Studies on the Regulation of the Gene Transfer Agent (GTA) of *Rhodobacter capsulatus*

by

SARAH MARIE FLORIZONE

B.Sc., University of British Columbia, 2002

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Microbiology and Immunology)

THE UNIVERSITY OF BRITISH COLUMBIA
May 2006

Abstract:

The gene transfer agent (GTA) of *Rhodobacter capsulatus* acts as a system of genetic exchange in this purple, non-sulphur, photosynthetic bacterium. GTA is a small bacteriophage-like particle that transfers random 4.5 kbp segments of the donor cell's genome to recipient cells where allelic replacement can occur. The structural gene cluster for GTA encodes 15 open reading frames, most of which are homologous to known phage genes. Three proteins have been found that regulate the expression of GTA genes: GtaI, CtrA and CckA. GtaI is involved in quorum sensing while CtrA and CckA are thought to be part of a sensor kinase/response regulator signalling pathway. I set out to elucidate this poorly understood regulatory pathway. My initial hypothesis was that there are other regulatory proteins besides GtaI, CtrA and CckA, as mutations in any of the genes disrupts, but does not completely abolish, GTA structural gene expression. However, during mutant screening, it became apparent that the reporter plasmid used to monitor GTA gene expression (pYP) caused a decrease in GTA transduction, which I termed the pYP effect. Promoter analysis was performed to determine which part of plasmid pYP was causing this effect and to study *cis*-active regulatory sequences of the GTA structural gene cluster. Sequencing the GTA promoter from 3 strains that exhibit different levels of GTA gene expression and transduction did not find any differences, suggesting a *trans*-acting factor causes the observed differences. 5' mRNA mapping was used to find the 5' end of the GTA transcript, which overlaps with the upstream gene. Capsid levels throughout the growth curve were measured for Y262 (GTA overproducer), B10 (wild type), YCKF2 (*ctrA*$^-$), and YKKR2 (*cckA*$^-$). The immunoblots from these strains showed interesting differences in capsid production that give us a better understanding of GTA production and release from the cell in different stages of growth. Although I was unable to provide a clear pathway for the regulation of GTA, the data in this thesis gives us a better understanding of the complexity of GTA production and regulation.

Table of Contents

List of Tables:

page #

List of Figures:

List of Abbreviations:

aa	amino acid
Ap (Ap^R)	ampicillin (ampicillin resistance)
BLP	bacteriophage-like particle
BSA	bovine-serum albumin
CFU	colony forming units
DMSO	dimethyl sulphoxide
ds	double stranded
GTA	gene transfer agent
HSL	homoserine lactone
kb	kilobase pair
kDa	kilodalton
Km (Km^R)	kanamycin (kanamycin resistance)
KU	Klett unit
LB	Luria-Bertani medium
Mb	Megabase pair
NSPPB	Nonsulfur, purple photosynthetic bacteria
OD	optical density
ORF	open reading frame
phage	bacteriophage
RCV	*Rhodobacter capsulatus* minimal growth medium
S	Svedberg unit
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
Tc (Tc^R)	tetracycline (tetracycline resistance)
Tn5	Transposon 5
YPS	yeast-peptone salt (rich media)

1 Introduction

1.1 Horizontal gene exchange

Horizontal gene transfer between bacteria is an important mechanism contributing to the evolution of bacterial genomes. It is a critical area of research to the medical community as the number of multidrug resistant bacterial strains continues to increase due to the over prescribing of antibiotics and the prevalence of virulence genes that can be readily transferred between bacteria (Davies, 1994). Horizontal gene transfer is thought to be integral to the evolution of pathogenic enteric bacterial strains as it has been recently found that horizontally acquired 'pathogenicity islands' are important for transferring virulence between pathogenic and non-pathogenic strains (Ochman *et al.*, 2000).

There are three main types of horizontal gene exchange in bacteria: transformation, conjugation and transduction, which are summarized in Figure 1. Transformation involves a competent (either natural or induced) host cell taking up naked DNA from the environment. DNA from different species can be taken up in this fashion, although a recognition sequence is required, and the DNA can be integrated into the chromosome by homologous recombination if it contains regions homologous to the hosts' DNA. Natural DNA uptake is thought to be advantageous to the cell on a nutritional basis as it can provide substrates for growth, or at least help to alleviate the demand for new nucleotides to be synthesized when nutrients are low, as opposed to being a mechanism solely for genetic exchange as previously thought (Redfield, 2001).

Conjugation involves the transfer of plasmids or mobile elements via cell-to-cell contact. There are various plasmids that carry mobilization elements, such as the F plasmid in *Escherichia coli*, the Ti plasmid in *Agrobacterium* and broad host-range antibiotic resistance (R)

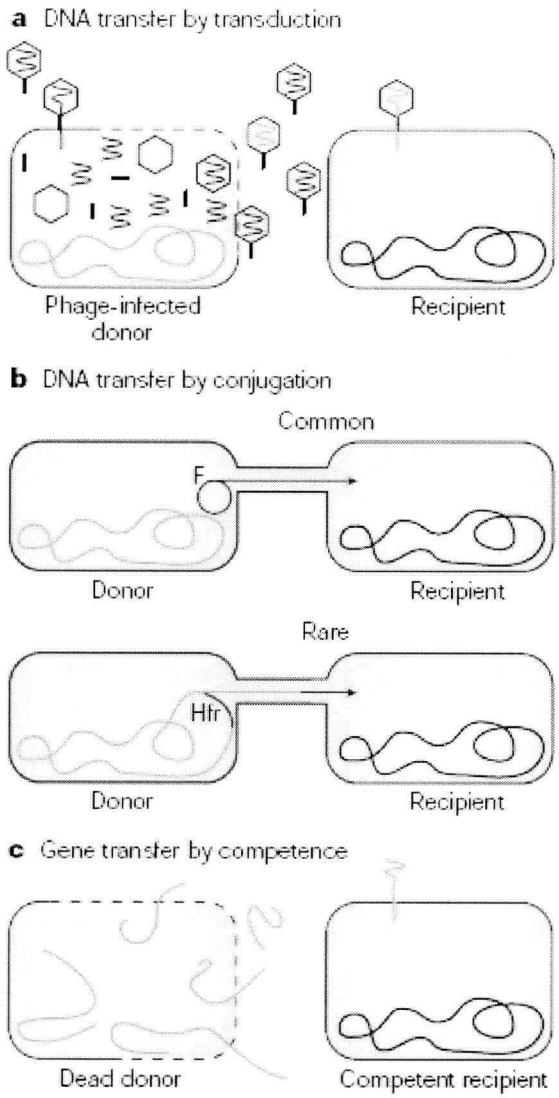

Figure 1: Three types of genetic exchange in bacteria: transduction, conjugation and transformation. Transduction is shown in a) in which a phage accidentally packages bacterial DNA and transfers it to a recipient cell upon phage infection. Conjugation is shown in b) in which genetic material is transferred via cell to cell contact. If the mobilization element is integrated into the chromosome, the entire chromosome can be transferred instead of the plasmid by itself. Transformation by competent cells is shown in c) in which a lysed cell's DNA is taken up by a recipient competent cell. This DNA can then be integrated into the chromosome or digested and used for nutrients. Figure taken from (Redfield, 2001).

plasmids in a variety of species (Griffiths, 2005). Very rarely, if the mobile element integrates into the chromosome, the entire chromosome may be transferred from the donor cell into the recipient cell (Campbell, 1996).

Transduction is the transfer of host genes between bacteria via a bacteriophage. Bacteriophage usually only transfer their own genes, however bacterial DNA can be accidentally incorporated into the bacteriophage by either specialized or generalized transduction. Specialized transduction occurs when host DNA encoding one or more specific genes is excised with the bacteriophage DNA, because of aberrant excision of a prophage that is integrated into the bacterial chromosome. Generalized transduction occurs when random fragments of host DNA are packaged into the head of the bacteriophage instead of the bacteriophage DNA. The genes that are incorporated or packaged by either method can then be transferred to other cells upon bacteriophage infection, and the DNA can be inserted into the chromosomal DNA by homologous recombination (Campbell, 1996). *Bartonella bacilliformis* is an example of an organism with virulence genes being transferred between strains by bacteriophage-like particle (BLP) transduction. *B. bacilliformis* is transmitted by the bite of the female sandfly, and causes Oroya fever (haemolytic type disease) (Barbian and Minnick, 2000).

1.2 Gene transfer agents in prokaryotes

A common property to all the BLPs (bacteriophage-like particles or gene transfer agents) described below is that they package an amount of DNA insufficient to encode a complete dsDNA phage, such as λ (Campbell, 1996; Hendrix, 1983). *Rhodobacter capsulatus* GTA packages 4.5 kb of DNA, whereas the gene cluster that encodes its structural components is 15 kb (Lang and Beatty, 2000). It was proposed that the selective advantage afforded to *R. capsulatus* by evolution of GTA was to prevent damage to the host cell by lysis due to infection

by a phage that gave rise to GTA (Redfield, 2001). There are several examples of "constitutive" generalized transduction in distantly related prokaryotes that all engage in genetic exchange between closely related species, so it would seem that there would be a selective advantage for these systems to have evolved. However, it remains to be seen whether the selective pressure was to escape infection and lysis by a phage, or for the purpose of genetic exchange.

Several examples of gene transfer agents that exhibit "constitutive" generalized transduction in prokaryotes have been discovered and their properties are listed in Table 1. *B. bacilliformis*, as mentioned above in Section 1.1, produces BLPs that are 80-nm, tail-less, capsid particles that package 14 kb of linear dsDNA, and is important to the medical community due to the disease *Bartonella* is responsible for producing. Dd1, from *Desulfovibrio desulfuricans*, packages 13.5 kb fragments of linear dsDNA into a tailed phage structure with head diameter of 43 nm and tail length of 7 nm, and like GTA (see below) is not induced by mitomycin. The spirochete *Brachyspira hyodystenteria* produces VSH-1, which packages 7.5 kb fragments of dsDNA into a tailed phage with a head diameter of 45 nm and a tail length of 64 nm, and is the only gene transfer agent that is inducible by mitomycin (Humphrey, 1997; Matson *et al.*, 2005). *Methanococcus voltae*, a methanogenic archaebacterium, produces VTA (for *voltae* transfer agent) which is the most similar bacteriophage-like particle to GTA in that it transfers 4.4 kb of linear dsDNA, is a tailed phage with head diameter of 40 nm and tail length of 61 nm and is not inducible by mitomycin.

1.3 Transduction and gene exchange in R. capsulatus

In the early 1970's Barry Marrs was searching for genetic exchange tools for studying photosynthesis in nonsulphur, purple, photosynthetic bacteria (NSPPB), when he discovered a novel genetic exchange vector that causes generalized transduction between *R. capsulatus* strains

Table 1: Summary of gene transfer agent properties.

Agent	Host	Size and nature of nucleic acid	Structure	Head diameter, tail length	Mitomycin induction	References
GTA	*Rhodobacter capsulatus*	4.5 kb linear dsDNA	Tailed phage	30 nm, 50 nm	No	(Marrs, 1974; Solioz and Marrs, 1977; Yen *et al.*, 1979)
Dd1	*Desulfovibrio desulfuricans*	13.6 kb linear dsDNA	Tailed phage	43 nm, 7 nm	No	(Rapp and Wall, 1987)
VSH-1	*Brachyspira hyodystenteriae*	7.5 kb linear dsDNA	Tailed phage	45 nm, 64 nm	Yes	(Humphrey, 1997)
VTA	*Methanococcus voltae*	4.4 kb linear dsDNA	Tailed phage	40 nm, 61 nm	No	(Bertani, 1999; Eiserling *et al.*, 1999)
BLP	*Bartonella bacilliformis*	14 kb linear dsDNA	Capsid particle	80 nm, NA	No	(Barbian and Minnick, 2000)

(Marrs, 2002). *Rhodobacter* species are found in very diverse locations including freshwater, marine, and hypersaline environments as well as soils and paddy fields (Imhoff and Truper, 1989; Imhoff, 1995). They can be grown in the lab under various conditions including anaerobically in the presence of light (photosynthetically), aerobically in the dark, or anaerobically in the dark if an alternative electron acceptor is present, such as dimethyl sulfoxide (DMSO) (Madigan *et al.*, 2003). It is interesting that a bacterium found in such varied environments would have a mechanism for exchanging genetic information, which could further the survival of the population by transfer of an allele that improves fitness when conditions are changed.

Marrs was interested in finding tools of genetic exchange to further the study of photosynthesis because the research at that time was limited to studying individual mutants, and so linkage of mutations (genetic maps) could not be obtained (Marrs, 2002). Marrs and his lab members obtained soil and pond samples from a park in Saint Louis from which they isolated NSPPBs by enrichment. From each isolate (i.e. A, B, C, etc.), they selected one spontaneous rifampicin- and one spontaneous streptomycin-resistant mutant, which they mixed in a pair-wise manner (i.e. A rif^r + B $strep^r$, A $strep^r$ + B rif^r, etc.) and allowed them to grow to stationary phase in liquid culture. They then spread the cultures on plates containing both streptomycin and rifampicin and looked for a double mutant that contained both antibiotic resistance genes. To ensure that the double mutant had not arisen by spontaneous mutation, they also plated each single mutant alone to see how often the second mutation would spontaneously occur. In 1974 Marrs published the first description of this novel genetic exchange system found in a NSPPB although he had little idea at that time what he had actually discovered (Marrs, 1974).

Regardless, this system was soon used to map genes needed for bacteriochorophyll and carotenoid photosynthetic pigment synthesis (Yen and Marrs, 1976).

Upon further analysis Marrs found that this system of genetic exchange was not a typical mechanism of DNA exchange as described in Section 1.1 (Marrs, 2002). DNase did not block the genetic transfer, which ruled out transformation of DNA from naturally lysed cells (Marrs, 1974). The genetic material was shed into the medium and the cells did not need to be in contact for the transfer to occur, which ruled out conjugation (Marrs, 1974). The genetic exchange vector, which was named as the gene transfer agent or GTA, behaved as discrete particles that sedimented in sucrose density gradients as a tight band at 70S, which is about the size of a ribosome and much smaller than any known dsDNA transducing phage (Solioz *et al.*, 1975). All genetic markers examined could be transferred and ~4 x 10^{-4} is the maximum fraction of recipients that can acquire a given genetic marker, which is similar to other generalized transduction systems (Solioz *et al.*, 1975; Solioz and Marrs, 1977). However, every GTA particle is thought to be capable of transduction, in contrast to generalized transduction systems, where only abnormal particles perform generalized transduction (Solioz *et al.*, 1975). The properties described above, along with there being no plaques observed and no apparent cell lysis upon production of the particles, rules out GTA as being a typical phage-mediated generalized transduction system (Marrs, 1974).

Yen *et al.* (Yen *et al.*, 1979) isolated a GTA overproducing mutant strain, Y262, which aided in the further characterization of GTA because of the larger amounts of GTA particles. Isolation of DNA from the GTA particles revealed that this DNA is linear, double stranded, approximately 4.5 kbp in length and randomly packaged, as restriction digests of the DNA isolated from GTA showed similar patterns to restriction digests of chromosomal DNA from *R.*

capsulatus (Yen *et al.*, 1979). An electron micrograph of GTA particles was obtained, shown in Figure 2, which showed that GTA resembles a small, tailed bacteriophage, with a head diameter of approximately 30 nm, and a tail length of approximately 50 nm (Yen *et al.*, 1979). This is in contrast to the dsDNA (48.5 kb) tailed phage λ, which has a head diameter of 50 nm and a tail length of 150 nm (Hendrix, 1983). Although most of the physical characteristics were discovered for GTA at this time, little genetic information was known about GTA, and it was still not known how production of GTA was regulated.

1.4 Structural genes encoding GTA in R. capsulatus

In 2000, through the use of a transposon-mutagenized library, Lang and Beatty (2000; 2001) found a gene cluster encoding the GTA structural genes within the *R. capsulatus* genome. Although the *R. capsulatus* genome includes a plasmid (http://www.ergo-light.com/ERGO/), the GTA genes are located in the single, 4.4 Mb circular chromosome, and are flanked by typical bacterial genes (Figure 3). The GTA structural gene cluster is ~15 kb in length, encodes 15 open reading frames (*orfg1* to *orfg15*), and appears to be a polycistronic operon under the control of a single promoter (Lang and Beatty, 2000, 2001). Several of the open reading frames (ORFs) in the GTA gene cluster are homologous to known phage genes, such as the *Streptomyces* phage ϕC31 portal (*orfg3*) and capsid (*orfg5*) proteins, the *E. coli* T4 phage DNA packaging protein (*orfg2*), and the *E. coli* HK97 phage prohead protease protein (*orfg4*) (Lang and Beatty, 2000). The GTA ORFs are similar only to phage structural genes and no homologues to phage replication or regulatory genes were found. An interesting characteristic of the structural gene cluster for GTA is that the codon usage matches the codon usage of highly expressed *R. capsulatus* photosynthesis genes, implying that this cluster has been a part of the *R. capsulatus* genome longer than two putative prophages that are found in the *R. capsulatus* genome, whose

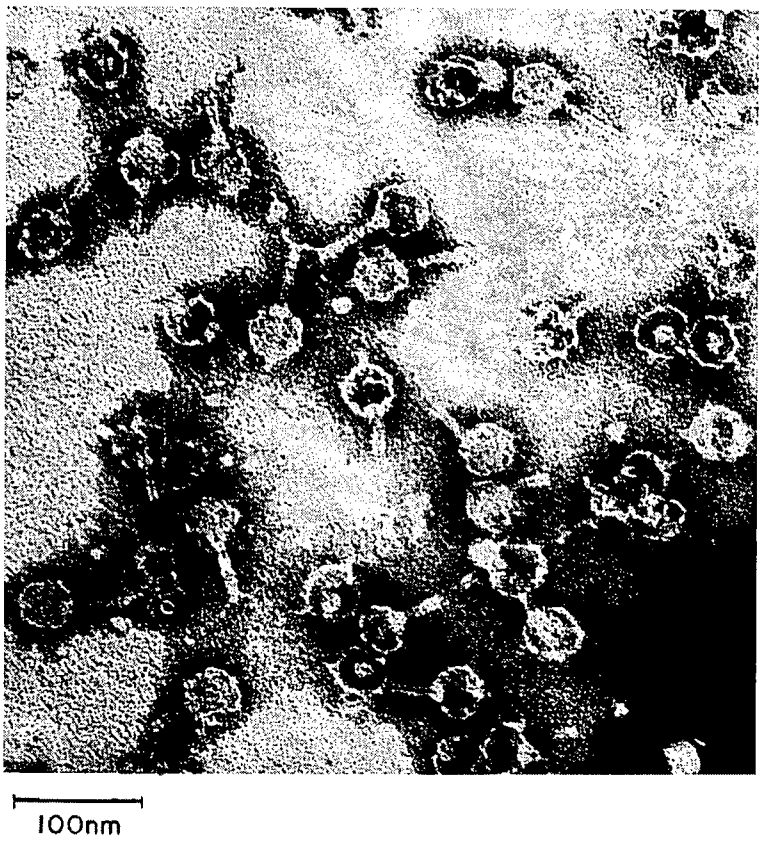

Figure 2: Electron micrograph of GTA negatively stained with phosphotungstic acid. Taken from (Yen *et al.*, 1979).

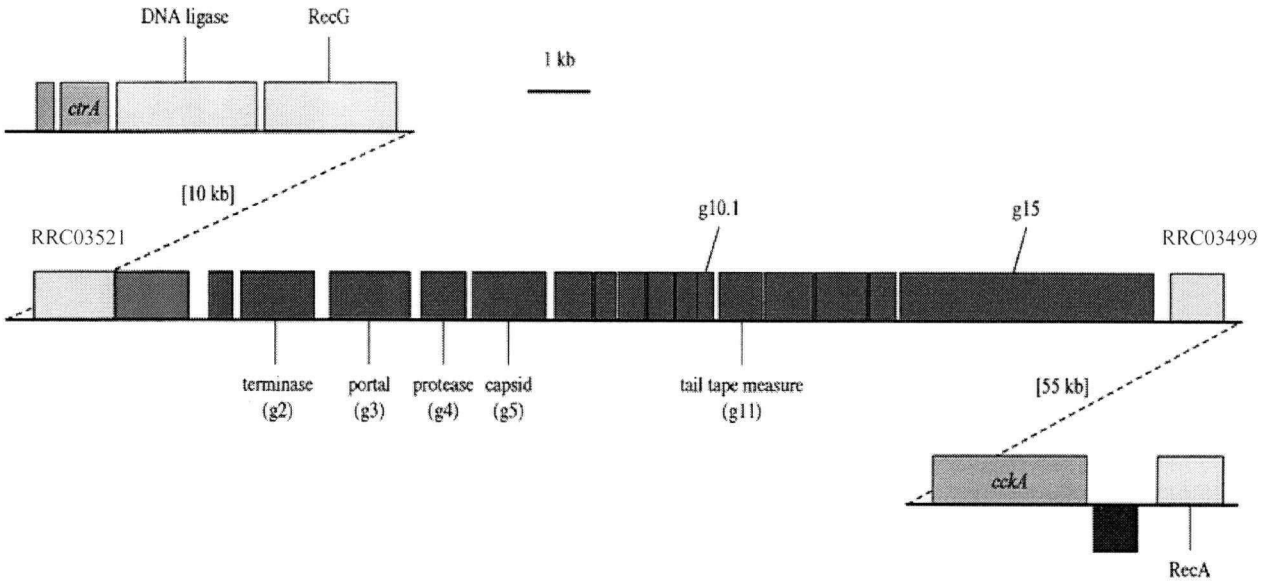

Figure 3: GTA structural gene cluster, *cckA, ctrA* and surrounding genes. Genes believed or known to be involved in GTA production are shown in blue, with GTA structural genes in dark blue and regulatory genes in light blue. Surrounding genes similar to motility (flagellar and chemotaxis) genes are shown in orange, genes similar to cellular genes of other known function are shown in yellow, genes similar to genes of unknown function are shown in red, and genes with no similarities are shown in purple. RRC03521 is annotated as a 3-oxoacyl- (acyl carrier protein) synthase while RRC03499 is a predicted serine acetyltransferase, Boxes drawn above the line represent genes oriented from left to right (5' to 3') and boxes drawn below the line represent genes oriented from right to left (3' to 5'). Modified from Lang and Beatty (2001).

codon usage is different (Lang and Beatty, 2000). Fifteen kb, however, is thought to be too short to encode a complete phage genome, as the complete genomes for HK97 and ϕC31 are approximately 40 kb. The head-to-tail structural regions of the HK97 and ϕC31 genomes are 19 kb and 17 kb respectively (Smith *et al.*, 1999), which are close in size to the GTA structural gene cluster. Since 15 genes is thought to be too few to encode a complete tailed dsDNA phage genome, it is possible that GTA uses regulatory genes from two prophages in the *R. capsulatus* genome (Haselkorn *et al.*, 2001), or that there are other genes involved in GTA production that are yet to be found.

1.5 GTA regulatory genes in R. capsulatus

In addition to the discovery of the GTA structural gene cluster, three genes have been found that encode proteins which regulate the expression of the GTA structural genes: GtaI, CtrA, and CckA. GtaI is involved in quorum sensing, while CtrA and CckA may be part of a sensor kinase (CckA)/response regulator (CtrA)-signalling pathway, as close homologues of these two proteins have been characterized in *Caulobacter crescentus* (Jacobs *et al.*, 2003; Lang and Beatty, 2000; Schaefer *et al.*, 2002).

The *gtaI* gene encodes a long-chain acyl-homoserine lactone (acyl-HSL) synthase. Acyl-HSLs are associated with quorum sensing, where acyl-HSLs are secreted by cells into the medium and then used as signals for gene expression when taken up by other cells in the population. When *gtaI* is mutated, GTA production decreases. Conversely, when exogenous acyl-HSL is added to the mutant culture, GTA production returns to normal levels. GTA levels are highest at stationary phase, when the cells are at the highest density, and so this is consistent with the discovery that quorum sensing regulates GTA activity (Schaefer *et al.*, 2002).

The *R. capsulatus ctr*A gene product acts to positively regulate expression of the GTA structural gene cluster and is predicted to encode a response regulator protein (Lang and Beatty, 2000, 2002). Lang and Beatty (2000) inserted a KIXX cartridge (which contains the *neo* gene encoding neomycin/kanamycin phosphotransferase) into *ctrA*, which reduced GTA gene expression to an undetectable level (>10^3-fold). The *C. crescentus* and *R. capsulatus ctrA* genes have 71% identity and since the predicted DNA recognition helix-turn-helix of *C. crescentus* CtrA is identical to the corresponding region of the *R. capsulatus* CtrA sequence, Lang and Beatty assumed that the binding sites would be the same as well. By comparing the known *C. crescentus* CtrA-binding consensus sequence to the sequence upstream of the *R. capsulatus ctrA* gene, two putative binding sites for CtrA were found, suggesting that CtrA is self-regulated. However, as no predicted CtrA binding sites were found upstream of the GTA structural gene cluster, it was thought that there must be another factor that interacts with the GTA gene cluster promoter region or a different binding site for CtrA is used. That is, perhaps CtrA activates expression of another gene, or transfers a phosphate to another protein, which then binds to the GTA structural gene promoter region to activate transcription (Lang and Beatty, 2000). It is also possible that CtrA binds to a non-canonical sequence in the GTA promoter, as appears to be the case for a minor fraction of CtrA-regulated genes in *C. crescentus* (Laub *et al.*, 2002).

The *R. capsulatus cckA* gene is also required for maximal expression of the GTA structural gene cluster, and is predicted to encode a sensor kinase protein which has 46 % identity (61% similarity) to the *C. crescentus* CckA protein (Lang and Beatty, 2000, 2002). In *C. crescentus*, the CckA protein phosphorylates (activates for DNA binding) CtrA (Jacobs *et al.*, 2003). Therefore it was assumed that in *R. capsulatus* CckA phosphorylates CtrA as part of a sensor kinase/response regulator-signalling pathway to induce GTA gene expression (Lang and Beatty, 2000). As explained in more detail below, when *cckA* is mutated, the level of GTA gene

expression and GTA activity decreases (50% decrease for expression of a plasmid-borne *lacZ* reporter driven by the GTA promoter, and 99% decrease for GTA transduction). But this decrease is not as low as in the *ctrA* mutant (85% decrease for expression of a plasmid-borne *lacZ* reporter driven by the GTA promoter and 100% decrease for GTA transduction). This shows that CckA is not as important as CtrA to GTA gene expression, and that induction of GTA gene expression is more complicated than a single system in which CckA senses a signal to cause autophosphorylation and transfer of phosphate to CtrA (Lang and Beatty, 2001). It is also possible that CckA acts in a different pathway than CtrA, and that there is more than one sensor kinase that activates CtrA.

The 550 bp sequence located 5' of the second gene in the GTA structural gene cluster, *orfg2*, was identified as the promoter region, on the basis of a translationally in-frame fusion of *orfg2* to the *E. coli lacZ* gene in plasmid pYP (Lang and Beatty, 2000). It was assumed that the promoter is located somewhere in the 550 bp region 5' of *orfg1*. Since there are no canonical CtrA binding sites in the promoter region of the GTA gene cluster, it was thought that there must be another factor that is acted upon by CtrA. It was suggested that since a sequence similar to an RNA polymerase σ-factor binding site (-10 promoter sequence) is found in the GTA structural gene promoter region, the activation of an alternative (stationary-phase) σ-factor, or other transcription factor, is regulated by CtrA. This other factor would then directly enhance GTA structural gene expression (Lang and Beatty, 2001).

1.6 Overview of the thesis

I hypothesized that there are GTA regulatory proteins other than GtaI and CckA, since mutations in the genes for these proteins reduce GTA structural gene transcription and transduction activity, but do not completely abolish both of these activities. I hypothesized that

there are other proteins involved in the CckA/CtrA pathway because: 1) a CtrA consensus binding site in the GTA promoter region has not been found; 2) the *cckA* knockout reduces GTA gene expression less severely than the *ctrA* knockout. I intended to identify these proteins that are in the same pathway as CtrA and CckA, or that act independently to regulate GTA structural gene expression (such as a sigma factor or transcription factor), and that the predicted properties of these proteins would elucidate how GTA is regulated in *R. capsulatus*.

My initial approach was to use the plasmid pYP (described in Materials and Methods and (Lang and Beatty, 2000) that drives *lacZ* expression from the GTA structural gene cluster promoter region to screen a *R. capsulatus* library of transposon mutants, and identify new genes needed to activate GTA gene expression. However, as my research progressed, I discovered multiple, interesting complexities that led me into new areas. I found and DNA-sequenced several mutations that reduced expression of the *lacZ* allele on pYP, and I also discovered that the pYP plasmid itself greatly reduced GTA production. In addition to the GTA structural gene promoter region, pYP contains *orfg1* and a segment of *orfg2* fused translationally in-frame to the *lacZ* reporter gene (see Materials and Methods). Thus it was not clear whether it was the extra copies (~5 per cell) of the GTA structural gene promoter region on pYP or the over-expression of *orfg1*, which reduced GTA expression in host cells. As the reasons for the effects of the transposon mutations on GTA gene expression were unclear, whereas the reduction in GTA production resulting from the presence of pYP in cells appeared to be more amenable to experimental analysis, I changed the focus of my research. The new focus was to better localize the GTA structural gene cluster promoter, and evaluate the sequences on pYP that reduce GTA production when that plasmid is present.

This thesis summarizes my initial work on using the plasmid pYP to identify transposon mutations of genes that appear to be needed for maximal expression of GTA structural genes.

Subsequent experiments were done on the GTA structural gene promoter region, to clarify whether the pYP effect arises from the expression of GTA genes (*orfg1* and *orfg2::lacZ*), in pYP, or the presence of activator binding sites in several copies relative to the chromosome. I also mapped the 5' ends of RNAs to two specific nucleotides in the promoter region, and propose -10 and -35 promoter sequences. Additionally, I used an immunoassay based on western blots of intact cells and culture supernatants probed with anti-GTA capsid protein antisera (Taylor, 2004), to evaluate the production of this protein (and by inference, production of GTA particles) in cultures of several strains of *R. capsulatus*.

The data I obtained improve the understanding of GTA gene expression in terms of cellular responses to culture conditions, and identify likely -10 and -35 regions of the GTA promoter.

2 Materials and Methods

2.1 Bacterial Strains, plasmids and growth conditions

The strains and plasmids used are described in Table 2. *E. coli* strains were grown in Luria-Bertani (LB) medium supplemented with 150 μg/ml ampicillin, 10 μg/ml tetracycline-HCl or 25 μg/ml kanamycin sulphate, as needed for selection for plasmids. *R. capsulatus* strains were grown aerobically in RCV minimal medium (Beatty and Gest, 1981) or photosynthetically in yeast extract/peptone/salts (YPS) medium (Wall *et al.*, 1975) supplemented with either 0.5 μg/ml tetracycline-HCl or 10 μg/ml kanamycin sulphate, as needed for plasmid or mutant selection.

B10 is a wild type strain of *R. capsulatus* that is rifampicin sensitive (Marrs, 1974) while SB1003 is a rifampicin resistant strain of *R. capsulatus* created in a B10 derivative strain, B100 by GTA transduction with a spontaneous rifampicin resistant strain, BB101 (Yen and Marrs, 1976, 1977). Y262 is a GTA overproducer strain made by nitrosoguanidine mutagenesis of a spontaneous streptomycin-resistant strain, BB103 which is a derivative of B10 (Yen *et al.*, 1979). The exact mutations that make Y262 an overproducer are not known, although it is assumed there are at least two separate mutations because two GTA transductions from Y262 were required to convert SB1003 into a GTA overproducer (Yen *et al.*, 1979). The strains YKKR2 (*cck::KIXX*) and YCKF2 (*ctrA::KIXX*), created in Y262(Lang, 2000; Lang and Beatty, 2000), were used as controls to compare to the GTA activity of the new mutants described later. The strain IKOI (*gtaI*), which was created in SB1003, is a knockout mutant deficient in quorum sensing and GTA activity (Schaefer *et al.*, 2002).

Table 2: Strains used in this study.

Strain	Genotype	Phenotype	Source
R. capsulatus			
B10	wild type	wild type	(Marrs, 1974)
Y262	Unknown mutations	GTA overproducer	(Yen *et al.*, 1979)
SB1003	*rif-10*	Rifampicin resistant, cured of phage	(Yen and Marrs, 1976)
YCKF2	*ctrA*::KIXX	Y262 background, no detectable GTA production	(Lang and Beatty, 2000)
YKKR2	*cckA*::KIXX	Y262 background, reduced GTA production	(Lang, 2000)
IKOI	*gtaI*::Ω	SB1003 background, reduced GTA production, quorum sensing mutant	(Schaefer *et al.*, 2002)
DW5	*puhA*⁻	Lacks RCH protein no photosynthetic growth.	(Wong *et al.*, 1996)
E. coli			
DH10B		Cloning strain	Invitrogen
S17-1		Cloning strain, capable of conjugation to *R. capsulatus*.	(Simon *et al.*, 1983)

Table 3: Plasmids used for this study.

Plasmid	Description and Markers	Source
pXCA601	Promoter probe vector, Tc^R	(Adams *et al.*, 1989)
pYP	Y262 GTA structural gene cluster promoter on a PstI to SalI fragment fused to *lacZ* in pXCA601, Tc^R	(Lang, 2000)
pYnP	Y262 GTA structural gene cluster *orfg2* lacking the promoter region (EcoRI-Sal fragment) in pXCA601, Tc^R	(Lang, 2000)
pSTU12	B10 structural gene cluster promoter on a PstI to HindIII fragment fused to *lacZ* in pXCA601, Tc^R	(Lang, 2000)
pX/S5	Subclone for creation of pYP	(Lang, 2000)
P9HSTU	Subclone for creation of pSTU12	(Lang, 2000)
pSMF001	Y262 GTA structural gene cluster promoter deletion fused to *lacZ* in pXCA601, Tc^R	This work
pSMF002	Y262 GTA structural gene cluster promoter deletion fused to *lacZ* in pXCA601, Tc^R	This work

The plasmid pXCA601 (Adams *et al.*, 1989) was used to evaluate promoter activity of various GTA structural gene promoter region fusions to *lacZ,* using translationally in-frame fusions to the 8th codon of *lacZ*, using a BamHI site. The inserts fused to *lacZ*in plasmids pYP and pYnP are shown in Figure 4. Plasmid pYP contains a translationally in-frame fusion of the GTA structural gene cluster *orfg2'* to '*lacZ* under the control of the putative GTA promoter, which is presumed to be located in the 550 bp region 5' of the start codon for *orfg*1 and 3' of the *Pst*I site near the 3' end of the gene directly upstream of the GTA gene cluster (RRC03520) (Lang and Beatty, 2000). pYnP contains the same fusion to *lacZ* as pYP, although its 5' sequence extends only up to the EcoRI site that is located 3' of the start codon for *orfg2,* (Figure 4). Plasmid pYnP acts as a control to determine the basal level of *lacZ* expression from the pXCA601 plasmid, but as there is a transcriptional terminator inserted on the 5' side of the *Bam*HI cloning sites in the *lacZ* gene, there is very little, if any, read through from other genes on the plasmid (Adams *et al.*, 1989). Two GTA structural gene cluster promoter deletions were made to determine a smaller portion of the 1.1 kb predicted promoter region that contains the entire promoter. These are described in greater detail below in Section 2.6.

2.2 Creation and screening of a transposon mutagenized R. capsulatus Y262 library containing pYP

A.S. Lang created a transposon-mutagenized library (Lang, 2000; Lang and Beatty, 2000) in Y262, which is a GTA overproducing strain of *R. capsulatus,* following the method of Simon *et al.* (1983). S17-1(pSUP2021) was used to transfer pSUP2021 into Y262 as this plasmid carries the Tn5 and cannot replicate in *R. capsulatus*. The resulting kanamycin resistant Y262 cells have Tn5 randomly inserted in their genome due to transposition events. The library contains approximately 2000 members and is stored in RCV medium (+20% glycerol) at -80°C.

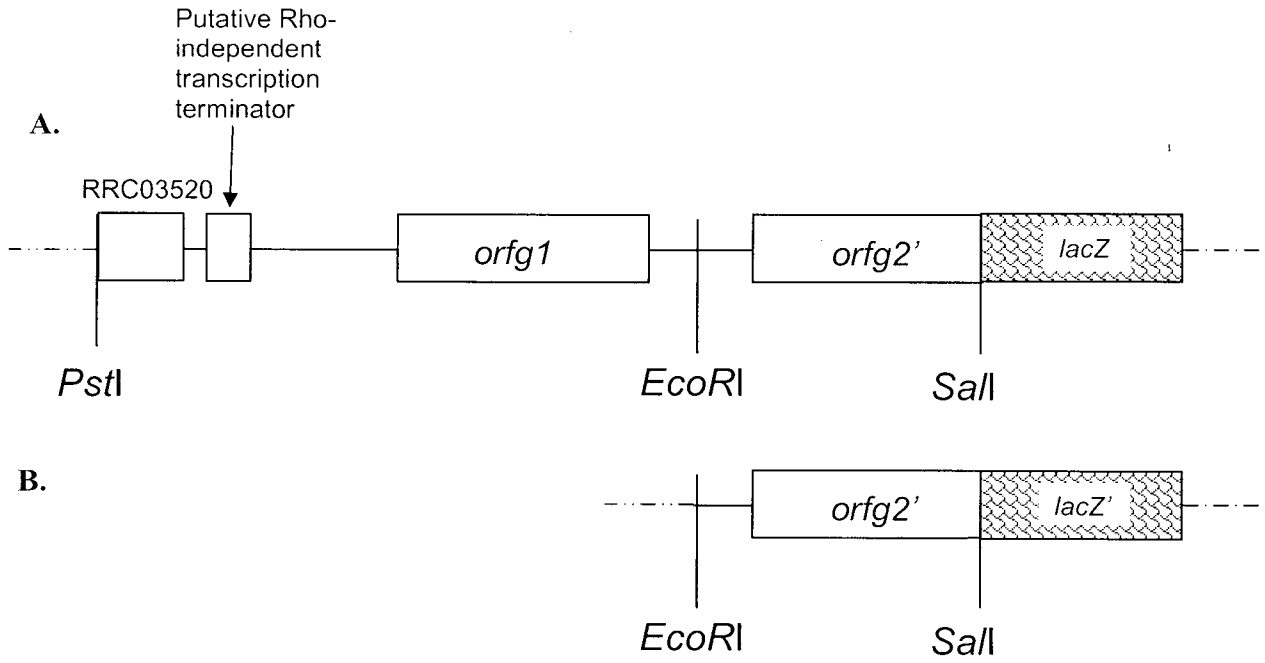

Figure 4: Overview of *R. capsulatus* gene sequences present in plasmids pYP (A) and pYnP (B).

Plasmid pYP (described above) was conjugated into the Y262::Tn5 library so the library could be screened using "blue/white" screening, which in a red (carotenoid) pigmented bacteria, appear as purple/red colonies. The Y262::Tn5 library containing pYP was grown in RCV medium and plated on RCV plates, containing X-gal (40 μg/ml), kanamycin sulphate (10 μg/ml) to select for Tn5 and tetracycline-HCl (0.5 μg/ml) to select for pYP, at approximately 300 colony forming units (CFUs) per plate. The plates were grown for two days under photosynthetic conditions and red colonies were picked for further analysis, as these will have low β-galactosidase activity and hence low expression of the *orfg2::lacZ* fusion on pYP. As red to purple is a much more subjective colour difference than white to blue, it was necessary to have various other tests to ensure that the mutation did in fact reduce GTA gene transcription and biological activity.

The first step was to make a stab plate of the colonies that appeared to have lower *lacZ* expression (red colonies) alongside control strains to compare levels of *lacZ* expression to ensure that it was not simply the size of the colony that made it appear lighter than those around it (Figure 5). The controls used were Y262(pYP) (normal expression; dark purple or blue colonies), YKKR2(pYP) (*cckA* mutant host with reduced expression; lighter purple colonies), YCKF2(pYP) (*ctrA* mutant host with low expression; reddish purple colonies), and Y262(pYnP) (no expression; red colonies). After this secondary screen, the potential mutants were then subjected to 3 more quantitative tests (described in detail below) to measure GTA gene expression: 1) β-galactosidase specific activity (*lacZ* assays); 2) GTA activity (GTA transduction bioassay); and 3) GTA capsid production (GTA immunoassay).

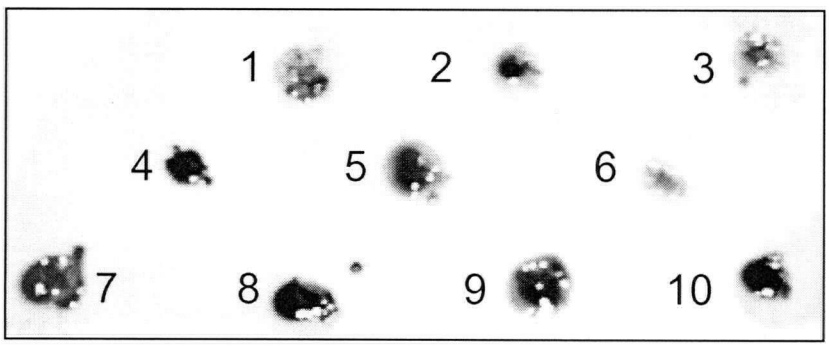

Figure 5: Stab plate containing X-gal of mutants and control strains with either plasmid pYP or pYnP. Colonies 1-6 are mutants PS1-PS6, all containing plasmid pYP; 7 is Y262(pYnP) where pYnP is a *lacZ* fusion to *orfg2* that lacks the promoter region; 8 is YCKF2(pYP) which is *ctrA*$^-$; 9 is YKKR2(pYP) which is *cckA*$^-$; and 10 is Y262(pYP) which is the parental strain of all the mutants.

2.3 *Methods of measuring GTA*

2.3.1 β-galactosidase assays measure transcription from the GTA promoter.

The β-galactosidase specific activities of strains containing GTA ORF fusions to *lacZ* were quantified using the same method as Lang and Beatty (Lang and Beatty, 2000). The cultures were harvested at late stationary phase, as determined by monitoring culture turbidity. The culture turbidity was measured by light scattering using a Klett-Summerson photometer (filter #66, red) (Lang, 2000). β-galactosidase activity was determined by a colourometric assay of *o*-nitrophenol-β-D-galactoside cleavage and is expressed as Miller units (Miller, 1992), which are proportional to the increase in *o*-nitrophenol per minute, as measured by absorbance at 420 nm, per mg of protein. The amount of protein in samples was estimated by Lowry assay (Miller, 1992). Each β-galactosidase assay was performed in triplicate and the absorbance of the colourometric reaction averaged prior to converting to Miller units.

2.3.2 Bioassays to determine frequency of gene transduction.

The amount of active GTA produced by the mutant cultures was measured by GTA bioassays. Cultures of mutant strains of *R. capsulatus* were grown to stationary phase in 17 ml anaerobic screw cap tubes, photosynthetically under high light at 30-35° in YPS medium with no antibiotics. An overnight culture of DW5, which is a *puhA*$^-$ (reaction center H gene knockout) strain that can no longer grow photosynthetically (Wong *et al.*, 1996), was grown aerobically in RCV at 30°C. The DW5 cells were pelleted by centrifugation and resuspended in 3.5 ml 1X G-buffer (10 mM Tris-HCl (pH 8.0), 1 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM NaCl, 500 μg/ml bovine serum albumin fraction V (Solioz *et al.*, 1975). To separate GTA from bacterial cells, the

cultures were filtered through 0.22 or 0.45 μm syringe filters (Millipore, Bedford, MA) into plastic tubes and kept on ice. Plastic tubes must be used as GTA sticks to glass and decreases the efficiency of the assay. Culture filtrate (0.1 ml), DW5 strain of *R. capsulatus* (0.1 ml), and 1X G buffer (0.4 ml) were combined and incubated for 1 to 1.5 hours at 35°C with slow shaking. After 1.5 hours, RCV medium (0.9 ml) was added to each tube and incubated for 3-4 more hours at 35°C. The cultures were pelleted in 1.5 ml tubes, the supernatant decanted, and the bacterial cells were resuspended in 0.1 ml RCV. The cells were then plated on RCV or YPS plates and incubated for 2-3 days in anaerobic Gas-Pak jars (BBL, BD scientific, Franklin Lakes, NJ) under photosynthetic conditions. As DW5 is *puhA*$^-$ and cannot grow photosynthetically, only DW5 cells that receive *puhA* by GTA mediated transduction will be able to grow photosynthetically. The numbers of *puhA*+ colonies were then counted on each plate. Control plates were also included that had only DW5 + G buffer (to ensure that reversions were not occurring) and donor culture filtrate + G buffer (to ensure that all cells were removed by filtration).

2.3.3 GTA immunoassays of the major capsid protein (*orfg5*).

The amount of capsid protein present either in the cell or in the supernatant of culture was measured by immunoblot. The antiserum used to visualize the capsid was obtained from rabbits inoculated with purified capsid protein (Taylor, 2004). The antiserum was crossreacted with sonicated M15 *E. coli* cells to decrease the non-specific binding observed with non-crossreacted antiserum (Taylor, 2004). Cell number was normalized based on the OD_{660} of the culture when it was harvested (equivalent of 100 μl of OD_{660}=0.071), to ensure that a similar amount of total protein was present, and samples were boiled in 3X loading dye (NEB,) for 5 minutes. 12% acrylamide SDS-PAGE gels with 5% stacking gels (acrylamide/bis 37.5:1) were used to separate the proteins and various cleavage products. Gels were run with the Mini-

PROTEAN II systems (Bio-Rad) according to manufacturer's protocols. The proteins were then transferred to nitrocellulose membranes (Pall Corporation, Pensacola, Florida) by electroblotting at 100 volts for 1.5 hours. The membranes were blocked in 0.5% skim milk/TBS-T for 1 hour at room temperature. The membranes were incubated overnight at 4°C on a rocking platform with anti-capsid antiserum (1 in 10000 dilution) in 0.5% skim milk/TBS-T. The membranes were rinsed in TBS-T and subsequently washed once for 15 minutes and twice for 5 minutes at room temperature. Incubation in secondary antibody (donkey anti rabbit, HRP, Amersham) was done at room temperature in 0.5% skim milk/TBS-T at a 1 in 3000 dilution. Homemade ECL was then used to visualize the membranes (Diaz *et al.*, 1998).

The control bands on gels is M, the His-tagged major capsid protein isolated from *E. coli*, expression plasmid pREP (Taylor, 2004) that was used to make the anti capsid antibodies. A pre-stained ladder (NEB) was also used to estimate protein size. This assay is useful because it enables us to see when capsid is produced in the cell, and when it is released from the cell. Although this assay cannot differentiate whether the capsid protein is part of transductionally active GTA, it yields the maximal amount of GTA that is potentially made.

2.4 Identification of disrupted genes

Genomic DNA was isolated from mutant cultures by multiple phenol/chloroform extractions, followed by ethanol precipitations (Sambrook *et al.*, 1989) and 1 μg of DNA was digested with *Bam*HI and *Eco*RI, or *Bam*HI alone, following the suppliers' protocol (Invitrogen, Carlsbad, CA). The digested chromosomal DNA was run on a 1% agarose gel in 0.5X TBE buffer (Sambrook *et al.*, 1989) at 40 volts for 16 hr. The DNA was transferred from the gel to a Biotrans nylon membrane (ICN, Irvine, CA) for 1 hr at 80 volts, in 0.5X TBE buffer. The membrane was then UV-crosslinked (BioRad, GS Gene linker) to bind the DNA, dried and

probed using the digoxigenin-dUTP (DIG) DNA labelling and detection kit (Roche Applied Science, Laval, Qc) to visualize the fragments that contain the Tn5 insertion. The hybridization and detection protocol supplied with the kit was used. The DNA used as a probe was isolated from pUC4KIXX as a *Sma*I fragment as described by Lang and Beatty (2000). This fragment contains the coding sequence of the *neo* gene which is present in the Tn5 transposon that was used to create the transposon mutant library. Once labelled, the probe was hybridized to the membrane for 16 hr at 68 °C and the hybridization pattern visualized by a colour-producing reaction that occurs on the DIG labelled probe. Figure 6 shows a Southern blot of 2 mutants, LM4 and D5, as well as Y262 as a control (no Tn5 insertion). As the banding pattern for LM4 and D5 is different, it is expected that these are two separate Tn5 insertions. There are two bands for each lane as the probe partially hybridizes to the inverted repeat on each side of the Tn5. BamHI cuts once in the middle of the Tn5 while EcoRI does not cut in the Tn5. If the banding pattern for both the BamHI alone and BamHI and EcoRI digested DNA had been the same, it would have meant that there was a BamHI site closer to the Tn5 than an EcoRI site, and pUC19 cut only with BamHI would have been used for cloning of the fragment.

Plasmid pUC19 (NEB) was used as a cloning vector and was digested with *Bam*HI and *Eco*RI, or *Bam*HI alone, following the suppliers' protocol (Invitrogen, Carlsbad, CA). Genomic DNA restriction digested as above was then ligated into digested pUC19 using T4 DNA ligase (Invitrogen, Carlsbad, CA). These ligated vectors were electroporated into DH10B electrocompetent cells (Invitrogen, Carlsbad, CA), and colonies selected on LB agar containing kanamycin sulphate to select for clones that contain the *neo* gene. Several colonies were picked and the plasmids were isolated using a Qiagen Miniprep kit (Qiagen Sciences, Maryland, BA). The University of British Columbia's Nucleic Acid and Protein synthesis unit (NAPS)

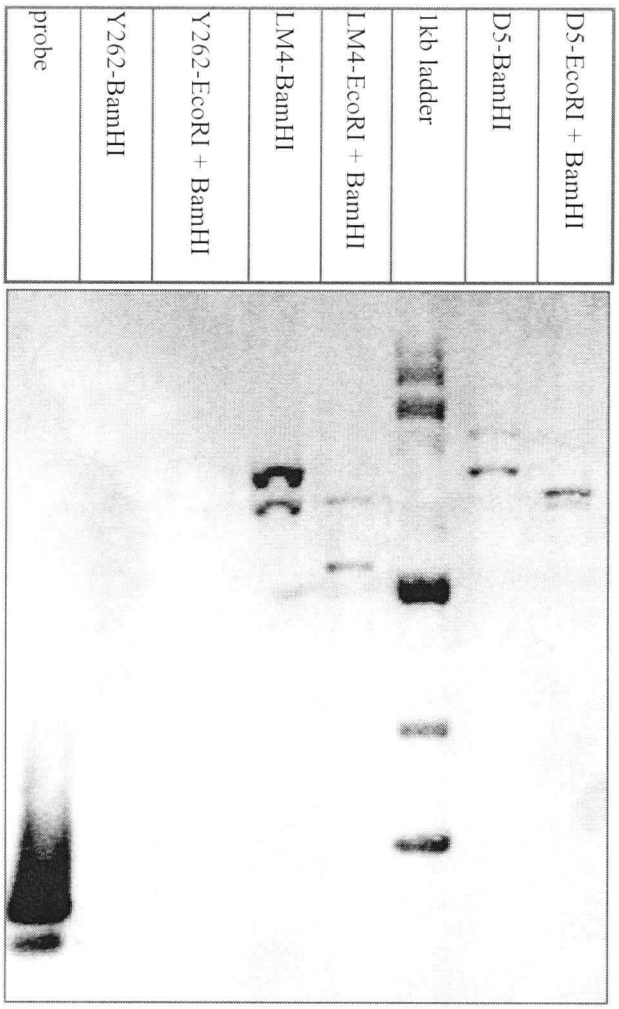

Figure 6: Southern Blot of digested mutants to determine size and restriction ends of fragments for cloning. Two mutants are shown, LM4 and D5, as well as Y262 as a negative control. There are no bands in either Y262 lane as there is no Tn5 insertion or *neo* gene found in that strain. As LM4 and D5 show different banding patterns from each other, it is expected that these are separate Tn5 insertions. There are two bands in each digest lane as BamHI cuts the Tn5 once, and the *neo* gene overlaps the inverted repeats at one end of the Tn5, allowing it to bind to both sides of the Tn5 (although weaker on the non-*neo* gene side). Both the LM4 and D5 fragments are BamHI to EcoRI fragments as opposed to BamHI only fragments, as the size of the bands decreases when cut with EcoRI (they would display the same banding pattern in both lanes if they were BamHI only fragments).

sequenced the plasmids using -47M13, a pUC19-sequencing primer, and TN5EXT, a Tn5 specific primer (Table 4). The sequences were then sent to the UBic BLAST and aligned to the *R. capsulatus* genome using Generic genome browser version 1.58 (supported by the U.B.C. Bioinformatics Center, University of British Columbia; http://luria.cmmt.ubc.ca/cgi-bin/gbrowse_img/rhodo).

2.5 Promoter deletions

To assess which part of pYP was causing the observed decrease in GTA production, fusions of the RR03520 to *orfg2* region (Figure 4), varying in length and position, were made in pXCA601. For *lacZ* in pXCA601 to be translationally in-frame, a BamHI site must be inserted into the coding sequence of a gene such that there is a GAT codon in the GGATCC BamHI site (Adams *et al.*, 1989). Although there was a discrepancy between annotations as to which of the two potential translational start sites was correct, both are in the same reading frame and so either will work for the fusion. Thus, to fuse *orfg1* to *lacZ*, primer G1Bam2 (Table 4) was used.

The first deletion (pSMF001) included sequences from the orfg1 fusion to a position between the predicted sigma binding site and the predicted rho-independent termination hairpin, using G1Pst1 (Table 4) as the second primer to insert a PstI site into the PCR product. The second deletion (pSMF002) was amplified using G1Pst2 (Table 4), which enabled the use of the native PstI site as in the original pYP. It was thought comparison of the effects of pSMF001 and pSMF002 to pYP and pYnP would allow us to differentiate between an effect caused by orfg1 and the 5' sequence. The fragments were PCR amplified from pYP (pSMF001) or pX/S5 (pSMF002) and digested by BamHI and PstI for cloning. To facilitate cloning small fragments (220 bp and 520 bp) into a large (20 kb), low copy number vector, the fragments were first subcloned into pUC19 in *E. coli*. The pUC19 fusions were then digested, dephosphorylated

Table 4: Primer sequences and descriptions

Name	Sequence (5'-3')	Description
-47M13	CGCCAGGGTTTTCCCAGTCACGAC	pUC19 primer (NEB)
TN5EXT	GAACGTTACCATGTTAGGAGGTC	Tn5 primer to IR ## bp from end.
G1Bam2	ACCCTGGATCCCCATGTCCATGTCT	Inserts BamHI site for fusion to *lacZ*.
G1Pst1	CGCCTGCAGCAACCCTGAATATAGC	Inserts PstI site 220bp upstream of orfg1 translation start.
G1Pst2	GCGAGCGTCTTTGTCAACC	Amplifies PstI site at the 3' end of RRC03520.
gtaPF	GACGCCCTATAACACCTATGTCAT	sequencing primer used to sequence B10 and Y262
gtaPR	ATCCGCACCCATTCCGCC	sequencing primer used to sequence B10 and Y262
GSP1	CTTGTCTGCGAAGTTTTTCA	R.A.C.E. primer for cDNA creation
GSP3	GCAATTTCCCGATAAAGCTCCTC	R.A.C.E. primer for amplification of cDNA

using Antarctic phosphatase (NEB) and ligated into digested pXCA601 without any isolation step to separate the fragments from pUC19. The ligations were then electroporated into DH10B electrocompetent cells, and screened by restriction digest to ensure the correct fragment was present.

Plasmids, pSMF001 and pSMF002 were conjugated into Y262 using the helper strain HB101(pRK2013) (Ditta *et al.*, 1985). The fusions were then tested by β-galactosidase assay and GTA immunoassay (as described above) to determine the effect they have on GTA gene transcription and GTA production.

2.6 5' R.A.C.E. analysis of mRNA

5' Rapid amplification of cDNA ends (R.A.C.E.) (Invitrogen, Carlsbad, CA) was used to determine the transcriptional start site for the GTA *orfg1*, and presumably the GTA gene cluster. The protocol supplied with the kit was used (Invitrogen, Carlsbad, CA). The RNA used for the procedure was isolated from both late log (280 KU, dense, but still growing exponentially) and early stationary (375 KU, dense, but growth is slowed considerably from log phase) phase cultures using RNeasy mini prep kit following the suppliers protocol (Qiagen, Maryland, BA?). Approximately 5 x 10^8 cells were used for the preps, which yielded ~25 ug of RNA. 3-4 ug of this was then used for R.A.C.E..

The RNA was converted to cDNA using superscript II (supplied with R.A.C.E. kit) and primer GSP1 (Table 4); whose position relative to the predicted translational start site for g1 is shown in Figure 16. GSP3 (Table 4, Figure 16) was used to amplify the cDNA with the kit supplied reverse primers.

3 Results

3.1 The search for GTA regulatory mutants

3.1.1 Screening of the transposon mutagenized library

Screening of the Y262::Tn5(pYP) library (described in Materials and Methods) yielded 60 potential mutants that appeared to have reduced β-galactosidase production driven by the GTA structural gene promoter, which were selected for further screening. Although there were a large number of suspected mutants, many of these were false positives due to the difficulty in distinguishing between "blue and white" in pigmented *R. capsulatus*. Figure 5 (Materials and Methods) demonstrates this difficulty by comparing several mutants and control strains on a stab plate which contained X-gal.

Further investigation of these 60 mutants was conducted through the use of more quantitative assays designed to test both transcription off the GTA promoter using the *orfg2*::*lacZ* fusion on pYP (Figure 4), and GTA transduction using a bioassay. The combined data from these two approaches led me to focus on 12 mutants that appeared to have the greatest reductions in GTA gene expression and transduction. The β-galactosidase assays were done first to determine mutants with low GTA gene expression from the GTA promoter driving expression of the pYP *orfg2:: lacZ* gene (data not shown). Tn5 insertions in genes whose products regulate GTA gene expression, either directly or indirectly, would show reduced β-galactosidase specific activity compared to Y262(pYP), the GTA overproducer parental strain. Mutants that had β-galactosidase activity of no more than 50% the β-galactosidase specific activity of Y262pYP were chosen to be further screened by the GTA bioassay. Figure 7 shows β-galactosidase expression for 6 mutants, as well as 4 control strains.

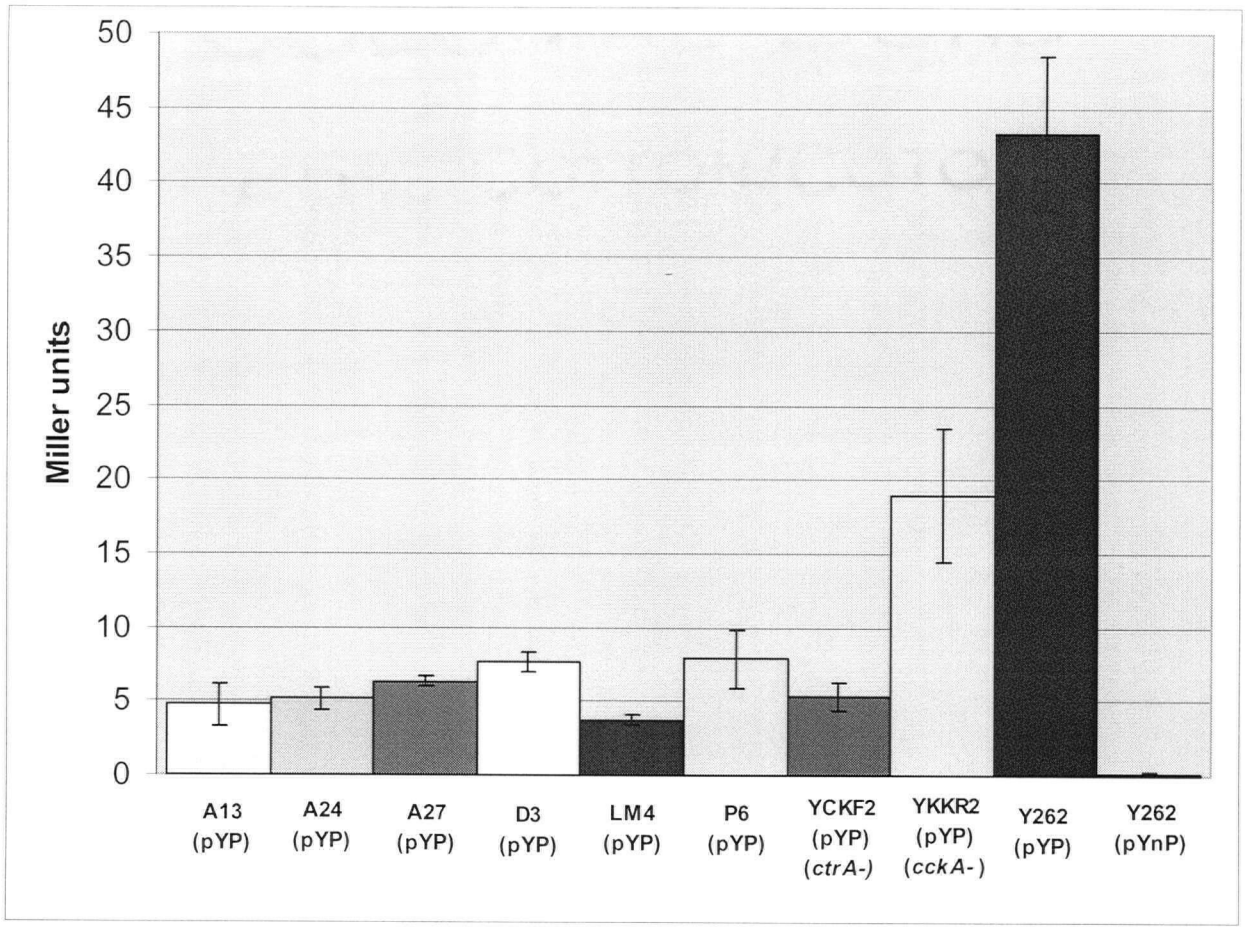

Figure 7: β-galactosidase assays performed in triplicate on 6 mutants (A13, A24, A27, D3, LM4, and P6) all containing the lacZ fusion plasmid pYP as well as 4 control strains: Y262(pYP), YCKF2(pYP), YKKR2(pYP), and Y262(pYnP). Y262 is the parental strain for all my Tn5 mutants, while YCKF2 is the *ctrA*⁻ strain and YKKR2 is the *cckA*⁻ strain. The error bars indicate standard deviation over 2 separate cultures and 2 or 3 dilutions of sonicated cells normalized to protein concentration.

Y262(pYP) had the highest level of *lacZ* expression, which was expected as Y262 is the GTA overproducer strain, and has been shown to contain the highest level of GTA mRNA (Lang, 2000). Y262(pYnP), which contains the expression plasmid with the *orfg2::lacZ* fusion minus the promoter region (Figure 4), had undetectable levels of *lacZ* expression; this was also expected as there is little read through on this reporter vector (Adams *et al.*, 1989). The *cckA*⁻ mutant, YKKR2(pYP), showed only a 50% decrease compared to Y262(pYP), while the *ctrA*⁻ mutant, YCKF2(pYP), and the 6 new mutants had similar levels of *lacZ* expression, which were approximately 10-15 % that of Y262(pYP).

The GTA bioassay measures GTA transduction activity based on the transfer of an essential allele (*puhA*) for growth under photosynthetic conditions to the DW5 *puhA* mutant (Wong *et al.*, 1996). Figure 8 shows GTA bioassay results for 37 mutants selected from the β-galactosidase assay secondary screen. The 16 mutants that had less than 10% of the GTA transduction of Y262 were chosen for further analysis. Figure 9 shows GTA bioassay results done in triplicate on the same 6 mutants as in Figure 7, and the variability in results. The most surprising result is that Y262(pYnP) produces more than 200% more GTA transductants than Y262(pYP), as it was expected that these two strains would produce similar amounts, as discussed further below (Section 3.2.2). The results for the Tn5 mutants show the variability with this transduction assay as the error is sometimes as large as the actual number (i.e. LM4(pYP)). The β-galactosidase data correspond loosely with the data received from the transduction assay, although differences do appear. For example, P6(pYP) had reduced β-galactosidase specific activity similar to the *ctrA* mutant YCKF2(pYP), although P6(pYP) had almost the same amount of GTA transduction activity as Y262(pYP).

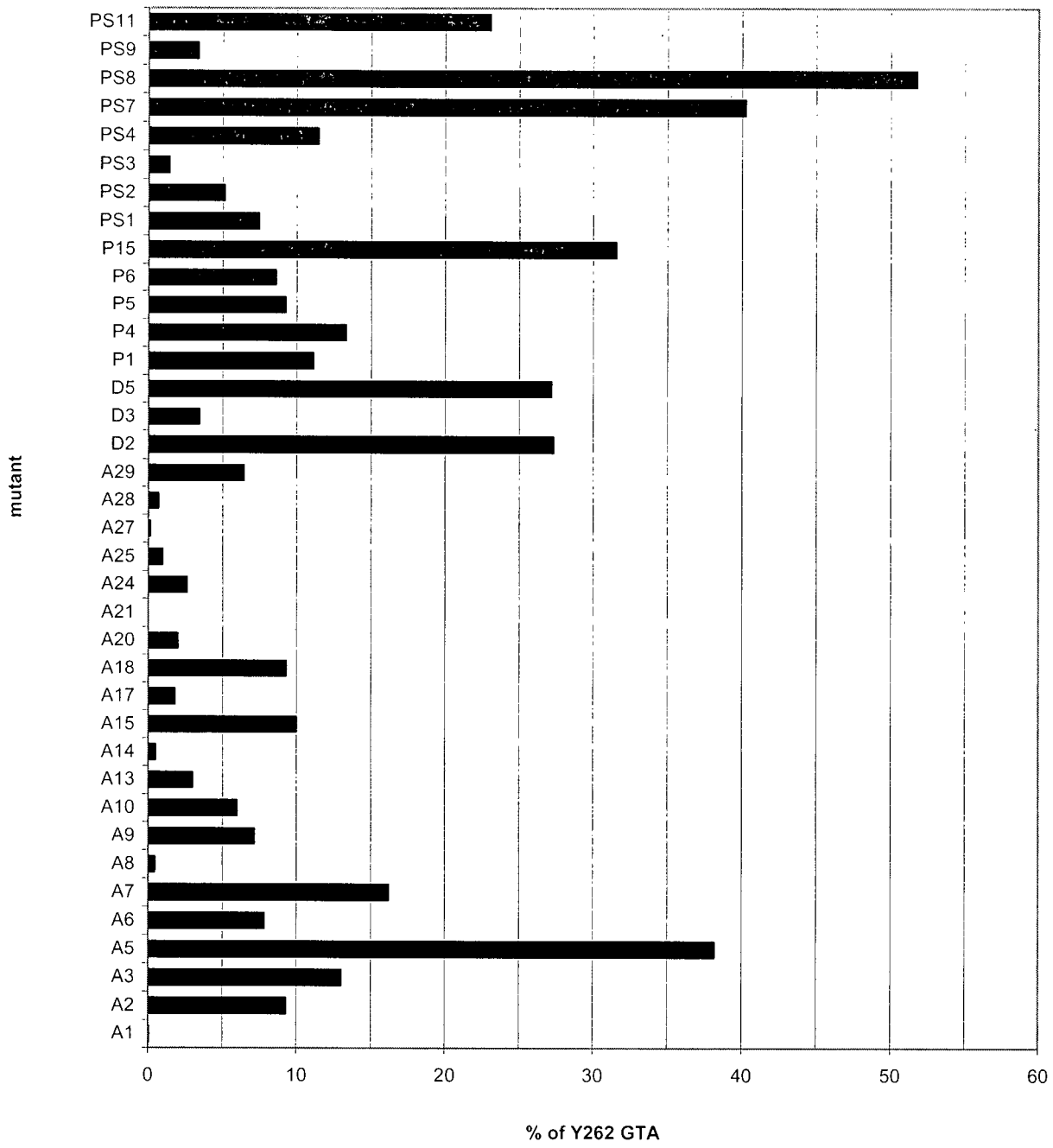

Figure 8: Overview chart of GTA transduction activity for 37 mutants. The values shown for each mutant are the % of Y262 GTA transduction (typically ~10^3 colonies).

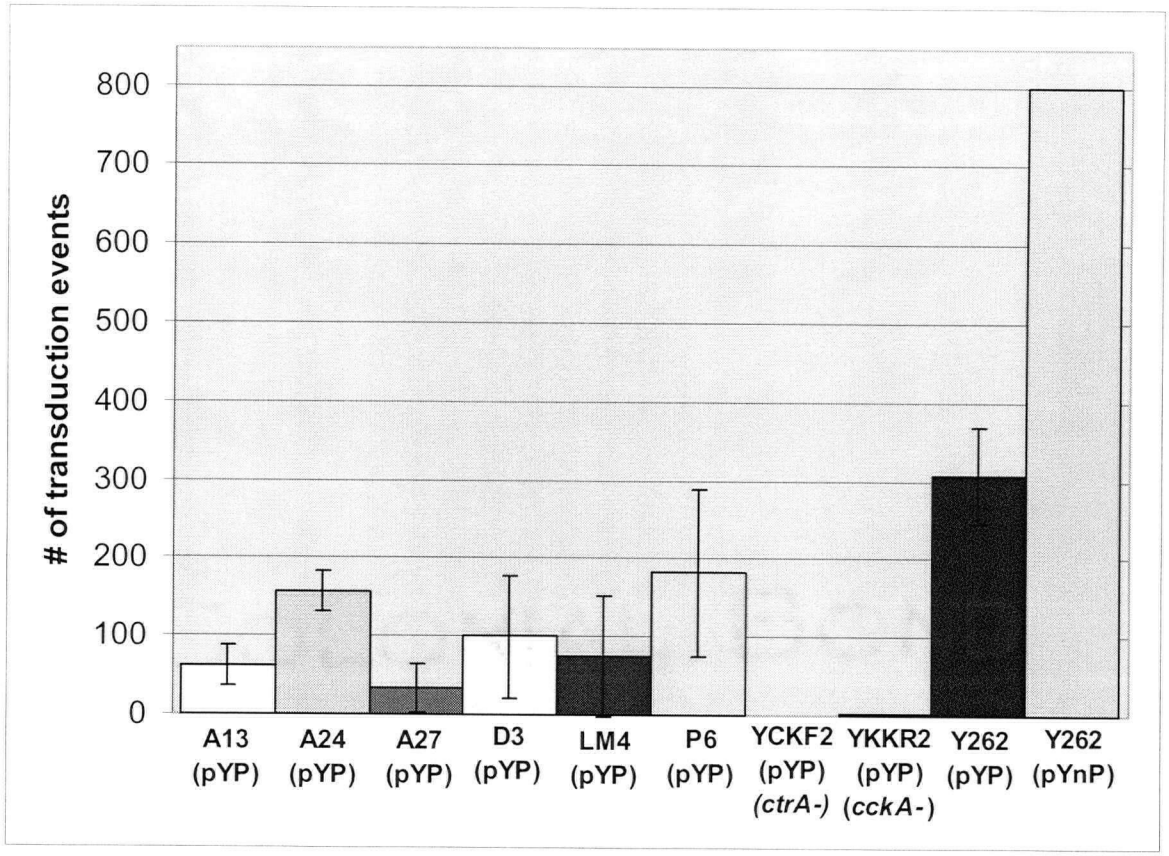

Figure 9: GTA assays performed in triplicate (described in Materials and Methods) on 6 mutants and 4 control strains. The # of transduction events is equal to the number of colonies of the photosynthesis incompetent mutant DW5 that arose on plates incubated under condition requiring photosynthetic growth, indicating that DW5 received the mutant allele from the GTA donor by GTA transduction. Error bars indicate standard deviation.

3.1.2 Studies on 16 mutants found to have reduced GTA expression

The Tn5 insertion point in each of 16 mutants was mapped using Southern blotting, followed by DNA sequencing and BLAST analysis of clearly unique mutants to determine which gene was disrupted, and therefore causing the observed effect. Figure 6 shows how Southern blots of restriction digested chromosomal DNA of the mutants determined if each insertion was unique, which reduced the number of mutants from 16 to 12, as 4 had similar banding patterns to others indicating that the transposon was inserted in the same spot. The DNA fragments containing the *neo* gene and flanking *R. capsulatus* sequences were then cloned into pUC19 and sequenced using a standard pUC19 primer and a primer to the inverted repeat segment of the Tn5 (Tn5EXT, Table 4). The genes which were disrupted in these twelve mutants are listed in Table 5.

While several of these transposon insertions were in hypothetical genes, some were in interesting putative genes. However, because of the great variety of Tn5-disrupted genes, none of which are obviously a sigma factor or a transcription regulator, it is difficult to interpret these results. It seems that many genes of uncertain function significantly affect GTA structural gene transcription, and transduction. Furthermore, some of these Tn5 insertions appear to be in a gene located 5' of other co-transcribed genes which could result in a phenotype due to a polar effect on an operon, and not solely due to the mutated gene. Therefore additional experiments would have to be done to unambiguously correlate specific genes to the observed phenotype, and to determine the genuine biological function of specific genes.

Table 5: 12 mutants and their corresponding disrupted genes and annotated gene function.

mutant	Disrupted ORF[1]	Annotated or predicted gene function[1]
A1	3513	Phage Prohead Protease -GTA gene cluster
A13*	2344	Sensory transduction protein/histidine kinase
A24*	337	Hypothetical protein W (LytR/AlgR family response regulator)
A25	2660	hypothetical protein (RecB family exonucleases)
A27*	1308	Hypothetical cytosolic protein
A28	4035	Membrane-bound Lytic Murein Transglycosylase B
D3*	4708	** (similarity to 2-component hybrid sensor and regulators)
LM4*	1241	Omega-amino acid-pyruvate aminotransferase
P1	4540	4-aminobutyrate aminotransferase
P4	2693	Type III Restriction-Modification System Methylation Subunit
P6*	2590	hypothetical protein in phage cluster
PS11	1930	Type I restriction-modification system restriction subunit

1. ORF number and annotation from http://www.ergo-light.com/ERGO/

* indicates GTA transduction and gene expression data for these 6 mutants shown in Figures 7 and 8.

** indicates no annotation for this gene.

3.2 Western blot analyses of GTA capsid protein levels

An anti-capsid (Orfg5) antiserum was used to probe western blots of intact cells and cell-free culture supernatants to evaluate the relative amounts of this capsid protein as an index of GTA gene expression and release of GTA particles from cells (Taylor, 2004). The antiserum was found to bind to a number of proteins (bands in blots of SDS-PAGE) in intact cells, but the use of the *orfg5* Tn5 knockout A1 (see Table 5) as a negative control simplified the analysis. Figure 10A shows a western blot that compares A1(pYP) and Y262, while Figure 10B gives the same western blot, but with the bands present in A1 subtracted from Y262, using Adobe Photoshop (© 1990-2002, Adobe Systems Incorporated). The His-tagged recombinant capsid protein expressed from the pREP expression plasmid in *E. coli* M15 cells, is included as an additional control.

As with homologous capsid proteins of dsDNA tailed phage (Duda *et al.*, 1995), this GTA capsid protein is made as a pro-protein of 42.2 kDa that is cleaved on the C-terminal side of a lysine residue to produce the 26 kDa mature capsid protein found in GTA particles (Lang and Beatty, 2000). This cleavage is thought to be done by the protease encoded by *orfg4* of the GTA structural gene cluster (Figure 3). In contrast, the recombinant His-tagged protein produced in *E. coli* is made as a 43.1 kDa pro-protein, but it is thought to be cleaved by a chemotrypsin-like protease to yield a 36.3 kDa species, which would explain the lower band in the immunoblots (Taylor, 2004). Mass spectrometry was done on both of these bands to establish their identity, which is in fact the product of *orfg5* (Taylor, 2004). Thus the recombinant protein yields two prominent capsid bands at 43.1 and 36.3 kDa, whereas intact cells of *R. capsulatus* yield the 42.2 kDa pro-protein and 26 kDa mature protein bands (Figure10), and cell-free culture supernatants contain the 26 kDa mature protein.

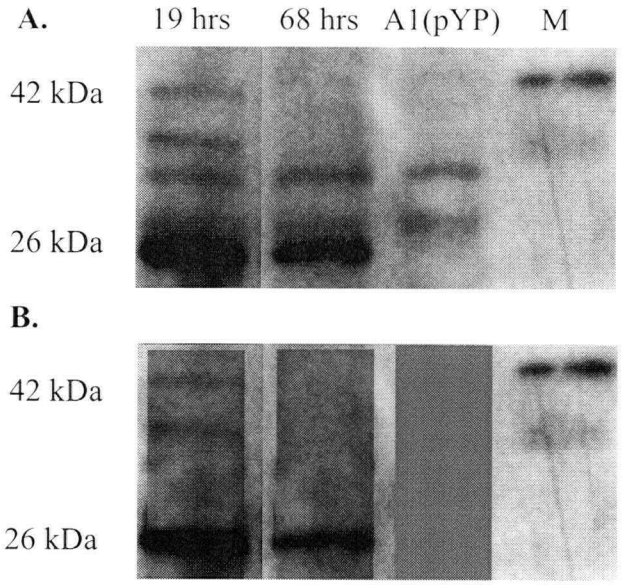

Figure 10: GTA immunoassays: resolving the extra bands. A, original immunoblots of Y262(pYP) whole cell samples at 19 and 68 hours of growth, A1(pYP) which is the Tn5 capsid mutant, and M (recombinant, His-tagged Orfg5 purified from *E. coli*). B, Photoshop altered version of immunoblots from A. By copying the A1(pYP) lane, inversing the colour, and overlaying it onto the other lanes, it is possible to cancel out bands that appear in both lanes. Bands assigned to the 42 kDa proprotein and 26 kDa mature Orfg5 capsid protein are labelled.

3.2.1 The pYP effect

While testing strains in the GTA anti-capsid (*orfg5*) immunoassay and the GTA bioassay, it was discovered that Y262(pYP) makes reduced amounts of GTA, although Y262(pYnP) makes amounts equal to Y262. Y262(pYnP) was cured of its plasmid by repeated subculture in the absence of tetracycline and tested to ensure that Y262 had not reverted, and plasmid pYP was re-introduced. However, the same results were obtained from this new Y262(pYP) strain. Figure 11 shows western blots of Y262(pYnP), the Y262 that was cured of pYnP and the reconstructed Y262(pYP), probed with the anti-capsid antiserum, which proves that it is pYP that causes the observed decrease in GTA production.

There are three general possibilities that could cause this decrease in the presence of plasmid pYP. The first possibility is that regulation of the GTA gene cluster is tightly regulated and additional copies of an activator-binding site on pYP could be binding an activator protein, resulting in less of this protein being available to bind to the chromosomal promoter. This hypothesis seems plausible as there is readily detectable *lacZ* expression in Y262(pYP) (Figure 7), which is regulated by the culture growth phase (Lang and Beatty, 2000); however, as there is reduced production of GTA in Y262(pYP), this would indicate reduced expression of the chromosomal GTA structural genes. Another possibility is that the expression of *orfg*1 on pYP, and/or the expression of the *orfg2* segment fused to *lacZ*, inhibits transcription of the GTA gene cluster. Lastly, overproduction of Orfg1 and/or the Orfg2::LacZ fusion protein from pYP may interfere with a posttranscriptional step in GTA production, such as assembly of the mature GTA particles. This is possible as there is some capsid pro-protein seen in the Y262(pYP) lane, although very little seems to be cleaved into mature capsid, and none is seen outside of the cell (Figure 11).

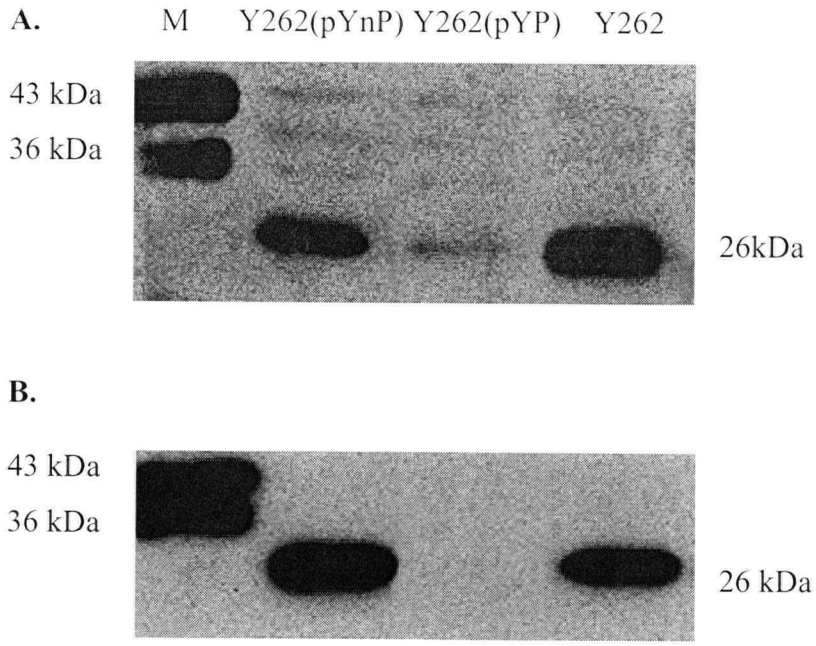

Figure 11: Immunoassay of Y262, Y262(pYP) and Y262(pYnP), probed with anti-capsid (*orfg5*) antisera. A. shows whole cells; B. shows cell free culture supernatant. The first lane (M) is the His-tagged capsid protein that was purified from *E. coli* M15 cells and used to create the α-capsid antiserum, with proteins running at 43.1 kDa and 36.3 kDa. These bands were attributed to the full-length, His-tagged Orfg5, and a cleavage product due to an unidentified *E. coli* protease (Taylor, 2004). For Y262(pYnP), Y262(pYP) and Y262, the prominent band in both the supernatant and the cell is 26 kDa, The 42.2 kDa Orfg5 protein was reported to be cleaved by an *R. capsulatus* protease, most likely the *orfg4* product, to yield the 26 kDa protein in the process of GTA particle maturation (Lang and Beatty, 2000).

3.2.2 Time course analysis of capsid production in Y262 and B10

A bioassay comparison of GTA transduction between Y262 (GTA over-producer) (Yen *et al.*, 1979) and B10 (wild type) (Marrs, 1974) indicated a difference of 3 orders of magnitude. However, nothing was known about the relative kinetics of the intracellular and extracellular production of GTA in these two strains. Therefore I used the capsid immunoassay to investigate these questions, over a growth curve in which Y262 and B10 grew identically (Figure 12A).

By comparing capsid production and release over the growth curve for these two strains, the difference is very noticeable as seen in Figure 12. In B10, both the capsid precursor and cleaved product appeared to be present in the cell in very low amounts starting at 21 hours after inoculation of the culture, and cell-free GTA was undetectable at all time points. However, in Y262, capsid production was observed in cells starting in late-log phase (19 hours post inoculation), with capsid release in later-log/ early stationary phase (21 hours). It is interesting to note that there appears to be less GTA present in the supernatant at 72 hours post inoculation, but that could be due to incorrect loading of the sample or a bubble between the gel and the membrane not allowing proper transfer instead of a decrease in the amount of cell-free GTA particles. These results confirm that the increase in GTA transduction by Y262 over B10 is due to an increase in the number of GTA particles produced in Y262. The relative intracellular amounts of the precursor and cleaved products in Y262 compared to B10 may be due to one or more mutations in Y262 that causes the over expression of a positive regulator of the GTA gene cluster, therefore allowing more GTA to be produced by each cell. Thus the difference in GTA transduction frequency between B10 and Y262 is due to an increase in GTA structural gene transcription/translation in Y262, and not due to increased release of GTA from cells.

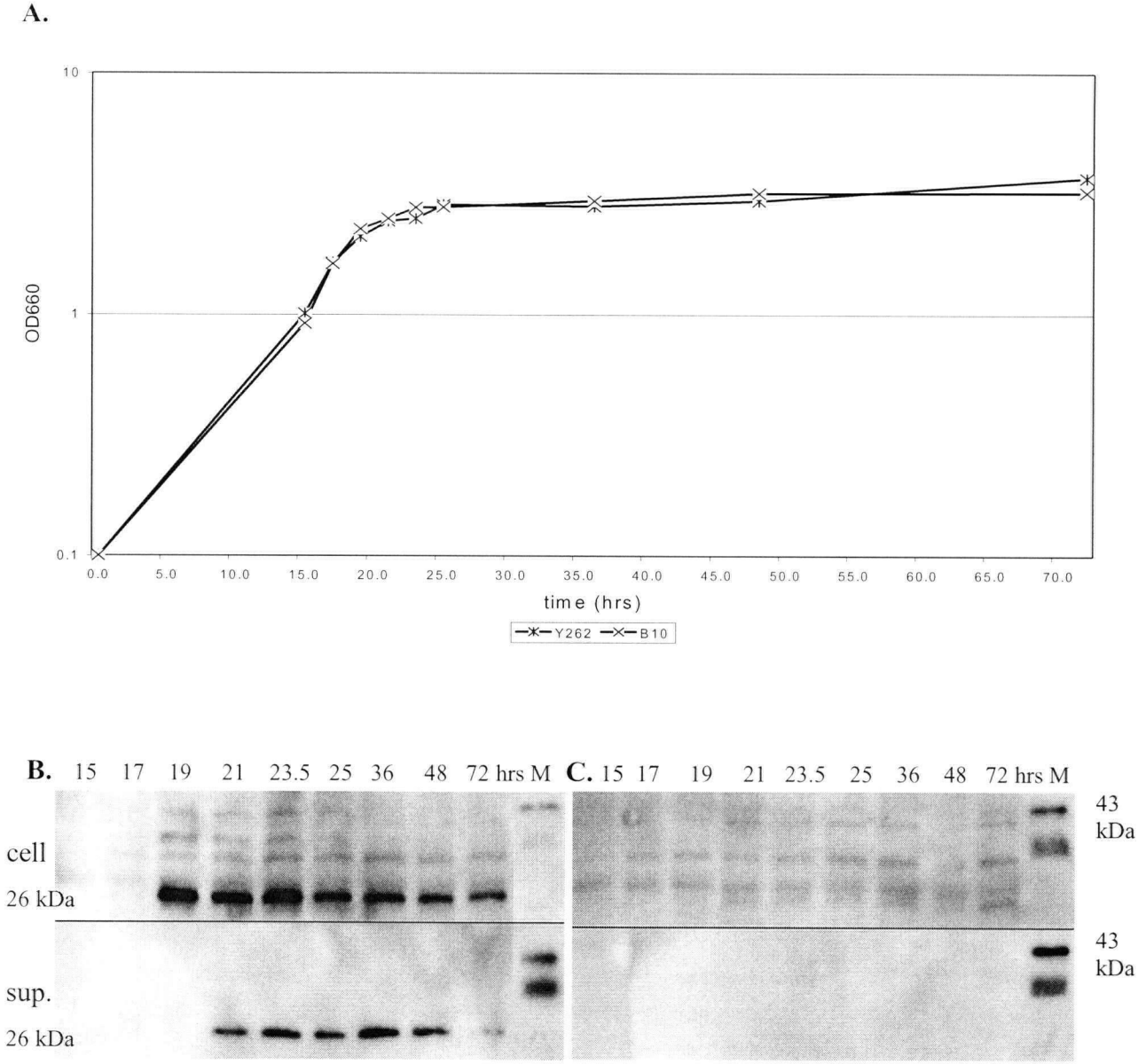

Figure 12: Time course of Y262 capsid production compared to B10. A, the growth curve and time points when samples were taken for Y262 and B10. B, immunoblots of Y262 cells (top) and cell free culture supernatant (bottom). C, immunoblots of B10 cells (top) and cell free culture supernatants (bottom). Anti capsid antiserum was used to visualize capsid, and therefore GTA in samples. Cell numbers loaded per sample were normalized as described in Materials and Methods.

3.2.3 Time course analysis of capsid production in *ctrA*⁻ and *cckA*⁻ strains

Recall that the *cckA* mutant, YKKR2, had reduced GTA transduction and β-galactosidase expression from pYP, although not as reduced as the *ctrA* mutant, YCKF2 (see Figures 7 and 9). By comparing these two regulatory mutant strains to the overproducing strain in western blots, it was possible to see differences between capsid production and release in these strains over a growth curve (Figure 13).

For YCKF2 (*ctrA*⁻), there appears to be very little, if any capsid produced in the cell (Figure 13 B)or released out of the cell (data not shown). However, for YKKR2, there appears to be a lot of both the capsid proprotein as well as the mature capsid protein within cells (Figure 13 C), however this mature capsid protein was never seen in the supernatant of the cultures (data not shown). When YKKR2 is compared side by side with Y262, and the same amounts of cells are loaded for each strain, there is more of the proprotein present in the cell, as well as mature capsid protein (Figure 13 D). This is consistent with the results obtained earlier that show that YKKR2(pYP) has only 50% of the β-galactosidase specific activity of Y262(pYP) and very little GTA transduction is detected for YKKR2 (Figures 7 and 9). From this is would appear that CckA is somehow involved in the maturation of the GTA or release of the GTA particles from the cell. It would be interesting to determine if the mature capsid protein seen in the cell is part of mature GTA particles, or if the reason that YKKR2 rarely shows GTA transduction is due to a deficiency in a regulatory element that promotes GTA assembly.

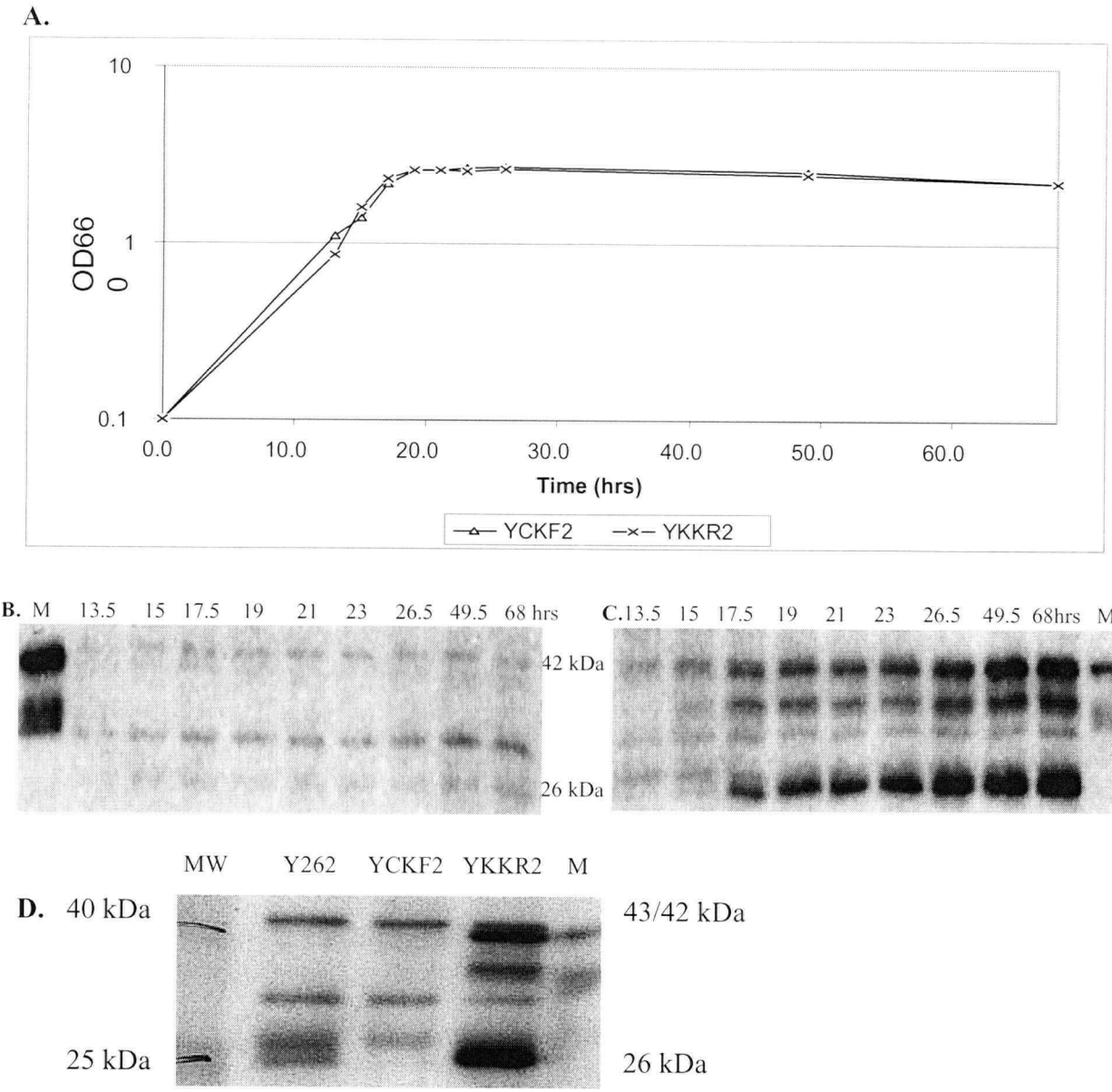

Figure 13: Time course of capsid production in *ctrA*⁻ and *cckA*⁻ mutant whole cell samples. A, growth curve and time points when samples were taken. B, immunoblots of YCKF2 (*ctrA*⁻) and C, YKKR2 (*cckA*⁻). D, a comparison of cell samples for Y262, YCKF2, and YKKR2 at the 68 hour time point. Only cell samples are shown as there is no capsid detected in the supernatant for either YCKF2 or YKKR2. Anti-capsid antiserum was used for the immunoblots to visualize capsid, and therefore GTA in the samples. Cell numbers loaded per sample were normalized as described in Materials and Methods.

3.2.4 Immunoblots of B10, SB1003 and IKOI

B10 is a wild type strain of *R. capsulatus* that produces low levels of GTA compared to the overproducer Y262 (see above). SB1003 is the strain of *R. capsulatus* used for genome sequencing (http://www.ergo-light.com/ERGO/). IKOI is a derivative of SB1003 which has the *gtaI* gene knocked out, which reduces GTA production compared to the parental strain (Schaefer *et al.*, 2002).

By comparing GTA immunoassay results for these three strains in Figure 14, it appears that SB1003 and IKOI produce and release more capsid than B10. The amounts of both the precursor and the mature capsid protein appear to be similar in B10 and SB1003 cells (Figure 14A), however the amounts of capsid found in the cell-free supernatant of the culture is very much higher in the SB1003 sample than the B10 sample. IKOI also has greater levels of capsid than B10 in the cell samples as well as the cell-free supernatant samples (Figure 14B). It appears that IKOI has more capsid precursor protein as well as more mature capsid protein than B10 in the cell, and it also has more mature capsid in the cell-free supernatant as well (Figure 14B). The capsid levels seen in IKOI samples are reminiscent of the levels seen in YKKR2 (*cckA*$^-$) where there seems to be a buildup of both capsid precursor and mature capsid protein within cells, but low levels of GTA transduction activity (Schaefer *et al.*, 2002).

The reason for these observed differences is unknown, as the only known difference between B10 and SB1003 is that SB1003 is rifampicin resistant (Yen and Marrs, 1976). SB1003 was created from a poorly characterized "phage-free" B10 derivative called B100, with rifampicin resistance being transferred by GTA-mediated transduction to B100 from another B10 derived strain (Solioz, 1975; Yen and Marrs, 1976, 1977). It has been observed that SB1003 produces greater amounts of photosynthetic complexes than B10 (J.T. Beatty, personal

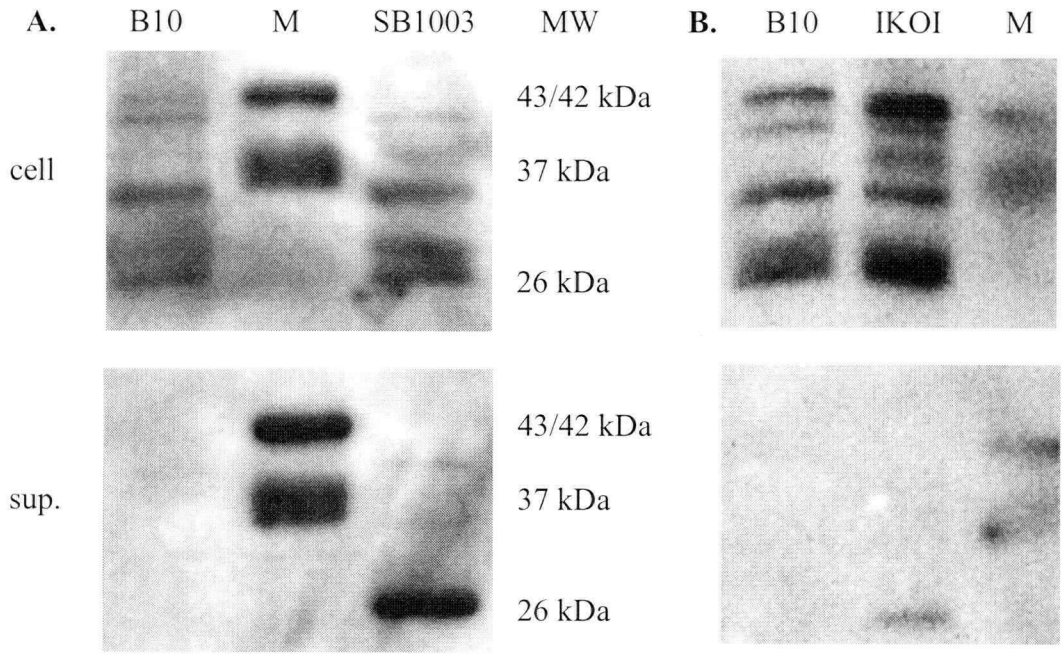

Figure 14: B10 capsid production compared to SB1003 and IKOI. A. shows immunoblots of B10 and SB1003 both cell (upper blot) and cell free supernatant (lower blot) samples. B. shows immunoblots of B10 and IKOI (*gtaI*) both cell (upper blot) and cell free supernatant (lower blot) samples. Anti-capsid antiserum was used for the immunoblots to visualize capsid, and therefore GTA in the samples. Cell numbers loaded per sample were normalized as described in Materials and Methods.

communication), but it is not clear whether this is due to the rifampicin-resistant allele or to another, unknown, genetic difference. Further experiments, such as transduction bioassays, would need to be performed to determine if the amounts of this capsid protein seen in the SB1003 immunoassay correspond to the numbers of functional GTA particles.

3.3 GTA promoter region analysis

3.3.1 DNA sequencing

Fragments of the GTA structural gene cluster promoter region were sequenced from two different strains (Y262 and B10) to determine whether the difference in GTA expression between those two strains might be due to a sequence change in the promoter region. These sequences were then compared to the genome-sequenced strain, SB1003 (http://www.ergo-light.com/ERGO/). Strain SB1003, like strain B10, was thought to not overexpress GTA genes, although my western blot analyses indicated that SB1003 produces more capsid protein both in the cell and in the cell free supernatant of cultures than B10 (see Section 3.2.3.3).

Plasmid pX/S5 (Lang, 2000) was used for sequencing the Y262 GTA structural gene cluster promoter region, as it contains 1.8 kb of sequence from an XhoI site (in RRC03520) to a SalI site (in *orfg2*, used to create plasmid pYP fusion to *lacZ*), which includes the GTA structural gene cluster promoter (Figure 15). Four primers were used to provide overlapping sequences, with 2 (gtaPR and gtaPF: sequences of primers in Table 4) priming in the promoter region, and the other two priming off the plasmid vector (M13 primers F and R).

Plasmid p9HSTU (Lang, 2000) was used for sequencing the B10 GTA structural gene cluster promoter region, in the same way as pX/S5, as p9HSTU contains 2.3 kb of sequence from a KpnI site (upstream of RRC03520) to a HindIII site (200 bp downstream of the SalI site

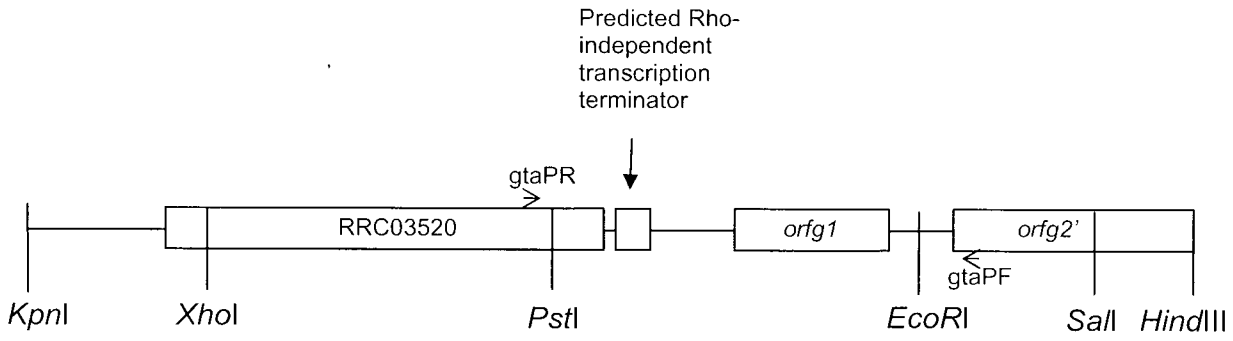

Figure 15: The GTA structural gene cluster promoter: overview. The position of sequencing primers gtaPR and gtaPF are shown with arrows. Further sequencing was done with M13 primers as these fragments (Y262 and B10) are in M13 derivative plasmids. For the Y262 promoter, XhoI to SalI was sequenced, and for B10, KpnI to HindIII was sequenced. For comparison to pYP and pYnP, see Figure 4.

used to create pYP), which includes the GTA promoter (see Figure 15).

Neither the B10 nor the Y262 strain had differences in their GTA gene cluster promoter region from the SB1003 sequence, nor from each other. A total of 1.8 overlapping kb, including part of *org2*, all of *orfg1* and part of RRC03520, was sequenced. Thus it appears that the reason for the overexpression of GTA in Y262 is due to one or more mutations in *trans*-active regulatory proteins, as opposed to a *cis*-active sequence in the GTA structural gene cluster promoter region. This also corresponds with the finding that when plasmid pYP is used for β-galactosidase assays in B10, the levels are not as high as in Y262(pYP) (Lang, 2000). These data also indicate that the increased levels of the GTA capsid protein detected in SB1003, relative to B10, are due to a trans-active factor.

3.3.2 Softberry's BPROM analysis

BPROM by Softberry is a computer program that was designed to search for bacterial promoters in DNA sequences based on sigma70 promoters, the major *E.coli* promoter class (http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb). The predicted -10 and -35 sites 5' of the GTA *orfg1* are shown in Figure 16. The -10 site has similarity to sites bound by sigma factor rpoD15 (Gilman *et al.*, 1981). BPROM also predicts -10 and -35 sites near the start codon for *orfg2* (Figure 15B), although the -10 sequence does not have a special similarity to any known sigma factor binding site. The problem with BPROM is that it is biased towards prokaryotes with low %GC, while *R. capsulatus* has an overall GC content of 65.5%. However the GC content of the predicted GTA structural gene promoter region is 54%, and the GC content of the *orfg1-orfg15* region is 69% (http://www.infobiogen.fr/services/analyseq/cgi-bin/freqsq_in.pl).

Figure 16: The GTA structural gene cluster promoter: sequence view.

Restriction site for noted enzymes are in purple.
Start codon for noted genes (A=Andrew Lang annotated, E=Ergo annotated) are in green.
Stop codon for noted genes are in red.
Putative Rho-independent transcription terminator is shown in brown.
Sequencing primers, name noted above starting with gta (sequences in Table 4) are in orange.
R.A.C.E. primers, name noted above starting with GSP (sequences in Table 4) are in blue.
5' mRNA ends determined by R.A.C.E. are highlighted in blue.
BPROM predicted -10 and -35 sites are highlighted in yellow.
Proposed -10 and -35 sites based on R.A.C.E. data are highlighted in light blue.
The underlined bases GGTTCA were changed to GGATCC (BamHI site) and the underlined bases GGTGCA were changed to TGCGCA (PstI site).

3.3.3 5' end mapping of GTA mRNA (R.A.C.E.)

Figure 16 shows the two RNA 5' ends that were determined for the mRNA transcript of orfg1 using R.A.C.E. (see Materials and Methods). The first is located 23 bp 5'(upstream) of the stop codon for RRC03520, which encodes a conserved hypothetical protein. The second is 111 bp upstream of the predicted start codon for orfg1, and overlapping the BPROM predicted -10 sequence.

In terms of localization of a promoter, these results may be interpreted in three general ways: 1) Both 5' ends arise from transcription initiation (i.e., there are two separate promoters); 2) The 5'-most end arises from transcription initiation, and the other end is the result of posttranscriptional cleavage of this transcript, or of premature termination of the reverse transcriptase reaction; 3) Both 5' ends result from mRNA cleavage, or premature termination of the reverse transcriptase, and the promoter is located closer to the PstI site in RRC03520 (Figure 16). I attempted to distinguish between these possibilities by fusing different lengths of the promoter region to *lacZ* on plasmid pXCA601, and measuring the relative amounts of β-galactosidase specific activities in Y262 cells that contain these plasmids (see below).

3.3.4 Promoter deletions

To determine the minimal sequence required to encode the entire promoter, two plasmids were made (pSMF001 and pSMF002, described in Material and Methods), which both differ from pYP by having the translationally in-frame fusion to *lacZ* at the 5th codon of *orfg1*.

The original pYP plasmid contained 1.1 kb of sequence, with the fusion to *lacZ* located 400 bp into the *orfg2* coding region, including *orfg1* and extending in the 5' direction to a PstI site located 320 bp into the upstream putative gene, RRC03520. The short *orfg1* fusion (160 bp, pSMF001) included the predicted sigma binding site but not the predicted rho-independent

termination hairpin (Figure 16). This deletion was made to evaluate the minimal sequence required for promoter activity. The longer *orfg1* fusion (530 bp, pSMF002) used the same PstI site as in pYP. These two *orfg1* fusions of varying length in pSMF001 and pSMF002 were used in attempts to localize the promoter to a smaller region. In parallel, pSMF001 and pSMF002 were used to investigate sequences that could cause the pYP effect (reduction of GTA expression; see Section 3.2.2).

These two plasmids were introduced into Y262, and the β-galactosidase specific activities of cells grown photosynthetically to late stationary phase were measured. Unfortunately neither of these fusions showed promoter activity that was detectable in the β-galactosidase assay (data not shown). There are three main reasons why this could be: 1) the fusions are not actually in-frame with the *lacZ* gene due to an error in the cloning procedure; 2) the annotated start codon for *orfg1* is not the actual start codon and therefore the fusions are not in frame with the correct start codon; 3) the promoter driving expression of the *orfg2::lacZ* fusion in pYP is located 3' of the 5th codon of *orfg1*. The DNA sequences of the *orfg1::lacZ* fusions in pSMF001 and pSMF002 were determined, and it was found that plasmid pSMF002 contained an insert of non-*R. capsulatus* DNA, whereas pSMF001 contained the expected sequence.

3.3.5 Summary of data and proposed promoter for GTA structural genes.

The R.A.C.E. data yield the most reliable data for localization of the GTA structural gene promoter, because the promoter deletions were not functional. Using the 5' most postion of the two mRNA 5' ends detected, it appears that the promoter for the GTA gene cluster is further upstream that what was previously thought, and is located within the upstream conserved hypothetical protein gene (RRC03520; see Figures 3 and 16). The R.A.C.E. data also show that

the Softberry BPROM software is not very accurate when predicting -10 and -35 sites in *R. capsulatus*, most likely due to a bias towards more AT-rich sites in *E. coli.*

3.3.6 Immunoassays of Y262(pSMF001)and Y262(pSMF002)

Even though the promoter deletions did not produce any detectable β-galactosidase specific activity, they were tested by GTA immunoassay to see if the sequences contained in them cause the pYP effect. Although pSMF002 was not the desired sequence, pSMF001 showed no apparent pYP effect (Figure 17). This result does not distinguish between an effect due to the promoter region, or one due to the Orfg2::LacZ fusion protein. However, the blot of Y262(pSTU12) gave more capsid in the cell than Y262(pYP) and these two plasmids differ only in the position of the Orfg2::LacZ fusion (Figure 16 and Table 3). Therefore it appears that it is whether or not the Orfg2::LacZ fusion protein is made (pYP vs. pYnP), or the exact nature of the fusion protein (pYP vs. pSTU12), that affects the level of GTA production (the pYP effect).

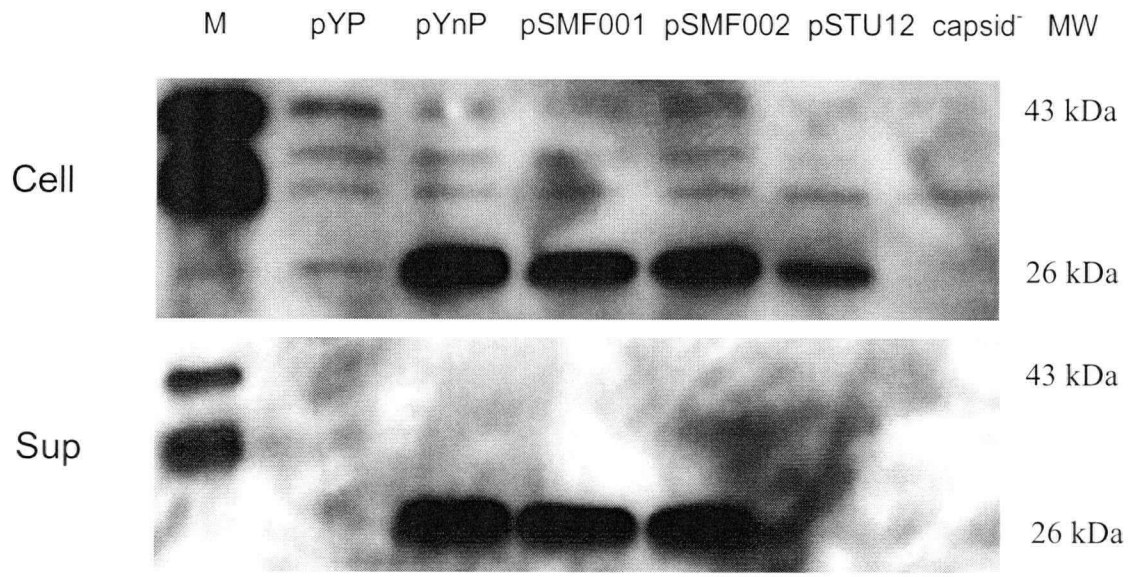

Figure 17: GTA immunoassay of pSMF001 and pSMF002 compared to controls. Cell samples are shown in (A), while supernatant samples are shown in (B). All plasmids are in Y262, with the capsid⁻ strain being A1(pYP) (Table 5), which has a transposon insertion in *orfg5* of the GTA structural gene cluster. Anti-capsid antiserum was used for the immunoblots to visualize capsid, and therefore GTA in the samples. Cell numbers loaded per sample were normalized as described in Materials and Methods.

4 Discussion

As regulation of GTA in *R. capsulatus* was so poorly understood, I undertook this research to attempt to elucidate the regulatory pathway or network that determines when GTA is made and released from the cell. Although I faced many problems, and was unsuccessful in determining a single "master element" that regulates GTA, I found a number of genes that appear to be needed for maximal expression of the GTA structural genes. I also studied the promoter region of the GTA structural gene cluster to locate the position of a plausible promoter sequence, and discovered interesting differences in GTA production between several *R. capsulatus* strains. These results are discussed below in more detail.

4.1 The surprising variety of genes needed for maximal production of GTA.

In theory, a promoter fusion to *lacZ* is a powerful method for finding regulatory mutants that no longer positively or negatively regulate a gene of interest. I found 12 genes that are listed in Table 5 that appear to regulate GTA in *R. capsulatus,* by using a plasmid construct with the promoter for the GTA structural gene cluster fused to *lacZ* to screen a library of transposon mutagenized cells.

The Tn5 insertion in mutant A1 mapped to *orfg5* of the GTA structural gene cluster. A1 was later used as a capsid$^-$ control in immunoblots to resolve the 'extra' bands seen with the anti-capsid antiserum. It was also a nice check to ensure that this system can find mutants that genuinely affect GTA production, although there appears to be some sort of feedback regulation as it would not be expected that the transcription of the GTA gene cluster (as measured by the β-

galactosidase activity of pYP) would be affected by a mutation in one of the chromosomally-located structural genes.

Four of the Tn5 mutations that were sequenced mapped to genes simply annotated as hypothetical proteins. One large problem I faced while deciphering which genes were involved in GTA regulation was a poorly annotated genome with a large number of genes annotated as hypothetical or conserved. However, upon closer inspection of one of the hypothetical genes (RRC02590, disrupted in mutant P6; Table 5), I found that RRC02590 appears to be in the middle of an apparent phage gene cluster. There are in fact 6 such putative phage gene clusters in the *R. capsulatus* genome, one of which is the GTA structural gene cluster. Two are annotated as prophages RcM1 and RcP1 (named based on the contigs they were originally found on (Haselkorn *et al.*, 2001)), and the other 3 appear to be sequences acquired by duplication of the other phage clusters, or the remnants of past prophages. The Tn5 insertion in mutant P6 (Table 5) was mapped to a putative phage gene cluster separate from the GTA gene cluster and the 2 prophages. This is extremely interesting since the GTA structural gene cluster does not appear to contain enough genes to encode a complete dsDNA phage, therefore GTA could be using prophage genes (or remnants) to provide essential functions.

Another of the hypothetical proteins (RRC02660, disrupted in mutant A25; Table 5) had similarity to a RecB family of exonucleases (e value of 3e-65). In *E. coli*, the RecB nuclease domain of the RecBCD nuclease-helicase generates single-strand regions at the ends of DNA duplexes and is though to be involved in horizontal gene transfer (Kowalczykowski *et al.*, 1994). Perhaps RRC02660 is a gene in *R. capsulatus* that encodes a protein needed for processing of the DNA during maturation of the GTA transducing particles.

Several of the Tn5 insertions mapped to genes annotated as sensory transduction or response regulator proteins, which was quite exciting. One of the hypothetical proteins

(RRC00337, disrupted in mutant A24; Table 5) had sequence similarity to a response regulator of the LytR/AlgR family (e value of 3e-08), which could directly or indirectly (as for CtrA) regulate the expression of the GTA structural genes. The Tn5 insertion in mutant A13 (RRC02344; Table 5) is in a putative sensory transduction/histidine kinase gene, and while the gene which contains the Tn5 insertion in mutant D3 is not annotated, I found that this gene has similarity to a 2-component hybrid sensor regulator protein (e value of 5.88e-06). It is possible that these proteins interact with CtrA, CckA or GtaI as part of a complicated regulatory network that leads to GTA structural gene transcription and/or particle assembly.

The Tn5 insertion in mutant P1 did not appear at first to be of interest as it was in an ORF (RRC04540) annotated to encode a 4-aminobutyrate aminotransferase; but upon further inspection I found a putative homoserine kinase gene located directly downstream of this gene, perhaps in an operon. As homoserine lactone has been determined to be part of the regulation of quorum sensing and, directly or indirectly, GTA expression (Schaefer *et al.*, 2002), it is conceivable that this is a kinase needed for phosphorylation of this homoserine lactone as part of GTA regulation.

Two of the Tn5 insertions mapped to homologues of restriction modification system subunits, one to a Type I system (RRC01930 in mutant PS11; Table 5) and the other to a Type III system (RRC02693 in mutant P4; Table 5). Although GTA particles may be filled with DNA using a headful mechanism whereby the DNA is inserted into the premade capsid head until the head is full, and then the DNA gets cut, it is possible that the DNA is cut using a semi-random restriction modification system that cuts the DNA prior to packaging it into the GTA particle head.

Although the annotated putative functions for the 12 genes listed in Table 5 sometimes do not correlate to what we believe should be regulating GTA activity, the similar phenotypes of

the diverse mutants indicate an extraordinary complexity of this system in *R. capsulatus*, consistent with the idea that this bacteriophage-like particle is regulated by the cell. This could demonstrate that GTA has been co-evolving with *R. capsulatus* for a long time to become deeply rooted in cellular processes, as opposed to being a recently acquired prophage that has become defective. My finding of multiple genes that appear to regulate different aspects of GTA activity also disputes our past assumption that GTA is regulated by a simple two-component system and a quorum sensing system, with just three key regulatory proteins (i.e. CtrA, CckA and GtaI). It is also interesting to note that some of the genes that I discovered seem to be involved with the regulation of GTA post-transcriptionally, and so it is possible that there is a feedback mechanism that prevents more GTA mRNA from being transcribed if there is a build-up of unmatured GTA particle components in the cell (as in the *cckA*⁻ and *gtaI*⁻ mutants).

4.2 The western blot approach

The GTA immunoassay developed by Taylor (2004) is a great tool for determining what is happening with GTA production. Although my experiments do not determine whether the mature (26 kDa) capsid protein (Orfg5) is part of a complete phage head, it can be assumed that if this capsid protein has been cleaved, which is the form seen outside of the cell, that it is indeed able to be incorporated into a complete phage head. In the *cckA*⁻ mutant, however, the mature (26 kDa) capsid protein appears to be made to high amounts in the cell (Figure 13), but it does not appear to be released from the cell, and so indicates that this mutation prevents assembly of the mature GTA particles, or release of GTA from the cell.

4.2.1 The pYP effect

While testing control strains for GTA activity with and without plasmid pYP, it was discovered that the presence of pYP reduced GTA transduction activity (measured by the GTA bioassay) and capsid production (measured by GTA immunoassay) in all strains. However, pYnP, a similar construct to pYP except for the absence of the promoter region for the GTA gene cluster, did not have the same effect on GTA transduction levels (Figure 9), or on capsid levels as measured by the GTA immunoassay (Figure 11). Therefore it must be the sequence present in pYP and absent from the pYnP construct that reduces GTA structural gene expression. The ongoing DNA sequence determination of pSMF001 and pSMF002 may clarify this question.

4.2.2 Time course analysis of capsid production in Y262 and B10

The production of GTA in *R. capsulatus* was previously determined to be at the highest amounts in stationary phase, using the bioassay which measures GTA transduction (Yen *et al.*, 1979). The transcription of the GTA gene cluster, measured by monitoring β-galactosidase specific activity using plasmid pYP (described in Material and Methods), was shown to increase throughout the growth curve and reach a maximum in late stationary phase (Lang, 2000). The immunoblots of capsid protein in the cell and in cell-free culture supernatant in Figure 12 support the bioassay data, and give us a little more insight into the production and release of GTA in Y262 as well as in B10.

By comparing the immunoblots of Y262 in Figure 12B to the growth curve shown in Figure 12A, it is apparent that GTA is made in late-log phase, and released very shortly (approximately 2 hours) after it accumulates in the cell. This improves our understanding of when GTA is released from the cell, and adds that the mature capsid protein, and presumably GTA, accumulates in the cell before being released, as opposed to being released as fast as it is

made. As we have no understanding of how GTA is released from the cell, this accumulation of GTA in the cell prior to release could mean three things: 1) that the accumulation of GTA is needed for it to be released, i.e. the release of GTA is regulated by the amount of GTA precursors in the cell; 2) that GTA is released by lysis of a small number of cells in the population and that this accumulation of GTA causes a few cells to lyse and release the GTA into the culture supernatant (recall that no plaques were observed with GTA, meaning that either no cell lysis occurs, or only a small subset of the cells lyse); 3) the assembly of the GTA particles takes a few hours to complete, and that is why we see mature capsid protein 2 hours prior to release of GTA into the culture supernatant. Although it is not possible to determine which of these options is correct, the immunoblot of Y262 over a time-course gives us a better view of when capsid is made and when GTA is released from the cell over the growth curve.

By comparing the amount of capsid seen in Y262 and B10, we are able to clearly see differences in capsid production, and therefore GTA production in these two strains. Previous work showed that B10 had less GTA transduction activity (Yen *et al.*, 1979), less of GTA *orfg4* transcript and lower β-galactosidase specific activity measured from the *lacZ* gene fusion on pYP than Y262(Lang, 2000). The immunoblot shown in Figure 12C supports these data, as there is very little capsid protein seen in the B10 cell samples, and no detectable capsid protein in the B10 cell-free culture supernatant samples. Therefore it can be concluded that the reason that Y262 has more GTA transduction activity than B10 is due to an increase in GTA structural gene transcription/translation in Y262, and not due to increased release of GTA from cells.

4.2.3 Time course analysis of capsid production in *ctrA*⁻ and *cckA*⁻ mutants

Two mutant strains, YCKF2 (*ctrA*⁻) and YKKR2 (*cckA*⁻), known to be defective in GTA transduction, were investigated over a time-course using the GTA immunoassay to attempt to elucidate why these strains have low/undetectable GTA transduction activity. From the immunoblots in Figure 13, it is possible to deduce possible mechanisms that are defective in these two mutants. It was previously thought that CtrA and CckA were involved in a two component signalling pathway similar to one seen in *C. crescentus*, but my results indicate that the pathway is more complex than a simple two component system, and it seems as though CtrA and CckA do not necessarily operate in the same pathway as previously thought.

From the immunoblots of YCKF2 in Figure 13B, it is clear that very little, if any capsid is made at any point over the growth curve for this mutant. This is consistent with the β-galactosidase data from YCKF2(pYP) shown in Figure 7, in which the levels of β-galactosidase were very low compared to Y262(pYP). This also confirms our previous assumption that CtrA is very important to the regulation of transcription of the GTA structural gene cluster, and that in the YCKF2 mutant lacking this protein, there is little, if any transcription or expression of proteins from the GTA gene cluster.

From the immunoblots of YKKR2 (*cckA* mutant) in Figure 13C, it is seen that large amounts of capsid, both the pro-protein and the mature, cleaved protein are produced. However, there is little mature capsid seen outside of the cell, which would explain why YKKR2's transduction frequency is almost non-existent. This also correlates with the β-galactosidase data in Figure 7 that shows that the transcription from the GTA promoter is only decreased by approximately 50% in YKKR2(pYP). These data indicate that CckA plays a role in post-transcriptional, and even post-translational processing of the GTA structural gene cluster, or in

the assembly or release of the mature GTA particles from the cell. I no longer assume that CckA and CtrA are inextricable parts of a two-component system, as two things would need to be occurring for that to work with the data I have shown: 1) CckA would also have to regulate some other part of GTA assembly or post-translationally processing as described above, as well as activate CtrA for there to be as much capsid seen in the cell, compared to the *ctrA* mutant; 2) there must be some other sensor kinase protein that can activate CtrA to promote transcription of the GTA structural gene cluster for there to be such a high level of both transcription of the GTA gene cluster (as shown by the β-galactosidase activity in Figure 7), and expression of the capsid protein (as shown by the immunoblot in Figure 13C) in the *cckA*$^-$ strain.

4.2.4 Immunoblots of B10, SB1003 and IKOI

Anther interesting complexity between strains is the paradox of B10 compared to SB1003 and IKOI (the *gtaI* mutant derived from SB1003). It was previously thought that B10 and SB1003 would have similar GTA levels, and that IKOI would have lower levels than both, as SB1003 is a B10 derivative, and IKOI (quorum sensing mutant) has been shown to be defective in GTA transduction. However, the levels of capsid seen in these 3 strains vary considerably, as shown in Figure 14.

Figure 14A compares SB1003 to B10, in both cell and cell-free culture supernatant samples. The levels of capsid seen in the cell appear similar, but the levels differ considerably in the cell-free culture supernatant samples. In B10, no capsid protein is seen, however in SB1003, there is quite a bit of the mature capsid seen. SB1003 is a rifampicin resistant derivative of B10, that was created by GTA transduction from a rifampicin resistant strain and a poorly characterized strain (B100) that was deemed to be cured of lytic phages (Solioz, 1975). It has been observed that SB1003 produces greater amounts of photosynthetic complexes than B10

(J.T. Beatty, personal communication), and so it is possible that the mutation that makes SB1003 rifampicin resistant (thought to be an amino acid change in the RNA polymerase-subunit that prevents rifampicin from binding to it) could also increase the affinity of the RNA polymerase for certain promoters, such as the photosynthesis gene *puf* operon and the GTA structural gene cluster.

Figure 14B compares B10 to IKOI, which is a quorum sensing mutant (*gtaI*) that was created in SB1003 (Schaefer *et al.*, 2002). Again, the levels of capsid compared to B10 are increased, however in this case it is the capsid levels in the cell that are increased the most dramatically. There is some mature capsid protein seen in the cell-free culture supernatant sample, although more than B10, and less than SB1003, which is consistent with the work of Schaefer *et al.* (2002), who reported that IKOI had reduced GTA transduction activity compared to SB1003. The intracellular levels of capsid are elevated in the IKOI strain, however, and are similar to those seen in Figure 13C for YKKR2 (*cckA*$^-$). There appears to be a buildup of both the capsid proprotein and the mature capsid protein, and it is conceivable that the extracellular mature capsid protein seen is due to lysis of cells in the culture rather than active release of the GTA. This is an interesting finding as it seems to support the theory that GTA release is due to lysis of a subset of cells in a population, perhaps due to a quorum sensing signal telling the cells that the culture is overcrowded. It has been observed that photosynthetically grown cultures of IKOI do not enter death phase as readily as Y262 cultures, (i.e. culture densities, as measured by OD660, stay constant for longer than 70 hours post inoculation when Y262 and B10 cultures have started to decline in density; M. Leung, personal communication).

Although these immunoblots do not give conclusive results as to what is happening in regards to GTA production and release, they give a better picture of what is occurring in regard

to capsid production and processing, especially in the mutant strains YCKF2 (*ctrA*$^-$), YKKR2 (*cckA*$^-$) and IKOI (*gtaI*$^-$), compared to control strains.

4.3 GTA structural gene cluster promoter analysis

Prior investigation of the promoter for the GTA structural gene cluster was limited to sequencing a short region (470 bp) and looking for a CtrA consensus binding site, although nothing of obvious significance was found in either (Lang, 2000). The promoter for the GTA gene cluster was thought to be located between the start codon for *orfg1* and the stop codon for the upstream gene (RRC03520), which is not thought to be involved in the GTA gene cluster, although it is transcribed in the same direction (Lang, 2000; Lang and Beatty, 2002).

I undertook this promoter analysis research because I was interested in discovering which part of the putative GTA structural gene cluster was responsible for the observed pYP effect that I discovered. DNA sequencing of the putative promoter region was done to discover if there were any differences between the various strains of *R. capsulatus* that we knew had different GTA transduction levels. Softberry's BPROM software was used to predict -10 and -35 sites, rapid amplification of cDNA ends (R.A.C.E.) was used to determine the 5' ends of the GTA *orfg1*, and presumably the entire GTA gene cluster. Promoter region deletions were made to attempt to discover the region of the pYP plasmid fusion that causes the pYP effect, and the minimal region needed for full promoter activity.

4.3.1 Sequence differences between the GTA promoter in Y262, B10 and SB1003

Lang (2000) hypothesized that a *cis*-acting regulatory sequence in the promoter region of the GTA gene cluster caused the observed GTA activity differences between Y262 and B10. Two promoter fusions, pYP and pSTU12, were made using DNA directly cloned from two strains that had differing GTA activities, Y262 and B10 respectively, and were tested by Lang for promoter activity in Y262 (Lang, 2000). It was found that the fusion plasmid containing the GTA promoter region from B10, pSTU12, produced unmeasurable levels of β-galactosidase, while the Y262 fusion plasmid, pYP, produced readily detectable β-galactosidase levels (Lang, 2000). It was previously thought that the reason for the difference in β-galactosidase specific activity between the B10 *orfg2'::'lacZ* fusion and the Y262 *orfg2'::'lacZ* fusion was due to a *cis*-acting regulatory sequence that was mutated in Y262 that either prevented a negative regulatory protein from binding or allowed a positive regulatory protein to bind more specifically, which in both cases would cause the up-regulation seen in Y262 compared to B10. However my DNA sequencing data show that there is no difference between strains Y262, B10 or SB1003 over a substantial region encompassing the predicted GTA structural gene cluster promoter. Therefore there must be one or more *trans*-acting factor(s) mutated in Y262 and SB1003 to cause the observed increase in production of GTA. To explain the reduced β-galactosidase specific activity seen from the B10 *orfg2'::'lacZ* fusion compared to the Y262 *orfg2'::'lacZ* fusion (Lang, 2000), it must be the difference between fusion sites that were used to create the two plasmids, pYP and pSTU12 that cause this difference. That is, although both plasmids contain exactly the same GTA structural gene promoter region, these plasmids differ in the sequence fused translationally in frame to *lacZ*. Thus the two hybrid β-galactosidase proteins

produced contain different N-terminal sequences, both in length and amino acid sequence composition, which probably accounts for the difference in β-galactosidase specific activity.

When Y262(pSTU12) was compared to Y262(pYP) using the immunoassay, there was reduced amounts of capsid seen in the supernatant and the cell samples of Y262(pSTU12), although they were not as reduced as in Y262(pYP) (Figure 17). As the sequences contained within the promoter region and *orfg1* differ only in the fusion joint of *orfg2* to *lacZ*, it must be the nature of the Orfg2::LacZ fusion that explains the difference in both β-galactosidase specific activity and GTA immunoassay capsid levels.

Another interesting finding was that there are equal levels of *ctrA* mRNA in Y262 and B10, whereas there is almost no detectable *org4* (from the GTA gene cluster) mRNA seen in B10 and a greater amount in Y262 (Lang, 2000). It was hypothesized that this could be explained by there being a mutation in the promoter that would allow a regulatory protein (i.e., CtrA) to bind with higher affinity to the GTA gene cluster promoter in Y262. However as there is no sequence difference in the GTA promoter region between these two strains, there must be a change in the expression or activation of CtrA that causes the difference in GTA transcript levels between the two strains, or a protein other than CtrA induces the transcription of the GTA structural gene cluster.

4.3.2 Predicted promoter region for GTA: R.A.C.E. and Softberry analysis

Softberry's BPROM software predicted a -10 and -35 site 100 bp upstream of the start codon for *orfg1*. Although BPROM is biased towards AT-rich organisms, which *R. capsulatus* is not, the putative promoter region for the GTA gene cluster is actually quite AT rich (46%) compared to the average over the genome (32%). The 5' end mapping of *orfg1* mRNA gave two 5'ends. This could be due to the reverse transcriptase reaction terminating prematurely, two

messages with the shorter being cleaved from the longer, or two different promoters initiating transcription. The shorter 5' end overlaps with the BPROM predicted -10 site. The longer 5' end overlaps gene RRC03520 by 23 bp, and so it is possible that this is a misannotated gene, or that the promoter for GTA is located within the coding sequence of this gene. Multiple promoters within gene coding regions have been documented in *R. capsulatus* (Wellington and Beatty, 1991; Wellington *et al.*, 1991). From my results, it is not possible to conclusively determine exactly where the promoter for GTA is located, although I suggest that the promoter is in fact further upstream than previously thought, because one of the RNA 5' ends was located 20 bp 5' of the stop codon of the upstream gene, RRC03520. Although more experiments could be preformed to provide conclusive evidence of the exact location of the promoter, I suggest -35 and -10 sites in Figure 16.

4.3.3 Promoter deletions

Neither the pSMF001 nor the pSMF002 *lacZ* fusion yielded β-galactosidase activity, although it was later discovered that the pSMF002 fusion was non-*R. capsulatus* DNA, and therefore would not be expected to have β-galactosidase activity. Since pSMF001 (*orfg1* fusion position shown in Figure 16) had no detectable β-galactosidase, it is possible that *orfg1* is not translated, although the R.A.C.E. data show that *orfg1* is transcribed. There could be something about the 5 Orfg1 amino acids fused to the N-terminus of β-galactosidase in pSMF001 that inhibits β-galactosidase activity in the fusion protein, but this is unlikely because of the extremely robust nature of such fusions (Silhavy and Beckwith, 1985). It is conceivable that the GTA structural gene promoter is located 3' of the *orfg1::lacZ* fusion joint in pSMF001. However, this seems unlikely because of the strongly conserved Orfg1-homologous sequences

(using the same reading frame as in the pSMF001 fusion of *orfg1* to *lacZ*) located 5' of Orfg2 homologues in several other bacteria that contain GTA-like gene clusters (Lang *et al.*, 2002). The most probable explanation is that pSMF001 does not contain the entire GTA structural gene cluster promoter, and therefore, does not transcribe the *orfg1::lacZ* fusion. To ensure that *orfg1* is transcribed/translated, and that the fusion point is indeed correct, pSMF002 will have to be made and tested. If pSMF002 has β-galactosidase activity, then it can be assumed that pSMF001 does not contain the entire GTA structural gene cluster promoter, however if pSMF002 does not have β-galactosidase activity, then either the fusion point is not to the correct start codon for *orfg1* or *orfg1* is not translated. However, the absence of β-galactosidase expression in pSMF001 is consistent with the promoter sequence suggested in Figure 16, which is absent from pSMF001.

4.3.4 GTA immunoassays of Y262(pSMF001) and Y262(pSMF002)

Although pSMF002 was determined to contain non-*R. capsulatus* DNA, pSMF001 was the expected fusion. When Y262(pSMF001) was compared to Y262(pYP), Y262(pYnP) and Y262(pSTU12), it was apparent that pSMF001 did not exhibit the pYP effect seen in both pYP and pSTU12 (Figure 17). However, pSTU12 is seen to have less of an effect on capsid production than pYP does. As Y262(pSTU12) has no detectable β-galactosidase specific activity, it can be assumed that this fusion makes a non-functional β-galactosidase enzyme, which leads us to believe that it is the fusion protein itself that causes the observed effect on capsid production. By comparing this to pSMF001, either pSMF001 does not contain the entire GTA structural gene cluster promoter, or the Orfg1::LacZ fusion protein is non-functional. Further deletions would have to be made to determine which of these predictions is correct.

4.4 Future Research

There are several paths that this research could take. For the twelve mutants discovered to have an effect on GTA production (Table 5), each gene would have to be knocked out with a translationally inframe (non-polar) deletion and complemented in *trans* with only the gene knocked out, transcribed using a promoter like the *puf* promoter as done for the *ctrA* and *cckA* knock outs (Lang and Beatty, 2000, 2001). Even so, it would be difficult to ascribe a biochemical function to these gene products, especially the hypothetical proteins.

Further studies of GTA production in SB1003, compared to B10 and Y262, could help elucidate key differences between these strains that are not known at this time. GTA transduction assays (bioassay) of SB1003 compared to Y262 would show if the capsid seen in the cell-free supernatant samples of SB1003 are transductionally active. Comparative genomics and transcriptomics between these strains would also be interesting to do, once a microarray for *R. capsulatus* is available. Such work may give a clue into the mutations that make Y262 a GTA over producing strain compared to B10, and the difference that I found between SB1003 and B10. Introducing plasmid pYP into SB1003 to measure transcription levels from the GTA structural gene cluster promoter would also be an interesting follow up to these immunoassays. If the levels of transcription are similar to the transcription levels in Y262, it could be that one of the mutations that makes Y262 a GTA overproducer is the same that makes SB1003 rifampicin resistant.

Further DNA sequence analysis and comparisons of known key GTA regulatory factor sequences between Y262, B10 and SB1003 could be used to design directed mutagenesis experiments to explore how the different phenotypes shown for GTA production and transduction activity relate to potential gene sequence differences.

Additional promoter deletions could be made to pYP, using various combinations of 5' sequences in fusions of both *orfg1* and *orfg2* to *lacZ*. A translationally in-frame (nonpolar) deletion of *orfg1* could also be created. Such experiments would determine if *orfg1* is a real gene, and whether *orfg2* is in fact the first real gene in this cluster. R.A.C.E. could also be done using primers that anneal within *orfg2* to see if there is another 5' mRNA closer to *orfg2* than the two that I discovered for *orfg1*, or if the 5' end will still map to the 5' ends that I found.

In conclusion, the study of the *R. capsulatus* GTA, although the best understood of all known generalized transducing BLPs, remains a challenging area for future studies. Hopefully my discoveries will advance this ongoing area of research.

5 References:

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

Barbian, K.D., and Minnick, M.F. (2000) A bacteriophage-like particle from *Bartonella bacilliformis*. *Microbiology-Uk* **146**: 599-609.

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

Bertani, G. (1999) Transduction-like gene transfer in the methanogen *Methanococcus voltae*. *Journal of Bacteriology* **181**: 2992-3002.

Campbell, J. (1996) *Biology*. Menlo Park, California: Benjamin/Cummings Publishing Company, inc.

Davies, J. (1994) Inactivation of antibiotics and the dissemination of resistance genes. *Science* **264**: 375-382.

Diaz, A.N., Sanchez, F.G., and Garcia, J.A.G. (1998) Phenol derivatives as enhancers and inhibitors of luminol-H2O2-horseradish peroxidase chemiluminescence. *Journal of Bioluminescence and Chemiluminescence* **13**: 75-84.

Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, X.W., Finlay, D.R., Guiney, D., and Helinski, D.R. (1985) Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. *Plasmid* **13**: 149-153.

Duda, R.L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D., and Hendrix, R.W. (1995) Structural transitions during bacteriophage HK97 head assembly. *Journal of Molecular Biology* **247**: 618-635.

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

Gilman, M.Z., Wiggs, J.L., and Chamberlin, M.J. (1981) Nucleotide sequences of two *Bacillus subtilis* promoters used by *Bacillus subtilis* sigma-28 RNA polymerase. *Nucleic Acids Research* **9**: 5991-6000.

Griffiths, A.J.F. (2005) *Introduction to Genetic Analysis*. New York: W.H. Freeman.

Haselkorn, R., Lapidus, A., Kogan, Y., Vlcek, C., Paces, J., Paces, V., Ulbrich, P., Pecenkova, T., Rebrekov, D., Milgram, A., Mazur, M., Cox, R., Kyrpides, N., Ivanova, N., Kapatral, V., Los, T., Lykidis, A., Mikhailova, N., Reznik, G., Vasieva,

O., and Fonstein, M. (2001) The Rhodobacter capsulatus genome. *Photosynth Res* **70**: 43-52.

Hendrix, R.W. (1983) *Lambda II.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

Humphrey, S., Stanton, TB, Jensen, NS, Zuerner, RL (1997) Purification and characterization of VSH-1, a generalized transducing bacteriophage of *Serpulina hyodysenteriae. Journal of Bacteriology* **179**: 323-329.

Imhoff, J.F., and Truper, H.G. (1989) The Purple Nonsulfur Bacteria. In *Bergey's Manual of Systematic Bacteriology.* Vol. 3. Staley, J.T., Bryant, M.P., Pfenning, N. and Holt, J.C. (eds). Baltimore: Williams and Wilkins, pp. 1658-1661.

Imhoff, J.F. (1995) Purple and Green Bacteria. In *Anoxygenic Photosynthetic Bacteria.* Blankenship, B.E., Madigan, M.T. and Bauer, C.E. (eds). Boston: Kluwer Academic Publishers, pp. 1-11.

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

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D., and Rehrauer, W.M. (1994) Biochemistry of homologous recombination in Escherichia coli. *Microbiology Review* **58**: 401-465.

Lang, A.S. (2000) The Gene Transfer Agent(GTA) of *Rhodobacter capsulatus.* In *Microbiology and Immunology* Vancouver: University of British Columbia, pp. 108.

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

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

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

Lang, A.S., Taylor, T.A., and Beatty, J.T. (2002) Evolutionary implications of phylogenetic analyses of the gene transfer agent (GTA) of *Rhodobacter capsulatus. Journal of Molecular Evolution* **55**: 534-543.

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

Madigan, M.T., Martinko, J.M., and Parker, J. (2003) *Brock biology of microorganisms.* Upper Saddle River, NJ: Prentice Hall/Pearson Education.

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

Marrs, B.L. (2002) The early history of the genetics of photosynthetic bacteria: a personal account. *Photosynthesis Research* **73**: 55-58.

Matson, E.G., Thompson, M.G., Humphrey, S.B., Zuerner, R.L., and Stanton, T.B. (2005) Identification of genes of VSH-1, a prophage-like gene transfer agent of Brachyspira hyodysenteriae. *Journal of Bacteriology* **187**: 5885-5892.

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.

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

Rapp, B.J., and Wall, J.D. (1987) Genetic Transfer in *Desulfovibrio desulfuricans*. *PNAS* **84**: 9128-9130.

Redfield, R.J. (2001) Do bacteria have sex? *Nature Reviews Genetics* **2**: 634-639.

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

Schaefer, A.L., Taylor, T.A., Beatty, J.T., and Greenberg, E.P. (2002) Long-Chain Acyl-Homoserine Lactone Quorum-Sensing Regulation of *Rhodobacter capsulatus* Gene Transfer Agent Production. *Journal of Bacteriology* **184**: 6515-6521.

Silhavy, T.J., and Beckwith, J.R. (1985) Uses of lac fusions for the study of biological problems. *Microbiol Rev* **49**: 398-418.

Simon, R., Priefer, U., and Pühler, A. (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. *Bio/Technology* **1**: 784-791.

Smith, M.C., Burns, R.N., Wilson, S.E., and Gregory, M.A. (1999) The complete genome sequence of the *Streptomyces* temperate phage straight phiC31: evolutionary relationships to other viruses. *Nucleic Acids Research* **27**: 2145-2155.

Solioz, M. (1975) The gene transfer agent of *Rhodopseudomonas capsulatus*. In *Biochemistry* Saint Louis: Saint Louis University, pp. 112.

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

Solioz, M., and Marrs, B. (1977) The gene transfer agent of *Rhodopseudomonas capsulata*: purification and characterization of its nucleic acid. *Archives of Biochemistry and Biophysics* **181**: 300-307.

Taylor, T.A. (2004) Evolution and Regulation of the Gene Transfer Agent (GTA) of *Rhodobacter capsulatus*. In *Microbiology and Immunology* Vancouver: University of British Columbia, pp. 67.

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

Wellington, C.L., and Beatty, J.T. (1991) Overlapping mRNA transcripts of photosynthesis gene operons in Rhodobacter capsulatus. *Journal of Bacteriology* **173**: 1432-1443.

Wellington, C.L., Taggart, A.K., and Beatty, J.T. (1991) Functional significance of overlapping transcripts of crtEF, bchCA, and puf photosynthesis gene operons in Rhodobacter capsulatus. *Journal of Bacteriology* **173**: 2954-2961.

Wong, D.K., Collins, W.J., Harmer, A., Lilburn, T.G., and Beatty, J.T. (1996) Directed mutagenesis of the *Rhodobacter capsulatus puhA* gene and *orf*214: pleiotropic effects on photosynthetic reaction center and light-harvesting 1 complexes. *Journal of Bacteriology* **178**: 2334-2342.

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

Yen, H.C., and Marrs, B. (1977) Growth of *Rhodopseudomonas capsulata* under anaerobic dark conditions with dimethyl sulfoxide. *Archives of Biochemistry and Biophysics* **181**: 411-418.

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