtaR protein negatively regulates transcription of the *gtaRI* operon and modulates gene transfer agent (RcGTA) expression in *Rhodobacter capsulatus*. *Mol. Microbiol.* **83**(4): 759-774.

Little, J. W. (1984). Autodigestion of lexA and phage lambda repressors. *Proc. Natl. Acad. Sci. USA* **81**(5): 1375-1379.

Lonetto, M., M. Gribskov and C. A. Gross (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**(12): 3843-3849.

Madigan, M. T. (1995). Microbiology of nitrogen fixation by anoxygenic photosynthetic bacteria. In *Anoxygenic photosynthetic bacteria*. C. E. Bauer (Ed). Boston, Kluwer Academic Publishers. **2:** 915-928.

Madigan, M. T. and D. O. Jung (2009). An overview of purple bacteria: systematics, physiology, and habitats. In *The Purple Phototrophic Bacteria*. C. N. Hunter, F. Daldal, M. C. Thurnauer and J. T. Beatty (Eds.). New York, Springer**:** pp. 1-15.

Marrs, B. L. (1974). Genetic recombination in *Rhodopseudomonas capsulata*. *Proc. Natl. Acad. Sci. USA* **71**: 971-973.

Mascher, T., J. D. Helmann and G. Unden (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* **70**(4): 910-938.

Maurer, R., B. J. Meyer and M. Ptashne (1980). Gene regulation at the right operator (OR) of bacteriophage λ: I. OR3 and autogenous negative control by repressor. *J. Mol. Biol.* **139**(2): 147-161.

McDaniel, L. D., E. Young, J. Delaney, F. Ruhnau, K. B. Ritchie and J. H. Paul (2010). High frequency of horizontal gene transfer in the oceans. *Science* **330**(6000): 50.

Merrick, M. J. (1993). In a class of its own — the RNA polymerase sigma factor σ^{54} (σ^{N}). *Mol. Microbiol.* **10**(5): 903-909.

Meyer, B. J., R. Maurer and M. Ptashne (1980). Gene regulation at the right operator (OR) of bacteriophage λ: II. OR1, OR2, and OR3: Their roles in mediating the effects of repressor and *cro*." *J. Mol. Biol.* **139**(2): 163-194.

Mooney, R. A., S. A. Darst and R. Landick (2005). Sigma and RNA Polymerase: An on-again, off-again relationship? *Mol. Cell* **20**(3): 335-345.

Nakamura, Y., T. Itoh, H. Matsuda and T. Gojobori (2004). Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nature Genet.* **36**: 760-766.

Ochman, H., J. G. Lawrence and E. A. Groisman (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**(6784): 299-304.

Olendzenski, L. and J. P. Gogarten (2009). Evolution of genes and organisms: the tree/web of life in light of horizontal gene transfer. *Ann. NY Acad. Sci.* **1178**: 137-145.

Osterberg, S., T. d. Peso-Santos and V. Shingler (2011). Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.* **65**(1): 37-55.

Paget, M. and J. Helmann (2003). The sigma70 family of sigma factors. *Genome Biol.* **4**(1): 203.

Pal, C., M. D. Macia, A. Oliver, I. Schachar and A. Buckling (2007). Co-evolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**(7172): 1079-1081.

Paolozzi, L. and P. Ghelardini (2006). The bacteriophage Mu. In *The bacteriophages*. R. Calendar (Ed). New York, Oxford University Press: pp. 469–496.

Pemberton, J. M., I. M. Horne and A. G. McEwan (1998). Regulation of photosynthetic gene expression in purple bacteria. *Microbiol.* **144**(2): 267-278.

Price, C. W. (2010). General stress response in *Bacillus subtilis* and related gram-positive bacteria. In *Bacterial Stress Response*. G. Storz and R. Hengge (Eds). Washington, D.C., ASM Press. pp. 179–197.

Quon, K. C., G. T. Marczynski and L. Shapiro (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84**: 83-93.

Quon, K. C., B. Yang, I. J. Domian, L. Shapiro and G. T. Marczynski (1998). Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. *Proc. Natl. Acad. Sci. USA* **95**(1): 120-125.

Rapp, B. J. and J. D. Wall (1987). Genetic transfer in *Desulfovibrio desulfuricans*. *Proc. Natl. Acad. Sci. USA* **84**: 9128-9130.

Rodriguez-Valera, F., A.-B. Martin-Cuadrado, B. Rodriguez-Brito, L. Pasic, T. F. Thingstad, F. Rohwer and A. Mira (2009). Explaining microbial population genomics through phage predation. *Nat. Rev. Micro.* 7(11): 828-836.

Rohwer, F. and R. Vega Thurber (2009). Viruses manipulate the marine environment. *Nature* **459**: 207-212.

Rosario, K. and M. Breitbart (2011). Exploring the viral world through metagenomics. *Curr. Opin. Virol.* **1**: 289-297.

Schaefer, A. L., T. A. Taylor, J. T. Beatty and E. P. Greenberg (2002). Long-chain acyl-homoserine lactone quorum-sensing regulation of *Rhodobacter capsulatus* gene transfer agent production. *J. Bacteriol.* **184**: 6515-6521.

Sciochetti, S. A., N. Ohta and A. Newton (2005). The role of polar localization in the function of an essential *Caulobacter crescentus* tyrosine kinase. *Mol. Microbiol.***56**(6): 1467-1480.

Sevcikova, B., B. Rezuchova, D. Homerova and J. Kormanec (2011). The anti-anti-sigma factor BldG is involved in activation of the stress response sigma factor σ^H in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **192**(21): 5674-5681.

Severinova, E., K. Severinov and S. A. Darst (1998). Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T4 AsiA. *J. Mol. Biol.* **279**(1): 9-18.

Sharma, A. K., A. C. Rigby and S. L. Alper (2011). STAS domain structure and function. *Cell. Physiol. Biochem.* **28**(3): 407-422.

Solioz, M. and B. Marrs (1977). The gene transfer agent of *Rhodopseudomonas capsulata*: Purification and characterization of its nucleic acid. *Arch. Biochem. Biophys.* **181**: 300-307.

Solioz, M., H. C. Yen and B. Marrs (1975). Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J. Bacteriol.* **123**: 651-657.

Stanton, T. B. (2007). Prophage-like gene transfer agents: Novel mechanisms of gene exchange for *Methanococcus*, *Desulfovibrio*, *Brachyspira*, and *Rhodobacter* species. *Anaerobe* **13**(2): 43-49.

Staroń, A. and T. Mascher (2010). General stress response in α-proteobacteria: PhyR and beyond. *Mol. Microbiol.* **78**(2): 271-277.

Staroń, A., H. J. Sofia, S. Dietrich, L. E. Ulrich, H. Liesegang and T. Mascher (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Mol. Microbiol.* **74**(3): 557-581.

Stock, A. M., V. L. Robinson and P. N. Goudreau (2000). Two-component signal transduction. *Annu. Rev. Biochem.* **69**(4): 183 - 215.

Strnad, H., A. Lapidus, J. Paces, P. Ulbrich, C. Vlcek, V. Paces and R. Haselkorn (2010). Complete genome sequence of the photosynthetic purple non-sulfur bacterium *Rhodobacter capsulatus* SB 1003. *J. Bacteriol.* **192**(13): 3545-3546.

Suttle, C. A. (2005). Viruses in the sea. *Nature* **437**(7057): 356-361.

Suttle, C. A. (2007). Marine viruses - major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**(10): 801-812.

Taylor, T. A. (2004). Evolution and regulation of the gene transfer agent (GTA) of *Rhodobacter capsulatus*. M.Sc. Thesis, University of British Columbia, Vancouver, British Columbia

Thomas, C. M. and K. M. Nielsen (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3**: 711-721.

Wall, J. D., P. F. Weaver and H. Gest (1975). Gene transfer agents, bacteriophages, and bacteriocins of *Rhodopseudomonas capsulata*. *Arch. Microbiol.* **105**: 217-224.

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* **28**(2): 127-181.

Whitworth, D. E. and P. J. A. Cock (2008). Two-component systems of the myxobacteria: structure, diversity and evolutionary relationships. *Microbiol.* **154**(2): 360-372.

Whitworth, D. E. and P. J. A. Cock (2009). Evolution of prokaryotic two-component systems: insights from comparative genomics. *Amino Acids* **37**(3): 459-466.

Winter, C., A. Smit, G. J. Herndl and M. G. Weinbauer (2004). Impact of virioplankton on Archaeal and Bacterial community richness as assessed in seawater batch cultures. *Appl. Environ. Microbiol.* **70**(2): 804-813.

Yen, H.-C., N. T. Hu and B. L. Marrs (1979). Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. *J. Mol. Biol.* **131**: 157-168.

Zhaxybayeva, O., J. P. Gogarten, R. L. Charlebois, W. F. Doolittle and R. T. Papke (2006). Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. *Genome Res.* **16**: 1099-1108.

Zinder, N. D. and J. Lederberg (1952). Genetic exchange in *Salmonella*. *J. Bacteriol.* **64**: 679-699.

.

Co-authorship Statement

Chapter 2 is a version of a manuscript published in the Journal of Bacteriology (Mercer, R.G., S.J. Callister, M.S. Lipton, L. Pasa-Tolic, H. Strnad, V. Paces, J.T. Beatty, and A.S. Lang (2010) Loss of the response regulator CtrA causes pleiotropic effects on gene expression but does not affect growth phase regulation in *Rhodobacter capsulatus*. Journal of Bacteriology 192: 2701-2710; DOI: 10.1128/JB.00160-10). Contributions to this chapter were made by collaborators from the United States, the Czech Republic and Canada. The basic concept of the project was designed by A.S. Lang, J.T. Beatty and S.J. Callister. The draft *R. capsulatus* genome used to create the microarrays was prepared by H. Strnad and V. Paces. The design of the research was carried out by A.S. Lang, S.J. Callister and myself. I prepared strains and cells for both transcriptome and proteome analyses. The proteomic work on prepared cells was performed by S.J. Callister, M.S. Lipton, L. Pasa-Tolic. I carried out the RNA isolation and preparation for the microarrays, as well as all transcriptomic data analysis and proteomic data analysis with the assistance of S.J. Callister. The manuscript was drafted and prepared by myself and A.S. Lang with subsequent editorial input from the other co-authors; the materials and methods for the proteomic section were prepared by S.J. Callister.

Chapter 3 is a version of a manuscript published in FEMS Microbiology Letters in 2012 (Mercer, R.G., M. Quinlan, A.R. Rose, S. Noll, J.T. Beatty, and A.S. Lang (2012) Regulatory systems controlling motility and gene transfer agent production and release in *Rhodobacter capsulatus*. FEMS Microbiology Letters 331: 53-62; DOI: 10.1111/j.1574-6968.2012.02553.x). Research in this chapter was proposed and designed by myself and

A.S. Lang. I constructed the *ctrA*, *sciP*, *ctrA/sciP* strains and complementation vectors for all mutants except *chpT*. The *chpT* mutant and cloning work for genetic complementation was carried out by M. Quinlan and A.R. Rose. The complete *cckA* knockout strain was constructed by S. Noll. I created the vectors containing the mutagenized *ctrAD51E* and *ctrAD51A* genes. I carried out all western blots, motility assays and viable cell counts. I performed the data analysis and prepared the manuscript with A.S. Lang, with subsequent editorial input from the other co-authors.

Chapter 4 is a manuscript in preparation for submission. The research was designed by myself with intellectual input from A.S. Lang. I carried out all mutant construction, experiments, practical aspects of the work, and data analysis. The manuscript was written by myself with editorial assistance from A.S. Lang.

Chapter 2 - Loss of the response regulator CtrA causes pleiotropic effects on gene expression in *R. capsulatus* but does not affect growth phase regulation

2.1. Introduction

The purple non-sulfur bacterium *Rhodobacter capsulatus* is a model organism for the study of various aspects of bacterial physiology, such as bioenergetics and N_2-fixation. It also engages in an unusual mechanism of genetic exchange, carried out by a bacteriophage-like element called the gene transfer agent (RcGTA) (Marrs, 1974; Yen *et al.*, 1979). The production of RcGTA is maximal in the stationary phase of growth of *R. capsulatus* cultures (Solioz *et al.*, 1975) and is regulated by at least 2 distinct signalling systems, one through quorum sensing of a long chain acyl-homoserine lactone (Schaefer *et al.*, 2002), and the other involving the response regulator protein CtrA (Lang & Beatty, 2000).

The CtrA protein was first characterized in *Caulobacter crescentus* (Quon *et al.*, 1996) where it is essential for viability and acts as a master regulator of the cell cycle (Skerker & Laub, 2004), controlling at least 25% (144 of 553) of the genes involved in cell cycle progression (Laub *et al.*, 2002). Despite sharing remarkable sequence identity (71%) with the CtrA protein from *C. crescentus*, the *R. capsulatus* protein has a very different role because it is not essential and does not appear to be involved in cell cycle processes. One function of CtrA in common between the two species is the regulation of expression of genes that encode the flagellum (Lang & Beatty, 2002; Quon *et al.*, 1996).

The *ctrA* genes of *Sinorhizobium meliloti* (Barnett *et al.*, 2001), *Brucella abortus* (Bellefontaine *et al.*, 2002), and *Ruegeria* sp. TM1040 (Miller & Belas, 2006) have also been studied. Similarly to *C. crescentus* and *R. capsulatus*, the *Ruegeria* CtrA controls motility (Miller & Belas, 2006). A search of the GenBank database reveals that convincing homologues which share >50% identity with the *C. crescentus* protein are present in all α-proteobacterial complete genome sequences, with the exception of *Pelagibacter ubique* (Giovannoni *et al.*, 2005).

A portion of the *R. capsulatus* SB1003 genome sequence was described previously (Haselkorn *et al.*, 2001; Vlcek *et al.*, 1997) and the complete annotated sequence is now available (NCBI Accession - NC_014034). The availability of this sequence has allowed me to identify the genes dysregulated by loss of CtrA through comparisons of transcriptomic and proteomic data from wild type and *ctrA* mutant cell cultures. Because RcGTA production is CtrA-dependent and changes over a culture growth phase, measurements of the transcriptome and proteome of cultures in the logarithmic and stationary phases of the population growth cycle were compared. Therefore, I have also analyzed growth phase differences in gene expression in *R. capsulatus*. The transcriptome results show that CtrA is an important regulator in *R. capsulatus* because it is required for proper expression of more than 225 genes, including those predicted to control motility, gene exchange, pilus and gas vesicle formation, as well as expression of many putative signal transduction proteins and transcriptional regulators. Of the genes dysregulated by the loss of CtrA, proteins for 58 were observed and 51 of these quantitatively validated the transcriptome data (88%). My analyses also

demonstrate that CtrA is not responsible for growth phase dependent gene expression in *R. capsulatus*, and that CtrA does not control *R. capsulatus* genes involved in cell cycle events such as DNA replication or cell division. Despite almost universal conservation of CtrA in the α-proteobacteria, a bioinformatic analysis of representative organisms suggests that CtrA has different functions that fall along phylogenetic lines in the different major taxonomic orders.

2.2. Materials and Methods

2.2.1. Bacterial strains and growth conditions

Cultures of *R. capsulatus* were grown anaerobically in YPS medium (Wall *et al.*, 1975) at 37°C with illumination from standard 60W incandescent light bulbs. YPS is a complex medium and was chosen because of its use in previous studies on CtrA and RcGTA (Lang & Beatty, 2000; Lang & Beatty, 2002; Schaefer *et al.*, 2002). Strains used were the genome-sequenced SB1003 (Yen & Marrs, 1976) and a *ctrA* mutant. The mutant was constructed from SB1003 by GTA transduction of a disrupted version of *ctrA* (Lang & Beatty, 2002) into the chromosome. Growth was monitored over time by turbidity, and samples were collected in the logarithmic growth phase and early in the stationary phase of growth (Figure 2.1).

Iron supplementation experiments were conducted by preparing YPS medium with the addition of the salt and trace element components of RCV medium [EDTA, $MgSO_4$, $CaCl_2$, $FeSO_4$, thiamine HCl, $MnSO_4$, H_3BO_3, $Cu(NO_3)_2$, $ZnSO_4$, and $NaMoO_4$] (Beatty & Gest, 1981), and also with the same components except for $FeSO_4$. All of these

were added together to prevent the iron from precipitating, which happens if $FeSO_4$ is added on its own. This addition increased the Fe^{2+} concentration from ~7 μM in YPS (7) to ~50 μM.

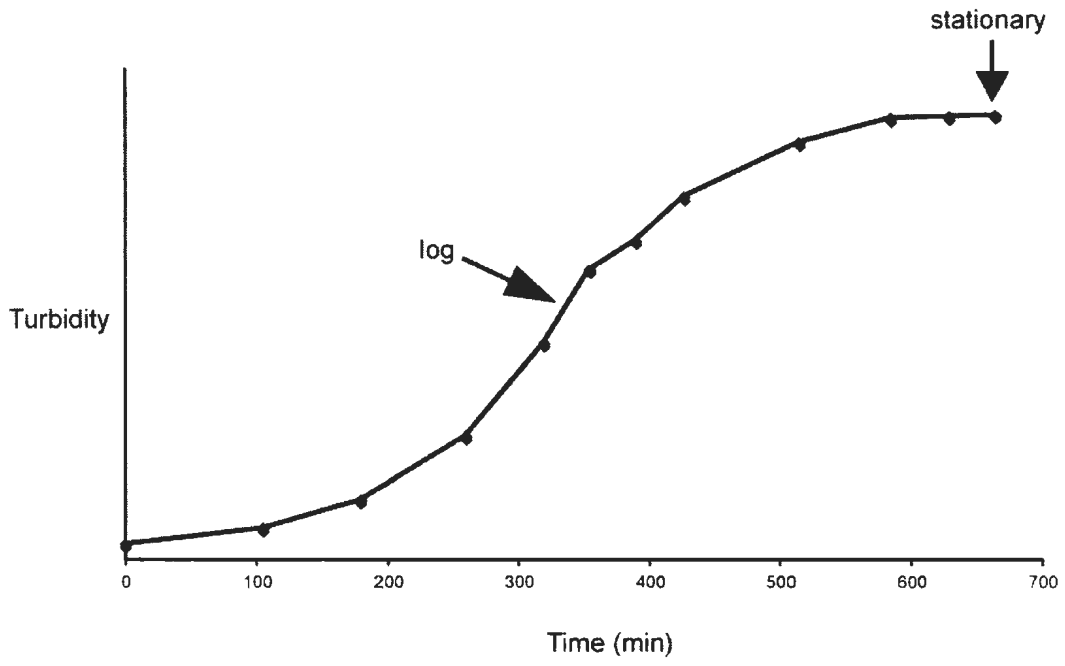


Figure 2.1. Growth curve of a representative experimental *R. capsulatus* culture. The time-points where cells were harvested for analyses are indicated.

2.2.2. RNA isolation and microarray analyses

RNA was extracted using the RNeasy Kit (Qiagen Inc., Mississauga, Canada) according to the manufacturer's recommendations. RNA was isolated from both strains in both growth phases from three growth experiments on different days, yielding 12 total samples. Whole-genome expression 100-3660 arrays were constructed by Affymetrix Inc. (Santa Clara, CA) through a custom design. The arrays were designed based on a

preliminary version of the complete genome sequence and contained oligonucleotide probes for 3635 open reading frames (ORFs) and 1452 intergenic regions greater than 90 bp; with each ORF represented by 11 probe pairs. RNA samples and arrays were processed at the Michael Smith Genome Science Centre (Vancouver, Canada) with cDNA synthesis, labeling and target hybridization performed as described in the Affymetrix Expression Analysis Technical Manual for prokaryotic samples.

Raw data from scanned arrays were Robust Multi-Array (RMA) normalized (Irizarry *et al.*, 2003). Using GeneSpring v7.2 (Agilent Technologies, Santa Clara, CA), the signal intensities were further normalized (50th percentile per chip) so that the median normalized signal intensity equals 1. Subsequent data visualization and analyses were performed using GeneSpring v7.2. Data filtering was set at a raw signal intensity >50 to remove low signal intensity features. Differential expression thresholds were set at a 2-fold limit, and gene lists were created by selecting those that differed by ≥2-fold for normalized signal intensity in one sample relative to the other within an individual replicate of RNA samples, and subsequent identification of the genes in common to 3 of 3 replicate experiments.

The relationship between the probe identification based on the preliminary genome annotation used to design the microarrays and the final genome sequence (NCBI Accession - NC_014034) is provided in Table S2.1.

2.2.3. Protein extraction and digestion

Global, soluble, and insoluble protein fractions were extracted from cell lysates using established protocols (Callister *et al.*, 2006). Proteins from the global and soluble

preparations were denatured and reduced by adding urea, thiourea and DTT to final concentrations of 7 M, 2 M, and 5 mM respectively. Samples were incubated at 60°C for 30 minutes, then diluted 10-fold with 100 mM NH_4HCO_3, pH 8.4. $CaCl_2$ was added to the diluted sample to a final concentration of 1 mM, and the sample was digested at 37°C using sequencing grade trypsin (Roche, Indianapolis, IN) at a ratio of 1 unit per 50 μg protein. Digested samples were then de-salted using a C-18 Solid Phase Extraction (SPE) column (Supelco, St. Louis, MO).

For the insoluble protein extract, the cell lysate was ultracentrifuged at 350000 × *g* for 10 minutes at 4°C. The resulting supernatant, containing soluble proteins was digested as described above. The insoluble pellet was washed in 100 mM NH_4HCO_3, pH 7.8 and suspended in 7 M urea, 2 M thiourea, 1% CHAPS in 50 mM NH_4HCO_3, pH 7.8. DTT was then added to a final concentration of 9.7 mM. The sample was incubated and digested as described above with the exception that 50 mM NH_4HCO_3, pH 7.8 was used for the 10-fold dilution. Following digestion, removal of salts and detergent was performed using a Strong Cation Exchange (SCX) column (Supelco, St. Louis, MO).

2.2.4. Tandem LC-MS analysis and reference peptide database generation

Peptides from each protein digest were fractionated (30 fractions each digest) using SCX and reverse-phase HPLC according to established protocols (Adkins *et al.*, 2006). Peptides from each fraction (10 μg) were analyzed on a quad column HPLC system coupled to a LTQ mass spectrometer (ThermoFisher Scientific, San Jose, CA). Reverse-phase separation of peptide digests occurred by way of columns (60 cm x 360 μm o.d. x 75 μm i.d. fused silica capillary tubing; Pacific Northwest National Laboratory,

Richland, WA) packed with 3 μm Jupiter C_{18} stationary phase (Phenomenex, Torrence, CA). The HPLC system was equilibrated with 100% mobile phase A (0.2% acetic acid and 0.05% TFA in water) at 10 kpsi. Mobile phase B (0.1% TFA in 90% acetonitrile/10% water) displaced mobile phase A 50 min after peptide injection, generating an exponential gradient. Split flow controlled the gradient speed operating under constant pressure (10 kpsi). Separated peptides were ionized (positive) using an ESI interface (Pacific Northwest National Laboratory) with chemically etched electrospray emitters (150 mm o.d. x 20 mm i.d) (Kelly *et al.*, 2006). The capillary temperature and ESI voltage were 200°C and 2.2 kV, respectively.

Selected parent ion peaks were fractionated using collision induced dissociation and tandem mass spectra were matched to theoretical spectra, for peptide sequence assignment, using the sequence search algorithm X!Tandem (Robertson & Beavis, 2004). A modified parameter file allowed for partial tryptic peptides and static modifications to pass the first round of X!Tandem matching. Theoretical spectra were generated from the translated *R. capsulatus* SB1003 genome sequence. Peptides of residue length 6 amino acids or greater, having an E-value $\leq -2(\log_{10})$ were placed in a reference peptide database along with their calculated theoretical masses and normalized elution times (Kiebel *et al.*, 2006) for matching to LC-MS measurements.

2.2.5. High mass accuracy LC-MS and label free quantification

A modified 9.4 tesla FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) was used to obtain relative quantitative proteome information. HPLC conditions were the same as reported above. Mass spectra from quadruplicate measurements (16 total

analyses) were processed with in-house developed software Decon2LS (Mayampurath *et al.*, 2008) based on the THRASH algorithm. The resulting mono-isotopic masses were clustered into LC-MS features based on the neutral mass, charge state, abundance, isotopic fit (i.e., quality of fit between recorded and simulated isotopic pattern), and spectrum number (relating to LC retention time). The assembled set of LC-MS features was then searched against the reference peptide database using VIPER (Monroe *et al.*, 2007). A tolerance window of mass measurement accuracy <5 ppm and normalized elution time error <1% was applied to ensure reliable peptide identification.

MS-peak intensities were used as a measure of the relative peptide abundances. The mean abundance of the LC-MS features was used and the relative abundances of constituent peptides were averaged to derive the relative abundance of the parent protein (Smith *et al.*, 2002). Peptide abundances were normalized to a common baseline using central tendency normalization (Callister *et al.*, 2006) and 'rolled up' to a protein abundance estimate using the Z-score rollup algorithm available in DAnTE (Polpitiya *et al.*, 2008). Software tools used for this proteomics analysis are publicly available at http://ncrr.pnl.gov/.

2.2.6. Identification of potential CtrA binding sites

Consensus CtrA full sites (TTAA-N7-TTAAC) and half-sites (TTAACCAT) were identified using the sequence search function in Artemis (Rutherford *et al.*, 2000).

2.2.7. Transcriptomic and proteomic data

Transcriptomic data from this work have been deposited in the NCBI GEO database (accession no. GSE18149); Proteomic data are available at http://omics.pnl.gov/.

2.3. Results

2.3.1. Nucleotide sequence of the *Rhodobacter capsulatus* genome

The genome of *R. capsulatus* strain SB1003 consists of a circular chromosome containing 3738958 bp and a circular plasmid of 132962 bp. The genome has a relatively high GC content (66.6%). The full genome annotation contains 3531 ORFs identified in the chromosome and 154 ORFs in the plasmid. Functions are assigned to 3100 ORFs (84.1%). The annotation is accessible at http://rhodo.img.cas.cz and the NCBI database (accession no: NC_014034).

2.3.2. Growth phase expression analyses

Transcriptomic and proteomic data were gathered during the logarithmic (log) and stationary phases of the growth of SB1003 cultures for comparative analyses. The proteomic analysis identified peptides corresponding to 1213 ORFs. The transcript comparisons identified 334 genes in the genome (9%) that show ≥2-fold increase or decrease in expression between the log and stationary phases of growth of SB1003. Of these, quantitative peptide data were obtained for 152 proteins (46%) in the proteome. Transcript levels were ≥2-fold higher (range 2 to 20-fold) in the log phase of growth relative to the stationary phase for 143 genes (Table S2.2). These genes are predicted to be involved in various processes, with most in metabolic categories (Figure 2.2).

Transcripts of genes encoding ribosomal proteins were as much as 5-fold more abundant in the log phase cultures. A number of flagellar and chemotaxis genes also are more highly expressed in the log phase of growth compared to stationary phase. Of the 143 genes more highly expressed in the log phase, observed peptides identified 78 of the corresponding proteins (55%). The relative abundance patterns of the proteins agree with the transcript data for 50 of the 78 (64%) genes (Table S2.2).

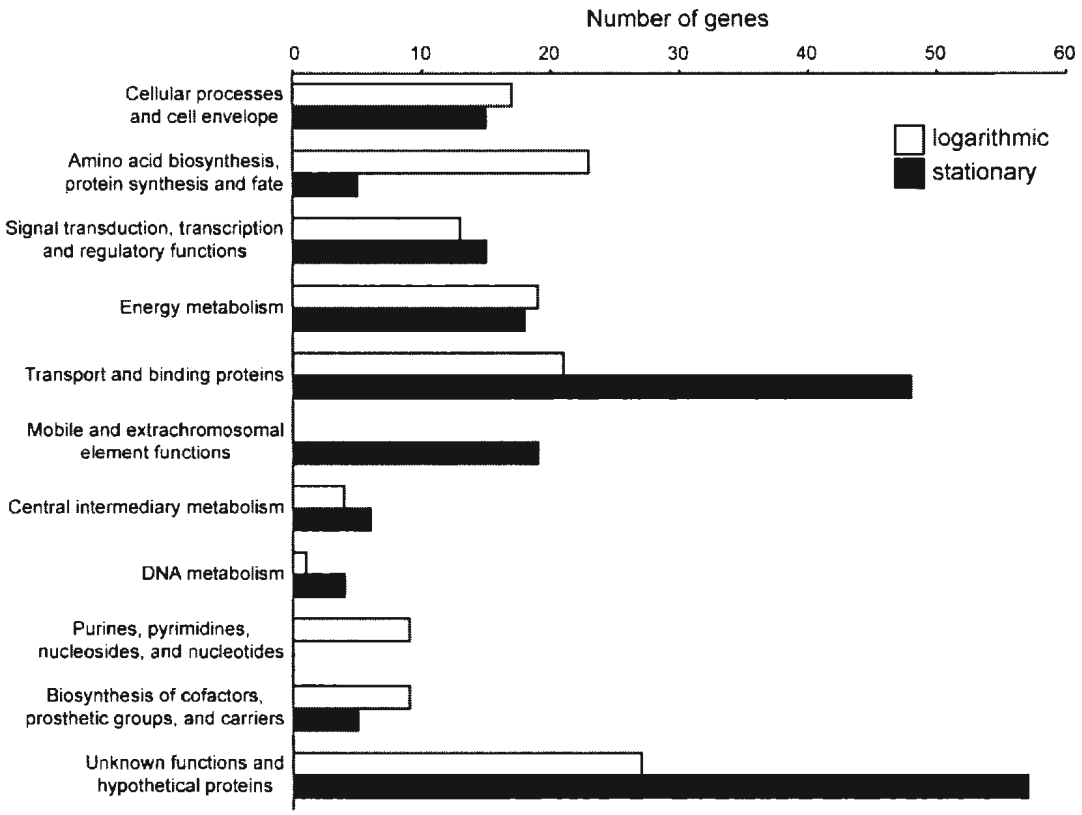


Figure 2.2. Growth phase differences in gene expression in *R. capsulatus*. The numbers of genes that display ≥2-fold higher transcript levels in the logarithmic or stationary phases that fall within each specified gene category are shown. The categories are based on the TIGR roles indicated in the genome annotation;

for simplicity we combined: cellular processes and cell envelope; amino acid biosynthesis, protein synthesis, and protein fate; signal transduction, transcription, and regulatory functions; and unknown functions and hypothetical proteins. The transcript and proteome data for all genes are in Table S2.2 (logarithmic phase) and Table S2.3 (stationary phase).

In the stationary phase, 192 genes showed ≥2-fold higher transcript levels relative to the log phase (Figure 2.2; Table S2.3). The proteome analysis detected 74 of the proteins encoded by the 192 genes more highly expressed in stationary phase (39%), and 66 (89%) showed a qualitative agreement with transcript levels (Table S2.3). There is a noticeable shift in the gene category representation numbers in stationary phase away from genes involved in biosynthetic functions, which is accompanied by expected increases in other categories (Figure 2.2). These data reveal that an adaptation of *R. capsulatus* to stationary phase in this medium is to increase the expression of a large number (48) of transport system genes, some of which were transcribed as much as 60-fold higher than in the log phase (Table S2.3). The transport category in stationary phase includes approximately 20 genes predicted to be involved in iron uptake, suggesting that the cells become iron-limited under these conditions. I conducted growth experiments to test if iron limitation is the cause of the stationary phase under these growth conditions. Increasing the iron concentration in the medium from ~7 μM to ~50 μM resulted in no change in growth rate or increased yield in the cultures (Figure 2.3). Therefore, although the cells do increase expression of iron acquisition genes in the stationary phase, iron limitation is not the cause of the stationary phase in these conditions. There are also large

increases in the mobile element category (from 0 to 19), mostly consisting of the RcGTA genes, and the unknown function and hypothetical category (from 28 to 58; Figure 2.2).

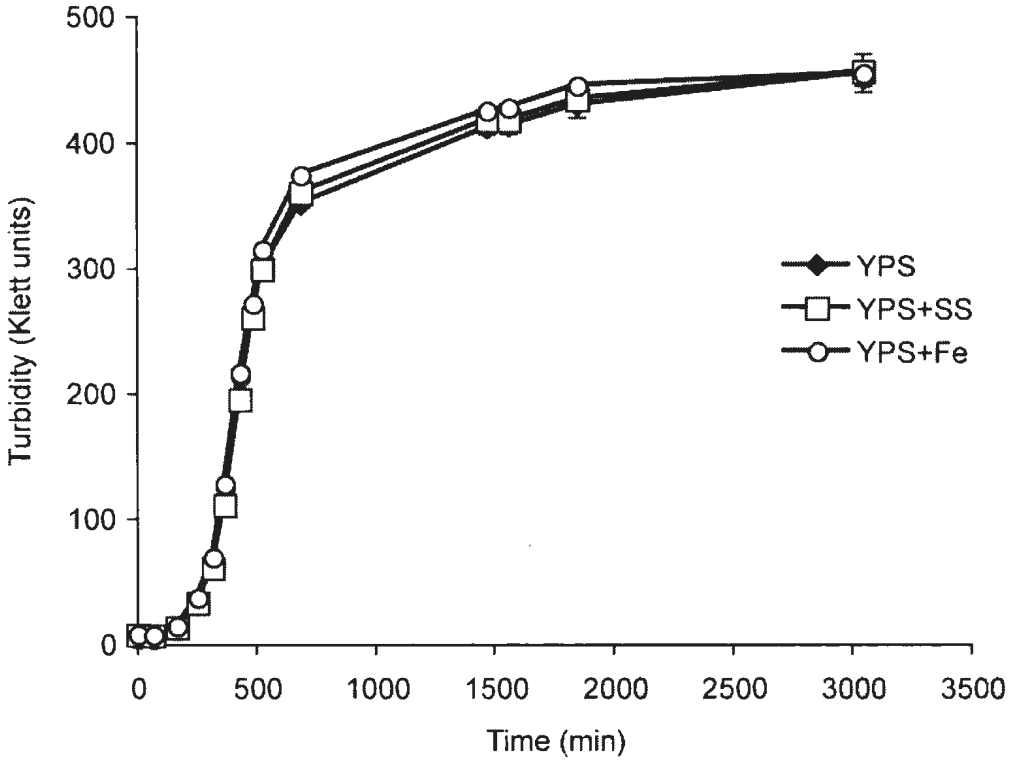


Figure 2.3. Effect of increased iron on *R. capsulatus* photosynthetic growth in YPS medium. *R. capsulatus* SB1003 was grown in YPS medium (YPS; filled diamonds), YPS medium supplemented with EDTA, $MgSO_4$, $CaCl_2$, $FeSO_4$, thiamine HCl, $MnSO_4$, H_3BO_3, $Cu(NO_3)_2$, $ZnSO_4$, and $NaMoO_4$ (YPS+Fe; open circles) and YPS supplemented with the same components except $FeSO_4$ (YPS+SS; open squares). Triplicate cultures were grown in each type of medium and the growth monitored by turbidity using a Klett-Summerson photoelectric colourimeter, and plotted as the mean ± SD.

From the *R. capsulatus* genome, I identified 8 genes that are predicted to encode σ factors (Table 2.1). The transcript levels are increased in the stationary phase relative to the log phase for all of these, and two show >2-fold increased expression in the stationary phase. These two genes, *rcc00458* and *rcc02637*, have 2.6- and 3.7-fold higher transcript levels, respectively. The protein encoded by *rcc00458* is the orthologue (the protein sequences are 72% identical) of the *R. sphaeroides* $RpoH_{II}$ stress response σ factor (Green & Donohue, 2006). A bioinformatics analysis indicates *rcc02637* may encode a protein belonging to the FecI-like group of ECF σ factors (ECF05-ECF10), which may indicate a function in iron acquisition (Staron *et al.*, 2009). Two other ECF σ factors show higher transcript levels in the stationary phase. Gene *rcc02291*, which encodes an EcfG-like stress response σ factor (ECF15) based on a bioinformatics analysis and its genomic context (Staron *et al.*, 2009), and gene *rcc02811* encodes RpoHI, which has been described as a heat shock σ factor (Emetz & Klug, 1998). The gene *rcc02724* is annotated as a σ factor and displays sequence similarity to other genes annotated as σ factors, although a BLAST analysis at the NCBI Conserved Domain Database does not identify any σ factor domains in the sequence. The gene *rcc00568*, encoding RpoN that regulates nitrogen fixation genes (Alias *et al.*, 1989; Jones & Haselkorn, 1989), has extremely low transcript levels under these growth conditions (<0.6 for all samples; Table 2.1) as expected (Cullen *et al.*, 1994) because the cells are not fixing nitrogen in this medium. Two genes, *rcc03323* and *rcc03324*, are predicted to encode an anti-anti-σ and anti-σ pair, and are 3.3- and 2.1-fold higher in the log phase, respectively; these two genes are also affected by the loss of CtrA (see Table 2.2 below).

Table 2.1. Sigma (σ) factor expression in *R. capsulatus*.

Gene	Product and function[a]	Transcript levels (logarithmic, stationary)		Transcript fold change[b]		
		SB1003	*ctrA*	SB1003 vs. *ctrA* (logarithmic)	SB1003 vs. *ctrA* (stationary)	Stationary vs. logarithmic (SB1003)
rcc00458	$RpoH_{II}$, stress response σ factor (Green & Donohue, 2006)	4.60, 12.02	3.81, 14.43	1.2	-1.2	2.6
rcc00568	RpoN, nitrogen fixation σ factor (Alias *et al.*, 1989; Cullen *et al.*, 1994; Jones & Haselkorn, 1989)	0.25, 0.33	0.31, 0.55	-1.2	-1.7	1.3
rcc00699	σ^{24} ECF σ factor	2.61, 4.25	2.64, 4.74	0	-1.1	1.6
rcc02291	EcfG-like σ factor (ECF15), stress response? (Staron *et al.*, 2009)	0.76, 1.29	0.82, 1.97	-1.1	-1.5	1.7
rcc02637	σ^{24} ECF σ factor	6.12, 22.76	8.45, 27.74	-1.4	-1.2	3.7

rcc02724	ECF σ factor (?)	0.35, 0.60	0.48, 0.60	-1.4	0	1.7
rcc02811	RpoHI, heat shock σ factor (Emetz & Klug, 1998)	8.75, 15.67	10.70, 20.88	-1.2	-1.3	1.8
rcc03054	RpoD, major vegetative σ factor (Pasternak *et al.*, 1996)	7.80, 10.80	9.35, 12.05	-1.2	-1.1	1.4

[a]The potential functions are discussed further in the text

[b]Positive values indicate higher transcript levels in the wild type and negative values indicate higher levels in the *ctrA* mutant

For a different perspective on the growth phase gene expression patterns, I also identified the 100 most highly transcribed genes, excluding rRNA genes, in both of the growth phases (Table S2.4). These represent the upper ~3% of genes in terms of the magnitude of hybridization signal. There is considerable overlap between the growth phases, with 57 genes shared between phases, and many of these genes are related to photosynthetic growth and electron transfer. The most abundant transcript segment observed in these experiments was *pucB* encoding the light-harvesting (LH) 2 β protein, followed by *pufB* encoding the LH1 β protein. A striking feature of the 86 genes that are not shared between the phases (i.e. 43 from each phase that are not in the top 100 for both phases) is that many of them are highly expressed in both growth phases, even though they were not in the top 100 in both growth phases. Of the 43 genes unique to the log phase top 100, less than half (19 genes; 44%) show ≥2-fold difference in expression between the growth phases, while for the stationary phase, only slightly more than half (23 genes; 53%) show ≥2-fold differences. Therefore, under the conditions investigated here, much of the stationary phase adaptation occurs through genes that are not the most highly expressed. The *ctrA* gene was among the top 100 genes in both phases, further supporting its role as an important regulator in *R. capsulatus*, as described below.

2.3.3. CtrA is a positive regulator of gene expression

The global regulatory activity of CtrA in *R. capsulatus* was determined using paired transcriptome and proteome measurements. Comparisons were made between wild type and *ctrA* mutant strains, and between the log and stationary phases of culture growth. These data demonstrate that regardless of the phase of growth, CtrA is almost exclusively

a positive regulator of gene expression. In the log phase, 172 genes had ≥2-fold lower transcript levels in the *ctrA* mutant, and 191 transcripts were ≥2-fold less abundant in the mutant in stationary phase (Figure 2.4). In these two groups of genes, 147 are shared between the growth phases, and therefore the total number of genes positively regulated by CtrA in the conditions investigated here is 216 (~6% of the genome). In contrast, only 11 genes show transcript amounts ≥2-fold greater in the *ctrA* mutant (Figure 2.4), accounting for <5% of regulated genes.

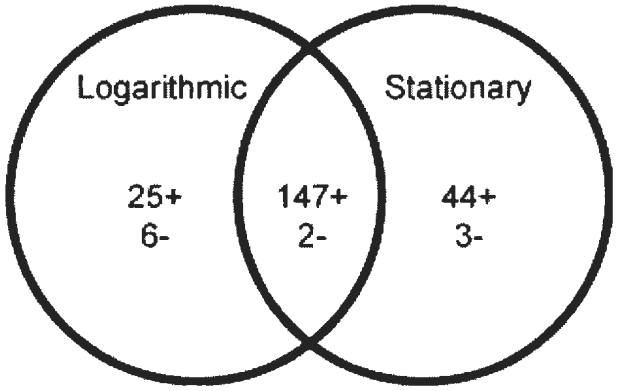


Figure 2.4. The effects of loss of CtrA as defined by whole genome transcript analyses. The Venn diagram shows the number of gene transcripts that were ≥2-fold down-regulated in the *ctrA* mutant compared to SB1003 (+) and the number of gene transcripts that were ≥2-fold up-regulated in the *ctrA* mutant compared to SB1003 (-). The numbers of genes found in the logarithmic and stationary growth phases and the amount of overlap between the two lists are indicated. The included genes were identified in 3/3 independent replicate experiments.

The relative gene expression data is presented as heat maps for the 216 genes negatively affected (Figure 2.5A) and 11 genes positively affected (Figure 2.5B) by the

loss of CtrA is presented as a heat map (the complete numerical transcript and proteome data for these genes is available in Table S2.5). The genes are organized into two groups in each growth phase based on their transcript levels in the *ctrA* mutant (Figure 2.5). I chose to divide the genes based on an intensity value of 1 because an intensity >1 indicates the gene is above the median observed for all genes in the genome. For the genes that are negatively affected by the loss of CtrA, group 1 contains genes that have a signal intensity <1 in the mutant, while those within group 2 have an intensity >1 in the mutant and therefore appear to be transcribed even in the absence of CtrA (Figure 2.5A). Almost all of these genes fall within the same cluster in both phases (Table S2.5). For the genes that are positively affected by the loss of CtrA, group 1 contains genes that have a normalized signal intensity <1 in the wild type, while those within group 2 have an intensity >1 in the wild type and therefore appear to be transcribed even in the presence of CtrA (Figure 2.5B). Unlike the trend observed for the negatively affected genes, only 5 of these 11 positively affected genes are in the same cluster in the different growth phases (Table S2.5). Overall, the range of transcript levels observed in the 227 genes in the absence of CtrA likely indicate that some of these genes are subject to additional mechanisms of regulation, and some may be only indirectly affected by the loss of CtrA.

Transcript levels of the *ctrA* gene were high in both growth phases for the wild type cells and there was an increase in the level of *ctrA* transcripts in the stationary phase (Table S2.4), in agreement with previous observations (Lang & Beatty, 2000). However, a comparison of the proteomes shows a slightly higher amount of CtrA in the log phase

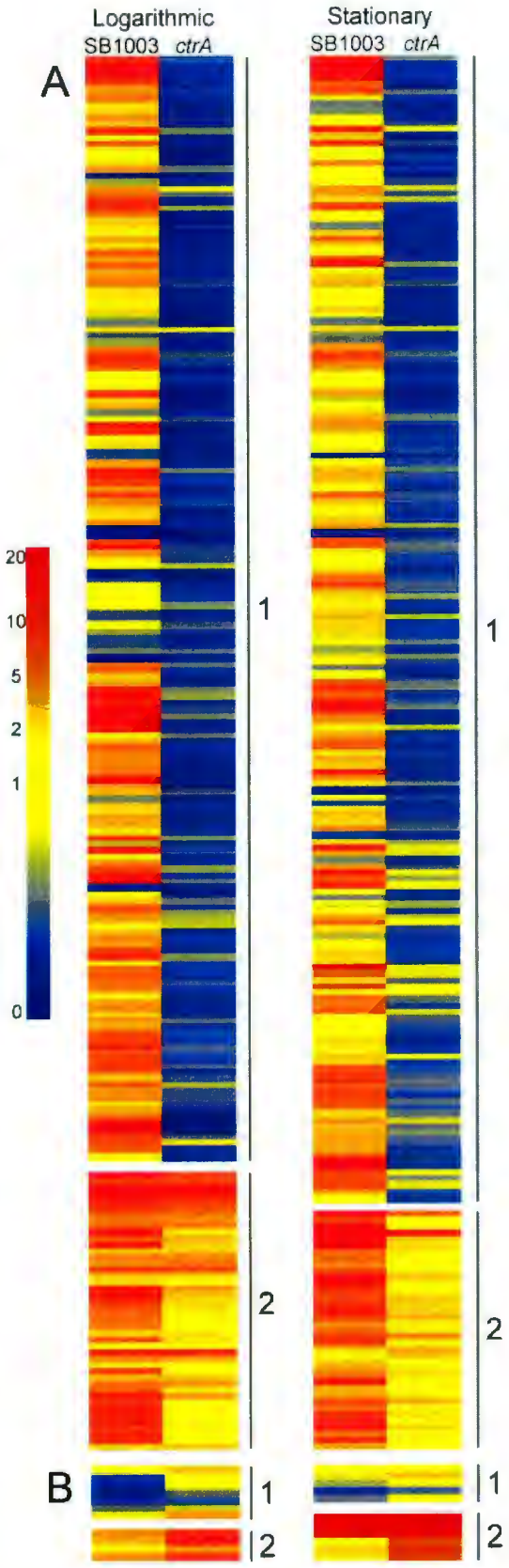


Figure 2.5. The effects of loss of CtrA on gene expression in *R. capsulatus*. The quantitative transcript heat map shows the 216 genes negatively affected ≥2-fold by loss of CtrA (A) and the 11 genes positively affected ≥2-fold by loss of CtrA (B). The genes are organized in 2 groups for each category and growth phase based on their transcript levels in the *ctrA* mutant. For genes negatively affected by loss of CtrA (A), group 1 contains genes that have a signal intensity <1 in the mutant, while those within group 2 have an intensity >1 in the mutant. For genes that are positively affected by the loss of CtrA (B), group 1 contains genes that have a signal intensity <1 in the wild type, while those within group 2 have an intensity >1 in the wild type. Genes are presented in their order of occurrence in the genome (Table S2.1) within each group. The scale is coloured to indicate relative amounts of the transcripts as indicated in the culture samples indicated above the columns. The numerical transcript and proteomic data for all 227 genes is in Table S2.5.

relative to the stationary phase. Additionally, because of the difficulty of detecting phosphorylated peptides, I was unable to detect the region of CtrA predicted to contain the conserved site of phosphorylation (Asp-51), and so I cannot comment on the relative levels of active protein in the two growth phases. Nevertheless, the decrease in transcript levels for 172 genes in the logarithmic phase caused by the loss of *ctrA* demonstrates that the CtrA signalling system does not require a stationary phase signal for activation, in contrast to a previous suggestion (Lang & Beatty, 2000).

2.3.3.1. Regulation of flagellar and chemotaxis genes

Previous work demonstrated that CtrA is required for expression of some flagellar genes in *R. capsulatus* (Lang & Beatty, 2002). The data presented in this chapter allow an investigation of the function of CtrA in the expression of all flagellar and chemotaxis genes. Genes predicted to be involved in flagellar and chemotaxis functions are a large proportion of the genes negatively affected by loss of CtrA, with 73 of the 216 genes (~34%) assigned to either flagellar or chemotaxis functions based on homology, or residing in apparent operons with flagellar or chemotaxis (*che*) genes. CtrA is essential for maximal expression of essentially all putative flagellum-dependent motility genes (Figure 2.6). The majority of the motility genes (65/73) fall within group 1 (Figure 2.5A) and therefore have extremely low (signal intensity <1) transcript levels in the *ctrA* mutant. These motility genes include duplications of at least some of the *che* genes in different locations (*rcc01759* - *rcc01767* and *rcc01352* - *rcc01358*) and 13 putative methyl-accepting chemotaxis proteins (Table S2.6); we can identify only one potential *che* gene, which contains several conserved Che protein domains (*rcc00782*), that is not CtrA-

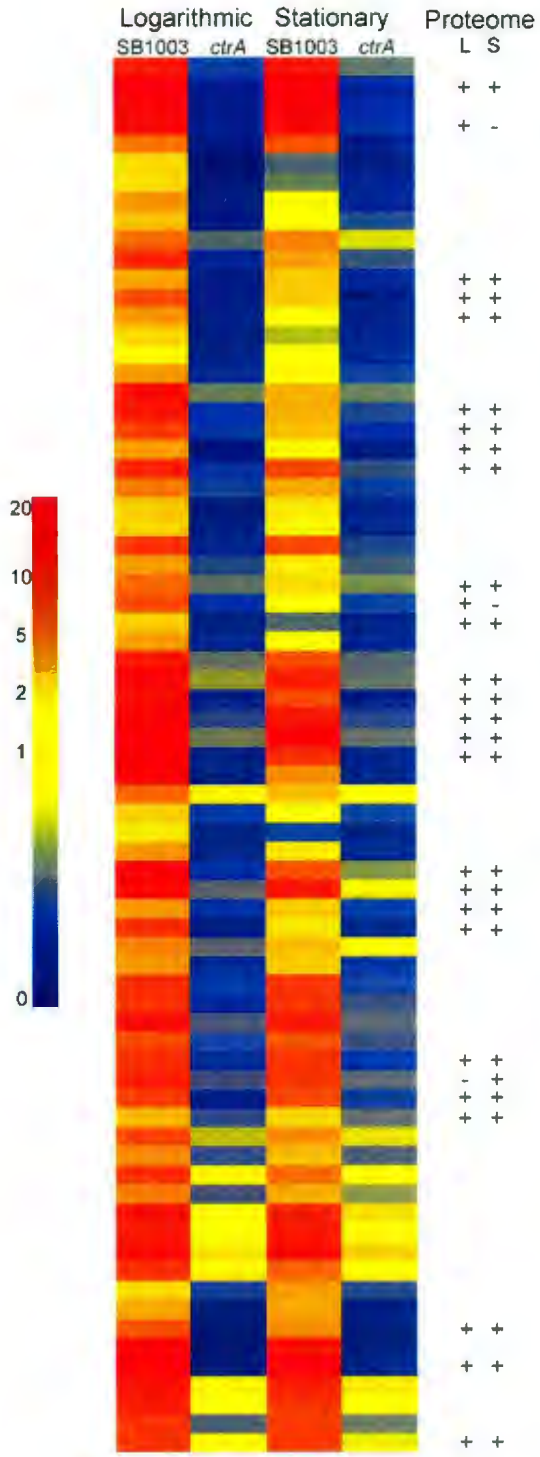


Figure 2.6. CtrA control of flagellum-dependent motility gene expression in *R. capsulatus*. The heat map shows the quantitative transcription data for the 73 putative flagellum-dependent motility genes identified as requiring CtrA for proper expression, by their gene order on the chromosome. For the peptides detected by proteomics, concordance/discordance of protein and transcription data are indicated with a "+" or "-", respectively, for the different growth phases (L, logarithmic; S, stationary). The scale is coloured to indicate relative amounts of the transcripts as identified in the culture samples shown above the columns. The numerical transcript and proteome data for these genes is in Table S2.6.

dependent for maximal transcript levels under the conditions investigated here. Of the 73 motility genes, 28 proteins (38%) were identified and represent 2.3% of all peptides detected. The qualititative protein data match the pattern found in the transcriptome for 25 of the 28 genes, and match the transcriptomic data in one of the two growth phases for the other 3 genes (Figure 2.6; Table S2.6).

2.3.3.2. Expression of the gene transfer agent gene cluster

CtrA was discovered in *R. capsulatus* because it is required for production of the gene transfer agent, RcGTA (Lang & Beatty, 2000). Transcript levels of the RcGTA genes increase in the stationary phase, and these genes/ORFs constitute most of the genes in the mobile element category (Figure 2.2). The proteome analysis detected the 3 RcGTA proteins encoded by transcript segments of the RcGTA gene cluster present at the highest levels, and there was a large increase in the amounts of these 3 proteins in the stationary phase of wild type cultures (Figure 2.7). The results support a model where CtrA is required for transcription of the RcGTA gene cluster (Lang & Beatty, 2000), and quorum sensing causes an increase in expression of the cluster in stationary phase (Schaefer *et al.*, 2002).

2.3.3.3. Other genes affected by loss of CtrA

There are 138 genes dysregulated in the *ctrA* mutant that are not involved in motility or part of the RcGTA gene cluster. Of these genes, 27 (20%) yielded peptides that were identified, and there was an agreement between the transcript and peptide data for 23 (85%) of the proteins (Table S2.5).

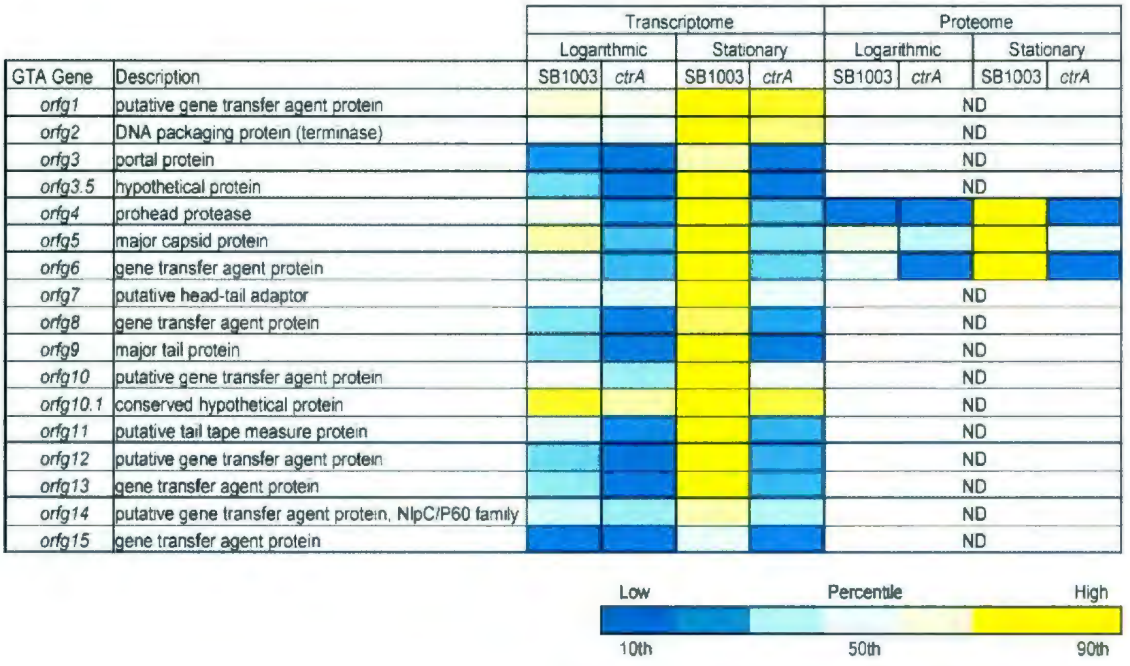


Figure 2.7. Expression of the RcGTA structural gene cluster. The requirement for CtrA and the effects of growth phase on RcGTA gene expression are shown. The scale is coloured to indicate relative abundance of the transcripts and proteins as indicated. ND, not detected. The RcGTA proteins encoded by *orfg3* and *orfg15* (*rcc01684* and *rcc01698*, respectively) were detected in the proteomic analyses but did not pass the threshold criteria for inclusion.

I identified genes encoding 23 putative (non-chemotaxis) signal transduction and transcription regulatory proteins that are affected by the loss of CtrA (Table 2.2). These include proteins containing conserved signalling domains such as PAS, EAL and GGDEF, phosophorelay histidine kinase and receiver, and transcription regulatory domains such as the helix-turn-helix motif. Although there were no large effects on the transcript levels of the 8 σ factor genes (Table 2.1) in the *ctrA* mutant, there were effects on the expression of 2 σ factor regulatory genes, *rcc03323* and *rcc03324* (Table 2.2).

Table 2.2. Signal transduction and transcriptional regulator proteins affected by the loss of CtrA.

Gene	Size of protein (aa)	Predicted function (signalling and regulatory domains[a])	Transcript fold change[b] (logarithmic, stationary)
rcc00180	109	Phosphorelay (HPT)	26.3, 25.3
rcc00181	409	Response regulator (REC, PP2C)	17.4, 10.8[c]
rcc00346	514	c-di-GMP signalling (GGDEF, EAL)	7.6, 4.7[c]
rcc00537	247	Response regulator (REC)	28.4, 23.7
rcc00620	610	Response regulator and c-di-GMP signalling (REC, GGDEF, EAL)	14.0, 6.3
rcc00621	619	Sensor kinase (CHASE4, SK)	5.7, 5.0
rcc00645	1245	c-di-GMP signalling (GGDEF, EAL, PAS)	7.7, 25.9[c]
rcc01849	253	Unknown (PAS)	3.3, 1.7
rcc02075	416	Unknown (PAS)	5.8, 3.6
rcc02539	627	c-di-GMP signalling (GGDEF, EAL)	8.1, 4.1
rcc02629	353	c-di-GMP signalling (GGDEF)	8.1, 2.8
rcc02675	170	Transcriptional regulator (HTH)	19.3, 5.9
rcc02856	412	Unknown (PAS)	13.9, 2.7
rcc02857	1158	c-di-GMP signalling (GGDEF, EAL, PAS)	12.5, 5.6
rcc03176	411	Unknown (PAS)	17.1, 10.0[c]

rcc03177	284	c-di-GMP signalling (EAL)	19.5, 17.6
rcc03301	1284	c-di-GMP signalling (GGDEF, EAL, PAS)	4.5, 3.0
rcc03323	115	Anti-anti-σ factor (STAS)	32.3, 8.7
rcc03324	163	Anti-σ regulatory factor (HATPase)	22.0, 9.4
rcc03452	743	Sensor kinase (SK, REC)	9.8, 3.0
rcp00117	280	c-di-GMP signalling (EAL)	24.9, 12.1[c]
rcp00137	412	Unknown (PAS)	9.0, 9.4

[a]Conserved domains were identified by BLAST analyses at the NCBI Conserved Domain Database; HPT, histidine phosphotransfer; REC, response-regulator receiver; PP2C, Serine/threonine phosphatase; GGDEF, diguanylate cyclase; EAL, c-di-GMP phosphodiesterase; CHASE4, extracellular sensory domain; SK, sensor kinase; PAS, Per-Arnt-Sim; HTH, Helix-turn-helix; STAS, anti-sigma factor antagonist; HATPase, histidine kinase-like ATPase

[b]Transcript fold change is wild type/mutant

[c]Protein detected in proteome and matches transcript trend; data available in Table S2.5

An unusual differential expression pattern was observed for a subset of these 138 genes: log phase transcript levels were similar in the wild type and *ctrA* mutant strains, but transcript levels were ≥2-fold lower in the mutant only in the stationary phase. The seven genes having this expression profile (*rcc00499* - *rcc00501* and *rcc00504* - *rcc00507*) fall within group 2 (Figure 2.5A), and are part of an 11-gene cluster that contains homologues of the genes that direct synthesis of the CtrA-dependent pili in *C. crescentus* (Skerker & Shapiro, 2000). There are recognizable homologues of the CpaB,

CpaC, CpaE and CpaF proteins, but the *R. capsulatus* genes in the locations corresponding to the *C. crescentus pilA*, *cpaA* and *cpaD* (*rcc00499*, *rcc00500* and *rcc00503*, respectively) do not have recognizable similarity to the *C. crescentus* sequences. Five other downstream genes contain conserved domains found in pilus assembly proteins (*rcc00506* - *rcc00510*). In the wild type strain, gene *rcc00499* has the third highest transcript level observed (Table S2.4); the gene organization context and this high level of expression support assigning the putative function of the encoded protein to be a pilin. Such Flp (fimbrial low-molecular-weight protein) pilus gene clusters are widely distributed in bacteria (Tomich *et al.*, 2007) and the pattern of partial gene cluster conservation described above has been documented as widespread in other members of the order Rhodobacterales (Slightom & Buchan, 2009).

Also of note is a cluster of 26 genes in the genome (*rcc01051* through *rcc01076*) in which 22 have >2-fold lower transcript levels in the *ctrA* mutant (Table S2.5). This cluster contains 10 genes predicted to encode gas vesicle proteins, most of which are in an apparent 12-gene operon (the largest gap between any of these 12 genes is 3 bases). There is another apparent 5-gene operon (*rcc01063* - *rcc01067*) in this cluster wherein *rcc01066* encodes a photoactive yellow protein (PYP) xanthopsin photoreceptor (Kyndt *et al.*, 2004). This gene organization has been noted previously, and it was hypothesized that *R. capsulatus* uses PYP to repress gas vesicle formation under high light intensity (Kyndt *et al.*, 2004). The finding that these genes are co-regulated in that they are dependent on CtrA for proper expression further strengthens the hypothesis that there is a connection between PYP and gas vesicle formation. However, the putative gas vesicle genes were

transcribed in our high light growth conditions, although it is possible they could be more highly expressed under lower light conditions.

2.3.4. CtrA binding sites

CtrA binding sites have been best characterized in *C. crescentus* and are characterized primarily by the nucleotide consensus sequences TTAA-N_7-TTAAC (full site) (Quon *et al.*, 1996) and TTAACCAT (half-site) (Laub *et al.*, 2002). However, it was recently shown that CtrA also binds to some TTAA sequences (Spencer *et al.*, 2009). It is thought that CtrA recognizes the same sites in *R. capsulatus* because of an identical amino acid sequence in the putative helix-turn-helix DNA sequence recognition domain (Lang & Beatty, 2000). It is also known (Bellefontaine *et al.*, 2002; Brassinga *et al.*, 2002) or thought (Barnett *et al.*, 2001; Hallez *et al.*, 2004) that the CtrAs of other species bind the same sequences. There are 12 sites in the *R. capsulatus* genome that perfectly match the full CtrA site (Table 2.3), and 8 of these are found 5' of genes that were affected by the loss of CtrA. The four genes not affected ≥2-fold include the *cckA* gene, but the quantitative protein data suggest that loss of *ctrA* does lead to lower amounts of CckA. I found 15 occurrences of the CtrA half-site, and 8 of these occur 5' of genes we identified as dysregulated by loss of CtrA (Table 2.3); however, 3 of these 8 genes also have the full sites. Also, some genes affected by loss of CtrA, such as *rcc00006* and *rcc00499*, have nearly perfect consensus full sites (TTAA-N_7-TTAAG and TTAA-N_7-TTTAC, respectively).

Together, these observations indicate that full CtrA binding sites, and variations thereof, are more relevant than the half-site as CtrA-dependent regulatory sequences in *R.*

capsulatus, at least under the conditions investigated here. The CtrA-dependent putative regulatory and signalling proteins (Table 2.2) may provide a link between the regulation of many of the genes affected by loss of CtrA that lack recognizable CtrA binding sites.

Table 2.3. Potential CtrA binding sites in the *R. capsulatus* genome.

Gene	Description	Location of predicted site[a]		Transcript fold change[b] (logarithmic, stationary)	Quantitative proteome data[c] (logarithmic, stationary)
		TTAA-N_7-TTAAC	TTAACCAT		
rcc00042	PAS/PAC sensor domain protein	-47		12.6, 14.9	ND[d]
rcc00222	DNA repair protein RadC		-30	7.6, 21.8	ND
rcc00280	hemolysin-type calcium-binding repeat family protein		-100	1.2, -1.1	ND
rcc00393	UvrABC system protein C	-372		-1.3, -1.7	ND
rcc00434	hypothetical protein	-57	-54	1.4, 3.9	ND
rcc00459	glutathione S-transferase domain protein		-49	-1.4, -1.3	0.97, -0.43
rcc00463	protein of unknown function UPF0102	-315		4.4, 13.7	ND

rcc00498	transglycosylase, Slt family		-249	-1.25, -1.1	ND
rcc00516	tyrosine phenol-lyase		-570	-1.7, -1.4	-1.43, -1.47
rcc00645	diguanylate cyclase/phosphodiesterase with PAS/PAC sensor	-68		7.7, 25.9	-1.14, 0.74
rcc00844	conserved hypothetical protein	-82	-90	21.3, 28.7	ND
rcc00845	hypothetical protein	-273	-118	5.2, 17.8	
rcc00872	alcohol dehydrogenase, zinc-binding domain protein		-193	-1.3, -1.1	-0.05, -0.50
rcc01237	glycosyl transferase, family 2	-166		2.4, 5.1	ND
rcc01324	type I restriction-modification system, R subunit		-475	-1.4. -1.1	ND
rcc01384	UvrABC system protein B	-65		-1.4, 0	-2.42, 1.25
rcc01462	Ribonucleoside-diphosphate reductase NrdJ	-491		-1.3, 1.6	-2.06, 1.23

rcc01749	histidine kinase CckA	-96		1.4, -1.1	1.33, 1.12
rcc01849	protein of unknown function DUF1457		-40	3.3, 1.7	ND
rcc01932	glycosyl transferase, family 4		-183	-1.7, -1.4	ND
rcc01940	hemolysin-type calcium-binding repeat family protein		-129	-5.0, -2.5	ND
rcc03000	conserved hypothetical protein		-241	2.0. 1.7	ND
rcc03209	protein containing DUF484	-46		2.6, 2.1	1.98, 3.09
rcc03214	protein of unknown function DUF1127		-131	-1.3, -2.5	ND

[a]Position of the last (3') base of the predicted binding site relative to the first base of the predicted start codon for the gene

[b]Positive values indicate higher transcript levels in the wild type and negative values indicate higher levels in the *ctrA* mutant

[c]Absolute difference in Z-score data for each protein with positive values representing higher amounts in wild type and negative values representing higher amounts in the *ctrA* mutant

[d]ND, not detected

2.4. Discussion

2.4.1. CtrA regulates diverse processes in *R. capsulatus*

I have identified 227 genes that are dysregulated by the loss of CtrA in *R. capsulatus*, and these findings demonstrate that CtrA is almost exclusively a positive regulator of gene expression. The transcriptomic and proteomic approach has identified putative pilus and gas vesicle formation genes and proteins as requiring CtrA for maximal expression. Additionally, I have validated previously described functions for CtrA in RcGTA production and flagellum biosynthesis (Lang & Beatty, 2002; Lang & Beatty, 2000). Therefore, it appears that CtrA is involved in the control of multiple types of extracellular structures (RcGTA, flagella and pili), and in the control of movement through distinct mechanisms (flagella and gas vesicles). Gas vesicles have not been reported in *R. capsulatus* and the proteins encoded by these genes were not detected in the proteomic work, but this may be due to the difficulty in solubilizing gas vesicle protein membranes (DasSarma & Arora, 1997). Pilus-like structures have been visualized on *R. capsulatus* cells, including on a *ctrA* mutant (Shelswell *et al.*, 2005). Transcripts of the putative pilus genes identified here were significantly reduced in the *ctrA* mutant in the stationary phase, but not in the log phase, and so it is possible these genes are involved in production of the previously observed structures. Additionally, I identified a number of potential signal transduction and transcription regulator proteins that are affected by the loss of CtrA. I expect that alterations in the expression of these genes contribute to the pleiotropic effects that result from loss of CtrA. The relative amounts of peptides were

measured for ~25% of the corresponding proteins and the quantitative proteomic data validated the CtrA-dependence for 88% of these. Therefore, I have observed a high degree of congruence between the RNA- and protein-derived gene expression data.

The transcript profiles indicate that almost all (including putative) genes involved in flagellum-dependent motility in *R. capsulatus* are CtrA-dependent. I found that 35 of the 73 motility genes were down-regulated in the *ctrA* mutant >20-fold relative to the wild type in the log phase of growth, showing a very strong dependence on CtrA. However, the patterns of regulation are not consistent amongst all of these genes. Some genes display significantly higher transcript levels in the wild type strain, but their transcripts are still present in the *ctrA* mutant. These genes are not absolutely dependent on CtrA for expression although CtrA is clearly required for maximal expression. Other genes that are affected by the loss of CtrA also show this same pattern and continue to be transcribed in the *ctrA* mutant. These may represent genes that are indirectly affected by loss of CtrA and not direct targets of CtrA regulation.

We identified putative CtrA binding sites 5' of 24 genes/operons in the *R. capsulatus* genome, and 13 of these are genes we identified as dysregulated in the transcriptomic analyses. The correlations between these putative CtrA sites and the changes in transcripts detected suggest that both the full and half-sites are important for regulation of transcription by CtrA. There are other partially conserved sites where CtrA may bind, as found for *C. crescentus* (Laub *et al.*, 2002; Quon *et al.*, 1996; Spencer *et al.*, 2009), but the genuine DNA-binding properties of the *R. capsulatus* CtrA remain to be determined. Regardless, it is likely that there are downstream regulatory proteins needed

for the control of many of the 227 identified genes, and we found 23 genes that encode putative signal transduction and/or transcription regulatory proteins affected by the loss of CtrA (not counting chemotaxis signal transduction genes). Other than general predictions about the functions of some of these (e.g. sensor kinase or c-di-GMP signalling), I cannot speculate on the specific roles of any of these proteins because we are unable to identify orthologues of these proteins that have a known, specific function in another species.

2.4.2. Growth phase changes in gene expression

The data in this paper show that the CtrA signalling system regulates RcGTA and other gene expression in the both log and stationary phases of growth, as opposed to a solely stationary phase-dependent regulation of gene expression. A portion (~15%) of the genes affected by loss of CtrA shows a ≥2-fold increase in transcript levels in stationary relative to log phase; this may represent the integration of other regulatory processes such as quorum sensing, which is known to positively regulate RcGTA gene expression in the stationary phase (Schaefer *et al.*, 2002). In addition to increasing the expression of the RcGTA genes, *R. capsulatus* responded to the stationary phase by shifting transcription away from biosynthesis functions and towards transport functions. In particular, the data indicate that cultures become iron-limited in the stationary phase of these growth conditions. It is known that iron is required for maximal photosynthesis pigment production in *R. capsulatus* (Cooper, 1963), and photosynthesis gene expression in the related bacterium *Rhodobacter sphaeroides* (Horne *et al.*, 1998). However, iron limitation is not the cause of the stationary phase under these growth conditions because increasing the iron concentration in the growth medium had no observable effect. There was also a

large increase in expression of genes in the unknown function and hypothetical protein category, reflecting a poor understanding of the stationary phase of growth in the α-proteobacteria in general.

Some of the stationary phase adaptations appear to be mediated through changes in σ factor expression, although there is no homologue of the stationary phase σ factor, RpoS (Hengge-Aronis, 2002; Loewen *et al.*, 1998), as found in *E. coli*. It also appears that some regulation of gene expression by CtrA is through σ factors, but this is mediated through regulation of σ-regulatory proteins rather than direct effects on σ factor gene expression.

2.4.3. CtrA functions vary across the α-proteobacteria

The bacterium *Ruegeria* sp. TM1040, in the marine *Roseobacter* group, is the closest relative of *R. capsulatus* in which CtrA and CckA have been studied (Miller & Belas, 2006). Similar to *R. capsulatus*, CtrA is not essential for viability and controls flagellum-dependent motility. Disruption of *ctrA* causes elongation of *Ruegeria* cells; however, we have found no effects of *ctrA* disruption on cell size or shape in *R. capsulatus*. The role of CtrA has been thoroughly characterized in *C. crescentus* (Laub *et al.*, 2002; Laub *et al.*, 2000) where 144 genes are affected by loss of CtrA (Laub *et al.*, 2000). Control of genes involved in flagellum synthesis, chemotaxis, and pilus synthesis by CtrA is conserved between *C. crescentus* and *R. capsulatus*. However, a major difference between the two organisms is that the *R. capsulatus* CtrA is not involved in controlling genes for cell division, DNA methylation, DNA replication and repair, ribosomes, RNA polymerase, or NADH dehydrogenase. Although CtrA does affect

expression of signal transduction proteins in both species, the genes are not orthologous. Also, there is no indication that the *R. capsulatus* CtrA is involved in regulating the initiation of DNA replication because there are no putative CtrA binding sites near the origin of replication.

A phylogenetic tree of CtrA sequences from species representing the major taxonomic orders within the α-proteobacteria for which complete genome sequences were available is shown in Figure 2.8, and mapped onto the tree are the known or predicted functions of CtrA in these species. Predicted functions are based on the presence of CtrA-binding sites located 5' of genes involved in the various processes. The relationships of the *ctrA* sequences match the relationships for these organisms as determined by other means (Gupta & Mok, 2007; Williams *et al.*, 2007), suggesting a vertical descent of the *ctrA* gene from the last common ancestor of the α-proteobacteria. Based on the pattern of CtrA functions across these species, I speculate that CtrA was involved in the regulation of the cell cycle and possibly motility in the last common ancestor of these groups. It seems that in the *Rhodobacterales* (and possibly the *Rhodospirillales*), a role for CtrA in the cell cycle has been lost, and *ctrA* is therefore no longer essential. A role of CtrA in motility has been retained in all groups other than the *Rickettsiales*. There do not appear to be flagellar genes in the *Rickettsiales*, and so these organisms appear to have lost these genes as a result of genome size reduction in adaptation to an obligately parasitic lifestyle (Batut *et al.*, 2004; Ogata *et al.*, 2001).

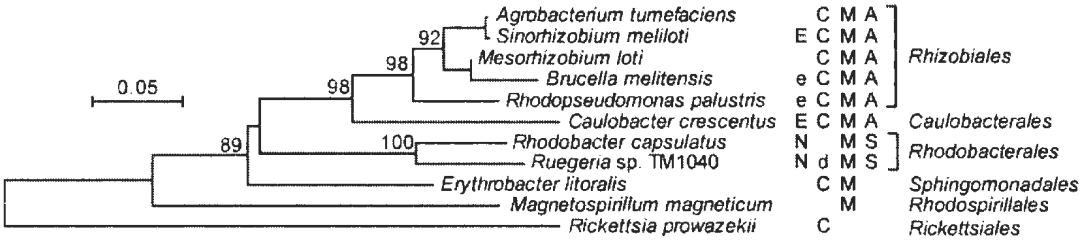


Figure 2.8. Correlation of CtrA function with phylogeny in the α-proteobacteria. The (putative) functions of CtrA in the species are indicated: E, essential; e, putatively essential based on the inability to create a mutant strain [*Brucella* (Bellefontaine *et al.*, 2002); *Rhodopseudomonas* (Lang and Beatty, unpublished)]; N, non essential; C, cell cycle; d, possible role in cell division; M, motility. The mode of cell division is indicated where known: A, asymmetrical; S, symmetrical. The indicated CtrA functions are based on previous publications where available (Barnett *et al.*, 2001; Bellefontaine *et al.*, 2002; Brassinga *et al.*, 2002; Haselkorn *et al.*, 2001;Kahng & Shapiro, 2001;Laub *et al.*, 2000; Loewen *et al.*, 1998; Marrs, 1974; Ogata *et al.*, 2001; Rutherford *et al.*, 2000; Schaefer *et al.*, 2002), or the presence of CtrA binding sites upstream of genes involved in the various processes. Note that not all of these species are known to be motile by means of flagella, only that they have predicted CtrA binding sites next to flagellar genes. Phylogenetic analyses were conducted in MEGA4 (Tamura *et al.*, 2007), with evolutionary relationships inferred by Neighbor-Joining (Saitou & Nei, 1987), and bootstrap values based on 10000 replicates shown next to well-supported branch points (Felsenstein, 1985). Evolutionary distances were computed using the JTT matrix *(Jones et al.*, 1992) and the scale bar indicates the number of substitutions per site. Accession numbers for the sequences used are: *A. tumefaciens*, NP_355385.1; *S. meliloti*, NP_386824.1; *M. loti*, NP_104064.1; *B. melitensis*, NP_539340.1; *R. palustris*, NP_946978.1; *C. crescentus*, NP_421829.1; *R. capsulatus*, AAF13177.1; *Ruegeria* sp. TM1040, YP_613394.1; *E. litoralis*, YP_459735.1; *M. magneticum*, YP_419992.1; *R. prowazekii*, NP_220465.

In *C. crescentus* the activity of CtrA is controlled through transcriptional, phosphorylation and proteolytic regulation (Bowers *et al.*, 2008). There is a complex system controlling the activation and turnover of CtrA, involving the histidine kinase proteins DivJ, DivL, PleC and CckA, the histidine phosphotransferase ChpT, the response regulators DivK and CpdR, the protease ClpXP, and a protein of unknown function RcdA. However, of these proteins, convincing homologues of only CckA and ClpXP can be found in the *R. capsulatus* genome. The *Caulobacter* model of a control system centred on CtrA has become a paradigm, but the results show that this system is not conserved in all α-proteobacteria, in agreement with Hallez *et al.* (2004). I suggest that there are (at least) two types of CtrA-centred systems: one (as in *Caulobacter* and the *Rhizobiales*) in which the CtrA protein is essential and that, along with CckA, Div and Ple proteins, forms a regulatory network needed for cell viability; the second (as in *Rhodobacter*) in which the CtrA protein is not essential for cell viability, and control of the cell cycle proceeds through a different regulatory network. These two different systems may relate to fundamental differences in cell division mechanisms between asymmetrically-dividing cells of organisms such as *Caulobacter*, and cells of organisms such as *Rhodobacter* that divide by symmetric binary fission (Hallez *et al.*, 2004).

2.5. References

Adkins, J. N., H. M. Mottaz, A. D. Norbeck, J. K. Gustin, J. Rue, T. R. Clauss, S. O. Purvine, K. D. Rodland, F. Heffron, and R. D. Smith (2006). Analysis of the *Salmonella typhimurium* proteome through environmental response toward infectious conditions. *Mol. Cell. Proteomics* **5**: 1450-1461.

Alias, A., F. J. Cejudo, J. Chabert, J. C. Willison, and P. M. Vignais (1989). Nucleotide sequence of wild-type and mutant *nifR4* (*ntrA*) genes of *Rhodobacter capsulatus*: identification of an essential glycine residue. *Nucl. Acids Res.* **17:** 5377.

Barnett, M. J., D. Y. Hung, A. Reisenauer, L. Shapiro, and S. R. Long. (2001). A homologue of the CtrA cell cycle regulator is present and essential in *Sinorhizobium meliloti*. *J. Bacteriol.* **183:** 3204-3210.

Batut, J., S. G. E. Andersson, and D. O'Callaghan (2004). The evolution of chronic infection strategies in the α-proteobacteria. *Nat. Rev. Microbiol.* **2:**933-945.

Beatty, J. T., and H. Gest (1981). Generation of succinyl-coenzyme A in photosynthetic bacteria. *Arch. Microbiol.* **129:** 335-340.

Bellefontaine, A.-F., C. E. Pierreux, P. Mertens, J. Vandenhaute, J.-J. Letesson, and X. D. Bolle (2002). Plasticity of a transcriptional regulation network among alpha-proteobacteria is supported by the identification of CtrA targets in *Brucella abortus*. *Mol. Microbiol.* **43:** 945-960.

Bovallius, Å. and B. Zacharias (1971). Variations in the metal content of some commercial media and their effect on microbial growth. *Appl. Microbiol.* **22:** 260-262.

Bowers, L. M., E. B. Shapland, and K. R. Ryan (2008). Who's in charge here? Regulating cell cycle regulators. *Curr. Opin. Microbiol.* **11:** 547-552.

Brassinga, A. K. C., R. Siam, W. McSween, H. Winkler, D. Wood, and G. T. Marczynski (2002). Conserved response regulator CtrA and IHF binding sites in the α-Proteobacteria *Caulobacter crescentus* and *Rickettsia prowazekii* chromosomal replication origins. *J. Bacteriol.* **184:** 5789-5799.

Callister, S. J., R. C. Barry, J. N. Adkins, E. T. Johnson, W. J. Qian, B. J. M. Webb-Robertson, R. D. Smith, and M. S. Lipton (2006). Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **5:** 277-286.

Callister, S. J., C. D. Nicora, X. Zeng, J. H. Roh, M. A. Dominguez, C. L. Tavano, M. E. Monroe, S. Kaplan, T. J. Donohue, R. D. Smith, and M. S. Lipton. 2006. Comparison of aerobic and photosynthetic *Rhodobacter sphaeroides* 2.4.1 proteomes. *J. Microbiol. Meth.* **67**: 424-436.

Cooper, R. (1963). The biosynthesis of coproporphyrinogen, magnesium protoporphyrin monomethyl ester and bacteriochlorophyll by *Rhodopseudomonas capsulata*. *Biochem. J.* **89:** 100-108.

Cullen, P. J., D. Foster-Hartnett, K. K. Gabbert, and R. G. Kranz (1994). Structure and expression of the alternative sigma factor, RpoN, in *Rhodobacter capsulatus*; physiological relevance of an autoactivated *nifU2-rpoN* superoperon. *Mol. Microbiol.* **11:** 51-65.

DasSarma, S., and P. Arora (1997). Genetic analysis of the gas vesicle gene cluster in haloarchaea. *FEMS Microbiol. Lett.* **153:** 1-10.

Emetz, D., and G. Klug (1998). Cloning and characterization of the *rpoH* gene of *Rhodobacter capsulatus*. *Mol. Gen. Genet.* **260:** 212-217.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evol.* **39:** 783-791.

Giovannoni, S. J., H. J. Tripp, S. Givan, M. Podar, K. L. Vergin, D. Baptista, L. Bibbs, J. Eads, T. H. Richardson, M. Noordewier, M. S. Rappe, J. M. Short, J. C. Carrington, and E. J. Mathur (2005). Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242-1245.

Green, H. A., and T. J. Donohue (2006). Activity of *Rhodobacter sphaeroides* $RpoH_{II}$, a second member of the heat shock sigma factor family. *J. Bacteriol.* **188:** 5712-5721.

Gupta, R., and A. Mok (2007). Phylogenomics and signature proteins for the alpha Proteobacteria and its main groups. *BMC Microbiol.* **7:** 106.

Hallez, R., A.-F. Bellefontaine, J.-J. Letesson, and X. De Bolle (2004). Morphological and functional asymmetry in α-proteobacteria. *Trends Microbiol.* **12:** 361-365.

Haselkorn, R., A. Lapidus, Y. Kogan, C. Vlcek, J. Paces, V. Paces, P. Ulbrich, T. Pecenkova, D. Rebrekov, A. Milgram, M. Mazur, R. Cox, N. Kyrpides, N. Ivanova, V. Kapatral, T. Los, A. Lykidis, N. Mikhailova, G. Reznik, O. Vasieva, and M. Fonstein. 2001. The *Rhodobacter capsulatus* genome. *Photosynth. Res.* **70**: 43-52.

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

Horne, I. M., J. M. Pemberton, and A. G. McEwan (1998). Manganous ions suppress photosynthesis gene expression in *Rhodobacter sphaeroides*. *Microbiol.* **144:** 2255-2261.

Irizarry, R., B. Bolstad, F. Collin, L. Cope, B. Hobbs, and T. Speed (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31:**e15.

Jones, R., and R. Haselkorn (1989). The DNA sequence of the *Rhodobacter capsulatus* *ntrA*, *ntrB* and *ntrC* gene analogs required for nitrogen fixation. *Mol. Gen. Genet.* **215:** 507–516.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *Comp. Appl. Biosci.* **8:** 275-282.

Kahng, L. S., and L. Shapiro (2001). The CcrM DNA methyltransferase of 560 *Agrobacterium tumefaciens* is essential, and its activity is cell cycle regulated. *J. Bacteriol.* **183:** 3065-3075.

Kelly, R. T., J. S. Page, Q. Luo, R. J. Moore, D. J. Orton, K. Tang, and R. D. Smith (2006). Chemically etched open tubular and monolithic emitters for nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **78:** 7796-7801.

Kiebel, G., K. Auberry, N. Jaitly, D. Clark, M. Monroe, E. Peterson, N. Tolić, G. Anderson, and R. Smith (2006). PRISM: a data management system for high-throughput proteomics. *Proteomics* **6**: 1783-1790.

Kyndt, J. A., J. K. Hurley, B. Devreese, T. E. Meyer, M. A. Cusanovich, G. Tollin, and J. J. Van Beeumen (2004). *Rhodobacter capsulatus* photoactive yellow protein: genetic context, spectral and kinetics characterization, and mutagenesis. *Biochem.* **43:** 1809-1820.

Lang, A. S., and J. T. Beatty (2002). A bacterial signal transduction system controls genetic exchange and motility. *J. Bacteriol.* **184:** 913-918.

Lang, A. S., and J. T. Beatty (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. USA* **97:** 859-864.

Laub, M. T., S. L. Chen, L. Shapiro, and H. H. McAdams (2002). Genes directly controlled by CtrA, a master regulator of the *Caulobacter* cell cycle. *Proc. Natl. Acad. Sci. USA* **99:** 4632-4637.

Laub, M. T., H. H. McAdams, T. Feldblyum, C. M. Fraser, and L. Shapiro (2000). Global analysis of the genetic network controlling a bacterial cell cycle. *Science* **290:** 2144-2148.

Loewen, P. C., B. Hu, J. Strutinsky, and R. Sparling (1998). Regulation in the *rpoS* regulon of *Escherichia coli*. *Can. J. Microbiol.* **44:** 707-717.

Marrs, B. L. (1974). Genetic recombination in *Rhodopseudomonas capsulata*. *Proc. Natl. Acad. Sci. USA* **71:** 971-973.

Mayampurath, A., N. Jaitly, S. Purvine, M. Monroe, K. Auberry, J. Adkins, and R. Smith (2008). DeconMSn: a software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra. *Bioinf.* **24:** 1021-1023.

Miller, T. R., and R. Belas (2006). Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ. Microbiol.* **8:** 1648-1659.

Monroe, M., N. Tolić, N. Jaitly, J. Shaw, J. Adkins, and R. Smith (2007). VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinf.* **23:** 2021-2023.

Ogata, H., S. Audic, P. Renesto-Audiffren, P. E. Fournier, V. Barbe, D. Samson, V. Roux, P. Cossart, J. Weissenbach, J. M. Claverie, and D. Raoult (2001). Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**: 2093-2098.

Pasternak, C., W. Chen, C. Heck, and G. Klug (1996). Cloning, nucleotide sequence and characterization of the *rpoD* gene encoding the primary sigma factor of *Rhodobacter capsulatus*. *Gene* **176:** 177-184.

Polpitiya, A. D., W.-J. Qian, N. Jaitly, V. A. Petyuk, J. N. Adkins, D. G. Camp, II, G. A. Anderson, and R. D. Smith. (2008). DAnTE: a statistical tool for quantitative analysis of -omics data. *Bioinf.* 24: 1556-1558.

Quon, K. C., G. T. Marczynski, and L. Shapiro (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84:** 83-93.

Robertson, C., and R. C. Beavis (2004). TANDEM: matching proteins with mass spectra. *Bioinf.* **20:** 1466-1467.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.-A. Rajandream, and B. Barrell (2000). Artemis: sequence visualization and annotation. *Bioinf.* 16: 944-945.

Saitou, N., and M. Nei (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406-425.

Schaefer, A. L., T. A. Taylor, J. T. Beatty, and E. P. Greenberg (2002). Long-chain acyl-homoserine lactone quorum-sensing regulation of *Rhodobacter capsulatus* gene transfer agent production. *J. Bacteriol.* **184:** 6515-6521.

Shelswell, K. J., T. A. Taylor, and J. T. Beatty (2005). Photoresponsive flagellum-independent motility of the purple phototrophic bacterium *Rhodobacter capsulatus*. *J. Bacteriol.* **187:** 5040-5043.

Skerker, J. M., and M. T. Laub (2004). Cell-cycle progression and the generation of asymmetry in *Caulobacter crescentus*. *Nat. Rev. Microbiol.* **2:** 325-337.

Skerker, J. M., and L. Shapiro (2000). Identification and cell cycle control of a novel pilus system in *Caulobacter crescentus*. *EMBO J.* **19:** 3223-3234.

Slightom, R. N., and A. Buchan (2009). Surface colonization by marine Roseobacters: integrating genotype and phenotype. *Appl. Environ. Microbiol.* **75:** 6027-6037.

Smith, R. D., G. A. Anderson, M. S. Lipton, L. Pasa-Tolic, Y. Shen, T. P. Conrads, T. D. Veenstra, and H. R. Udseth (2002). An accurate mass tag strategy for quantitative and high-throughput proteome measurements. *Proteomics* 2: 513-523.

Solioz, M., H.-C. Yen, and B. Marrs (1975). Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J. Bacteriol.* **123:** 651-657.

Spencer, W., R. Siam, M.-C. Ouimet, D. P. Bastedo, and G. T. Marczynski (2009). CtrA, a global response regulator, uses a distinct second category of weak DNA binding

sites for cell cycle transcription control in *Caulobacter crescentus*. *J. Bacteriol.* **191:** 5458-5470.

Staron, A., H. J. Sofia, S. Dietrich, L. E. Ulrich, H. Liesegang, and T. Mascher (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Mol. Microbiol.* **74:** 557-581.

Tamura, K., J. Dudley, M. Nei, and S. Kumar (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24:** 1596 - 638 1599.

Tomich, M., P. J. Planet, and D. H. Figurski (2007). The *tad* locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol.* **5:** 363-375.

Vlcek, C., V. Paces, N. Maltsev, J. Paces, R. Haselkorn, and M. Fonstein (1997). Sequence of a 189-kb segment of the chromosome of *Rhodobacter capsulatus* SB1003. *Proc. Natl. Acad. Sci. USA* **94:** 9384-9388.

Wall, J. D., P. F. Weaver, and H. Gest (1975). Gene transfer agents, bacteriophages, and bacteriocins of *Rhodopseudomonas capsulata*. *Arch. Microbiol.* **105:**217-224.

Williams, K. P., B. W. Sobral, and A. W. Dickerman (2007). A robust species tree for the Alphaproteobacteria. *J. Bacteriol.* **189:** 4578-4586.

Yen, H.-C., N. T. Hu, and B. L. Marrs (1979). Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. *J. Mol. Biol.* **131:** 157-168.

Yen, H.-C., and B. Marrs (1976). Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. *J. Bacteriol*. **126:** 619-629.

Chapter 3 - Regulatory systems controlling motility and gene transfer agent production and release in *Rhodobacter capsulatus*

3.1. Introduction

One of the most common modes of signal transduction in bacteria is through histidyl-aspartyl phosphorelay systems (Stock *et al.*, 2000). These systems can instigate changes in gene expression and behaviour in response to a variety of environmental and intracellular stimuli. These phosphorelays involve histidine protein kinase and response regulator proteins, and can also include additional histidine phosphotransfer proteins. One well-studied phosphorelay controls the cell cycle in the α-proteobacterium *Caulobacter crescentus*. This regulatory network centres around the response regulator CtrA (Quon *et al.*, 1996), whose activity is controlled through the histidine kinase CckA (Jacobs *et al.*, 2003), a histidine phosphotransferase ChpT (Biondi *et al.*, 2006), as well as a helix-turn-helix transcription factor, SciP (Gora *et al.*, 2010; Tan *et al.*, 2010). The role of the CckA-ChpT phosphorelay is to activate CtrA, by phosphorylation on an aspartate residue, which elicits changes in the expression of genes related to the cell cycle (Brown *et al.*, 2009). CtrA~P also activates transcription of *sciP*, followed by SciP repression of *ctrA* and at least 58 CtrA targets, such as flagellar and chemotaxis genes (Tan *et al.*, 2010). This signalling system is partially conserved in many genera of α-proteobacteria, but the exact functions and components of the system vary between species (Barnett *et al.*, 2001; Bird & MacKrell, 2011; Bellefontaine *et al.*, 2002; Brilli *et al.*, 2010; Hallez *et al.*, 2004; Lang & Beatty, 2000; Lang & Beatty, 2002; Mercer *et al.*, 2010; Miller & Belas, 2006).

The CtrA and CckA homologues found in *Rhodobacter capsulatus* were shown to be involved in regulation of motility (Lang & Beatty, 2002) and production of RcGTA (Lang & Beatty, 2000), a bacteriophage-like particle known as a gene transfer agent. It appears that the most conserved function of the CtrA and CckA proteins in disparate species is related to motility (Bird & MacKrell, 2011; Brilli *et al.*, 2010; Lang & Beatty, 2002; Mercer *et al.*, 2010; Miller & Belas, 2006; Quon *et al.*, 1996). Unlike *C.crescentus*, the CckA and CtrA proteins are not essential in regulation of the *R. capsulatus* cell cycle, but CtrA is required for the proper expression of more than 225 genes (Mercer *et al.*, 2010). However, it is not known whether phosphorylated or unphosphorylated CtrA is the active form of the protein in this species.

Recently, Brilli *et al.* (2010) analyzed 37 α-proteobacterial genomes and identified orthologues of the 14 genes involved in CtrA-dependent cell cycle regulation in *C.crescentus*. Their bioinformatic analyses of possible CtrA networks further strengthened some of the previous work indicating that CtrA regulation and function has a patchwork of conservation in different α-proteobacteria, and they identified a possible *chpT* orthologue in *Rhodobacter*. To further understand the CtrA network in *R.capsulatus*, I have analyzed the motility and RcGTA production phenotypes of strains lacking the putative CtrA regulators SciP and ChpT in comparison with *ctrA* and *cckA* mutants. I also investigated the effects of CtrA phosphorylation state using a phosphomimetic protein, CtrAD51E, and a version of the protein that is unable to be phosphorylated, CtrAD51A. Similar CtrA mutants have been used in *C. crescentus* and *Rhodospirillum centenum* to study CtrA activities (Bird & MacKrell, 2011; Domian *et al.*,

1999; Jacobs *et al.*, 1999; Ryan *et al.*, 2002; Siam & Marczynski, 2003;). The CtrAD51E protein mimics CtrA~P *in vivo* (Domian *et al.*, 1997; Siam & Marczynski, 2003) and the CtrAD51A mutant serves as a constitutively unphosphorylated form (Ryan *et al.*, 2002).

3.2. Materials and Methods

3.2.1. Bacterial strains, plasmids, and culture conditions

The experimental strains, plasmids, and PCR primers used for this study are listed in Table 3.1. *R. capsulatus* was grown at 35°C in anaerobic photoheterotrophic conditions in complex YPS medium (Wall *et al.*, 1975) or aerobically in RCV medium (Beatty & Gest, 1981) supplemented with appropriate antibiotics when necessary: kanamycin (10 μg ml^{-1}) and tetracycline (0.5 μg ml^{-1}). *Escherichia coli* was grown in LB medium at 37°C and supplemented with the appropriate antibiotics when necessary: ampicillin (100 μg ml^{-1}), kanamycin (25 μg ml^{-1}), and tetracycline (10 μg ml^{-1}).

Table 3.1. Experimental strains, plasmids, and primers used in this study.

Strains, plasmids, and PCR primers	Description	References
***R. capsulatus* strains**		
SB1003	Genome-sequenced strain	(Strnad *et al.*, 2010; Yen & Marrs, 1976)
ctrA	SB1003 with disrupted *ctrA*	(Mercer *et al.*, 2010)
chpT	SB1003 with disrupted *chpT*	This work
cckA	SB1003 with disrupted *cckA*	This work
sciP	SB1003 with disrupted *sciP*	This work
ctrA/sciP	SB1003 with disrupted *ctrA* and *sciP*	This work
DW5	SB1003 Δ*puhA*	(Wong *et al.*, 1996)
Plasmids		
pRK767	Broad host range plasmid	(Gill & Warren, 1988)
p*ctrA*	*ctrA* and 181 bp of 5' sequence in KpnI site of pRK767	This work
p*chpT*	*chpT* and 544 bp of 5' sequence in the T4 polymerase-blunted KpnI site of pRK767	This work
p*cckA*	*cckA* and 482 bp of 5′sequence in the KpnI site of pRK767	This work
p*sciP*	*sciP* and 501 bp of 5' sequence in the KpnI site of pRK767	This work
p*ctrA/sciP*	*ctrA* and *sciP*, including intergenic region, and 501 bp of 5' sequence in the KpnI site of pRK767	This work
p*D51E*	*ctrAD51E* and 181 bp of 5' sequence in KpnI site of pRK767	This work
p*D51A*	*ctrAD51A* and 181 bp of 5' sequence in KpnI site of pRK767	This work
Primers	5'-3'[a]	
chpT-F	CAGCACGAGATGGTCGTAAA	
chpT-R	GCACGATCCGTGCGTAAG	
sciP-F	TCCTTGAACTTCACCCGTTC	
sciP-R	TCGGTGCAGTAGACGTTGAG	
ctrA-ComF	GGTGGTACCAATTCGATATCTGGG	
ctrA-ComR	ATGGTACCAGAAGCGGAGACTCA	
cckA-ComF	ATGGTACCGTCTTCGATCTGGTGCTGGT	
cckA-ComR	AGGGTACCACGCTTTCGCACAGGATTAC	
sciP-comF	ATGGTACCCTTGAACTTCACCCGTTC	

sciP-comR	ATGGTACCGGTGCAGTAGACGTTGAG
D51E-F	CTGATCCTTCTCGAACTGAACCTTCCCGA C
D51E-R	GTCGGGAAGGTTCAGTTCGAGAAGGATCA G
D51A-F	CTGATCCTTCTCGCTCTGAACCTTCCCGAC
D51A-R	GTCGGGAAGGTTCAGAGCGAGAAGGATC AG
cckA-p1	GCCCGTCAGGGCGCGTCAGCGGGTGTTGG CGGGTGTCGGGGCTGGCTTAACAGAGGCG GCGGATCTGGCTGCAAGGG
cckA-p2	TATGTTGTGTGGAATTGTGAGCGGATAAC AATTTCACACAGGAAACAGCTTCGCCGCC AAGCTTCATGATCGAGCCC
cckA-p3	TGCTTGTGCTTGGCCTTGCGGCCGGGACG GCGGGATATTTCGTGCAGGACTGCAGCAG CACGTGTTGACA
cckA-p4	TCAGCGAGAAGGGCTTGGGCAGGAAGAC CGAATTCGGGGTCGGCGGACGGTCAGAA GAACTCGTCAAGAAGG

[a] Underlined nucleotides indicate sequences used for cloning, site-directed mutagenesis, or recombineering steps

The ORFs encoding the predicted ChpT (*rcc03000*) and SciP (*rcc01662*) homologues were amplified by PCR from the genome of *R. capsulatus* strain SB1003 using the primers chpT-F and chpT-R, and sciP-F and sciP-R, respectively (Table 3.1). The amplified products were cloned into pGEM-T-Easy (Promega, Madison, WI) according to the manufacturer's guidelines. The genes were disrupted by insertion of a ~1.4-kb SmaI fragment of the kanamycin resistance-encoding KIXX cartridge (Barany, 1985) at specific restriction sites within the ORFs. The *chpT* ORF was disrupted at the NruI site 79 bp from the start of the 633 bp predicted ORF, and the *sciP* ORF was disrupted at a SmaI site 61 bp from the start of the 276 bp predicted ORF. Another genomic fragment containing *sciP* and adjacent *ctrA* was amplified using primers sciP-

comF and ctrA-comR for disrupting both genes. The KIXX cartridge replaced a 644 bp SmaI/BamHI fragment, deleting the last 215 bp of *sciP* and the first 331 bp of the 711 bp *ctrA*. Plasmids containing disrupted versions of the genes were conjugated to *R. capsulatus* from *E. coli* C600 (pDPT51) (Taylor *et al.*, 1983). Mutant strains were generated by GTA transfer of the disrupted versions of the genes into the chromosome of SB1003 (Scolnik & Haselkorn, 1984). PCR with the original amplification primers was used to confirm the resulting kanamycin resistant strains contained only the disrupted genes.

Plasmid recombineering (Noll *et al.*, 2009) was used for the generation of the *cckA* deletion construct. Primers cckA-p1 and cckA-p2 (Table 3.1) were used to amplify a 4,351-bp region encoding *cckA* (*rcc01749*) plus ~1 kb of flanking sequence on each side. Gel-purified PCR fragments were recombined into pUC19 (Vieira & Messing, 1982) and parental plasmids were selectively linearized by SalI treatment. Primers cckA-p3 and cckA-p4 (Table 3.1) were designed to PCR-amplify kanamycin resistance cassette A002 (Gene Bridges, Germany) with 50-bp tails homologous to *cckA* bases 53-103 and 2202-2252, respectively. λ-Red recombination resulted in replacement of ~91% of plasmid-encoded *cckA* with the kanamycin resistance marker, yielding pUCΔ*cckA*. The resulting plasmid was used to generate the *cckA* mutant strain as described above for the *chpT*, *sciP* and the *ctrA/sciP* double mutants.

Trans complementation of wild type genes under control of their native upstream sequences was performed with the low copy number broad host range plasmid, pRK767 (Gill & Warren, 1988). The complementing fragments were amplified from the genome

with appropriate primers (Table 3.1). Site-directed mutagenesis was performed with the QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) to create pRK767-borne mutant *ctrA* genes with their native promoter region encoding a CtrA phosphomimetic protein, CtrAD51E (Domian *et al.*, 1997), and a CtrA protein that is unable to be phosphorylated, CtrAD51A (Ryan *et al.*, 2002), using primers D51E-F/D51E-R and D51A-F/D51A-R, respectively (Table 3.1). The mutagenesis created single bp substitutions that resulted in a glutamate (D51E) or alanine (D51A) in place of the conserved aspartate (D51) phosphorylation site; the presence of the mutations were confirmed by sequencing. These plasmids and the empty pRK767 control were transferred to *R. capsulatus* via conjugation using *E. coli* S17-1 (Simon *et al.*, 1983).

3.2.2. Gene transfer bioassays

RcGTA packages random fragments of the *R. capsulatus* genome and transfers these to recipient cells. A gene transfer bioassay was used to measure production and release of RcGTA particles. This assay quantifies the transfer of an essential photosynthesis gene, *puhA*, to a Δ*puhA* mutant strain, DW5 (Table 3.1). Cultures were grown in photoheterotrophic conditions for 45 hours, at which point they are ~35 hours into the stationary phase of growth. These cultures were filtered using 0.45-μm PVDF syringe filters and filtrates assayed for RcGTA activity by mixing 0.1 ml of filtrate with DW5 cells in a total volume of 0.6 ml GTA buffer (Solioz *et al.*, 1975). After incubation for 1 hour, 0.9 ml of RCV broth was added and the mixtures incubated for an additional 4 hours with shaking at 200 rpm. The samples were plated on YPS agar, incubated in anaerobic phototrophic conditions to select for transfer of the *puhA* marker, and colony

numbers were counted after 48 hours. RcGTA activity was calculated as a ratio relative to paired wild type RcGTA activity in 3 replicate experiments. Statistically significant differences in RcGTA activities were identified by one-way analysis of variance (ANOVA) in R (Chambers *et al.*, 1992).

3.2.3. Western blotting

Western blots targeting the RcGTA major capsid protein (~32 kDa) were performed on the same cultures used for RcGTA activity assays. For each culture, 0.5 ml of culture was centrifuged at >13,000 x *g* for 1 min to pellet the cells, and 0.4 ml of the resulting supernatants were carefully collected into a separate tube. The cell pellets were resuspended in 0.5 ml of TE buffer. These samples, 5 µl of cells and 10 µl of supernatants, were mixed with 3x SDS-PAGE sample buffer, boiled for 5 minutes at 98°C, and electrophoretically separated on a 10% SDS-PAGE gel. Proteins were transferred to a nitrocellulose membrane by electro-blotting in transfer buffer (48 mM Tris Base, 39 mM glycine, 20% methanol [v/v]). The presence of approximately equivalent total protein levels within supernatant and cell sample groups was verified by staining the blotted membrane with Ponceau-S. The membranes were rinsed and blocked with a 5% (w/v) skim milk solution in TBST (20 mM Tris, 137 mM NaCl, 0.1% Tween-20 [v/v]; pH 7.5) for 1 hour at room temperature with gentle rocking. The membranes were rinsed with TBST and incubated overnight at 4°C with a primary antibody (1:1000 dilution in TBST) specific for the RcGTA major capsid protein (Agrisera, Sweden) (Fu *et al.*, 2010). The membranes were washed 3 times in TBST, for 5 minutes each, and incubated with peroxidase-conjugated anti-rabbit IgG (Santa Cruz Biotechnology, USA)

(1:5000 dilution in TBST) for 1 hour at room temperature. The membranes were rinsed 3 times with TBST for 5 minutes each, and bands detected by chemiluminescence using the SuperSignal West Femto Reagent Kit (Thermo Fisher Scientific, Canada). Images were captured on an Alpha Innotech U400 camera, and then inverted and adjusted for brightness and contrast with image processing software.

3.2.4. Motility assays

Motility assay tubes (Krieg & Gerhardt, 1981) were made with 0.35% agar YPS and the stabs were incubated phototrophically at 35°C. Tubes were photographed after 2 days of growth and the images adjusted for brightness and contrast with image processing software. The average diameters of turbid growth in the stab tubes were measured as a ratio relative to the paired wild type in 3 replicate experiments. Statistically significant differences in culture motility were identified by one-way analysis of variance (ANOVA) in R (Chambers *et al.*, 1992).

3.2.5. Viable cell counts

Viable cells counts were performed on the same cultures used for each paired bioassay and western blot experiment. Serial dilutions were plated and colony-forming units (cfu) were calculated to determine the number of viable cells for each culture. Each mutant strain was compared to the wild type in 3 biological replicates. Statistically significant differences in viable cell numbers were identified by one-way ANOVA in R (Chambers *et al.*, 1992).

3.3. Results

3.3.1. Phenotypes of putative CtrA regulatory mutants

The *R. capsulatus* SB1003 (Strnad *et al.*, 2010) regulatory gene orthologues discussed in this work are *rcc01663* (*ctrA*), *rcc01662* (*sciP*), *rcc03000* (*chpT*), and *rcc01749* (*cckA*); all four genes are predicted to be transcribed as independent mRNAs based on genomic context (Strnad *et al.*, 2010) and transcriptome data (Mercer *et al.*, 2010). I compared the *R. capsulatus* CtrA, CckA, ChpT, and SciP sequences to the *C. crescentus* orthologues, and the regions of similarities and the conserved protein domains identified (Marchler-Bauer *et al.*, 2010) are shown in Figure 3.1. Strains with disruptions in the *chpT* and *sciP* genes were made to test whether these proteins were involved in regulation of motility and RcGTA production, as found for CtrA and CckA (Lang & Beatty, 2000, Lang & Beatty, 2002). Additionally, a new *cckA* mutant was constructed because the original *R. capsulatus* mutant strains (Lang & Beatty, 2000, Lang & Beatty, 2002) retained ~70% of the *cckA* coding sequence undisrupted before the insertional mutation site (between the HA and REC domains; Figure 3.1), possibly allowing for expression of a partially functional protein. I also created a *ctrA/sciP* double mutant.

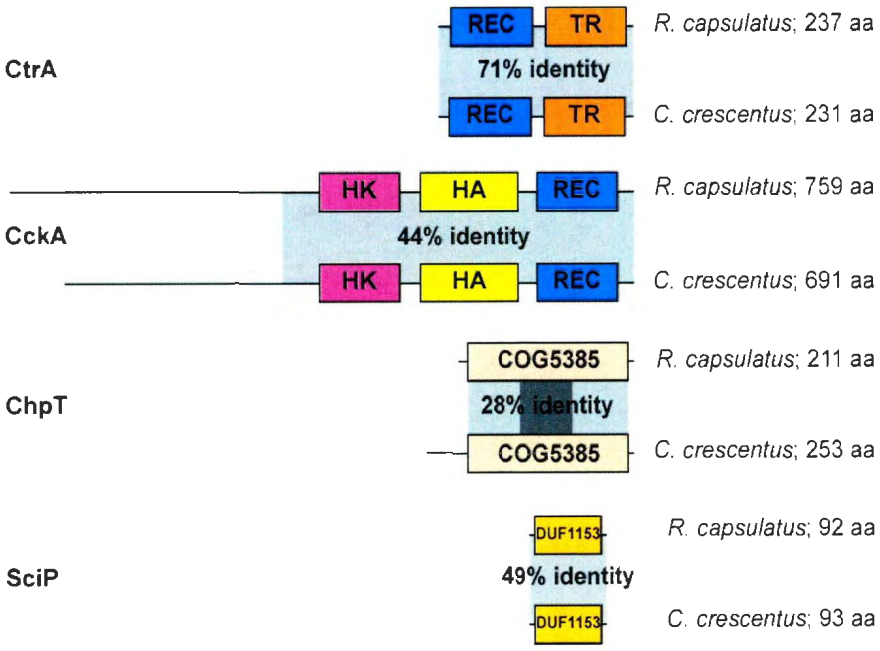


Figure 3.1. Signalling proteins involved in *R. capsulatus* motility and/or RcGTA production. The *R. capsulatus* proteins are compared to their *C. crescentus* homologues, with the % identities indicated for the regions of homology highlighted in gray. For the ChpT proteins, the dark gray highlights a 16-amino acid region with high sequence identity (11/16, 68%) centred on the putative histidine residue involved in phosphotransfer (Biondi *et al.*, 2006). Conserved domains identified by BLAST analyses at the NCBI Conserved Domain Database are shown as coloured boxes and labeled as follows: REC, response regulator receiver; TR, response regulator transcriptional regulator; HK, histidine kinase; HA, histidine kinase-like ATPase; COG, clusters of othologous groups; DUF, conserved domain of unknown function. Protein lengths are indicated on the right.

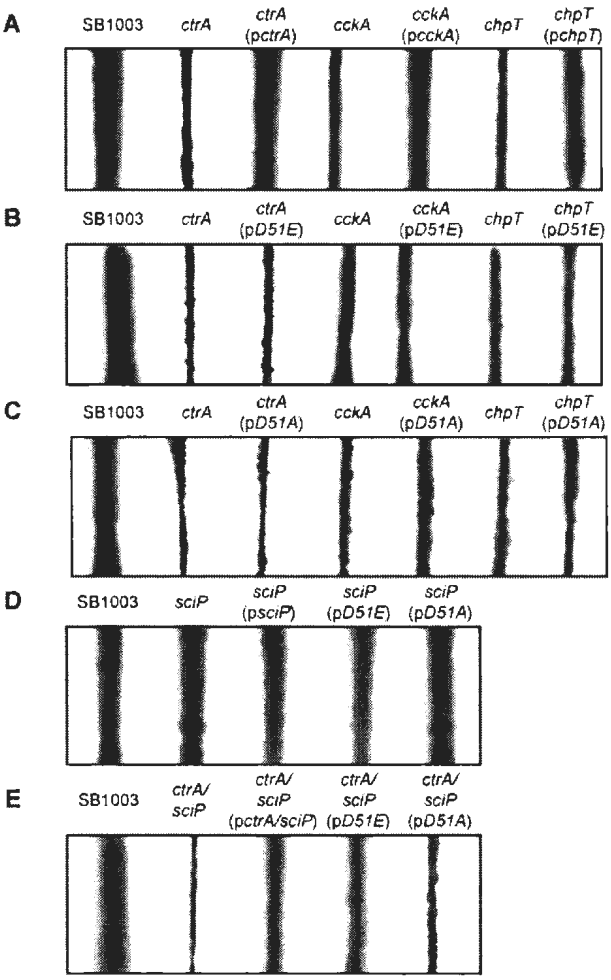


Figure 3.2. Effects of regulatory mutations on *R. capsulatus* motility. Motility assays were carried out in soft agar stabs and the distance of growth outwards from the centre stab line indicates the relative motility of the strain. A. The motility of mutant and complemented *ctrA*, *cckA* and *chpT* strains. B. The effects of CtrAD51E on culture motility in *ctrA*, *cckA* and *chpT* mutants. C. The effects of CtrAD51A on culture motility in *ctrA*, *cckA* and *chpT* mutants. D. Motility of *sciP* strains. E. Motility of *ctrA/sciP* strains. The swimming diameters were measured from 3 replicate assays to determine a ratio of motility relative to SB1003. Statistically significant differences were identified by analysis of variance (ANOVA) and with p-values for all pair-wise comparisons are reported in Supplemental Table S3.1.

Flagellar motility of the *cckA*, *chpT*, *sciP*, and *ctrA/sciP* mutants was assayed using soft agar stabs and compared to wild type strain SB1003 and the *ctrA* mutant (Figure 3.2). Motility in both the *chpT* and *cckA* mutants was reduced, but not as severely as for the *ctrA* and *ctrA/sciP* strains, while *sciP* disruption had no observable effect. Complementation *in trans* of *chpT* and *cckA* restored motility. Wild type *ctrA* does restore

motility in the *ctrA* mutant, but neither *ctrAD51E* nor *ctrAD51A* were able to restore motility in the *ctrA*, *cckA*, or *chpT* mutants. The *ctrAD51E* gene was able to partially restore motility in the *ctrA/sciP* double mutant (Figure 3.2E). Tests for significant differences in swimming distances were performed and all ANOVA results are available in Table S3.1.

RcGTA gene transfer activity was assayed for the *ctrA*, *cckA*, *chpT*, *sciP*, and *ctrA/sciP* mutants (Figure 3.3A). This was paired with analyses of RcGTA capsid protein levels in both cell and culture supernatant samples by western blotting (Figure 3.3B-F). As expected, the *ctrA* and *ctrA/sciP* mutants had no detectable RcGTA activity (Figures 3.3A) or capsid protein expression (Figure 3.3B and 3.3F), with gene transfer activity and protein levels at least partially restored by complementation with the plasmid-based genes. Both the *cckA* and *chpT* mutants demonstrated a nearly complete loss in RcGTA activity (Figure 3.3A). These findings initially suggested that a loss in either ChpT or CckA resulted in a decrease of RcGTA expression, possibly due to the loss of phosphorelay to CtrA. However, western blot analysis of the cultures demonstrated that both *cckA* and *chpT* mutants were expressing the RcGTA capsid protein at wild type levels, but the protein was not detected in the culture supernatants (Figure 3.3B). The extracellular levels of the major capsid protein and RcGTA activity were restored to the mutants upon complementation with the plasmid-borne genes. The gene transfer activity of the *sciP* mutant was lower than wild type (Figure 3.3A) but this difference was not statistically different (Supplemental Table S3.2).

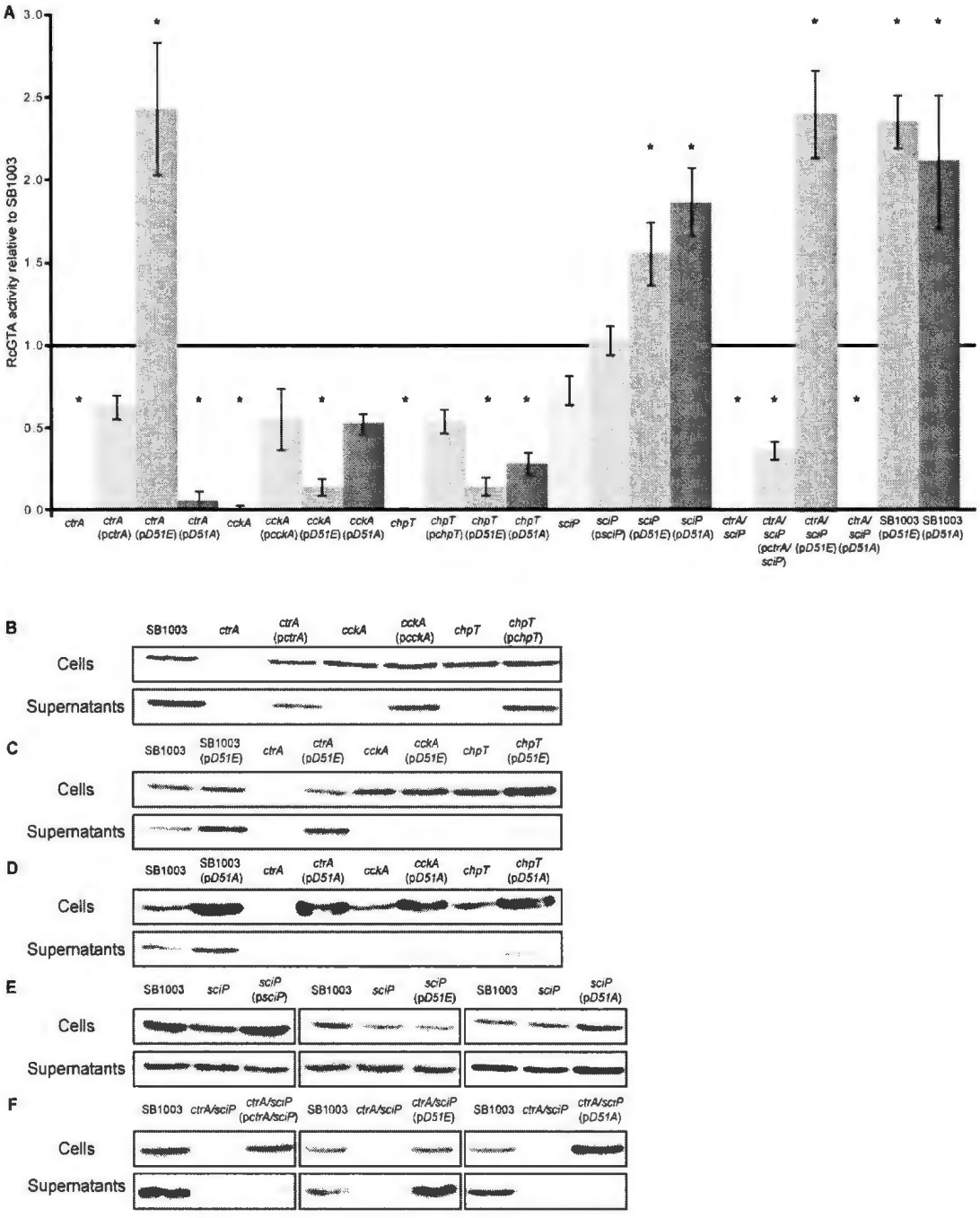


Figure 3.3. RcGTA gene transfer activity and gene expression in regulatory system mutants. A. The ratio of gene transfer activity for each indicated strain relative to the parental strain, SB1003. The gene transfer

activity was determined as an average relative to SB1003 in 3 replicate bioassays and the bars represent the standard deviation. RcGTA production levels that differed significantly from the wild type ($p<0.01$) were identified by ANOVA and are indicated by an asterisk (*). P-values for all pair-wise comparisons are reported in Supplemental Table S3.2. B-F. Western blots for detection of the RcGTA major capsid protein in the cells and culture supernatants of indicated strains. Blots were performed on all replicate gene transfer bioassay cultures (in A) and one representative set of blots is shown.

Introduction of the *ctrAD51E* allele restored RcGTA expression and increased activity in the *ctrA* and *ctrA/sciP* mutant >2-fold relative to wild type (Figure 3.3A). An increase in activity was also observed in both the wild type (2.4-fold) and *sciP* mutant (1.6-fold) strains containing *ctrAD51E*. CtrAD51E increased RcGTA activity and extracellular capsid protein levels slightly in the *cckA* and *chpT* mutants (Figure 3.3C). The *ctrAD51A* gene yielded surprising results as all strains expressing this version of CtrA showed a large increase in capsid protein levels inside the cells relative to wild type (Figure 3.3D). The wild type and *sciP* mutant, both of which carry a wild type copy of *ctrA*, also demonstrated significant increases in RcGTA activity (Figure 3.3A). However, unlike the CtrAD51E protein, activities in the *ctrA* and *ctrA/sciP* mutants remained very low (Figure 3.3A), which agreed with observed low extracellular capsid levels (Figure 3.3D and 3.3F). Introduction of the *ctrAD51A* allele caused an increase in RcGTA activity and extracellular capsid levels in both the *cckA* and *chpT* mutants (Figure 3.3A and 3.3D).

3.3.2. Viability of mutant strains in the stationary phase of growth

Viable cell counts were performed with the different strains on the same cultures used for the gene transfer bioassays and western blots. None of the strains were affected for growth rate and all reached the same approximate cell density at stationary phase as determined by culture turbidity (data not shown). The *ctrA/sciP*, *chpT*, and *cckA* mutations were found to have no significant effect on the number of colony forming units (Figure 3.4). Unexpectedly, the *ctrA* mutant showed a significant increase (1.6-fold; $p<0.01$) in colony forming units relative to wild type (Figure 3.4). Conversely, the *sciP* mutant was found to have a significant decrease (~0.5 of wild type; $p<0.01$) in colony forming units (Figure 3.4). All ANOVA results are available in Table S3.3. The introduction of the *ctrAD51E* and *ctrAD51A* genes had no effect (Figure 3.5).

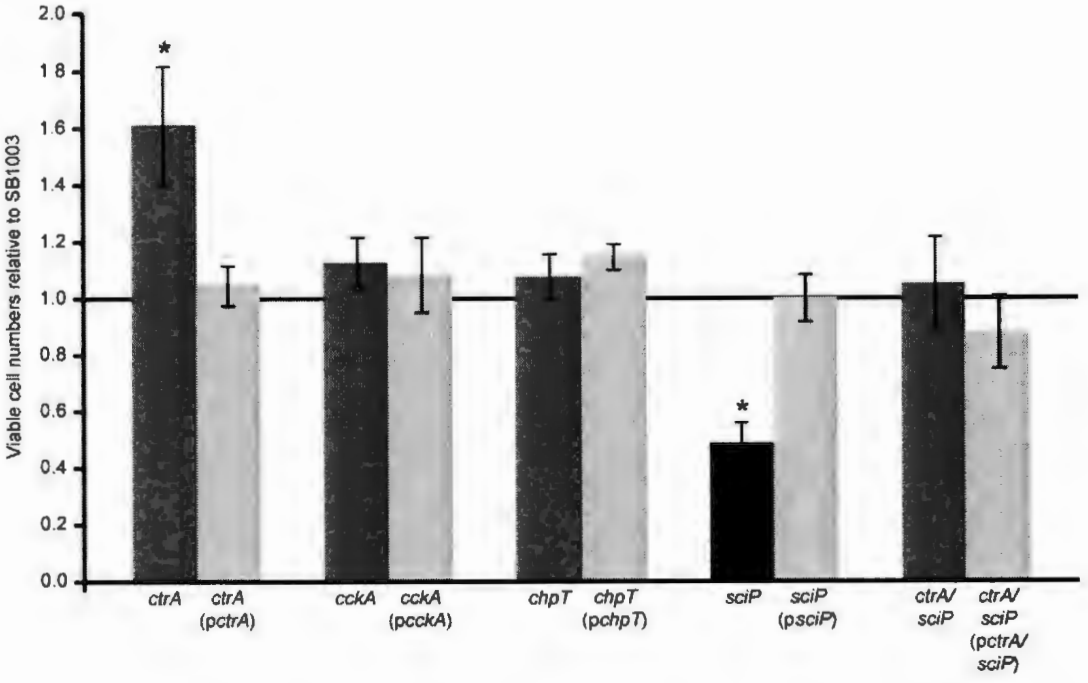


Figure 3.4. Effects of regulatory system mutations on colony forming unit numbers in stationary phase. The ratios of viable cells per ml relative to SB1003 are shown. Ratios were determined as an average over 3 biological replicates with the same cultures used for the RcGTA gene transfer activity assays and western blots. Error bars represent standard deviation, and statistically significant differences ($p<0.01$) were identified by ANOVA and are indicated by an asterisk (*). All p-values are reported in Table S3.3.

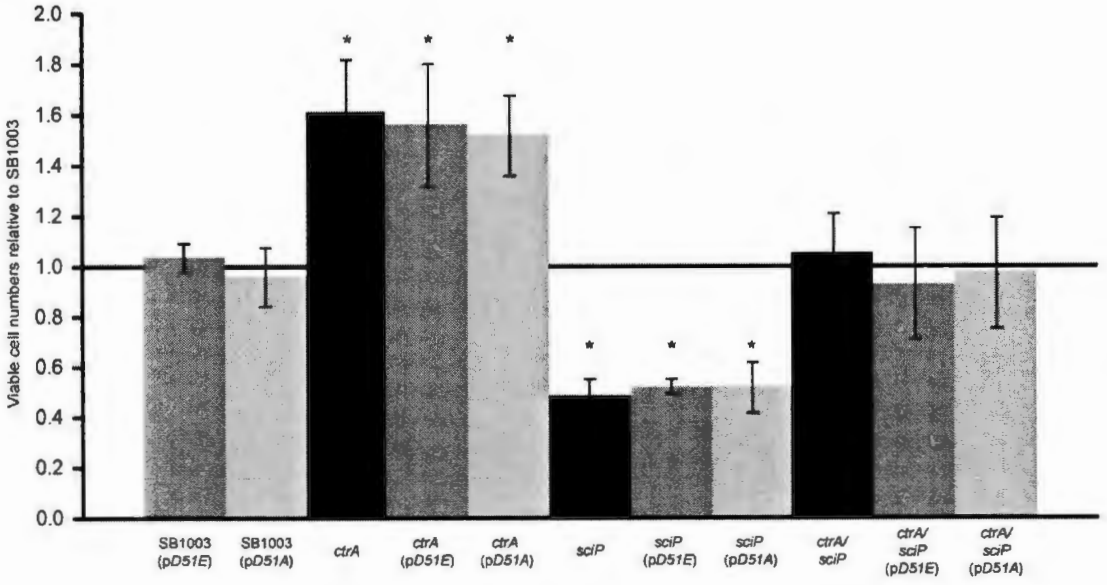


Figure 3.5. Effects of *ctrAD51E* and *ctrAD51A* on colony forming unit numbers in stationary phase. The ratios of viable cells per ml relative to SB1003 are shown. Ratios were determined as an average over 3 biological replicates with the same cultures used for the RcGTA gene transfer activity assays and western blots. Error bars represent standard deviation, and statistically significant differences ($p<0.01$) were identified by ANOVA and are indicated by an asterisk (*). All p-values are reported in Table S3.4.

3.4. Discussion

The experiments with *R. capsulatus* mutant strains lacking putative orthologues of proteins involved in a pathway controlling CtrA activity in *C. crescentus* show that there is a connection between these proteins and CtrA control of motility and RcGTA production (Figure 3.6). However, the results conflict with a model that phosphorylation of CtrA via CckA and ChpT activates both RcGTA and motility gene expression. The *chpT* and *cckA* mutations have negative effects on motility and production of RcGTA, both of which are also controlled by CtrA (Figures 3.2 and 3.3) (Lang & Beatty, 2000;

Lang & Beatty, 2002; Mercer *et al.*, 2010). However, while the phenotypes of the *cckA* and *chpT* mutants are similar to each other, they differ from the *ctrA* mutant (Figures 3.2 and 3.3). The *cckA* and *chpT* mutants retain RcGTA gene expression, but are affected for RcGTA release. Also, both *ctrAD51E* and *ctrAD51A*, which encode proteins that mimic phosphorylated and unphosphorylated CtrA, respectively, activate expression of the RcGTA capsid gene but only *ctrAD51E* leads to release in a *ctrA* mutant. Therefore, a phosphorelay to CtrA via CckA-ChpT is not required for RcGTA gene expression but CckA, ChpT, and CtrA~P are necessary for RcGTA release. Furthermore, *ctrAD51E* could not fully restore gene transfer activity in the *cckA* and *chpT* mutants indicating that CckA-ChpT and CtrA~P are independently required for proper release of RcGTA. This suggests that CckA-ChpT act on an additional response regulator, similar to the *C. crescentus* proteins which activate CpdR (Iniesta *et al.*, 2006) It is possible that this second response regulator is required in conjunction with CtrA~P for release of RcGTA. Despite the data not supporting the exact regulatory model as seen in *C. crescentus*, I hypothesize that the phosphorelay from CckA-ChpT to CtrA does occur and is responsible for both the release of RcGTA and culture motility. The ability of unphosphorylated CtrA to activate expression of RcGTA helps explain the phenotype of both *cckA* and *chpT* mutants.

SciP is a transcriptional regulator and an inhibitor of CtrA-dependent transcription in *C. crescentus* (Gora *et al.*, 2010; Tan *et al.*, 2010). The *sciP* gene is co-conserved with *ctrA* across the α-proteobacteria (Gora *et al.*, 2010) and its transcription is dependent upon CtrA in *R. capsulatus* (Mercer *et al.*, 2010). Inactivation of *sciP* did not have an

observable effect on motility or RcGTA gene expression and release (Figures 3.2 and 3.3). Nevertheless, our data indicate SciP is involved in control of motility. Neither of the site-directed mutant forms of CtrA restored motility in the *ctrA* mutant (Figure 3.2), and we hypothesized this was due to *sciP* activation by CtrAD51E and resulting over-repression of the CtrA-dependent flagellar motility genes by SciP. The difference between motility of *ctrA* (p*D51E*) and *ctrA/sciP* (p*D51E*) validates this hypothesis and implicates SciP as a negative regulator of the flagellar motility genes. The inability of *ctrAD51A* to affect motility in *ctrA/sciP* indicates it is CtrA~P that is required for transcription of the motility genes. It is also known that *C. crescentus* CtrAD51E does not bind DNA with the same affinity as CtrA~P *in vitro* and might only have partial activity relative to CtrA~P (Siam & Marczynski, 2003). Irregularities in complementation of swarming motility in a *ctrA* mutant by D51A and D51E versions of CtrA have also been observed in *R. centenum* (Bird & MacKrell, 2011).

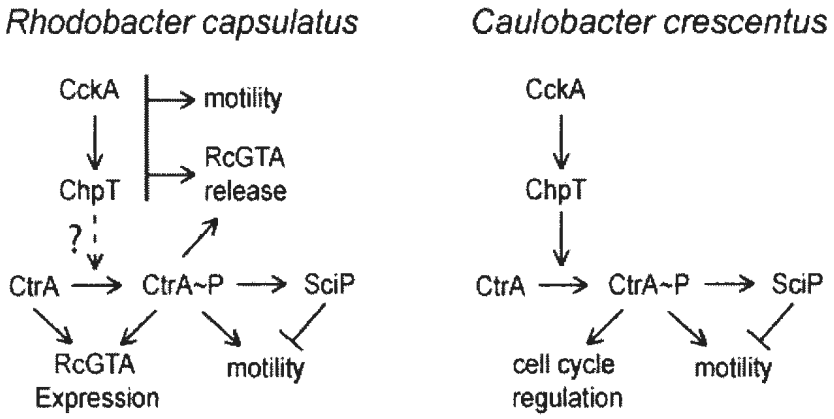


Figure 3.6. Proposed model for activities of the *R. capsulatus* CtrA signal transduction system components discussed in this paper. The activities and functions are indicated for *R. capsulatus* in comparison to the *C. crescentus* model. Arrows indicate positive effects and bars indicate negative effects.

Interestingly, it was found that independent *ctrA* and *sciP* mutations affected the number of viable cells in stationary phase cultures. The available data do not indicate that CtrA plays a role in cell-cycle regulation in *R. capsulatus*, but there is a significant increase in the number of viable cells relative to wild type in the *ctrA* mutant (Figure 3.4). Loss of CtrA causes a ≥2-fold decrease in the transcription of more than 200 genes in these conditions (Mercer *et al.*, 2010), and this could have an impact on these viable cell numbers. In contrast, disruption of *sciP* resulted in a significant decrease in viable cells in the stationary phase. Neither of these mutant strains was affected for growth rate or culture turbidity. This is the first instance where loss of an *R. capsulatus* homologue of a member of the *C. crescentus* CtrA network negatively affects cell viability. The reasons for these changes in stationary phase viable cell numbers remain to be determined. However, one hypothesis is that without *sciP*, CtrA~P augments the release of RcGTA

and overexpression of flagellar genes, perhaps compromising the integrity of the cells. This negative effect would presumably lead to a decrease in the number of viable cells in a *sciP* mutant. Conversely, a *ctrA* mutant would not be releasing any RcGTA particles via the predicted lysis mechanism, which may increase the number of viable cells relative to wild type.

The data support the involvement of CckA, ChpT, and SciP in a regulatory system related to CtrA function in *R. capsulatus* (Figure 3.6). SciP function as a negative regulator of motility is conserved between *R. capsulatus* and *C. crescentus*. The data does not allow me to conclude there is a phosphorelay from CckA-ChpT to CtrA, but there is clear co-involvement of these proteins in regulation of motility and RcGTA release. The reduction, but not complete loss, of motility and RcGTA gene transfer activity in the *cckA* and *chpT* strains could also reflect alternative sources for CtrA phosphorylation. RcGTA release, but not gene expression, is dependent on CtrA phosphorylation. Although it is CtrA~P that binds many regulatory sequences in *C. crescentus* (Reisenauer *et al.*, 1999; Siam & Marczynski, 2000), the unphosphorylated protein is also active (Spencer *et al.*, 2009), and other response regulators have been shown to both activate and repress a variety of genes in unphosphorylated forms, including RegA in *R. capsulatus* (Bird *et al.*, 1999). There are no predicted CtrA-binding sites upstream of either the motility or RcGTA genes (Lang & Beatty, 2000; Mercer *et al.*, 2010), which presumably reflects indirect control of transcription initiation of these genes by CtrA.

3.5. References

Barany, F. (1985). Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. *Gene* **37**: 111-123.

Barnett, M.J., D.Y. Hung, A. Reisenauer, L. Shapiro and S.R. Long (2001). A homolog of the CtrA cell cycle regulator is present and essential in *Sinorhizobium meliloti*. *J. Bacteriol.* **183**: 3204-3210.

Beatty, J.T. and H. Gest (1981). Generation of succinyl-coenzyme A in photosynthetic bacteria. *Arch. Microbiol.* **129**: 335-340.

Bellefontaine, A.-F., C.E. Pierreux, P. Mertens, J. Vandenhaute , J.-J. Letesson and X.D. Bolle (2002). Plasticity of a transcriptional regulation network among alpha-proteobacteria is supported by the identification of CtrA targets in *Brucella abortus*. *Mol. Microbiol.* **43**: 945-960.

Biondi, E.G., S.J. Reisinger , J.M. Skerker, M. Arif , B.S. Perchuk, K.R. Ryan and M.T. M.T. Laub (2006). Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**: 899-904.

Bird, T. & A. MacKrell (2011). A CtrA homolog affects swarming motility and encystment in *Rhodospirillum centenum*. *Arch. Microbiol.* **193**: 451-459.

Bird, T.H., S. Du and C.E. Bauer (1999). Autophosphorylation, phosphotransfer and DNA-binding properties of the RegB/RegA two-component regulatory system in *Rhodobacter capsulatus*. *J. Biol. Chem.* **274**: 16343–16348.

Brilli, M., M. Fondi, R. Fani, A. Mengoni, L. Ferri, M. Bazzicalupo and E. Biondi (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Systems Biol* **4**: 52.

Brown, P.J.B., G.G. Hardy, M.J. Trimble and Y.V. Brun (2009). Complex regulatory pathways coordinate cell-cycle progression and development in *Caulobacter crescentus*. *Adv. Microbial Physiol.* **54**: 1-101.

Chambers, J.M., A.E. Freeny and R.M. Heiberger (1992). Analysis of variance; designed experiments. In *Statistical Models in S.* Chambers, J.M. and T.J. Hastie (Eds.). Pacific Grove, Wadsworth & Brooks/Cole.

Domian, I.J., K.C. Quon and L. Shapiro (1997). Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell* **90**: 415-424.

Domian, I.J., A. Reisenauer and L. Shapiro (1999). Feedback control of a master bacterial cell-cycle regulator. *Proc. Natl. Acad. Sci. USA* **96**: 6648-6653.

Fu, Y., D. MacLeod, R. Rivkin, F. Chen, A. Buchan and A. Lang (2010). High diversity of *Rhodobacterales* in the subarctic North Atlantic Ocean and gene transfer agent protein expression in isolated strains. *Aquat. Microb. Ecol.* **59**: 283-293.

Gill, P.R. Jr. and G.J. Warren (1988). An iron-antagonized fungistatic agent that is not required for iron assimilation from a fluorescent rhizosphere pseudomonad. *J Bacteriol.* **170**: 163-170.

Gora, K.G., C.G. Tsokos, Y.E. Chen, B.S. Srinivasan, B.S. Perchuk and M.T. Laub (2010). A cell-type-specific protein-protein interaction modulates transcriptional activity of a master regulator in *Caulobacter crescentus*. *Mol. Cell* **39**: 455-467.

Hallez, R., A.-F.Bellefontaine, J.-J.Letesson and X. De Bolle (2004). Morphological and functional asymmetry in α-proteobacteria. *Trends Microbiol.* **12**: 361-365.

Iniesta, A. A., P. T. McGrath, A. Reisenauer, H. H. McAdams, and L. Shapiro (2006). A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc. Natl. Acad. Sci. USA* **103:**10935-10940.

Jacobs, C., I.J. Domian, J.R. Maddock and L. Shapiro (1999). Cell cycle-dependent polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division. *Cell* **97**: 111-120.

Jacobs, C., N. Ausmees, S.J. Cordwell, L. Shapiro and M.T. Laub (2003). Functions of the CckA histidine kinase in *Caulobacter* cell cycle control. *Mol. Microbiol.* **47**: 1279-1290.

Krieg, N.R. and P. Gerhardt (1981). Solid culture. In *Manual of Methods for General Bacteriology.* Gerhardt, P., R.G.E. Murray, R.N. Costilow, E.W. Nester, W.A. Wood, N.R. Krieg and G.B. Phillips (Eds.). Washington, DC, American Society for Microbiology, pp. 143-150.

Lang, A.S. and J.T. Beatty (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. USA* **97**: 859-864.

Lang, A.S. and J.T. Beatty (2002). A bacterial signal transduction system controls genetic exchange and motility. *J. Bacteriol.* **184**: 913-918.

Marchler-Bauer, A., S. Lu, J.B. Anderson , C. Farideh, M.K. Derbyshire, C. D. Weese-Scott, J. H. Fong, L.Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, A Tasneem, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, and S.H. Bryant (2010). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**: D225–D229.

Mercer, R.G., S.J. Callister, M.S. Lipton, L. Pasa-Tolic, H. Strnad, V. Paces, J.T. Beatty and Lang, A.S. (2010). Loss of the response regulator CtrA causes pleiotropic effects on gene expression but does not affect growth phase regulation in *Rhodobacter capsulatus*. *J. Bacteriol.* **192**: 2701-2710.

Miller, T.R. and R. Belas (2006). Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ. Microbiol.* **8**: 1648-1659.

Noll, S., G. Hampp, H. Bausbacher, N. Pellegata and H. Kranz (2009). Site-directed mutagenesis of multi-copy-number plasmids: Red/ET recombination and unique restriction site elimination. *Biotechniques* **46**: 527-533.

Quon, K.C., G.T. Marczynski and L. Shapiro (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84**: 83-93.

Reisenauer, A., K. Quon and L. Shapiro (1999). The CtrA response regulator mediates temporal control of gene expression during the *Caulobacter* cell cycle. *J. Bacteriol.* **181**: 2430-2439.

Ryan, K.R., E.M. Judd and L. Shapiro (2002). The CtrA response regulator essential for *Caulobacter crescentus* cell-cycle progression requires a bipartite degradation signal for temporally controlled proteolysis. *J. Mol. Biol.* **324**: 443-455.

Scolnik, P.A. and R. Haselkorn (1984). Activation of extra copies of genes coding for nitrogenase in *Rhodopseudomonas capsulata*. *Nature* **307**: 289-292.

Siam, R. and G.T. Marczynski (2000). Cell cycle regulator phosphorylation stimulates two distinct modes of binding at a chromosome replication origin. *EMBO J.* **19**: 1138-1147.

Siam, R. and G.T. Marczynski (2003). Glutamate at the phosphorylation site of response regulator CtrA provides essential activities without increasing DNA binding. *Nucleic Acids Res.* **31**: 1775-1779.

Simon, R., U. Priefer and A. Pühler (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Bio/Tech.* **1**: 37-45.

Solioz, M., H.-C. Yen and B. Marrs (1975). Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J. Bacteriol.* **123**: 651-657.

Spencer, W., R. Siam, M.-C. Ouimet, D.P. Bastedo and G.T. Marczynski (2009). CtrA, a global response regulator, uses a distinct second category of weak DNA binding

sites for cell cycle transcription control in *Caulobacter crescentus*. *J. Bacteriol.* **191**: 5458-5470.

Stock, A., V. Robinson and P. Goudreau (2000). Two-component signal transduction. *Annu. Rev. Biochem.* **69**: 183-215.

Strnad, H., A. Lapidus, J. Paces, P. Ulbrich, C. Vlcek, V. Paces and R. Haselkorn (2010). Complete genome sequence of the photosynthetic purple nonsulfur bacterium *Rhodobacter capsulatus* SB 1003. *J. Bacteriol.* **192**: 3545-3546.

Tan, M.H., J.B. Kozdon, X. Shen, L. Shapiro and H.H. McAdams (2010). An essential transcription factor, SciP, enhances robustness of *Caulobacter* cell cycle regulation. *Proc. Natl. Acad. Sci. USA* **107**: 18985-18990.

Taylor, D.P., S.N. Cohen, W.G. Clark and B.L. Marrs (1983). Alignment of the genetic and restriction maps of the photosynthetic region of the *Rhodopseudomonas capsulata* chromosome by a conjugation-mediated marker rescue technique. *J. Bacteriol.* **154**: 580-590.

Vieira, J. and J. Messing (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**: 259-268.

Wall, J.D., P.F. Weaver and H. Gest (1975). Gene transfer agents, bacteriophages, and bacteriocins of *Rhodopseudomonas capsulata*. *Arch. Microbiol.* **105**: 217-224.

Wong, D.K.-H., W.J. Collins, A. Harmer, T.G. Lilburn and J.T. Beatty (1996). Directed mutagenesis of the *Rhodobacter capsulatus puhA* gene and pleiotropic effects on

photosynthetic reaction center and light-harvesting I complexes. *J. Bacteriol.* **178**: 2334-2342.

Yen, H.-C. and B. Marrs (1976). Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. *J. Bacteriol.* **126**: 619-629.

Chapter 4 – Regulators of σ factor activity modulate RcGTA expression and affect cell physiology in *Rhodobacter capsulatus*

4.1. Introduction

The regulation of gene expression in bacteria is tightly controlled at the level of transcription in order to allow rapid adaptation to environmental changes. Transcription of bacterial genes is dependent on a multi-protein complex called RNA polymerase (RNAP) that is targeted to a sequence-specific promoter region located 5' of the gene transcription start site. The recruitment of the RNAP core complex and initiation of transcription at the promoter element is dependent on a dissociable subunit called a sigma (σ) factor, which interacts with the core enzyme to form the RNAP holoenzyme (Borukhov & Nudler, 2003). The availability of σ factors to interact with the RNAP and direct transcription is in part dependent on antagonistic proteins called anti-σ factors (Helmann, 1999). An anti-σ factor is defined by the ability to bind and inhibit its cognate σ factor from competing for core RNAP (Hughes & Mathee, 1998). These negative regulators can prevent the expression of specific genes by sequestering the σ factor, preventing it from interacting with the RNAP and promoter sequence. The number of anti-σ factors identified has increased dramatically over the last decade and it is now recognized that these proteins act as important regulators of a wide range of cellular processes including sporulation, stress response, flagellar biosynthesis and phage replication (Campbell *et al.*, 2008; Hughes & Mathee, 1998; Österberg *et al.*, 2011). Anti-σ factors can also be negatively regulated by antagonist proteins called anti-anti-σ factors,

which prevent the cognate σ factor inhibition and thereby promote target gene transcription (Campbell *et al.*, 2008; Helmann, 1999; Sharma *et al.*, 2011). Using anti-anti-σ factors in conjunction with anti-σ factors provides robust control of cognate σ factor availability to associate with the RNAP core enzyme. The use of anti- and anti-anti-σ factor regulation is well illustrated by *S. coelicolor*, which encodes 45 anti-σ factor and 15 anti-anti-σ factor homologues, presumably for regulating the large number of σ factors (up to 63) encoded in the genome (Bentley *et al.*, 2002).

One of the best-studied examples of σ factor regulation by anti- and anti-anti-σ factors is in the regulation of the general stress response factor, σ^B, in the gram-positive bacterium *Bacillus subtilis* (Hecker *et al.*, 2007). In response to both environmental and energy stress, σ^B binds RNAP and directs the transcription of >150 genes associated with resistance to various stressors (Marles-Wright & Lewis, 2007; Price *et al.*, 2001). The *sigB* gene is co-transcribed in an 8 gene operon with 7 other genes (*rsbR*, *S*, *T*, *U*, *V*, *W*, *sigB* and *rsbX*) encoding Rsb proteins (**r**egulators of **s**igma **B**), which control the availability of σ^B to associate with RNAP (Mittenhuber, 2002; Wise & Price, 1995). Under non-stress inducing conditions, the anti-σ factor RsbW binds and sequesters σ^B, preventing expression of target genes (Benson & Haldenwang, 1993). An anti-anti-σ factor, RsbV, is an interacting antagonist of RsbW and can form a complex in an unmodified form (Dufour & Haldenwang, 1994). RsbW is a kinase of RsbV, where phosphorylation during normal, exponential growth inactivates the RsbV antagonist and allows RsbW to bind σ^B (Dufour & Haldenwang, 1994). During a stress response, such as a drop in cellular ATP levels, additional Rsb proteins can affect the phosphorylation state

of the RsbV protein (Alper *et al.*, 1996; Zhang & Haldenwang, 2005). The phosphatase RsbU stimulates the release of σ^B by dephosphorylating RsbV (Yang *et al.*, 1996), which in turn inhibits RsbW from sequestering σ^B. This regulatory mechanism is known as 'partner switching' (Alper *et al.*, 1994) and has been found in diverse prokaryotic species for regulating alternative σ factor activity (Staroń & Mascher, 2010).

The activity of RsbU is itself controlled by three proteins, RsbR, RsbS and RsbT, which form a supramolecular complex called the stressosome (Pane-Farre *et al.*, 2005). The stressosome acts to integrate a diverse array of signals to activate the σ^B stress response (Pane-Farre *et al.*, 2005) and control the activity of the downstream regulatory module involving RsbU-RsbV-RsbW (Hecker *et al.*, 2007). The complete *rsb* operon and *sigB* are conserved in closely related species of *B. subtilis*, such as *B. licheniformis* and *B. halodurans*. The RsbV-RsbW-σ^B module is conserved in other related species such as *B. cereus*, but there are variations in the regulatory components (van Schaik *et al.*, 2005). In *B. cereus* the phosphatase is RsbY, which is structurally different from RsbU, and there is a hybrid sensor histidine kinase/response regulator protein, RsbK, that senses and integrates multiple signals (van Schaik *et al.*, 2005). RsbK can activate the RsbY phosphatase activity on RsbV~P (de Been *et al.*, 2010). This RsbK sensing mechanism is markedly different from the RsbRST stressosome and is found in other species of the *Bacillus* genus (de Been *et al.*, 2011).

Rsb and σ^B homologues have also been identified in various other gram-positive species and play key regulatory roles in the stress response and other cellular processes, such as differentiation in spore-forming bacteria (Hecker *et al.*, 2007). Similar to *B.*

cereus, these other gram positive species (e.g. *Staphylococcus aureus, Mycobacterium tuberculosis* and *Streptomyces coelicolor*) lack *rsbRST* genes encoding the stressosome proteins (Mittenhuber, 2002). The homologues of *rsbV* and *rsbW* are usually found encoded together, alongside a cognate σ factor. However there are variations such as *S. coelicolor*, which uses BldG (RsbV homologue) and ApgA (RsbW homologue) to control differentiation and antibiotic production, but where no cognate σ factor has been identified to date (Parashar *et al.*, 2009).

Anti- and anti-anti-σ factor sequences can be highly variable, even within species. Despite sharing only ~30% amino acid sequence identity, the *B. subtilis* RsbW/RsbV paralogs SpoIIAB/SpoIIAA have analogous partner-switching function in regulating the activity of σ^F and σ^G, which are responsible for the expression of genes required for sporulation (Duncan & Losick, 1993; Evans *et al.*, 2003). Recent literature searches could not find any report of studies on Rsb orthologues in gram-negative species.

The production of the gene transfer agent, RcGTA, of the gram-negative α-proteobacterium *R. capsulatus* has been hypothesized to be a stress response upon entry into stationary phase (Lang & Beatty, 2000). It is currently unknown which σ factor acts to recruit the RNAP to the promoter element of the RcGTA gene cluster, and what signal(s) might drive this process. *R. capsulatus* encodes 8 identifiable putative σ factors in its genome: 1 major vegetative sigma factor, RpoD; two σ^{32} family proteins, RpoHI and RpoHII; the nitrogen fixation σ^{54} factor, RpoN; two σ^{24} (RpoE-like) ECF σ factors; and 1 putative ECF-G σ factor (Chapter 2; Mercer *et al.*, 2010). While the RpoHI, RpoHII and RpoE σ factors have been thoroughly studied in *Rhodobacter sphaeroides* for

their role in response to photooxidative and heat stress (Green & Donohue, 2006; Nuss *et al.*, 2009, 2010), the regulons of each *R. capsulatus* σ factor are poorly understood. Although the response regulator CtrA has been found to be required for RcGTA expression, its role as a positive regulator of these genes is not fully understood and it has been hypothesized that CtrA may regulate the expression of a σ factor (Lang & Beatty, 2000). A transcriptome analysis of a *ctrA* mutant did not identify any differentially regulated sigma factors relative to SB1003 (Chapter 2; Mercer *et al.*, 2010). However, a pair of genes encoding putative anti-σ and anti-anti-σ factors, annotated as *rsbW* and *rsbV*, respectively, was identified to be regulated (directly or indirectly) by CtrA (Chapter 2; Mercer *et al.*, 2010). A more in-depth look into the large number of genes affected by the loss of CtrA identified another potential *rsb* gene homologues predicted to encode a putative phosphatase, RsbU. The identification of this set of putative *rsb* homologues led me to the hypothesis that CtrA affects RcGTA gene expression by regulating the activity of a σ factor through this Rsb system.

In this study, I have investigated the effects of mutations of the *rsb* homologues on RcGTA expression in *R. capsulatus*. I have also looked at the effects on culture viability, motility and colony morphology in the mutant strains. Disruptions were created in alternative σ factors to attempt to identify the protein involved in initiation of RcGTA gene transcription. I have also investigated protein interactions between RsbW and RsbV using *in vitro* pull-downs and bacterial two-hybrid analysis. The results of this study have provided evidence that the RsbW, RsbV and RsbU homologues may play a substantial

role in several physiological processes in *R. capsulatus*, including regulating the expression of the RcGTA gene cluster.

4.2. Material and methods

4.2.1. Bacterial strains, plasmids and culture conditions

The experimental strains, plasmids, and PCR primers used for this study are listed in Tables 4.1, 4.2, and 4.3, respectively. Cultures of *R. capsulatus* and *Escherichia coli* were grown in the same conditions and media, supplemented with appropriate antibiotics, as previously described in Chapter 3. Spectinomycin was supplemented at concentrations of 50 μg ml^{-1} for *E. coli* and 20 μg ml^{-1} for *R. capsulatus*.

Table 4.1. Experimental strains used in this study

Strain	Description	References
R. capsulatus strains		
SB1003	Genome sequenced strain	(Yen & Marrs, 1976; Strnad *et al.*, 2010)
rsbW	SB1003 with disrupted *rsbW*	This work
rsbV	SB1003 with disrupted *rsbV*	This work
rsbVW	SB1003 with disrupted *rsbV* and *rsbW*	This work
rsbU	SB1003 with disrupted *rsbU*	This work
rpoHI	SB1003 with disrupted *rpoHI*	This work
rpoHI-ps	SB1003 with disrupted *rpoHI* maintained under anaerobic, photosynthetic conditions	This work
rpoHII	SB1003 with disrupted *rpoHII*	This work
rpoHI/II	SB1003 with disrupted *rpoHI* and *rpoHII*	This work
SB699	SB1003 with disrupted *rcc00699*	This work
SB2291	SB1003 with disrupted *rcc02291*	This work
SB2724	SB1003 with disrupted *rcc02724*	This work
SB2637	SB1003 with disrupted *rcc02637*	This work
DW5	SB1003 Δ*puhA*	Wong *et al.*, 1996
E. coli strains		
BL21 (DE3)	Host for expression of recombinant proteins	New England Biolabs

BTH101	Reporter strain (*cya*$^-$) for bacterial two-hybrid assay	Karimova *et al.*, 1998

Open reading frames (ORFs) of the putative *rsb* homologues and σ-factors were amplified by PCR from the genome of *R. capsulatus* strain SB1003 and cloned into pGEM-T-Easy according to the manufacturer's guidelines. The genes were disrupted by insertion of a ~1.4 kb SmaI fragment of the KIXX cartridge (Barany, 1985) described in Chapter 3. The *rsbV* (*rcc03323*) and *rsbW* (*rcc03324*) ORFs were amplified using the primers rsbVW-F and rsbVW-R (Table 4.3). The *rsbV* gene was disrupted by insertion at a NruI site located 76 bp into the 348 bp ORF. The KIXX cartridge was inserted into a T4-polymerase blunted BlpI site, located 274 bp into the 492 bp *rsbW* ORF. A double disruption of *rsbVW* was created by replacing a 535 bp NruI/BlpI segment with the KIXX cartridge. The 1230 bp ORF predicted to encode the *rsbU* (*rcc00181*) homologue was amplified using the primers rsbU-F and rsbU-R (Table 4.3) and was disrupted at a MscI site located 307 bp into the gene.

Amplicons of the *R. capsulatus rpoHI* (*rcc02811*) and *rpoHII* (*rcc00458*) genes were amplified using primer rpoHI-F and rpoHI-R, and rpoHII-F and rpoHII-R, respectively (Table 4.3). The 900 bp *rpoHI* ORF was disrupted at a BamHI site located 323 bp from the start of the gene. A 507 bp StuI fragment of the 833 bp *rpoHII* ORF was replaced by the KIXX cartridge. The ORF encoding the putative ECF σ-factor *rcc02291* (570 bp) was amplified using primers 2291-F and 2291-R (Table 4.3) and disrupted by insertion at a StuI site located 133 bp into the gene. The 594 bp ORF *rcc02724* encoding another putative ECF σ-factor was amplified using primers 2724-F and 2724-R (Table

4.3) and disrupted by inserting KIXX into a BsaBI site located 221 bp from the start of the gene. The ORFs *rcc00699* (545 bp) and *rcc02637* (585 bp) encoding putative σ^{24} ECF sigma factors were amplified using primers 699-F and 699-R, and 2637-F and 2637-R, respectively (Table 4.3). The KIXX cartridge was inserted into a StuI site 376 bp in *rcc00699* and an AfeI site located 176 bp from the start of *rcc02637*. Disruptions were not made for the major vegetative σ-factor, *rpoD* (*rcc3054*), or the nitrogen fixation σ-factor, *rpoN* (*rcc00568*). A separate *rpoHI* disruption using a 2-kb spectinomycin resistance-encoding cassette (omega) (Prentki & Krisch, 1984) was constructed to allow creation of an *rpoHI/II* double mutant strain.

Table 4.2. Experimental plasmids used in this study

Plasmids	Description	References
pRK767	Broad host range plasmid	Gill & Warren, 1988
p*rsbW*	*rsbW* and 91 bp of 5' sequence of modified *rsbV* in a T4 polymerase-blunted KpnI site of pRK767	This study
p*rsbV*	*rsbV* and 91 bp of 5' sequence in a T4 polymerase-blunted KpnI site of pRK767	This study
p*rsbVW*	*rsbV* and *rsbW* and 91 bp of 5' sequence in a T4 polymerase-blunted KpnI site of pRK767	This study
p*rsbU*	*rsbU* and 220 bp of 5' sequence in a KpnI site in pRK767	This study
pX2	RcGTA *orfg2'::'lacZ* fusion as a PstI/BamHI fragment in pXCA601	Hynes *et al.*, 2012
pX2NP	Promoterless RcGTA *orfg2'::'lacZ* fusion as a PstI/BamHI fragment in pXCA601	Hynes *et al.*, 2012
p*Xg2Δp*	RcGTA *orfg2'::'lacZ* fusion with deletion of bp -129 to -100 from the *g1* start codon as a PstI/BamHI fragment in pXCA601	This study
p*Xg2Δs*	RcGTA *orfg2'::'lacZ* fusion with deletion of bp -73 to -46 from the *g1* start codon as a PstI/BamHI fragment in pXCA601	This study
pET15W	Expression vector with IPTG-inducible T7 promoter for expression of N-terminal 6x-histidine tagged RsbW	This study

pET15V	Expression vector with IPTG-inducible T7 promoter for expression of N-terminal 6x-histidine tagged RsbV	This study
pET26W	Expression vector with IPTG-inducible T7 promoter for expression of C-terminal 6x-histidine tagged RsbW	This study
pET15H	Expression vector with IPTG-inducible T7 promoter for expression of N-terminal 6x-histidine tagged RpoHI	This study
pKNT25	Derivative of pSU40 containing a multi-cloning sequence 5' of the T25 fragment (first 224 amino acids of the adenylate cyclase of *Bordetella pertussis*)	Karimova *et al.*, 1998
pKT25	Derivative of pSU40 containing a multi-cloning sequence 3' of the T25 fragment (first 224 amino acids of the adenylate cyclase of *Bordetella pertussis*)	Karimova *et al.*, 1998
pUT18	Derivative of pUC19 containing a multi-cloning sequence 5' of the T18 fragment (amino acids 225-399 of the adenylate cyclase of *B. pertussis*)	Karimova *et al.*, 1998
pUT18c	Derivative of pUC19 containing a multi-cloning sequence 3' of the T18 fragment (amino acids 225-399 of the adenylate cyclase of *B. pertussis*)	Karimova *et al.*, 1998
pKT25-zip	pKT25 containing an in-frame fusion of T25 fragment to the leucine zipper of GCN4	Karimova *et al.*, 1998
pUT18c-zip	pUT18c containing an in-frame fusion of T18 fragment to the leucine zipper of GCN4	Karimova *et al.*, 1998
pKNT-rsbW	pKNT containing an in-frame fusion of *rsbW* to the N-terminus of the T25 fragment	This study
pKT-rsbW	pKT containing an in-frame fusion of *rsbW* to the C-terminus of the T25 fragment	This study
pUT18-rsbW	pUT18 containing an in-frame fusion of *rsbW* to the N-terminus of the T18 fragment	This study
pUT18c-rsbW	pUT18c containing an in-frame fusion of *rsbW* to the C-terminus of the T18 fragment	This study
pKNT-rsbV	pKNT containing an in-frame fusion of *rsbV* to the N-terminus of the T25 fragment	This study
pKT-rsbV	pKT containing an in-frame fusion of *rsbV* to the C-terminus of the T25 fragment	This study
pUT18-rsbV	pUT18 containing an in-frame fusion of *rsbV* to the N-terminus of the T18 fragment	This study
pUT18c-rsbV	pUT18c containing an in-frame fusion of *rsbV* to the C-terminus of the T18 fragment	This study
pKNT-rpoD	pKNT containing an in-frame fusion of *rpoD* to	This study

	the N-terminus of the T25 fragment	
pKT-rpoD	pKT containing an in-frame fusion of *rpoD* to the C-terminus of the T25 fragment	This study
pUT18-rpoD	pUT18 containing an in-frame fusion of *rpoD* to the N-terminus of the T18 fragment	This study
pUT18c-rpoD	pUT18c containing an in-frame fusion of *rpoD* to the C-terminus of the T18 fragment	This study
pKNT-rpoHI	pKNT containing an in-frame fusion of *rpoHI* to the N-terminus of the T25 fragment	This study
pKT-rpoHI	pKT containing an in-frame fusion of *rpoHI* to the C-terminus of the T25 fragment	This study
pUT18-rpoHI	pUT18 containing an in-frame fusion of *rpoHI* to the N-terminus of the T18 fragment	This study
pUT18c-rpoHI	pUT18c containing an in-frame fusion of *rpoHI* to the C-terminus of the T18 fragment	This study
pKNT-699	pKNT containing an in-frame fusion of *rcc00699* to the N-terminus of the T25 fragment	This study
pKT-699	pKT containing an in-frame fusion of *rcc00699* to the C-terminus of the T25 fragment	This study
pUT18-699	pUT18 containing an in-frame fusion of *rcc00699* to the N-terminus of the T18 fragment	This study
pUT18c-699	pUT18c containing an in-frame fusion of *rcc00699* to the C-terminus of the T18 fragment	This study
pKNT-2637	pKNT containing an in-frame fusion of *rcc02637* to the N-terminus of the T25 fragment	This study
pKT-2637	pKT containing an in-frame fusion of *rcc02637* to the C-terminus of the T25 fragment	This study
pUT18-2637	pUT18 containing an in-frame fusion of *rcc02637* to the N-terminus of the T18 fragment	This study
pUT18c-2637	pUT18c containing an in-frame fusion of *rcc02637* to the C-terminus of the T18 fragment	This study

Conjugation of the plasmids containing disrupted versions of the genes to *R. capsulatus*, followed by GTA transfer into the chromosome of SB1003, were carried out as previously described in Chapter 3. The resulting kanamycin and kanamycin/spectinomycin resistant strains (Table 4.1) were confirmed to contain the disrupted genes by PCR using the original amplification primers (Table 4.3).

In trans complementation was performed on mutant strains shown to have effects on RcGTA production. Wild type genes with their native upstream sequences were placed on the low copy, broad host range plasmid, pRK767 (Gill & Warren, 1988). A wild type fragment of *rsbV* and *rsbW* was amplified using primers rsbVcF and rsbVW-R (Table 4.3). Primers rsbVcF and Anti-anti-R (Table 4.3) were used to amplify the wild type *rsbV* fragment. The *rsbW* complement sequence contained an in-frame deletion of the majority of *rsbV*, replacing bp 24 to bp 272 with a KpnI site. This was created by joining 2 fragments, amplified with rsbVcF and rsbVdR, and rsbVdF and rsbVW-R, via a primer-embedded KpnI site. The complementation vectors (Table 4.2) were conjugated to *R. capsulatus* using *E. coli* S17-1 (Simon *et al.*, 1983).

Table 4.3. Primers used in this study

Primer	Sequence (5' – 3')[1]	References
rsbVW-F	GGCCAGATAGTGCTTTTTCG	This study
rsbVW-R	CCCGAGTCCACAGATTGTTT	This study
rsbU-F	GCGGTACCTTCCCAATCGATCATAGTT	This study
rsbU-R	ATGGTACCCTGCGCGAGCCTTAA	This study
rpoHI-F	GGCATCCTGAAGATCACCTC	This study
rpoHI-R	GCTCTCATGCAGGGTCAGAT	This study
rpoHII-F	TGACGATCCACAGATCGAAG	This study
rpoHII-R	CGCCAGTTCATAACCCATCT	This study
2291-F	AGATGGAAGGCGTGAAACAG	This study
2291-R	GCCAGGAGGACAACAGGTAA	This study
2724-F	CACGACCTTTCCCTGACAAT	This study
2724-R	CTGATCCCACGTGACCTTCT	This study
699-F	GAGACCATGTCCGAGACCAG	This study
699-R	GCGATCAGCCAAAGGTAAAT	This study
2637-F	CAGATCCGCATCTTCGAACT	This study
2637-R	AGGGTTTCGGCCAGTCTAAC	This study
rsbVcF	GCGCATCGGTCTTCCTTCATCATT	This study
rsbVdF	TAGGTACCGCGTGATTCCGCATAAAG	This study
rsbVdR	TAGGTACCCGGATGGATTCCGTTTTC	This study
GTA-F1	CGGCTGCAGACCGATCCGG	Hynes *et al.*, 2012

GTA-F2	ATACTGCAGCATGGACATGGGGTTCAA	Hynes *et al.*, 2012
GTA-R1	AGGATCCACGTCGCGCACCTGAT	Hynes *et al.*, 2012
GTA-DP-F	GCGGTACCTAAGGCATGCTAGGAGAGG	This study
GTA-DP-R	TCGGTACCGTGCTATATTCAGGGTTGCA	This study
GTA-DS-F	GCGGTACCCTTGCCCACCTCTCCTA	This study
GTA-DS-R	ATGGTACCTCATTTCGCTCGTGCGG	This study
Anti-S-F	GACTATCATATGTTGGCTGATCGTCCGCA	This study
Anti-S-R	TGGATCCGTTACCCGTTGCAGTTCC	This study
Anti-anti-F	GGCTGGCATATGAATCTTTATGCGGAATC	This study
Anti-anti-R	AGGATCCTCAGCCAACATGGCGC	This study
Anti-SC-F	GATCCATGGATGTTGGCTGATCGTCC	This study
Anti-SC-R	ACTCGAGACTGACGGCGCCGG	This study
RpoH1-F	GCACTATCATATGTCGAGCTATGCCAACCT	This study
RpoH1-R	AACTATCATATGGCTCTCATGCAGGGTCAG	This study
AS-AF	ATGGTACCTATGTTGGCTGATCGTCCGC	This study
AS-AR	TGGTACCGGACTGACGGCGCCGG	This study
AAS-AF	CGGGTACCGATGAATCTTTATGCGGAATC	This study
AAS-AR	TGGTACCGGGCCAACATGGCGCAG	This study
rpoD-F	ATGGTACCTATGGCCGCCAAGGACATC	This study
rpoD-R	ATGGTACCGGCTGGTCGAGGAAGCT	This study
rpoH-AF	GCGGTACCTATGTCGAGCTATGCCAACCT	This study
rpoH-AR	ATGGTACCGGGCCCGGCAGGCTCAT	This study
2637-AF	GTGGTACCTATGGAGATGGCCTTCGACG	This study
2637-AR	ATGGTACCGGGTCATGGCCATACCC	This study
699-AF	TGGTACCTATGGCGGATGCGGGAC	This study
699-AR	ATGGTACCGGTCCTTCCAGACACTCCC	This study

[1]Underlined sequences indicate restriction sites used for cloning.

4.2.2. Gene transfer bioassays

Gene transfer bioassays were carried out as previously described in Chapter 3, section 3.2.2 of this thesis.

4.2.3. Western blotting

Western blots targeting the RcGTA major capsid protein (~32 kDa) were carried out on the same cultures used for RcGTA activity assays as previously described in Chapter 3, section 3.2.3 of this thesis.

4.2.4. Motility assays

Motility assays were carried out as previously described in Chapter 3, section 3.2.4 of this thesis.

4.2.5. Viable Cell Counts

Viable cell counts were performed as previously described in Chapter 3, section 3.2.5 of this thesis.

4.2.6. β-galactosidase reporter fusion constructs

In-frame fusions of RcGTA *orfg2* to the *E. coli lacZ* gene were constructed using PstI/BamHI fragments cloned into the promoter probe vector pXCA601 vector (Adams *et al.*, 1989). Fragments 2 (pX2) and 2NP (pX2NP) were amplified by PCR using primers GTA-F1 and GTA-R1, and GTA-F2 and GTA-R1, respectively (Table 4.3). Fragments 2.1 and 2.2 were amplified using primers GTA-F1 and GTA-DP-R, and GTA-DP-F and GTA-R1, respectively (Table 4.3). Fragment *g2Δp* (pX2Δp) was created by ligating 2.1 and 2.2 via a primer-embedded KpnI restriction site, resulting in a deletion of the sequence from -129 to -100 5' of RcGTA *orfg1* (Table 4.2). Fragments 2.3 and 2.4 were amplified using GTA-F1 and GTA-DS-R, and GTA-DS-F and GTA-R1, respectively (Table 4.3). The fragment *g2Δs* was made by combining 2.3 and 2.4 via a primer-embedded KpnI restriction site, resulting in a deletion of the sequence from -73 to -46 5' of *orfg1* (Table 4.2). All in-frame fusions were confirmed by sequencing (The Centre for Applied Genomics, Toronto, ON), and the plasmids were transferred to *R. capsulatus* strains by conjugation using *E. coli* S17-1 (Simon *et al.*, 1983).

Strains of *R. capsulatus* containing the fusion constructs listed in Table 4.2 were grown in conditions identical to those for RcGTA activity assays (Chapter 3, section 3.2.2). Cells were permeabilized for 15 minutes using 15% (v/v) isopropyl alcohol and washed using Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 1 mM $MgSO_4$, 10 mM KCl, 50 mM β-mercaptoethanol; pH 7) (Miller, 1992). The substrate, fluorescein di-β-D-galactopyranoside (FDG) (Sigma-Aldrich), was dissolved in H_2O:DMSO:ethanol (8:1:1). The cells were resuspended in Z buffer with substrate at a final concentration of 0.1 mg ml^{-1} and incubated for 1 hour at room temperature. The samples were then diluted 1:200 in Z buffer and analyzed by flow cytometry capturing 100000 events. The mean sample fluorescence was obtained from gated cells from two biological replicates.

4.2.7. Expression and purification of recombinant proteins from *E. coli*

For expression of recombinant N-terminal 6x-histidine tagged proteins, *rsbW* and *rsbV* were independently cloned as NdeI/BamHI fragments into the pET15b vector (Novagen, Darmstadt, Germany), using primers Anti-S-F and Anti-S-R, and Anti-anti-F and Anti-anti-R, respectively (Table 4.3). This resulted in RsbW and RsbV protein coding sequences in-frame with an N-terminal 6x-histidine tag. A C-terminal 6x-histidine tagged sequence of RsbW was also created using the primers Anti-SC-F and Anti-SC-R (Table 4.3) and cloned as a NcoI/XhoI fragment into the pET26b vector (Novagen, Darmstadt, Germany). Primers RpoH1-F and RpoH1-R (Table 4.3) were used to amplify *rpoHI* as an NdeI fragment for cloning in-frame to an N-terminal 6x-histidine tag in pET15b. The plasmids, pET15W, pET15V, pET26W and pET15H (Table 4.2), were sequenced to

confirm the *R. capsulatus* sequences were in-frame with the histidine tags and then transformed into *E. coli* BL21(DE3) (New England Biolabs).

Overnight starter cultures were used to inoculate 200 ml of LB broth containing the appropriate antibiotics, followed by incubation for 1 hour at 37°C with shaking at 250 rpm. Expression of the recombinant proteins was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM followed by growth at 37°C for 4 hours with shaking at 250 rpm. Cell pellets of induced cultures were resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 0.1% (v/v), Benzonase® nuclease (Qiagen), 1 mg ml^{-1} lysozyme (w/v); pH 8) and incubated on ice for 30 minutes. The lysates were centrifuged at 14000 x *g* for 30 minutes at 4°C and supernatants were mixed 4:1 (v:v) with Ni-NTA agarose (Qiagen) and incubated at 4°C with shaking at 200 rpm for 1 hour. The samples were loaded into polypropylene columns, washed twice with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole; pH 8) and the fusion proteins eluted in 1 ml aliquots of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole; pH 8). The purified proteins were dialyzed into a coupling buffer (20 mM sodium phosphate buffer, 500 mM NaCl; pH 7.5) and quantified using a ND-1000 Nanodrop spectrophotometer.

4.2.8. In-gel digestion and peptide extraction for LC-MS/MS sequencing

Purified recombinant protein samples were mixed with 3x SDS-PAGE sample buffer, boiled for 5 minutes at 98°C, and separated by electrophoresis on a 10% SDS-PAGE gel. The gels were stained with Coomassie Blue [0.25% (w/v) Coomassie Brilliant Blue R-250 in methanol:H_2O:acetic acid (5:4:1)] for 30 minutes, destained

[methanol:H_2O:acetic acid (5:4:1)] and recombinant protein bands of predicted sizes were cut out using a clean scalpel. The gel slices were washed first with water, followed by 100 mM NH_4HCO_3, then acetonitrile, with samples being vortexed for 10 minutes, centrifuged at 3000 x *g* and supernatants decanted after each wash step. The samples were dried in a vacuum centrifuge for 5 minutes before adding a sufficient amount of 10 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 to cover the gel slices. After incubation for 45 minutes at 56°C, the samples were centrifuged at 3000 x *g* and the supernatant decanted. The solution was replaced by 55 mM iodoacetamide in 100 mM NH_4HCO_3 and the samples incubated in the dark at room temperature for 30 minutes with occasional vortexing. The samples were centrifuged at 3000 x *g* and all liquid was removed by aspiration. The samples were washed in 100 mM NH_4HCO_3 with vortexing for 10 minutes followed by centrifuging at 3000 x *g* and removal of the supernatant. This wash procedure was repeated once with acetonitrile and twice with 50% (v/v) acetonitrile. The samples were vacuum centrifuged for 15 minutes before the addition of sequencing grade trypsin (12 ng μl^{-1}) in trypsin digestion buffer (Promega). The samples were sealed and incubated overnight at 37°C. After treatment with 5% formic acid (v/v) with vortexing, the samples were centrifuged at 3000 x *g* and supernatants collected in a separate tube. This extraction step was repeated using 1% (v/v) formic acid in 5% acetonitrile (v/v), 1% formic acid in 60% acetonitrile, and 1% formic acid in 99% acetonitrile. The supernatants from each of these extractions were collected together in one tube and vacuum centrifuged. The dried extracts were sequenced by LC-MS/MS at the Genomic and Proteomic (GaP) facility at Memorial University.

4.2.9. *In vitro* pull-down assays

In vitro interaction assays were carried out by separately conjugating 50 μg of 6x-His-RsbW and RsbW-6x-His to NHS-activated beads (GE Life Sciences) according to the manufacturer's guidelines. The conjugated beads were washed several times with 100 mM Tris-HCl (pH 8.0) then resuspended as a 50% (v/v) slurry in the same solution. A sub-sample of conjugated bead slurry was resuspended in a binding buffer [10 mM Tris-HCl (pH 8.0), 200 mM NaCl, 5% (v/v) glycerol, 0.5 mM DTT, and 0.5% (v/v) triton X-100] and either 6x-His-RsbV, 6x-His-RpoHI, or control protein (chicken egg white lysozyme) was added to a final concentration of ~1 μM. The interaction was incubated on ice for 30 minutes with occasional mixing before adding 0.5 ml of binding buffer. The beads were allowed to sediment by gravity and the supernatant was removed. This washing step with 0.5 ml of binding buffer was repeated 3 times to remove all non-bound proteins. The beads were resuspended in 30 μl of 3x SDS-PAGE buffer, boiled for 5 minutes at 98°C, and 20 μl of the sample was analyzed on a 10% SDS-PAGE gel. To control for specific binding between recombinant fusion proteins, additional control reactions were performed. First, non-conjugated beads were treated with 100 mM Tris-HCl (pH 8.0) and then incubated with test proteins to ensure adequate blocking of bead active sites. As well, conjugated 6x-His-RsbW and RsbW-6x-His were incubated with chicken egg white lysozyme (Sigma) to ensure specific interactions between the experimental test proteins.

4.2.10. Bacterial two-hybrid assay

Bacterial two-hybrid analyses for determining protein interactions were carried out as described by Karimova *et al.* (1998) using the bacterial adenylate cyclase-based two hybrid, or BACTH, system (EUROMEDEX, Souffelweyersheim, France). Primers AS-AF and AS-AR, and AAS-AF and AAS-AR (Table 4.3), were used to amplify *rsbW* and *rsbV* by PCR, respectively. The *rpoD* and *rpoHI* σ-factor-encoding genes were amplified using rpoD-F/rpoD-R and rpoH-AF/rpoH-AR, respectively (Table 4.3). Putative ECF-σ-factor-encoding genes *rcc02637* and *rcc00699* were amplified using 2637-AF and 2637-AR, and 699-AF and 699-AR, respectively (Table 4.3). All amplicons were cloned as KpnI fragments into all 4 BACTH vectors: pKNT25, pKT25, pUT18 and pUT18c (Table 4.2). All pairwise combinations of bait (*rsbW*) and prey (*rsbV*, *rpoD*, *rpoHI*, σ^{2637} and σ^{699}) recombinant vectors were co-transformed into *cya*$^-$ *E. coli* strain BTH101 and plated on LB agar supplemented with 100 μg ml^{-1} ampicillin, 50 μg ml^{-1} kanamycin, 40 μg ml^{-1} 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and 0.5 mM IPTG. Positive control vectors encoding interacting fragments of a leucine zipper protein, pKT25-zip and pUT18C-zip (Table 4.2), were also co-transformed. All plates were incubated for 48 hours at 30°C.

For quantitative determination of β-galactosidase activity, 3 biological replicate co-transformants were picked for each interaction to inoculate fresh LB broth containing appropriate antibiotics and 0.5 mM IPTG. Overnight cultures grown at 37°C were diluted 1:5 in LB broth and OD_{600} was determined. The cells were permeabilized with one drop (~30 μl) of 0.1% SDS and 2 drops of chloroform then mixed in a 1:1 ratio with PM2

buffer (70 mM Na_2HPO_4, 30 mM NaH_2PO_4, 1 mM $MgSO_4$, 0.2 mM $MnSO_4$; pH 7) containing 100 mM 2-mercaptoethanol. The cells were incubated for 5 minutes at 28°C and one volume of 0.4% *o*-nitrophenol-β-D-galactopyranoside (ONPG) substrate was added to 4 volumes of cells in PM2 buffer. After sufficient colour development, the reaction was stopped by addition of 2 volumes of 1 M $NaHCO_3$. The OD_{420} and OD_{550} were obtained for each sample and β-galactosidase activity was calculated as units mg^{-1} of dry weight bacteria.

4.3. Results

4.3.1. Bioinformatic analyses and genomic context of *rsb* homologues

The RsbW homologue in *R. capsulatus* possesses a conserved HATPase domain and has 25 to 30% amino acid sequence identity with *B. subtilis* and *B. cereus* anti-σ factors RsbW and SpoIIAB, respectively. The RsbV homologue has a conserved STAS domain commonly found in anti-anti-σ factors, and shares 22-30% amino acid sequence identity with RsbV and SpoIIAA proteins in *Bacillus* spp. A BLAST search of the NCBI GenBank database revealed that homologues of the *R. capsulatus* RsbW and RsbV are found in other members of the Rhodobacterales order in the class α-proteobacteria. Very few homologues were found in other proteobacterial lineages, and the next most conserved sequences outside the α-proteobacteria were in gram-positive species in the phylum Firmicutes. The RsbU protein in *R. capsulatus* shares 27% sequence identity with the orthologous protein in *B. cereus* and possess a conserved phosphorelay REC domain and a PP2C phosphatase superfamily domain.

An analysis of orthologous neighbourhood regions using the IMG database [http://img.jgi.doe.gov/cgi-bin/w/main.cgi; Markowitz *et al.*, 2011] shows that the *R. capsulatus rsbV* and *rsbW* genes are conserved in two-gene operons in some related species (Fig. 4.1A). Some species, such as *Rhodopseudomonas palustris*, also have an *rsbU* homologue in a 3 gene operon with *rsbV* and *rsbW*, which is how these genes are commonly organized in studied gram-positive *Bacillus* and *Staphylococcus* spp. (Hecker *et al.*, 2007). In more distantly related species, such as the δ-proteobacterium *Pelobacter*, the *rsb* genes are encoded along with a putative cognate σ factor (σ^{24}).

4.3.2. *rsb* mutant phenotypes

Insertional disruptions of the *rsb* homologues in *R. capsulatus* demonstrated that the proteins encoded by several of these genes are involved in controlling the expression of the RcGTA genes. The *rsbW* mutant showed a significant (p-value <0.05) increase in RcGTA activity of 2.85-fold relative to SB1003 (Fig. 4.2A), which agreed with an increase in RcGTA capsid protein levels inside and outside the cells (Fig. 4.2B). This mutant was also found to be non-motile (Fig. 4.3), but had no observable differences in viable cell number or colony morphology relative to SB1003 (Fig. 4.4 and Fig. 4.5). Complementation with the wild type *rsbW* with its native upstream region, and an in-frame deletion of *rsbV*, did not restore RcGTA activity or capsid levels to wild type (Fig. 4.2), but did restore culture motility (Fig. 4.3). However, complementation of the *rsbW* mutant with both *rsbV* and *rsbW* resulted in wild type RcGTA capsid levels and gene transfer activity (Fig. 4.2).

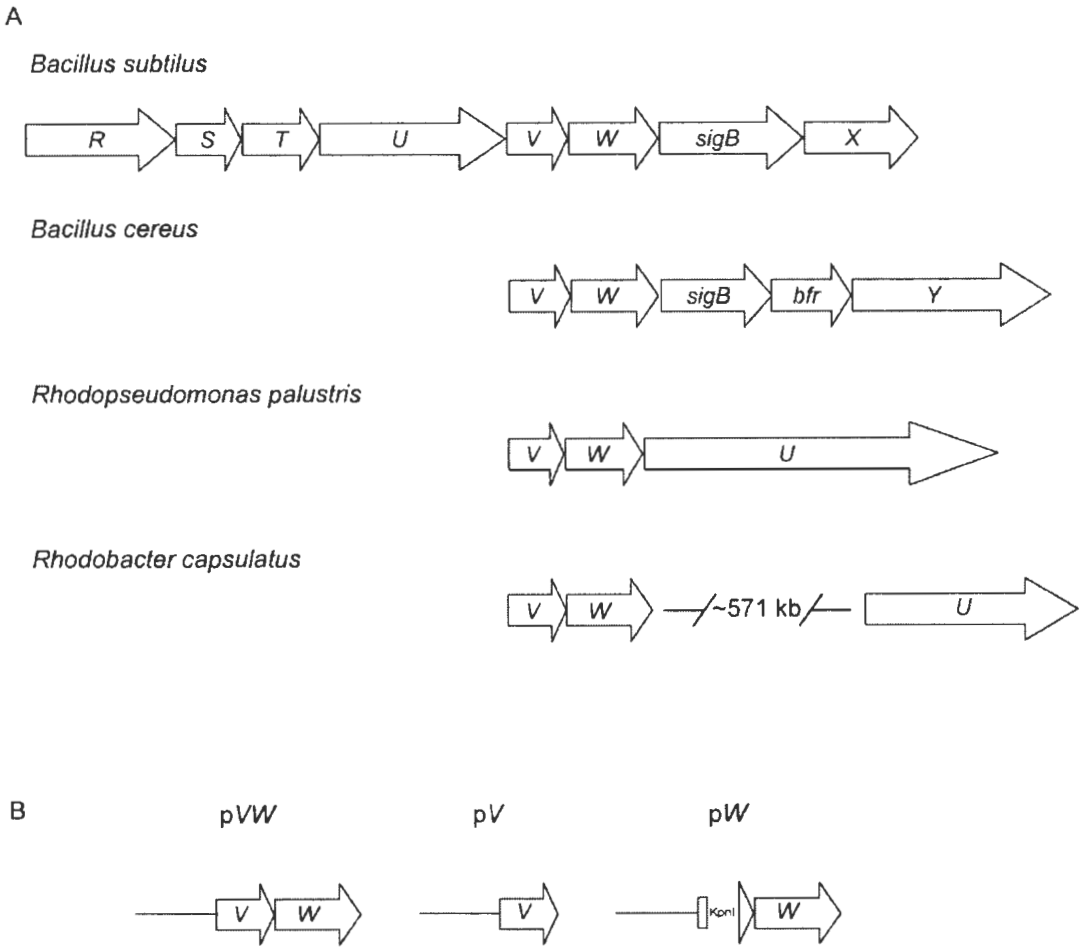


Figure 4.1. Genomic arrangements of *rsb* genes. A. The *rsbV* and *rsbW* genes are found in an 8 gene operon with *rsbRSTY, sigB* and *rsbX* in *B. subtilis*. *B. cereus* lacks *rsb* genes upstream of *rsbV* and a bacterioferritin (*bfr*) gene is found between *sigB* and the PP2C serine phosphatase in this system, *rsbU*. In *R. palustris*, *rsbVWU* are organized in a 3-gene operon while in *R. capsulatus*, only *rsbV* and *rsbW* are found together with *rsbU* located elsewhere in the genome. B. The p*V*, p*W* and p*VW* plasmids used for genetic complementation. All vectors possess the native upstream sequences and ORF coding sequences except p*W* which has an in-frame deletion of >70% of the *rsbV* coding sequence replaced by a KpnI restriction site.

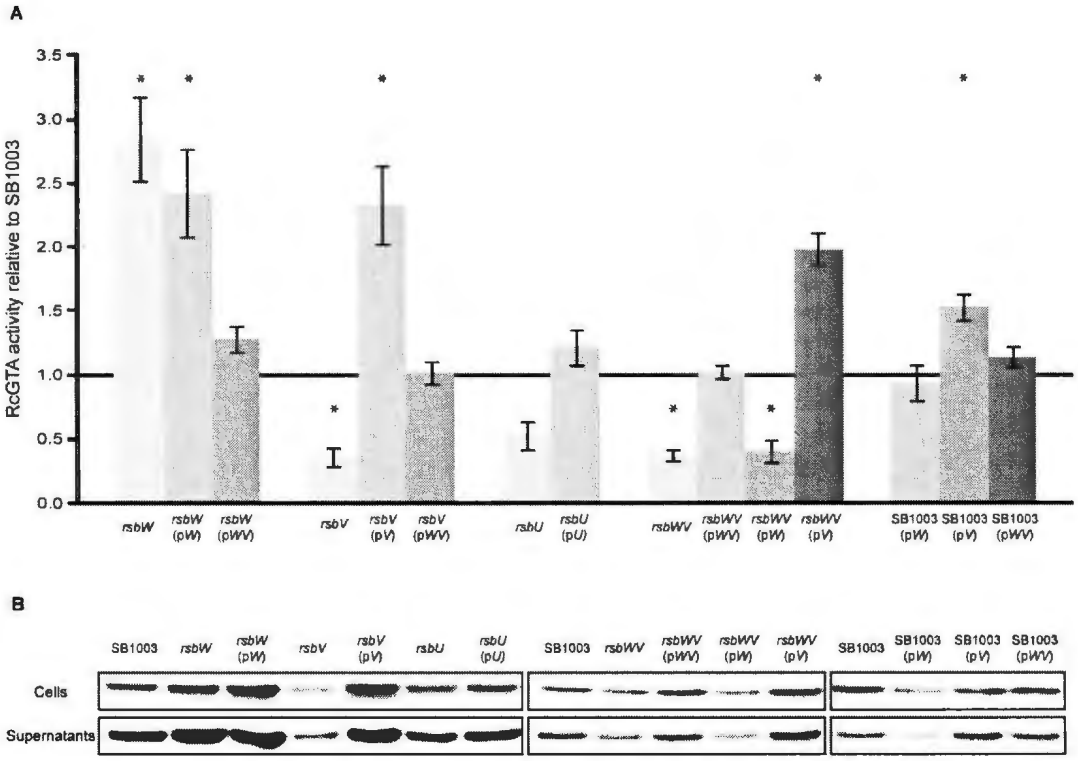


Figure 4.2. RcGTA gene transfer activity and gene expression in *rsb* mutants and SB1003 complemented strains. A. The ratio of gene transfer activity for each indicated strain relative to the parental strain, SB1003. The gene transfer activity was determined as an average relative to SB1003 in 3 replicate bioassays and the error bars represent the standard deviation. RcGTA production levels that differed significantly from the wild type were identified by analysis of variance (ANOVA) and are indicated by an asterisk (*$p<0.05$). P-values for all pair-wise comparisons are reported in Supplemental Table S4.1. B. Western blots for detection of the RcGTA major capsid protein in the cells and culture supernatants of indicated strains. Blots were performed on all replicate gene transfer bioassay cultures (in A) and one representative set of blots is shown.

The *rsbV* and *rsbU* mutants had similar phenotypes, with both strains having lower RcGTA activity. However, only the *rsbV* mutant was found to have a considerable

decrease in capsid protein levels relative to wild type (Fig. 4.2). The decrease in gene transfer activity in the *rsbU* mutant was not found to be as statistically significant (p-value = 0.08). Both strains showed a reproducible decrease in viable cells per ml in late stationary cultures (Fig. 4.4), but only the *rsbU* mutant was found to be statistically different (p-value <0.05). Complementation of *rsbU* restored gene transfer activity and the number of viable cells in stationary phase to wild type (Fig. 4.2 and Fig. 4.4). Complementation of the *rsbV* mutant with *rsbV* resulted in a significant increase in RcGTA activity (p-value <0.05) and capsid levels, similar to the level observed for *rsbW* and *rsbW* (p*W*) strains (Fig. 4.2), while complementation with both *rsbV* and *rsbW* genes restored the strain to wild type levels. A statistically significant increase in gene transfer activity (p-value <0.05) and capsid levels were also observed in a SB1003 strain carrying an additional copy of the *rsbV* gene on a plasmid (Fig. 4.2). Heterogeneous colony morphologies were also noted on plates of serially diluted late stationary phase cultures of the *rsbV* and *rsbU* mutants, and ~25% of colonies were found to be undulate and flattened instead of the circular and slightly raised wild type phenotype (Fig. 4.5). These unusual colonies could generate *de novo* photosynthetic cultures that gave rise to both normal and unusual colonies with approximately the same percentage. Complemented strains *rsbU* (p*U*), *rsbV* (p*V*), and *rsbV* (p*VW*) also generated this sub-population of unusual colonies.

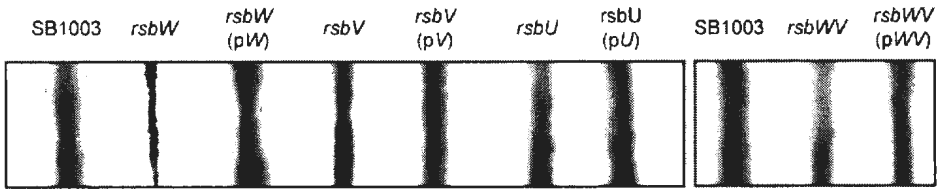


Figure 4.3. Effects of *rsb* mutations on *R. capsulatus* motility. Motility assays were carried out in soft agar stabs and the distance of growth outwards from the center stab line indicates the relative motility of the strain.

The double *rsbVW* mutant had an unexpected phenotype. RcGTA activity resembled that of the *rsbV* and *rsbU* mutants and not the *rsbW* mutant (Fig. 4.2). The *rsbVW* mutant was motile and showed a significant decrease in viable cells (Fig. 4.3 and 4.4). The strain also produced the unusual colony morphology phenotype (Fig. 4.5), which was not eliminated by complementation. Introduction of both *rsbV* and *rsbW* on a plasmid restored wild type RcGTA activity and capsid levels (Fig. 4.2), while complementation with only *rsbW* did not change the mutant phenotype (Fig 4.2, Fig. 4.4). The *rsbVW* (p*V*) strain had significantly increased RcGTA activity (p-value <0.05) and capsid levels, similar to the *rsbV* (p*V*) and SB1003 (p*V*) strains. Viable cell numbers of *rsbVW* (p*VW*) and *rsbVW* (p*V*) were not significantly different from wild type (p-value >0.05) (Fig. 4.4).

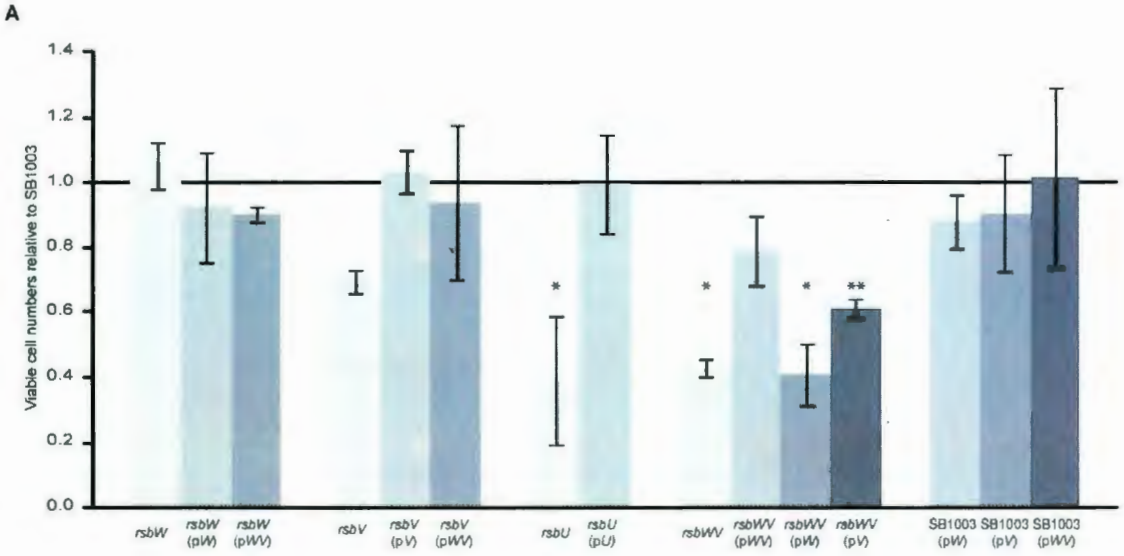


Figure 4.4. Effects of *rsb* mutations and introduction of plasmid-borne *rsb* genes on colony forming unit numbers in stationary phase. The ratios of viable cells per ml relative to SB1003 are shown. Ratios were determined as an average of 3 biological replicates with the same cultures used for the RcGTA gene transfer activity assays and western blots. Error bars represent standard deviation, and statistically significant differences were identified by analysis of variance (ANOVA) and are indicated by an asterisk (*p<0.05, **p<0.01). All p-values are reported in Supplemental Table S4.2.

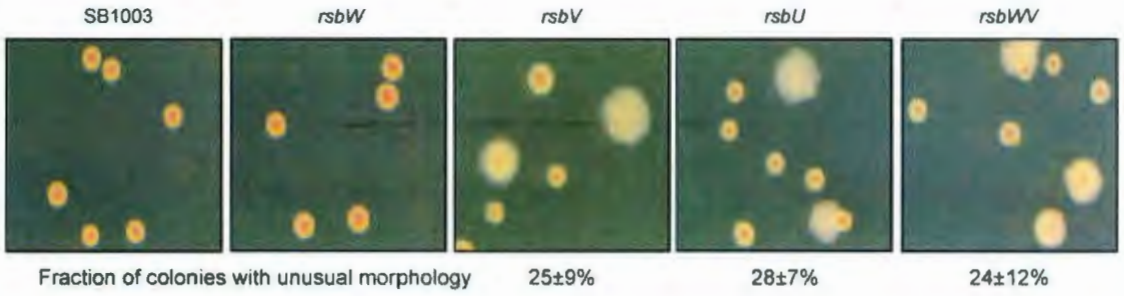


Figure 4.5. Colony morphology in *rsb* mutants. Stationary phase cultures that were used for determining cell viability yielded noticeable difference in colony morphology for *rsbV*, *rsbU* and *rsbVW* strains

compared to SB1003 and *rsbW*. The fraction of colonies with the unusual morphology indicated was determined by counting and calculating the fraction of unusual colonies over total colonies.

4.3.3. σ factor gene disruptions in *R. capsulatus*

To determine which σ factor was responsible for targeting RNAP to the promoter of the RcGTA gene cluster, I attempted to make genetic disruptions of all putative *R. capsulatus* σ factor-encoding genes (Table 2.1). Two exceptions were the nitrogen-fixation σ^{54}-encoding *rpoN*, which is not expressed under RcGTA producing conditions (Chapter 2; Mercer *et al.*, 2010), and the essential major housekeeping σ^{70}-encoding *rpoD*. Disruptions of ORFs *rcc00458* (*rpoHII*), *rcc02291* and *rcc02724* produced viable strains that were not affected for RcGTA activity (data not shown) and their disruptions were confirmed by PCR. Attempts to create mutants of *rcc00699* and *rcc02637* resulted in putative mutants that were resistant to kanamycin, although the insertional disruptions could not be confirmed by PCR.

A disruption of the ORF predicted to encode the RpoHI σ factor, which was confirmed by PCR, initially demonstrated a significant loss (p-value < 0.05) of nearly 90% RcGTA production (Fig. 4.6A). However, this strain displayed some other behaviours that were indications of more general problems. It demonstrated a prolonged lag phase before entering exponential growth although it eventually did reach the same stationary phase culture density as wild type (Fig. 4.6B). While the *rpoHI* mutant had equivalent viable cell numbers in the logarithmic phase, there was a ~65% drop in stationary phase viable cell numbers relative to SB1003 (Fig. 4.6C). Additionally, the reduced RcGTA activity (Fig. 4.6A) did not match with the extracellular capsid levels in

the stationary phase (Fig. 4.6D). In the related species *R. sphaeroides*, RpoHI has an overlapping regulon with RpoHII in response to photooxidative and heat stress (Green & Donohue, 2006). To determine whether the residual RcGTA activity and stationary phase capsid protein levels was due to redundancy in regulatory function between the two paralogous proteins, a strain was generated in which both *rpoHI* and *rpoHII* were disrupted. The phenotype of the *rpoHI/II* mutant was nearly identical to that of the individual *rpoHI* mutant strain (Fig. 4.6). Also, due to its role in response to photooxidative stress in *R. sphaeroides*, a new *rpoHI* mutant strain, *rpoHI-ps*, was created and maintained completely under anaerobic phototrophic conditions. The gene disruptions were confirmed by PCR but, unexpectedly, this strain showed no differences in growth, RcGTA activity, or viable cell numbers in logarithmic or stationary phases (Fig. 4.6). When the *rsbV* and *rsbW* mutants were generated under these conditions and treated in the same way, there were no differences in phenotypes from the original mutant strains exposed to aerobic conditions (data not shown).

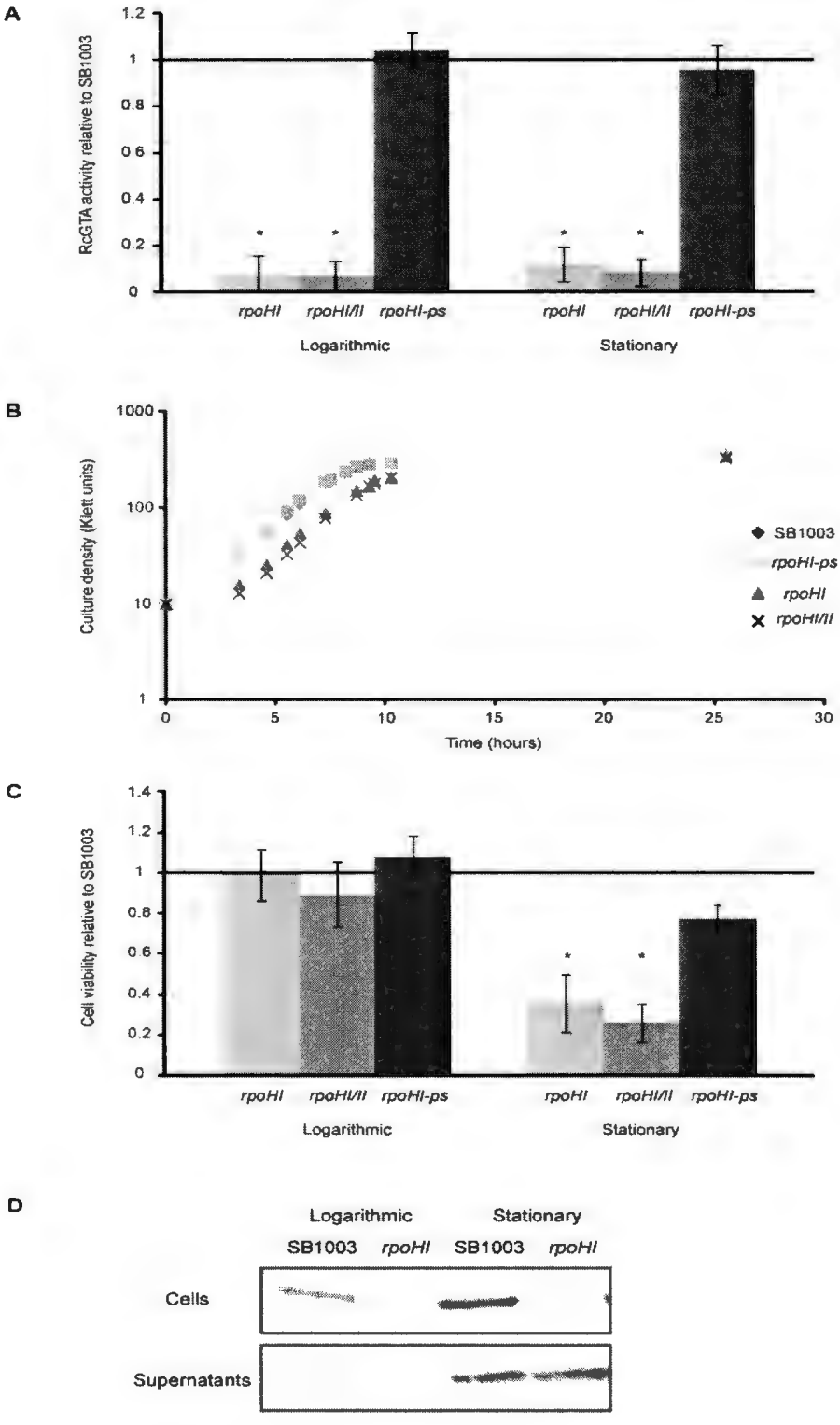

A
RcGTA activity relative to SB1003
1.2
1
0.8
0.6
0.4
0.2
0
rpoHI
rpoHI/II
rpoHI-ps
rpoHI
rpoHI/II
rpoHI-ps
Logarithmic
Stationary
B
Culture density (Klett units)
1000
100
10
1
0
5
10
15
20
25
30
Time (hours)
SB1003
rpoHI-ps
rpoHI
rpoHI/II
C
Cell viability relative to SB1003
1.4
1.2
1
0.8
0.6
0.4
0.2
0
rpoHI
rpoHI/II
rpoHI-ps
rpoHI
rpoHI/II
rpoHI-ps
Logarithmic
Stationary
D
Logarithmic
Stationary
SB1003
rpoHI
SB1003
rpoHI
Cells
Supernatants

Figure 4.6. Mutation of *rpoHI* in *R. capsulatus*. A. Ratio of RcGTA activity of mutant strains relative to SB1003. Gene transfer activity was determined as an average of 3 biological replicates relative to SB1003 in both logarithmic and stationary growth phases. Error bars represent standard deviation and strains found to be statistically different from SB1003 were determined by an analysis of variance (ANOVA) and are indicated by an asterisk (* $p < 0.05$). All p-values for pair wise comparisons are reported in Supplemental Table S4.3. B. Growth curves for SB1003 and mutant strains. Changes in culture turbidity from equivalent starting cultures was monitored over time using a Klett-Semmerson photoelectric colourimeter, and plotted on a log scale. C. Colony forming units of mutant strains in stationary phase cultures. The ratio of viable cells per ml in stationary phase mutant cultures relative to SB1003 was determined as an average of 3 biological replicates. Error bars represent standard deviation and strains found to be significantly different from SB1003 are indicated by an asterisk (*$p < 0.05$). All p-values for pair wise comparisons are reported in Supplemental Table S4.4. D. Representative western blots for detection of the RcGTA major capsid protein in the cells and culture supernatants of SB1003 and *rpoHI* in logarithmic and stationary phase samples. Blots were performed using the same cultures used for gene transfer activity assays.

4.3.4. The RcGTA promoter and RcGTA gene expression in *rsb* mutant cells

The DNA sequence upstream of *orfg1* of the RcGTA gene cluster was analyzed using BPROM (SoftBerry), which is a promoter recognition program for bacterial σ^{70} promoter sequences. This identified the -35 and -10 sequences of a putative 'rpoD17' binding site (Fig. 4.7A). The 'rpoD17' site is a class of σ^{70} promoters with a 17 nt spacer region (Robison *et al.*, 1998). Plasmid-borne *lacZ* fusion constructs to RcGTA *orfg2* were used to investigate whether this putative promoter sequence was required for RcGTA gene expression (Fig. 4.7B). I used flow cytometry to quantify fluorescence corresponding to β-galactosidase activity in late stationary phase cultures carrying the fusion constructs. Strains of SB1003 separately carrying the pX2 (possessing the native 5'

promoter region sequence of RcGTA) and pX2NP (containing no upstream regulatory sequence) plasmids were found to have mean fluorescence signals of 14.03 and 3.2, respectively (Fig. 4.7B, C and D). The plasmid pX2Δp is the same as pX2 except the putative 'rpoD17' promoter sequence located at -129 to -109 relative to the putative *orfg1* start codon has been deleted and replaced by a KpnI restriction site. The mean fluorescence of SB1003 carrying pX2Δp was 2.8, approximately the same as SB1003 (pX2NP) (Fig. 4.7B, C and D). To ensure that it was not simply disruption of any upstream sequence that was affecting RNAP binding and/or transcription initiation, a second plasmid, pX2Δs, was constructed containing a deletion of a putative RNA stem loop structure located -74 to -51 from the putative *orfg1* start codon. This sequence was also replaced by a KpnI site and the mean fluorescence of SB1003 (pX2Δs) was very similar to SB1003 (pX2) (Fig. 4.7C and D). Because removal of the stem loop did not result in a loss in *orfg2*'::'*lacZ* expression, the predicted 'rpoD17' site likely represents the previously uncharacterized promoter of RcGTA.

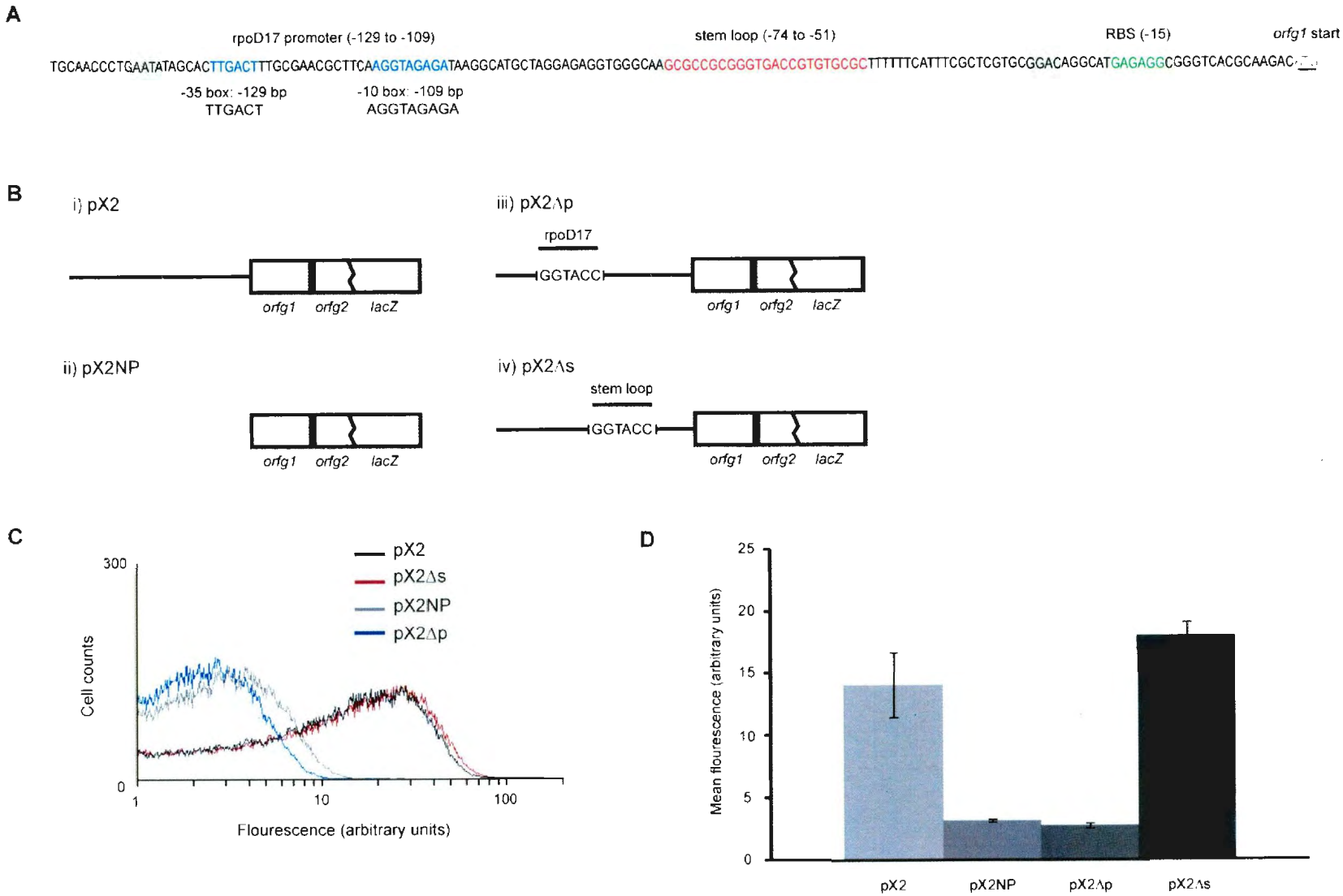

A
rpoD17 promoter (-129 to -109)
stem loop (-74 to -51)
RBS (-15)
orfg1 start
TGCAACCCTGAATATAGCACTTGACTTTGCGAACGCTTCAAGGTAGAGATAAGGCATGCTAGGAGAGGTGGGCAAGCGCCCGCGGGTGACCGTGTGCGCTTTTTTCATTTCGCTCGTGCGGACAGGCATGAGAGGCGGGTCACGCAAGAC
-35 box: -129 bp
TTGACT
-10 box: -109 bp
AGGTAGAGA
B
i) pX2
orfg1 orfg2 lacZ
iii) pX2Δp
rpoD17
GGTACC
orfg1 orfg2 lacZ
ii) pX2NP
orfg1 orfg2 lacZ
iv) pX2Δs
stem loop
GGTACC
orfg1 orfg2 lacZ
C
pX2
pX2Δs
pX2NP
pX2Δp
300
0
Cell counts
1
10
100
Flourescence (arbitrary units)
D
25
20
15
10
5
0
Mean flourescence (arbitrary units)
pX2
pX2NP
pX2Δp
pX2Δs

Figure 4.7. Analysis of predicted RcGTA promoter region. A. The sequence upstream of RcGTA *orfg1*. The -35 and -10 sequences of the putative 'rpoD17' promoter are indicated by blue text while the sequence for the putative RNA stem loop is in red text. A putative RBS sequence has been indicated by green text at -15 bp from the *orfg1* start codon, coloured brown and underlined. B. Representation of *orfg2'::'lacZ* fusion constructs. The plasmids pX2 (i) contains the native upstream sequence while the negative control, pX2NP (ii), contains no upstream sequence. The experimental plasmids, pX2Δp (iii) and pX2Δs (iv), have the predicted 'rpoD17' promoter and RNA stem loop sequences replaced by a KpnI site, respectively. C. RcGTA expression measured with reporter gene fusions in SB1003. Gene expression is represented by β-galactosidase activity and was determined by flow cytometry recording 100000 events. The average fluorescence was determined by replicate assays and error bars represent standard deviation. D. Representative histogram of RcGTA gene expression from reporter gene fusions in SB1003. Strains are indicated by black (pX2), red (pX2Δs), grey (pX2NP) and blue (pX2Δp) lines.

To determine the effects of *rsb* mutations on RcGTA gene expression, the pX2 and pX2Δp plasmids were introduced into *rsbW*, *rsbV* and *rsbU*. The fluorescence relative to the same plasmids in SB1003 agreed with the results of the gene transfer activity assays and western blots (Fig. 4.8). The *rsbW* mutant showed a 2 to 4-fold increase in fluorescence, corresponding to an increase in RcGTA *orfg2* expression (Fig. 4.8A and D). The *rsbV* and *rsbU* mutants demonstrated a decrease in mean fluorescence of 0.44 and 0.3-fold respectively (Fig. 4.8B, C and D). The mutant strains carrying the pX2Δp vector had nearly identical mean fluorescence as SB1003 (pX2Δp) (Fig. 4.8A, B and C).

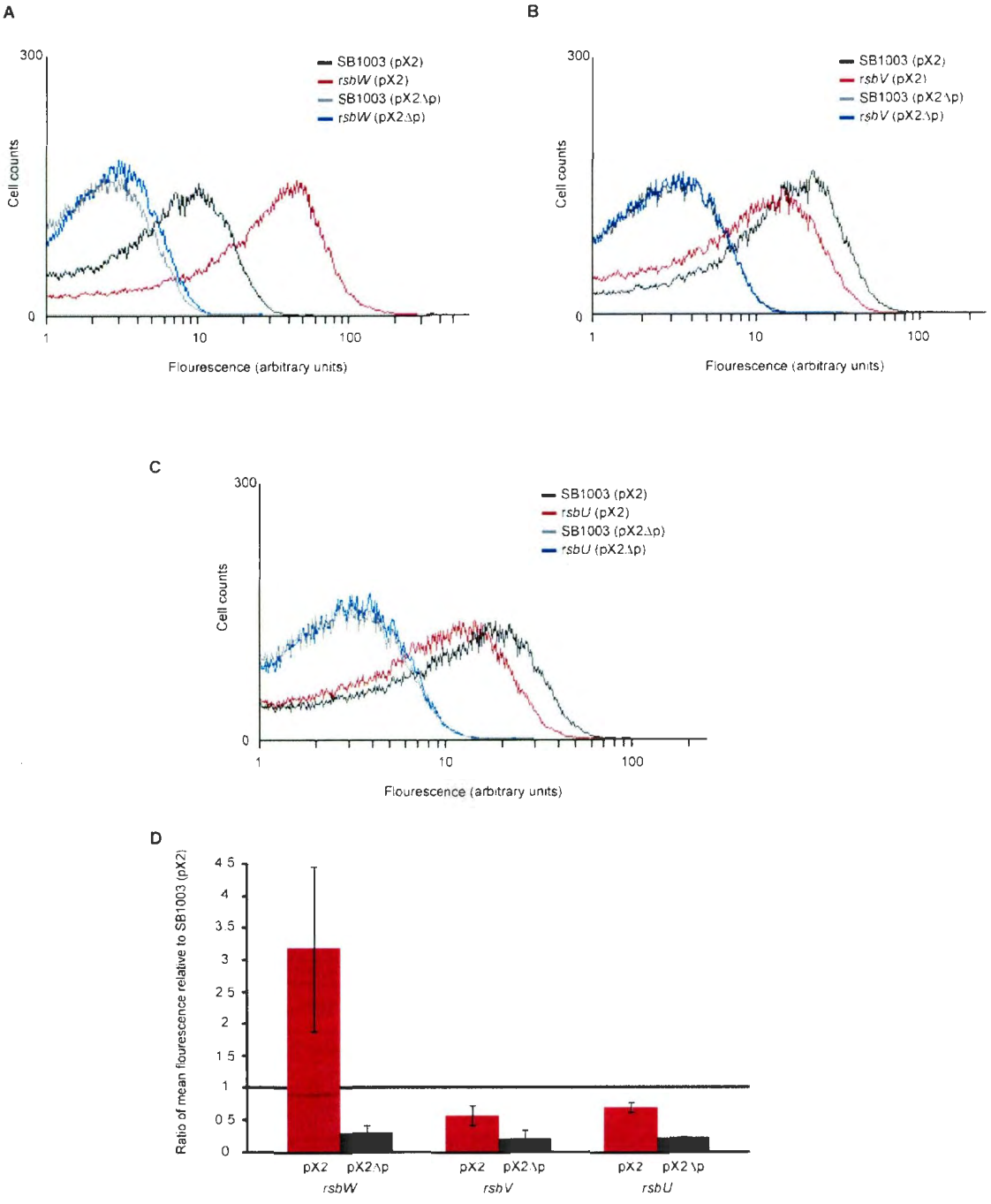


Figure 4.8. RcGTA reporter gene fusion expression in *rsb* mutants. A. Representative histograms of SB1003 and *rsbW* strains carrying either pX2 or pX2Δp fusion constructs. B. Representative histograms of SB1003 and *rsbV* strains carrying either pX2 or pX2Δp fusion constructs. C. Representative histograms of

SB1003 and *rsbU* strains carrying either pX2 or pX2Δp fusion constructs. For all histograms, different strains are represented by colours black, red, grey and blue as indicated in the individual legends. D. Ratio of mean fluorescence of *rsb* mutants carrying reporter fusions relative to the wild type. The ratio of average fluorescence of the indicated strains was determined by replicate assays relative to SB1003 (pX2) and the error bars represent standard deviation.

4.3.5. RsbW interacts with RsbV

Because RsbV is predicted to directly interact with RsbW as an antagonist, I used *in vitro* pull-downs and bacterial two-hybrid analysis to test for interactions between these proteins. Recombinant *R. capsulatus* RsbV and RsbW proteins were purified from *E. coli* by affinity chromatography. The purified samples were subject to in-gel trypsin digestion followed by peptide extraction and LC-MS/MS to confirm the identities of the recombinant proteins. Recombinant RsbW proteins (~20 kDa) carrying a 6x-His-tag on the N- or C-terminus were conjugated to NHS-activated sepharose beads and tested against recombinant 6x-His-RsbV (~15 kDa) and a control protein (lysozyme). The N-terminal 6x-His-RsbW immobilized on the beads was able to bind and pull-down 6x-His-RsbV but not the control protein (Fig. 4.9). In the control reaction, recombinant 6x-His-RsbV did not bind to the blocked sepharose beads that were first treated with buffer (Fig. 4.9).

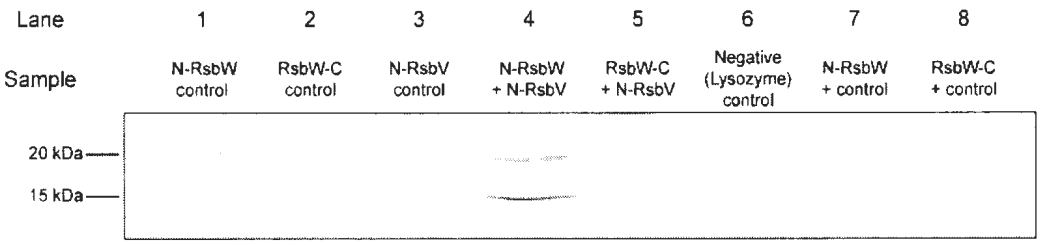


Figure 4.9. *In vitro* protein interactions between RsbW and RsbV. Pull down assays using bead-conjugated recombinant RsbW supplemented with recombinant RsbV or control protein (lysozyme). Both N- and C-terminal 6x-His-tagged RsbW proteins were conjugated and tested against N-terminal 6x-His-tagged RsbV (Lanes 4 and 5). Lysozyme (14.3 kDa) was tested against conjugated beads (Lanes 7 and 8) to determine protein binding specificity. Conjugated control beads (Lanes 1 and 2) were not supplemented with test protein while non-conjugated bead controls (Lanes 3 and 6) were blocked by 100 mM Tris. The gel was stained with Coomassie blue.

To confirm the specific interaction between RsbV and RsbW, a bacterial two-hybrid analysis was used. The vectors pKNT-rsbV and pUT18c-rsbW were co-transformed into the reporter strain BTH101 and β-galactosidase activity (units mg^{-1}) was determined in triplicate transformants alongside the appropriate controls (Fig. 4.10). The average activity of the interacting pair was found to be 1440 units mg^{-1} while all negative controls were found to have β-galactosidase activities between 101 and 147 units mg^{-1} (Fig. 4.10), consistent with previous findings using this BACTH assay (Karimova *et al.*, 1998). The positive control with interacting leucine zipper fragments had an average activity of 5694 units mg^{-1} (Fig. 4.10). The results of the *in vitro* pull-down (Fig. 4.9) and bacterial two-hybrid analysis (Fig. 4.10) show that RsbV interacts with RsbW.

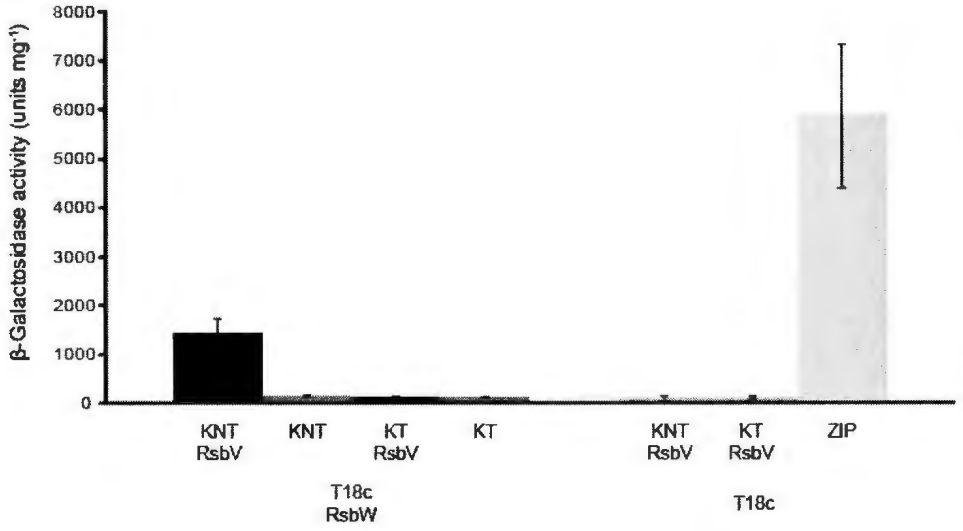


Figure 4.10. *In vivo* bacterial two hybrid analysis of protein interactions between RsbW and RsbV. β-galactosidase activity was determined for interactions between RsbW fused to the C- terminal end of the *E. coli* adenylate cyclase T18 fragment (T18c-RsbW) and RsbV fused to the N- and C-terminal ends of the T25 fragment of the same enzyme (KNT-RsbV and KT-RsbV, respectively). Adenylate cyclase fragments with no protein fused (KNT, KT and T18c) and interacting leucine zippers (ZIP) were used as controls. The activity was calculated as an average of 3 replicates and error bars represent standard deviation. The numerical data is presented in Table S4.5.

4.3.6. RsbW:σ interaction assays

The initial results described above with the *rpoHI* mutant suggested that RpoHI might be the σ factor responsible for RcGTA expression, and therefore it might interact with RsbW. *In vitro* pull-down assays using recombinant 6x-His-RpoHI as bait for bead conjugated 6x-His-RsbW were inconclusive (Fig. 4.11). A faint band corresponding to 6x-His-RpoHI (~36 kDa) can be seen in the experimental lane (2), with bead conjugated 6x-His-RsbW but a very faint band can also be seen in the negative control lane (1) with non-conjugated beads. Although the RpoHI band in the experimental lane is more

intense, this may not represent a true interaction. This presumably indicates that the negative control beads were not sufficiently washed to remove non-specifically bound 6x-His-RpoHI.

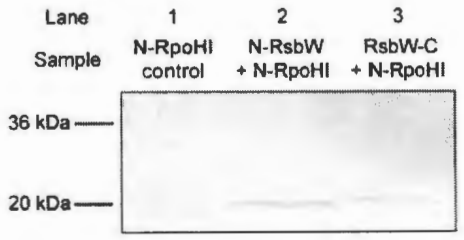


Figure 4.11. *In vitro* protein interactions between RsbW and RpoHI. Pull-down assays using bead-conjugated recombinant RsbW supplemented with recombinant RpoHI (~36 kDa). Both N- and C-terminal 6x-His-tagged RsbW proteins were conjugated and tested against N-terminal 6x-His-tagged RpoHI (Lanes 2 and 3). The non-conjugated bead control (Lane 1) beads were blocked by 100 mM Tris and supplemented with RpoHI to control for adequate washing. The gel was stained with Coomassie blue.

To try and identify a σ factor interaction with the putative anti-σ factor RsbW, I used bacterial two hybrid analysis with *rsbW* and the σ factors of interest cloned into the BACTH vectors in all conformations. Along with *rpoD* and *rpoHI*, the putative σ factor genes *rcc00699* and *rcc002637* were also used because viable mutants containing disruptions of these genes could never be confirmed. No positive interactions were observed in any transformants (Fig. 4.12). The β-galactosidase activity assay was carried out on the strains containing pUT18c-rsbW and the (putative) σ factors cloned into pKT25 or pKNT25. The bacterial two-hybrid analysis between RsbW and the *R. capsulatus* σ factors tested did not provide any positive results (Fig. 4.12). The negative

controls were calculated to have β-galactosidase activities ranging from 124.2 to 201.7 units mg^{-1}. Activities in experimental interactions ranged from 131.0 to 310.7 units mg^{-1} and the leucine zipper positive control was 7338.9 units mg^{-1}.

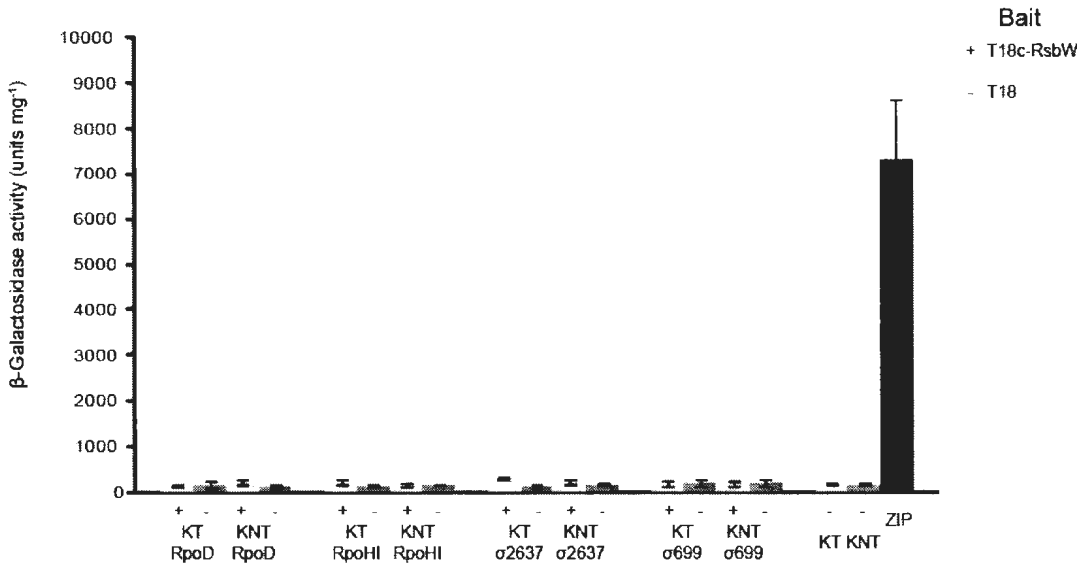


Figure 4.12. *In vivo* bacterial two hybrid analysis of protein interactions between RsbW and *R. capsulatus* σ factors. β -galactosidase activity was determined for interactions between bait protein T18c-RsbW (+) and σ factor preys indicated (fused to the N- or C- terminal ends of the T25 fragment of adenylate cyclase). T18 (-), KNT, KT controls were used to assess specific interactions and T25 and T18 fragments fused to interacting leucine zippers served as a positive control. The activity was calculated as an average of 3 replicates and error bars represent standard deviation. The numerical data are presented in Table S4.6.

4.4. Discussion

In several gram positive species, including *Bacillus* spp., the Rsb proteins control the activity of the alternative σ factor, σ^B, which is responsible for redirecting the

expression of genes involved in response to stress and entry into the stationary phase (Hecker *et al.*,2007). While many *rsb* homologues have been identified in gram-positive species, the general stress response in α-proteobacterial species has been found to be under the control of the ECF-σ^G, controlled by the anti-σ factor NepR, and the anti-anti-σ factor, PhyR (Alvarez-Martinez *et al.*, 2007; Bastiat *et al.*, 2010; Francez-Charlot *et al.*, 2009; Gourion *et al.*, 2008, 2009; Sauviac *et al.*, 2007). The results of this study suggests that *R. capsulatus* homologues of Rsb proteins play a role in controlling RcGTA gene expression, while also having pleiotropic effects on cell physiology. This represents the first instance in which Rsb homologues have been found to be active in α-proteobacteria. A BLAST analysis of the *rsbVW* genes of *R. capsulatus* revealed almost all Rsb homologues are restricted to members of the order Rhodobacterales in the class α-proteobacteria, with a few species of γ- and δ-proteobacteria containing recognizable Rsb homologues. The next most similar homologues outside of the proteobacteria are all in gram positive species, where the Rsb proteins were first characterized in species such as *Bacillus subtilis*, *Staphylococcus aureus*, and *Listeria monocytogenes* (Hecker *et al.*, 2007). The amino acid sequence identities between the RsbW and RsbV proteins of *R. capsulatus* and *Bacillus* spp. are ~30%, suggesting that these proteins are true homologues.

Mutation of the putative phosphatase-encoding *rsbU* resulted in a decrease in RcGTA expression, which agrees with a predicted function for the putative RsbU protein in this regulatory pathway. In *Bacillus*, the RsbU phosphatase dephosphorylates RsbV-P, thereby allowing the anti-anti-σ to interact with RsbW and promote target gene

expression by the cognate σ^B (Yang *et al.*, 1996). Removal of this protein would result in an increase in RsbV-P and RsbW inhibition of cognate σ factor activity. The *R. capsulatus* RsbV protein has two serine residues, S56 and S57, at approximately the same region as found in *Bacillus* RsbV where 1 of the 2 serves as the site of phosphorylation by the serine-kinase, RsbW (Dufour & Haldenwang, 1994; Eymann *et al.*, 2007). Since no homologues of stressosome-associated proteins RsbRST were identified in *R. capsulatus*, the RsbU phosphatase may be stimulated by phosphorelay signals from some other, currently unidentified signalling pathway. RsbK-like hybrid sensor kinases have been previously identified in proteobacteria and cyanobacteria, however their function in regulating σ factor activity is not conserved (de Been *et al.*, 2011). Identification of a histidine kinase, or perhaps a stressosome-like complex unique to the α-proteobacteria, would be essential to determine the signal to which this putative σ factor regulatory system responds. However, upstream signals from additional components may not be required to activate the RsbU phosphatase. The Rsb system of *S. aureus*, where homologues of RsbU, RsbV and RsbW have been identified, does not possess any stressosome components, and induced expression of RsbU is sufficient to stimulate the downstream σ-dependent transcription (Pane-Farre *et al.*, 2009).

The *rsbU*, *rsbV* and *rsbVW* mutants all had similar phenotypes, with effects on RcGTA gene expression, stationary phase cell viability, and colony morphology. The similarities in *rsbV* and *rsbU* mutant phenotypes support the hypothesis that these proteins are working in a common pathway. Since the *rsbW* mutant had no observable differences in stationary phase cell viability or colony morphology, these effects are

likely the result of mutation of the *rsbV* gene. Re-introducing *rsbV* into strains lacking the gene resulted in an increase in viable cell numbers relative to wild type, supporting the hypothesis that RsbV plays a role in maintaining cell viability in the stationary phase. The results from the *rsbV* (p*V*) and *rsbVW* (p*VW*) strains indicate that the insertional disruption of *rsbV* may have had a polar effect on the *rsbW* ORF. If this is the case and the *rsbV* disruption is polar, then the *rsbV* mutant strain is actually a double mutation of *rsbV* and *rsbW*. Unexpectedly, the *rsbVW* mutant does not resemble the *rsbW* mutant phenotype, suggesting a dominant effect of the *rsbV* mutation over *rsbW*. Considering that the *rsbW* strain is not motile, it is unknown why both the *rsbV* and *rsbVW* strains retain culture motility, but perhaps a secondary mutation alleviated the loss of *rsbW* in these already compromised strains. The decrease in RcGTA activity and capsid protein levels in *rsbV* can be restored by complementation, indicating that RsbV is a positive regulator of RcGTA production, but it clearly plays several other key regulatory roles in *R. capsulatus*. Further genetic studies, including construction of an in-frame deletion of *rsbV* instead of the insertional disruption, will be required to further elucidate this protein's full function.

Removal of the anti-σ factor, RsbW, led to increased production of RcGTA, presumably due to the increased availability of a cognate σ factor to compete for RNAP core enzyme. The RcGTA activity assays, western blots, and *orfg2'*::'*lacZ* fusion results with the *rsbW* mutant strain agree with this predicted anti-σ factor function. The loss of motility in *rsbW* is similar to how loss of the anti-σ factor FlgM can affect culture motility in *Yersinia pseudotuberculosis* (Ding *et al.*, 2009) and *Rhodobacter sphaeroides*

(Wilkinson *et al.*, 2011). FlgM acts on the flagellar-specific σ factor, FliA, as an essential checkpoint in flagellar biosynthesis and transcription of class III (or class IV in *R. sphaeroides*) flagellar genes (Karlinsey *et al.*, 2000; Poggio *et al.*, 2005). FlgM ensures proper assembly of the hook-basal body (HBB) structure, and dissociates from FliA when the HBB is complete, allowing expression of structural components necessary to assemble the remainder of the flagellum. Without FlgM, mis-timing of the expression of class III genes by FliA results in failure to properly assemble flagella and defective motility (Wilkinson *et al.*, 2011). The motility phenotypes of the *rsbW* and complemented strains give strength to the notion that RsbW is acting similarly with respect to culture motility. The results presented in chapter 2, along with a previous publication (Lang & Beatty, 2002), have already demonstrated a link between RcGTA production and motility, but *rsbW* is the first regulatory mutant strain created where RcGTA gene expression persists but motility has been abolished. It is not clear why *rsbW* (p*W*) maintained elevated RcGTA levels relative to SB1003. An initial hypothesis was that the RsbW protein was not translated from p*W*, which contains an in-frame deletion of *rsbV* (Figure 4.1B); however, restoration of motility in *rsbW* (p*W*) argues against this. This inability to restore RcGTA to wild type levels using *in-trans* complementation may be related to the 'dominant' effect RsbV seems to have in this system, as reflected in the phenotypes of the *rsbVW* double mutant. It is possible that RsbW interacts with multi-protein complexes, and disrupting the stoichiometry affects protein-protein interactions and produces unexpected results.

The results of this study show that RsbV interacts with RsbW both *in vitro* and *in vivo*, and if RsbW truly is an anti-sigma factor, RsbV presumably antagonizes this function. It is also possible that RsbV may interact with several other, as-of-yet unidentified anti-σ factors encoded in the *R. capsulatus* genome. Multi-partner interactions of anti-anti-σ factors have been observed in *Streptomyces coelicolor*, where 1 of the 15 RsbV homologues, BldG, interacts with two anti-σ factors, ApgA (an RsbW homologue) (Parashar *et al.*, 2009) and UshX (Sevcikova *et al.*, 2010). These interactions result in the anti-anti-σ factor playing a role in antibiotic production, morphological differentiation, and osmotic stress response (Bignell *et al.*, 2000; Parashar *et al.*, 2009). Another anti-σ factor of *S. coelicolor*, RsfA, interacts with two anti-anti-σ factors (Kim *et al.*, 2008). As well, the pathogen *Mycobacterium tuberculosis* utilizes the anti-σ factor UsfX to control the activity of σ^F, where UsfX is regulated by 3 anti-anti-σ factors (Sachdeva *et al.*, 2009). Therefore, it is not unreasonable to believe that RsbW may also have other interacting partners in addition to RsbV.

Despite the observed phenotypes, I cannot conclude that *R. capsulatus* RsbW functions as an anti-σ factor in this species because *in vitro* pull downs and bacterial two-hybrid analysis did not provide any strong evidence of protein interactions with the σ factor proteins tested. However, this could be due to experimental conditions as expression of *R. capsulatus* σ factors in *E. coli* may yield insoluble proteins as found with *R. sphaeroides* RpoD and RpoE (Anthony *et al.*, 2003; Newman *et al.*, 1999), which would disrupt the two-hybrid assays. It is also possible that the recombinant *R. capsulatus* proteins are interacting with native *E. coli* proteins, which could also interfere with the

two-hybrid assays. Structural interaction studies in *E. coli* have led to hypotheses that currently unknown small regulatory molecules affect the binding between the anti-σ factor Rsd (**R**egulator of **s**igma **D**) and σ^{70} (Hofmann *et al.*, 2011). The interaction of *R. capsulatus* RsbW with cognate σ factors may require co-factors and specific interactions might not occur without supplementing an experiment appropriately.

This study rules out the involvement of the EcfG-like (*rcc02291*), RpoHII (*rcc00458*), and putative ECF (*rcc02724*) σ factors in RcGTA expression because disruptions in these genes did not affect gene transfer activity. I hypothesize that the effect on RcGTA production in my original *rpoHI* mutant was due to non-specific effects. RpoHI is the closest homologue to σ^{B} in *R. capsulatus*, and this protein has been thoroughly studied in the related species *R. sphaeroides* where it is involved in responding to heat and photooxidative stress. Neither the RcGTA-like gene products nor the homologues of *R. capsulatus rsbV/W* appear to be expressed in *R. sphaeroides* (Mark Gomelsky, personal communication). It is unknown whether RpoHI functions the same in *R. capsulatus*, and the only work prior to this study suggested it was essential for growth at 32°C (Emetz & Klug, 1998). A very weak interaction may be argued between RsbW and RpoHI based on my *in vitro* experiments, which could be due to RsbW having a broad interaction range with purified σ factors under non-competing conditions. It could also represent a weak transient interaction that occurs *in vivo* between the proteins in *R. capsulatus*. The expression of the Rsb proteins and their effect on RcGTA is so far a unique characteristic of *R. capsulatus*, and further work must be done to determine what role RpoHI plays.

The identification of the RcGTA promoter sequence required for expression of *orfg1*and *orfg2* may help identify a σ factor responsible for RNAP recruitment and transcription initiation. The site is designated as an 'rpoD17' site, which is the most common type of promoter sequence for RpoD in *E. coli* (Robison *et al.*, 1998). The specificity of conservation of these sites across proteobacterial lineages is unclear, although the BPROM prediction software does recognize different 'RpoD', 'RpoS', and various other transcription factor binding sites.

The effect of *rsbV* and *rsbU* on colony morphology and culture viability may implicate these proteins as regulators of a σ factor with a large regulon, such as RpoD. As well, there is no FliA homologue in *R. capsulatus*, and therefore the transcription of flagellar genes is under the control of a different σ factor, which is predicted to interact with RsbW by my data. BLAST results using the *R. sphaeroides* FliA protein to search the *R. capsulatus* genome yields hits to RpoD and RpoHII, but I found the *rpoHII* mutant strain is motile. Some, but not all, motility (flagellar and chemotaxis) operons have putative RpoD-like promoters. I hypothesize that RsbW interacts with RpoD as an essential coordination for synthesis of flagella and production of RcGTA.

The Rsd protein of *E. coli* participates in a partner switch with the major vegetative σ factor, RpoD, and the alternative σ factor, RpoS, upon entry into the stationary phase (Hofmann *et al.*, 2011). This interaction is highly dependent on a balance between concentration and affinity of Rsd interacting partners. The *Bacillus* RsbW paralogue SpoIIAB also interacts with two σ factors, controlling the activity of both σ^F and σ^G, which are involved in the sporulation stress response (Duncan & Losick, 1993;

Evans *et al.*, 2003). It is possible that *R. capsulatus* RsbW may interact with RpoD and RpoHI (or another σ factor), similar to how Rsd and SpoIIAB switch partners, in response to an inducing signal. Transient interactions could be very weak and depend highly on cellular abundances of alternative σ factors competing for core RNAP.

While the results of this study have not identified the still elusive RcGTA-specific σ factor, I have identified a putative σ factor regulatory system affecting RcGTA production and a promoter sequence upstream of *orfg1* which is required for expression. These components of RcGTA regulation are important for understanding how *R. capsulatus* controls production of this unusual genetic exchange element. *Bacillus* species use the Rsb system in response to both stress and entry into stationary phase and it has long been hypothesized that *R. capsulatus* produces RcGTA in stationary phase as a stress response. The *rsb* genes are all up-regulated in stationary phase relative to logarithmic phase in SB1003 (Chapter 2; Mercer *et al.*, 2010). I hypothesize that CtrA modulates RcGTA expression in stationary phase through the Rsb proteins, which control the activity of one or more important σ factors in *R. capsulatus*.

4.5 References

Adams, C. W., M. E. Forrest, S. N. Cohen and J. T. Beatty (1989). Structural and functional analysis of transcriptional control of the *Rhodobacter capsulatus puf* operon. *J. Bacteriol.* **171**(1): 473-482.

Alper, S., A. Dufour, D. A. Garsin, L. Duncan and R. Losick (1996). Role of adenosine nucleotides in the regulation of a stress-response transcription factor in *Bacillus subtilis*. *J. Mol. Biol.* **260**(2): 165-177.

Alper, S., L. Duncan and R. Losick (1994). An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in *B. subtilis*. *Cell* **77**(2): 195-205.

Alvarez-Martinez, C. E., R. F. Lourenço, R. L. Baldini, M. T. Laub and S. L. Gomes (2007). The ECF sigma factor σ^T is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. *Mol. Microbiol.* **66**(5): 1240-1255.

Anthony, J. R., H. A. Green, T. J. Donohue and G. Sankar Adhya and Susan (2003). Purification of *Rhodobacter sphaeroides* RNA polymerase and its sigma factors. In *Methods in Enzymology*. S. Adhya and S. Garges (Eds). Academic Press. Volume **370**: 54-65.

Barany, F. (1985). Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. *Gene* **37**: 111-123.

Bastiat, B., L. Sauviac and C. Bruand (2010). Dual control of *Sinorhizobium meliloti* RpoE2 sigma factor activity by two PhyR-type two-component response regulators. *J. Bacteriol.* **192**(8): 2255-2265.

Benson, A. K. and W. G. Haldenwang (1993). *Bacillus subtilis* sigma B is regulated by a binding protein (RsbW) that blocks its association with core RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **90**(6): 2330-2334.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**(6885): 141-147.

Bignell, D. R. D., J. L. Warawa, J. L. Strap, K. F. Chater and B. K. Leskiw (2000). Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. *Microbiol.* **146**(9): 2161-2173.

Borukhov, S. and E. Nudler (2003). RNA polymerase holoenzyme: structure, function and biological implications. *Curr. Opin. Microbiol.* **6**(2): 93-100.

Campbell, E. A., L. F. Westblade and S. A. Darst (2008). Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr. Opin. Microbiol.* **11**(2): 121-127.

de Been, M., C. Francke, R. J. Siezen and T. Abee (2011). Novel σ^B regulation modules of Gram-positive bacteria involve the use of complex hybrid histidine kinases. *Microbiol.* **157**(1): 3-12.

de Been, M., M. H. Tempelaars, W. van Schaik, R. Moezelaar, R. J. Siezen and T. Abee (2010). A novel hybrid kinase is essential for regulating the B-mediated stress response of *Bacillus cereus*. *Environ. Microbiol.* **12**(3): 730-745.

Ding, L., Y. Wang, Y. Hu, S. Atkinson, P. Williams and S. Chen (2009). Functional characterization of FlgM in the regulation of flagellar synthesis and motility in *Yersinia pseudotuberculosis*. *Microbiol.* **155**(6): 1890-1900.

Dufour, A. and W. G. Haldenwang (1994). Interactions between a *Bacillus subtilis* anti-sigma factor (RsbW) and its antagonist (RsbV). *J. Bacteriol.* **176**(7): 1813-1820.

Duncan, L. and R. Losick (1993). SpoIIAB is an anti-sigma factor that binds to and inhibits transcription by regulatory protein sigma F from *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* **90**(6): 2325-2329.

Emetz, D. and G. Klug (1998). Cloning and characterization of the *rpoH* gene of *Rhodobacter capsulatus*. *Mol. Gen. Genet.* **260**(2): 212-217.

Evans, L., J. Clarkson, M. D. Yudkin, J. Errington and A. Feucht (2003). Analysis of the interaction between the transcription factor sigmaG and the anti-sigma factor SpoIIAB of *Bacillus subtilis*. *J. Bacteriol.* **185**(15): 4615-4619.

Eymann, C., D. Becher, J. Bernhardt, K. Gronau, A. Klutzny and M. Hecker (2007). Dynamics of protein phosphorylation on Ser/Thr/Tyr in *Bacillus subtilis*. *Proteomics* **7**(19): 3509-3526.

Francez-Charlot, A., J. Frunzke, C. Reichen, J. Z. Ebneter, B. Gourion and J. A. Vorholt (2009). Sigma factor mimicry involved in regulation of general stress response. *Proc. Natl. Acad. Sci. U.S.A.* **106**(9): 3467-3472.

Gill, P. R., Jr and G. J. Warren (1988). An iron-antagonized fungistatic agent that is not required for iron assimilation from a fluorescent rhizosphere pseudomonad. *J. Bacteriol.* **170**: 163–170.

Gourion, B., A. Francez-Charlot and J. A. Vorholt (2008). PhyR is involved in the general stress response of *Methylobacterium extorquens* AM1. *J. Bacteriol.* **190**(3): 1027-1035.

Gourion, B., S. Sulser, J. Frunzke, A. Francez-Charlot, P. Stiefel, G. Pessi, J. A. Vorholt and H.-M. Fischer (2009). The PhyR-σEcfG signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **73**(2): 291-305.

Green, H. A. and T. J. Donohue (2006). Activity of *Rhodobacter sphaeroides* RpoHII, a second member of the heat shock sigma factor family. *J. Bacteriol.* **188**(16): 5712-5721.

Hecker, M., J. Pané-Farré and V. Uwe (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu. Rev. Microbiol.* **61**(1): 215-236.

Helmann, J. D. (1999). Anti-sigma factors. *Curr. Opin. Microbiol.* **2**(2): 135-141.

Hofmann, N., R. Wurm and R. Wagner (2011). The *E. coli* anti-sigma factor Rsd: Studies on the specificity and regulation of its expression. *PLoS ONE* **6**(5): e19235.

Hughes, K. T. and K. Mathee (1998). The anti-sigma factors. *Annu. Rev. Microbiol.* **52**(1): 231-286.

Hynes, A. P., R. G. Mercer, D. E. Watton, C. B. Buckley and A. S. Lang (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol. Microbiol.* **85**(2): 314-325.

Karimova, G., J. Pidoux, A. Ullmann and D. Ladant (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. U.S.A.* **95**(10): 5752-5756.

Karlinsey, J. E., S. Tanaka, V. Bettenworth, S. Yamaguchi, W. Boos, S.-I. Aizawa and K. T. Hughes (2000). Completion of the hook–basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and *fliC* transcription. *Mol. Microbiol.* **37**(5): 1220-1231.

Kim, E. S., J. Y. Song, D. W. Kim, K. F. Chater and K. J. Lee (2008). A possible extended family of regulators of sigma factor activity in *Streptomyces coelicolor*. *J. Bacteriol.* **190**(22): 7559-7566.

Lang, A. S. and J. T. Beatty (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl Acad. Sci. USA* **97**: 859-864.

Lang, A. S. and J. T. Beatty (2002). A bacterial signal transduction system controls genetic exchange and motility. *J. Bacteriol.* **184**: 913-918.

Markowitz, V. M., I. M. Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, I. Anderson, A. Lykidis, K. Mavromatis, N. N. Ivanova and N. C. Kyrpides (2011). The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res.* **38**(Database issue): D382-90.

Marles-Wright, J. and R. J. Lewis (2007). Stress responses of bacteria. *Curr. Opin. Struct. Biol.* **17**(6): 755-760.

Mercer, R. G., S. J. Callister, M. S. Lipton, L. Pasa-Tolic, H. Strnad, V. Paces, J. T. Beatty and A. S. Lang (2010). Loss of the response regulator CtrA causes pleiotropic effects on gene expression but does not affect growth phase regulation in *Rhodobacter capsulatus*. *J. Bacteriol.* **192**(11): 2701-2710.

Miller, J. H. (1992). *A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria.* Plainview, NY, Cold Spring Harbor Laboratory Press.

Mittenhuber, G. (2002). A phylogenomic study of the general stress response sigma factor sigmaB of *Bacillus subtilis* and its regulatory proteins. *J. Mol. Microbiol. Biotechnol.* **4**(4): 427-452.

Newman, J. D., M. J. Falkowski, B. A. Schilke, L. C. Anthony and T. J. Donohue (1999). The *Rhodobacter sphaeroides* ECF sigma factor, σ^E, and the target promoters *cycA* P3 and *rpoE* P1. *J. Mol. Biol.* **294**(2): 307-320.

Nuss, A. M., J. Glaeser, B. A. Berghoff and G. Klug (2010). Overlapping alternative sigma factor regulons in the response to singlet oxygen in *Rhodobacter sphaeroides*. *J. Bacteriol.* **192**(10): 2613-2623.

Nuss, A. M., J. Glaeser and G. Klug (2009). RpoHII activates oxidative-stress defense systems and is controlled by RpoE in the singlet oxygen-dependent response in *Rhodobacter sphaeroides*. *J. Bacteriol.* **191**(1): 220-230.

Österberg, S., T. d. Peso-Santos and V. Shingler (2011). Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.* **65**(1): 37-55.

Pane-Farre, J., B. Jonas, S. W. Hardwick, K. Gronau, R. J. Lewis, M. Hecker and S. Engelmann (2009). Role of RsbU in controlling SigB activity in *Staphylococcus aureus* following alkaline stress. *J. Bacteriol.* **191**(8): 2561-2573.

Pane-Farre, J., R. J. Lewis and J. Stulke (2005). The RsbRST stress module in Bacteria: A signalling system that may interact with different output modules. *J. Mol. Microbiol. Biotechnol.* **9**(2): 65-76.

Parashar, A., K. R. Colvin, D. R. D. Bignell and B. K. Leskiw (2009). BldG and SCO3548 interact antagonistically to control key developmental processes in *Streptomyces coelicolor*. *J. Bacteriol.* **191**(8): 2541-2550.

Poggio, S., A. Osorio, G. Dreyfus and L. Camarena (2005). The flagellar hierarchy of *Rhodobacter sphaeroides* is controlled by the concerted action of two enhancer-binding proteins. *Mol. Microbiol.* **58**(4): 969-983.

Prentki, P. and H. M. Krisch (1984). *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29**: 303-313.

Price, C. W., P. Fawcett, H. Cérémonie, N. Su, C. K. Murphy and P. Youngman (2001). Genome-wide analysis of the general stress response in *Bacillus subtilis*. *Mol. Microbiol.* **41**(4): 757-774.

Robison, K., A. M. McGuire and G. M. Church (1998). A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J. Mol. Biol.* **284**(2): 241-254.

Sachdeva, P., R. Misra, A. K. Tyagi and Y. Singh (2009). The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J.* **277**(3): 605-626.

Sauviac, L., H. Philippe, K. a. Phok and C. Bruand (2007). An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *J. Bacteriol.* **189**(11): 4204-4216.

Sevcikova, B., B. Rezuchova, D. Homerova and J. Kormanec (2010). The anti-anti-sigma factor BldG is involved in activation of the stress response sigma factor σ^H in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **192**(21): 5674-5681.

Sharma, A. K., A. C. Rigby and S. L. Alper (2011). STAS Domain Structure and Function. *Cell. Physiol. Biochem.* **28**(3): 407-422.

Simon, R., U. Priefer and A. Pühler (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Bio/Tech* **1**: 37-45.

Staroń, A. and T. Mascher (2010). General stress response in α-proteobacteria: PhyR and beyond. *Mol. Microbiol.* **78**(2): 271-277.

Strnad, H., A. Lapidus, J. Paces, P. Ulbrich, C. Vlcek, V. Paces and R. Haselkorn (2010). Complete genome sequence of the photosynthetic purple nonsulfur bacterium *Rhodobacter capsulatus* SB 1003. *J. Bacteriol.* **192**(13): 3545-3546.

van Schaik, W., M. H. Tempelaars, M. H. Zwietering, W. M. de Vos and T. Abee (2005). Analysis of the role of RsbV, RsbW, and RsbY in regulating σ^B activity in *Bacillus cereus*. *J. Bacteriol.* **187**(16): 5846-5851.

Wilkinson, D. A., S. J. Chacko, C. Venien-Bryan, G. H. Wadhams and J. P. Armitage (2011). Regulation of flagellum number by FliA and FlgM and role in biofilm formation by *Rhodobacter sphaeroides*. *J. Bacteriol.* **193**(15): 4010-4014.

Wise, A. A. and C. W. Price (1995). Four additional genes in the *sigB* operon of *Bacillus subtilis* that control activity of the general stress factor sigma B in response to environmental signals. *J. Bacteriol.* **177**(1): 123-33.

Wong, D. K.-H., W. J. Collins, A. Harmer, T. G. Lilburn and J. T. Beatty (1996). Directed mutagenesis of the *Rhodobacter capsulatus puhA* gene and pleiotropic effects on photosynthetic reaction center and light-harvesting 1 complexes. *J. Bacteriol.* **178**: 2334-2342.

Yang, X., C. M. Kang, M. S. Brody and C. W. Price (1996). Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes Dev.* **10**(18): 2265-2275.

Yen, H.-C. and B. Marrs (1976). Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. *J. Bacteriol.* **126**: 619-629.

Zhang, S. and W. G. Haldenwang (2005). Contributions of ATP, GTP, and redox state to nutritional stress activation of the *Bacillus subtilis* σ^B transcription factor. *J. Bacteriol.* **187**(22): 7554-7560.

Chapter 5 - Summary

The results presented in this thesis further elucidate the complex network of regulation controlling RcGTA gene expression and production in *R. capsulatus*. While research had determined the requirement of CtrA for expression of the RcGTA genes (Lang & Beatty, 2000), it is evident that several other pathways, including additional phosphorelay proteins and regulators of σ factor activity, are required to properly control both expression and release of this unusual genetic exchange element. The work presented in this thesis represents the first full view of the CtrA regulon. It is also the most thorough investigation yet undertaken of the effects of different phosphorelay proteins on RcGTA production. This is also the first time Rsb homologues have been indentified to be active in α-proteobacteria, and I have shown that they play a substantial role in modulating RcGTA expression while also having pleiotropic effects on *R. capsulatus* physiology.

The role of CtrA in this species has been revealed to be quite substantial, as loss of this important regulator affects the proper expression of >200 genes in the *R. capsulatus* genome. These include all motility genes and a number of putative signal transduction and regulatory protein-encoding genes (Chapter 2; Mercer *et al.*, 2010). The identification of putative signal transduction genes affected by the loss of CtrA suggests that many downstream regulatory pathways likely contribute to the observed overall dysregulation of many *R. capsulatus* genes in both the logarithmic and stationary phases

of growth when CtrA is absent. Several of these genes are predicted to encode putative diguanylate cyclase and phosphodiesterase enzymes, responsible for cycling the secondary signalling molecule, c-di-GMP (Schirmer & Jenal, 2009). This molecule affects many regulatory systems, such as those controlling the cell cycle, antibiotic production, virulence, motility and adhesion in different bacteria (Hengge, 2009; Mills *et al.*, 2011). Investigations into the role of c-di-GMP in this species and its affect on RcGTA production will be an important area for future research extending from this thesis.

The expression of a gene cluster predicted to encode gas vesicles was also affected by the loss of CtrA, despite these cellular organelles never being identified in this species previously. These organelles affect cellular buoyancy which, like the bacterial flagellum and pilus, contributes to bacterial motility (Jarrell & McBride, 2008). The large number of conserved hypothetical genes indentified in Chapter 2 shows how little is still known about the role of CtrA and the functional genome in this species. Future work using the data generated in the transcriptome analyses will allow for further elucidation of the functions of CtrA and why this regulator is important in controlling RcGTA.

As a prominent regulator of gene expression in *R. capsulatus*, it would be useful to determine the full range of CtrA binding sites in the promoter regions of target genes. Despite sequence similarity with the *C. crescentus* protein, it is possible CtrA binds more variable sequences in *R. capsulatus*. Thorough investigations using DNase footprinting,

electrophoretic mobility shift assays (EMSAs) and/or chromatin immunoprecipitation (ChIP) assays would be essential in properly defining CtrA binding in this species and ultimately determine the direct CtrA regulon.

Chapter 3 of this thesis investigated the phosphorylation of CtrA in *R. capsulatus*. The results suggested that both phosphorylated and unphosphorylated CtrA is capable of activating expression of the RcGTA gene cluster. However, both motility and release of RcGTA were dependent on phosphorylation of the protein, which I predict does occur through the CckA-ChpT phosphorelay, and possibly other sources. Recently, Hynes *et al.* (2012) identified a putative lysis mechanism required for release of RcGTA. Analysis of preliminary expression data from the complete *cckA* mutant revealed this putative *N*-acetylmuramidase lysozyme was under the control of CckA, but not CtrA (Mercer, unpublished observation). This agrees with the finding that the CckA-ChpT phosphorelay was independently required for proper release of RcGTA, separately from CtrA~P. The identification of a putative second response regulator protein that is part of this phosphorelay system may provide a better understanding of the requirement for both CckA-ChpT and CtrA~P for RcGTA release.

Along with release of RcGTA, only phosphorylated CtrA was capable of stimulating culture motility. In contrast, the expression of the RcGTA genes is solely dependent on the presence of a CtrA protein, irrespective of phosphorylation state. Because both motility and a lysis mechanism seem to be regulated more specifically,

CtrA must control RcGTA gene expression in a different manner. CtrA may interact with a secondary regulatory protein that is required for RcGTA expression. The activity of this hypothetical second regulator may be altered based on the structural conformation of CtrA, where phosphorylation of CtrA promotes dimerization of the protein (Spencer *et al.*, 2009). RcGTA gene expression and release may be timed to coordinate with higher levels of CtrA phosphorylation in the later phase of growth. This process occurs in *C. crescentus*, where both spatial and temporal regulation of the phosphorelay pathway affects the abundance and gradient levels of CtrA~P in the cell (Brown *et al.*, 2008; Chen *et al.*, 2009).

As previously mentioned, I believe the phosphorelay pathway from CckA-ChpT to CtrA is conserved in this species. However, the results of Chapter 3 do not definitively support this, and clearly the proteins have additional branching roles similar to their *C. crecentus* orthologues. Future *in vitro* phosphorylation assays, site-directed mutagenesis of CckA and genetic disruptions of multiple members of the signalling pathway will provide a clearer picture of this complex regulatory pathway.

Regardless of CtrA phosphorylation, the protein is required for the proper expression of hundreds of genes with diverse functions in *R. capsulatus*. Amongst these is a pair of putative σ factor regulatory proteins that play a role in RcGTA expression, motility, colony morphology and cell viability in stationary phase (Chapter 4). These proteins are homologues of a well studied Rsb partner switching system from *Bacillus*

spp., and are the first identified to be actively playing a regulatory role in α-proteobacteria. In members of *Bacillus* and other gram positive species, the anti-σ factor, RsbW, and the anti-anti-σ factor, RsbV, are involved in controlling the activity of the general stress response σ factor, σ^B (Hecker *et al.*, 2007). The complexity of the Rsb systems varies from species to species, with some species carrying several other Rsb proteins that affect the activity of the RsbV-RsbW partner switching module and, subsequently, the cognate σ factor.

The general stress response in the α-proteobacteria utilizes a different σ factor regulatory system, which involves a response regulator-like anti-anti-σ factor, PhyR, which is activated by phosphorylation. PhyR negatively regulates the anti-σ factor, NepR, promoting the transcriptional activity of the cognate σ factor (Gourion *et al.*, 2008, 2009; Staroń & Mascher, 2010). Like the gram-positive Rsb systems, components and regulatory partners vary from species to species. These proteins have been found to control the activity of EcfG-like σ factors of *C. crescentus* (Alvarez-Martinez *et al.*, 2007) and *S. meliloti* (Sauviac *et al.*, 2007). While RcGTA expression and production have long been thought to be a stress response of *R. capsulatus*, they are not under the control of the same system involving PhyR-like proteins. Disruptions of a Phy-R-like homologue (*rcc02289*) and the Ecf-G-like σ (*rcc02291*) did not have any effects on RcGTA production in *R. capsulatus* (Appendix 5; Chapter 4).

Instead, the RsbW, RsbV and RsbU homologues in *R. capsulatus* appear to be some of the major regulators affecting RcGTA expression. Despite some unusual characteristics, the mutant phenotypes agreed with the predicted function of these proteins. As well, the RsbV and RsbW proteins do interact both *in vitro* and *in vivo* in an *E. coli* system. The *rsb* genes are downregulated in stationary phase relative to the logarithmic stage in SB1003 (Appendix 1, Table S2.3). This agrees with the notion that these regulators of RcGTA are less abundant in the phase of growth for which RcGTA production is highest. The biggest question arising from the work presented in Chapter 4 is which σ factor are these proteins controlling? The remaining potential candidate σ factors that may be part of this regulatory network are *rcc00699* (RpoE-like), *rcc02637* (RpoE-like) and RpoD (*rcc03054*). Attempts to identify the cognate σ factor by genetic disruptions, *in vitro* pull downs, and bacterial two hybrid analyses proved unsuccessful (Chapter 4, Section 4.3). This could be due to experimental factors that differ from what is happening *in vivo*. It would not be possible to disrupt essential *R. capsulatus* σ factors and determine any regulatory effect on RcGTA based solely on genetic mutation. However, it is interesting that a σ factor required for RcGTA expression and promoting horizontal genetic exchange appears essential to the organism's viability. The bacterial two hybrid analyses did not produce a positive interaction between RsbW and any of the σ factors tested. This could be due to a uniquely required co-factor or regulatory signal present in *R. capsulatus* that is not present in the *E. coli* system. Overexpression of certain

σ factors in *E. coli* produces insoluble aggregates, which may prevent interactions between recombinant fusion proteins. The next step in trying to identify the RcGTA σ factor would be by purifying the 2 RpoE-like σ factors and RpoD to investigate *in vitro* interactions with both RsbW and the RcGTA promoter element. An alternative explanation could lie in the existence of a currently unrecognized multi-protein σ factor, similar to what was unexpectedly discovered in *B. subtilis* (MacLellan *et al.*,2009).

The closely related, but notably different, *R. sphaeroides* utilizes RpoE, along with RpoHI and RpoHII, to control gene expression in response to singlet oxygen, photooxidative and heat stress (Nuss *et al.*, 2009, 2010). It is now known from this thesis work that RpoHII does not seem to play a substantial role in stress response (under the conditions tested) in *R. capsulatus* as determined by phenotypic similarities of the mutant compared to wild type. However, judging by the odd characteristics of the *rpoHI* mutant in *R. capsulatus*, it is likely this protein provides some essential stress response function. The *rpoHI* mutation's effect on viability resembles that of the *rsbV* and *rsbU* mutations, but the unusually slow growth, delayed production of the RcGTA capsid protein, and phenotype of a strictly photoheterotrophic *rpoHI* mutant imply that *R. capsulatus* carrying this mutation is compromised for some important stress response, possibly related to photooxidative and/or heat stress as is the case of RpoHI in *R. sphaeroides*.

I believe that the Rsb regulatory proteins are affecting the activity of either RpoD or several *R. capsulatus* σ factors. An Rsb-like stress response system designed to

regulate either the major vegetative σ factor or several other important σ factors to control RcGTA would represent a significant discovery in the field of molecular microbiology. Anti-σ factors controlling σ^{70} proteins have previously been identified in other proteobacterial species, such as AsiA and Rsd of *E. coli* (Hofmann *et al.*, 2011; Severinova *et al.*, 1998). Another possibility is that the Rsb proteins act in some multi-component structure that has key interactions with several σ factors. Anti-σ factors have been shown to interact with several σ factor partners to control a variety of physiological processes (Evans *et al.*, 2003; Hofmann *et al.*, 2011).

An interesting finding by Hynes *et al.* (2012) determined that ~3% of cells in a *R. capsulatus* population are responsible for 95% of RcGTA production. This leads to the hypothesis that only a small sub-population of cells is redirecting a σ factor to highly express RcGTA genes. This heterogeneous population would have different regulatory pathways activated to produce RcGTA and promote cell lysis. While this finding is incredibly interesting and may help to explain how the cells use a seemingly essential σ factor to drive RcGTA transcription, it also requires re-evaluation of regulatory mutant strains and their phenotypes. It is possible that the signalling pathways and protein activity in the other ~97% of cells in a population are uniquely different from the RcGTA producing sub-population in order to preserve the overall culture viability. Mutations in these key regulatory genes would be population-wide instead of only affecting the RcGTA producing cells. Further work on these key regulators and how their function in a

sub-population may differ from the majority of cells is crucial to furthering our understanding of RcGTA regulation.

The work I have presented in this thesis will provide future research with the framework and direction to further study regulation of RcGTA gene expression and production. In addition to what has previously been known (Fig. 5.1A), I propose a new model of RcGTA regulation (Fig. 5.1B) in this species. By investigating the control of RcGTA gene expression and the role of CtrA in this species, I have determined the involvement of additional signalling and regulatory components required for proper gene expression and production of this fascinating genetic exchange element.

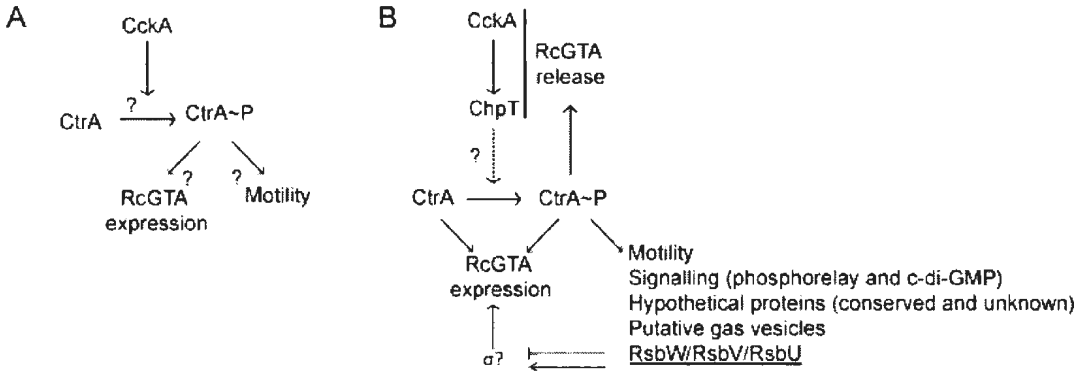


Figure 5.1. The regulation of production of RcGTA in *R. capsulatus*. A. Previous model of RcGTA regulation. B. New model of RcGTA regulation. Arrows indicate positive activity (direct, indirect or predicted); lines with horizontal bars indicate negative activity.

5.1. References

Alvarez-Martinez, C. E., R. F. Lourenço, R. L. Baldini, M. T. Laub and S. L. Gomes (2007). The ECF sigma factor σ^T is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. *Mol. Microbiol.* **66**(5): 1240-1255.

Brown, P. J. B., G. G. Hardy, M. J. Trimble, Y. V. Brun and K. P. Robert (2008). Complex regulatory pathways coordinate cell-cycle progression and development in *Caulobacter crescentus*. In *Adv. Microb. Physiol.*, R.K. Poole (Ed). Academic Press. **Volume 54:** 1-101.

Chen, Y. E., C. G. Tsokos, E. G. Biondi, B. S. Perchuk and M. T. Laub (2009). Dynamics of two phosphorelays controlling cell cycle progression in *Caulobacter crescentus*. *J. Bacteriol.* **191**(24): 7417-7429.

Evans, L., J. Clarkson, M. D. Yudkin, J. Errington and A. Feucht (2003). Analysis of the interaction between the transcription factor sigmaG and the anti-sigma factor SpoIIAB of *Bacillus subtilis*. *J. Bacteriol.* **185**(15): 4615-4619.

Gourion, B., A. Francez-Charlot and J. A. Vorholt (2008). PhyR is involved in the general stress response of *Methylobacterium extorquens* AM1. *J. Bacteriol.* **190**(3): 1027-1035.

Gourion, B., S. Sulser, J. Frunzke, A. Francez-Charlot, P. Stiefel, G. Pessi, J. A. Vorholt and H.-M. Fischer (2009). The PhyR-σEcfG signalling cascade is involved in

stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **73**(2): 291-305.

Hecker, M., J. Pané-Farré and V. Uwe (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu. Rev. Microbiol.* **61**(1): 215-236.

Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Micro.* **7**(4): 263-273.

Hofmann, N., R. Wurm and R. Wagner (2011). The *E. coli* anti-sigma factor Rsd: Studies on the specificity and regulation of its expression. *PLoS ONE* **6**(5): e19235.

Hynes, A. P., R. G. Mercer, D. E. Watton, C. B. Buckley and A. S. Lang (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol. Microbiol.* **85**(2): 314-325.

Jarrell, K. F. and M. J. McBride (2008). The surprisingly diverse ways that prokaryotes move. *Nat. Rev. Micro.* **6**(6): 466-476.

Lang, A. S. and J. T. Beatty (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. USA* **97**: 859-864.

MacLellan, S. R., V. Guariglia-Oropeza, A. Gaballa and J. D. Helmann (2009). A two-subunit bacterial σ-factor activates transcription in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* **106**(50): 21323-21328.

Mercer, R. G., S. J. Callister, M. S. Lipton, L. Pasa-Tolic, H. Strnad, V. Paces, J. T. Beatty and A. S. Lang (2010). Loss of the response regulator CtrA causes pleiotropic effects on gene expression but does not affect growth phase regulation in *Rhodobacter capsulatus*. *J. Bacteriol.* **192**(11): 2701-2710.

Mills, E., I. S. Pultz, H. D. Kulasekara and S. I. Miller (2011). The bacterial second messenger c-di-GMP: mechanisms of signalling. *Cell. Microbiol.* **13**(8): 1122-1129.

Nuss, A. M., J. Glaeser, B. A. Berghoff and G. Klug (2010). Overlapping alternative sigma factor regulons in the response to singlet oxygen in *Rhodobacter sphaeroides*. *J. Bacteriol.* **192**(10): 2613-2623.

Nuss, A. M., J. Glaeser and G. Klug (2009). RpoHII activates oxidative-stress defense systems and is controlled by RpoE in the singlet oxygen-dependent response in *Rhodobacter sphaeroides*. *J. Bacteriol.* **191**(1): 220-230.

Sauviac, L., H. Philippe, K. a. Phok and C. Bruand (2007). An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *J. Bacteriol.* **189**(11): 4204-4216.

Schirmer, T. and U. Jenal (2009). Structural and mechanistic determinants of c-di-GMP signalling. *Nat. Rev. Micro.* **7**(10): 724-735.

Severinova, E., K. Severinov and S. A. Darst (1998). Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T4 AsiA. *J. Mol. Biol.* **279**(1): 9-18.

Spencer, W., R. Siam, M.-C. Ouimet, D. P. Bastedo and G. T. Marczynski (2009). CtrA, a global response regulator, uses a distinct second category of weak DNA binding sites for cell cycle transcription control in *Caulobacter crescentus*. *J. Bacteriol.* **191**(17): 5458-5470.

Staroń, A. and T. Mascher (2010). General stress response in α-proteobacteria: PhyR and beyond. *Mol. Microbiol.* **78**(2): 271-277.

Appendix 4

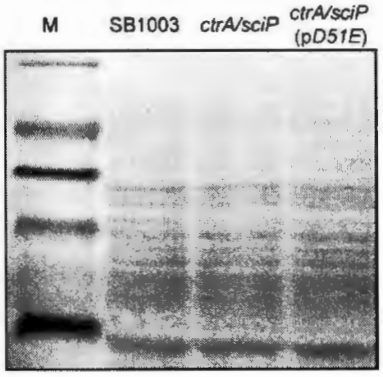


Figure A1 – Representative membrane stained with Ponceau-S for western blot normalization. Pictured is Ponceau stain of cell samples from western blot panel in Fig. 3.3F. Ponceau stains were used for all western blots in thesis to ensure equivalent loading of samples and successful equivalent transfer of proteins from the SDS-PAGE gel to the nitrocellulose membrane. M - molecular weight marker.

Appendix 5

Table A1. List of genes disrupted by insertional mutagenesis and the effect on RcGTA production in the corresponding mutant strains generated from this thesis work

Gene	Annotation	RcGTA Production	Referenced in
rcc00042	PAS/PAC sensor domain protein	~300% of wt (?)	n/a
rcc00142	conserved hypothetical protein	wt	n/a
rcc00180	Hpt domain protein	wt	n/a
rcc00181	response regulator receiver domain/protein phosphatase 2C domain protein	20% of wt (E)	Chapter 4
rcc00458	RNA polymerase σ^{32} factor (*rpoHII*)	wt	Chapter 4
rcc00537	response regulator receiver protein	wt	n/a
rcc00555	protein of unknown function DUF847	X% of wt (R)	Hynes *et al.*, 2012
rcc00699	RNA polymerase σ^{24} factor	Wt (*)	Chapter 4
rcc01335	GAF modulated σ^{54} specific transcriptional regulator, Fis family	wt	n/a
rcc01662	protein of unknown function DUF1153 (*sciP*)	wt	Chapter 3 (Mercer *et al.*, 2012)
rcc01663	cell cycle transcriptional regulator CtrA	0% of wt	Chapters 1 and 2 (Mercer *et al.*, 2010,

			2012)
rcc01749	signal transduction histidine kinase (*cckA*)	2% of wt (R)	Chapter 3 (Mercer *et al.*, 2012)
rcc02289	two-component response regulator receiver protein	wt	n/a
rcc02291	RNA polymerase sigma factor, σ^{70} family, ECF subfamily	wt	Chapter 4
rcc02459	transcriptional regulator, XRE family	wt	n/a
rcc02489	methylated-DNA--[protein]-cysteine S-methyltransferase Ada	wt	n/a
rcc02637	RNA polymerase sigma factor, σ^{70} family, ECF subfamily	wt (*)	Chapter 4
rcc02675	transcriptional regulator, MarR family	wt	n/a
rcc02724	RNA polymerase sigma factor, σ^{70} family, ECF subfamily	wt	Chapter 4
rcc02811	RNA polymerase σ^{32} factor (*rpoHI*)	1% of wt (*)	Chapter 4
rcc03000	conserved hypothetical protein (*chpT*)	1% of wt (R)	Honours student (M. Quinlan, 2010-2011)
rcc03052	conserved domain protein	wt	n/a
rcc03145	transcriptional regulator, AsnC/Lrp family	wt	n/a

rcc03176	PAS/PAC sensor domain protein	wt	n/a
rcc03177	EAL domain protein	wt	n/a
rcc03186	ATPase, AAA family	wt	German exchange student (K. Grebe, 2010)
rcc03207	conserved hypothetical protein	wt	Honours student (A. White, 2011-2012)
rcc03208	tyrosine recombinase XerC	wt	Honours student (A. White, 2011-2012)
rcc03209	protein containing DUF484	wt	Honours student (A. White, 2011-2012)
rcc03322	GAF domain protein	wt	n/a
rcc03323	anti-σ-factor antagonist (*rsbV*)	20% of wt (E)	Chapter 4
rcc03324	anti-σ regulatory factor (*rsbW*)	~300% of wt (E)	Chapter 4
rcc03401	band 7 protein	wt	n/a
rcc03452	sensor histidine kinase/response regulator receiver protein (*rsbK*)	wt	Chapter 4

(?) – Cause of phenotype unknown at this time

(*) – Mutation not confirmed or phenotype predicted to be due to a suppressor mutation

(E) – Phenotype predicted to be due to altered expression of RcGTA

(R) – Phenotype predicted to be due to altered release of RcGTA

Appendix 6 - other scientific contributions

Hynes, A.P., R.G. Mercer, D.E. Watton, C.B. Buckley, and A.S. Lang. DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Molecular Microbiology* 85: 314-325