Scientific Research Publishing

Advances in Microbiology, 2018, 8, 558-577
http://www.scirp.org/journal/aim
ISSN Online: 2165-3410
ISSN Print: 2165-3402

# CtrA Is Nonessential for Cell Cycle Regulation in *Rhodobacter sphaeroides*

**Lin Lin, Abha Choudhary, Anish Bavishi, Madhusudan Choudhary***

Department of Biological Sciences, Sam Houston State University, Huntsville, Texas, USA
Email: *mchoudhary@shsu.edu

**How to cite this paper:** Lin, L., Choudhary, A., Bavishi, A. and Choudhary, M. (2018) CtrA Is Nonessential for Cell Cycle Regulation in *Rhodobacter sphaeroides*. *Advances in Microbiology*, **8**, 558-577.
https://doi.org/10.4236/aim.2018.87037

**Received:** May 9, 2018
**Accepted:** July 27, 2018
**Published:** July 30, 2018



## Abstract

The bacterial cell cycle consists of a series of genetically coordinated biochemical and biophysical events. In *Caulobacter crescentus*, CtrA is an essential cell cycle regulator that modulates many cell cycle processes. In the present study, the role of the CtrA was investigated in *Rhodobacter sphaeroides* 2.4.1 by employing genetic, molecular, and bioinformatic approaches. Examination of the *ctrA*-null mutant revealed that the loss of CtrA did not affect growth characteristics and cell morphology in *R. sphaeroides* when grown under aerobic or photosynthetic growth conditions but slower growth was noticed in the anaerobic-dark-DMSO condition. Phylogenetic analyses demonstrated that CtrA has diversified its role in major lineages of *α-Proteobacteria* and has possibly been involved in adaptation to variable lifestyles. Analysis of the CtrA binding sites in the *R. sphaeroides* genome suggests that CtrA may regulate 127 genes involving different cellular processes. Protein homology searches revealed that only a small number of *ctrA*-regulated genes are homologous across *C. crescentus*, *R. capsulatus*, and *R. sphaeroides*. Comparison of the functions of putative *ctrA*-regulated genes in *C. crescentus*, *R. capsulatus*, and *R. sphaeroides* revealed that all three species possessed broad pathway control across a variety of cluster of orthologous gene functions (COGs). However, interestingly, it seems that the essentiality of CtrA in *C. crescentus* may depend more on the selective control that it exerts on a few critical cell cycle genes and pathways that are not controlled by CtrA in a similar fashion in *R. capsulatus* and *R. sphaeroides*.

## Keywords

CtrA, *Rhodobacter sphaeroides*, Cell Cycle, Cell Cycle Regulation, Symmetric Cell Division, Motility

## 1. Introduction

CtrA (cell cycle transcriptional regulator) proteins form a subgroup within the

OmpR family of winged helix turn helix (HTH) response regulators [1] [2]. Although not present in some bacteria such as *Pelagibacter ubique* [3], the ctrA gene is only found in *α-Proteobacteria* [4] and these regulators are highly conserved among almost all of the species of the seven major orders of this group: *Caulobacterales*, *Parvularculales*, *Rhizobiales*, *Rhodobacterales*, *Rhodospirillales*, *Rickettsiales*, *and Sphingomonadales*. CtrA plays an essential role in cell cycle regulation in *Caulobacter crescentus*, a model bacterial organism used to study the asymmetric cell cycle [5] and it directly or indirectly regulates over 25% of 553 cell cycle related genes in *C. crescentus* [5] [6], including genes involved in DNA replication, DNA methylation, cell division, flagellar biogenesis, and polar morphogenesis [6] [7] [8].

Like in *C. crescentus*, CtrA homologs have been shown to play essential or important roles in cell viability and cell cycle regulation in *Sinorhizobium meliloti* [4], *Brucella abortus* [9] [10], *Agrobacterium tumefaciens* [4], and *Rickettsia prowazekii* [4] [11]. However, CtrA was shown to be nonessential in both *Rhodobacter capsulatus* [12] [13] and *Ruegeria sp*. TM 1040 [14]. Both microarray and proteomic expression analysis of a *ctrA* mutant in *R. capsulatus* revealed that CtrA is not essential for its cell cycle effects but it is an important regulator that controls over 225 genes, including those involved in motility, gene exchange, gas vesicle formation, and synthesis of transcriptional regulators and signal transduction proteins [12] [15]. While *R. capsulatus* and *Rugeria sp.* TM1040 divide symmetrically, other species like *A. tumefaciens*, *B. abortus*, *C. crescentus*, and *S. meliloti* divide asymmetrically. These findings could suggest that CtrA does not possess as essential a role in cell cycle regulation in *Rhodobacterales* as it does in other orders such as *Caulobacterales* and *Rhizobiales*.

*R. sphaeroides* exhibits a wide range of metabolic versatility [16] and unlike most prokaryotes, including *C. crescentus* and *R. capsulatus*, the genome of *R. sphaeroides* is multi-partite and comprised of two circular chromosomes and five endogenous plasmids [17]. *R. sphaeroides* undergoes symmetric cell division as found in *R. capsulatus* and *Rugeria sp.* TM1040, even though it possesses a more complex genome structure. On the other hand, due to its complex genome structure, *R. sphaeroides* may possess different evolutionary features compared to bacteria whose genomes consist of a single primary chromosome. The hypothesis that CtrA may have functionally diverged in *R. sphaeroides* is investigated in this study and the role of CtrA will be compared with results previously described among other *α-Proteobacteria* species.

## 2. Materials and Methods

### 2.1. Bacterial Strains and Growth Conditions

Plasmids, bacterial strains, and oligonucleotide primers used in this study are listed in **Table 1**. *R. sphaeroides* 2.4.1 was grown aerobically in Sistrom's minimal medium at 30˚C [18]. Photosynthetic cultures were grown under low light

**Table 1.** Bacterial strains, plasmids and oligonucleotides.

| Strain/plasmid/ oligonucleotide | Phenotype/Genotype/Description | Source or reference |
|---|---|---|
| **Bacterial strains** | | |
| *R. sphaeroides* 2. 4. 1 | Wild type | Van Niel (1944) |
| *R. sphaeroides ΔctrA* | 2. 4. 1 derivative with in-frame deletion of the DNA sequence between codon 7 and 231 from the start codon of ctrA | This study |
| *E. coli* DH5α-*phe* | (Φ80d*lacZ*¢*M*15) ¢*lacU*169 *recA*1 *endA*1 *hsdR*17 *supE*44 11 *thi*1 *gyrA*96 *relA*1 *phe*::*Tn*10*d*Cm | (10) |
| *E. coli* S17-1 | Pro⁻ Res⁻ Mod⁺ *recA*; integrated plasmid RP4-Tc::Mu Km::Tn7 | Simon *et al.* (1983) |
| SHSU0001 | DH5α-*phe* containing pLO1*ΔctrA* plasmid | This study |
| SHSU0002 | E. coli S17-1 containing pLO1*ΔctrA* plasmid | This study |
| **Plasmids** | | |
| pLO1 | Km$^r$; *sacB*RP4-*oriT* ColE1-*oriV* | Lenz *et al.* |
| pLO1*ΔctrA* | pLO1::0.5 fragment of *ΔctrA* | This study |
| pRK415 | Tc$^r$; Mob$^+$*lacZ* IncP | Keen *et al.* (1988) |
| pRK415*ctrA* | pRK415::0.8 Kb fragment *ctrA* gene of *R. sphaeroides* | This study |
| **Oligonucleotides** | | |
| *ctrA*-upstream F1 | 5'-GCTCTAGAGCCGCGGAAGGTGTTCTCGTAA-3' | This study |
| *ctrA*-upstream R2 | 5'-TCAGGCGCCGATGGCGAACCGCTCCACCAGCA GTATTCTCAT-3' | This study |
| *ctrA*-downstream F3 | 5'-ATGAGAATACTGCTGGTGGAGCGGTTCGCCAT CGGCGCCTGA-3' | This study |
| *ctrA*-downstream R4 | 5'-GCTCTAGAGCGCGCCCCAGGCATAGGCGAA-3' | This study |
| *ctrA*-sequencing F1 | 5'-AATTCGCCCTGTGGCGGTCG-3' | This study |
| *ctrA*-sequencing R2 | 5'-TCTGCCAGCCGTGCAAGCTC-3' | This study |
| *ctrA*-qPCR-1F1 | 5'-GAATACTGCTGGTGGAGGATG-3' | This study |
| *ctrA*-qPCR-1R2 | 5'-ATCGTAATCGTAGAGCTTGGC-3' | This study |
| *ctrA*-qPCR-2F1 | 5'-ACGATTACATGACCAAGCCC-3' | This study |
| *ctrA*-qPCR-2R2 | 5'-TTCGCATCGAGATTAACGAGG-3' | This study |
| *ctrA*-qPCR-3F1 | 5'-ATCACCTTTACGGAGGCATG-3' | This study |
| *ctrA*-qPCR-3R2 | 5'-ACCGTCTCGATGTAGCTCTC-3' | This study |
| *rpoZ*-qPCR F1 | 5'-ATCGCGGAAGAGACCCAGAG-3' | This study |
| *rpoZ*-qPCR R2 | 5'-GAGCAGCGCCATCTGATCCT-3' | This study |
| *ctrA-2F* | 5'-AAGCTTTTAACCATTGGGCATCATTG-3 | This study |
| *ctrA-2R* | 5'-TCTAGAGGCAATCCGCCCGAG-3 | This study |

intensity (3 watts/m$^2$) or medium light intensity (15 watts/m$^2$). In addition, the *R. sphaeroides* culture was also grown under an anaerobic and dark condition

with DMSO (60 μmol/L). All *Escherichia coli* strains were grown at 37˚C in Luria-Bertani (LB) medium. Kanamycin (Km) was used to select for the presence of pLO1 plasmids in *E. coli* and *R. sphaeroides*, at a concentration of 50 μg/mL. For molecular analyses, standard methods were used for plasmid preparation, restriction endonuclease digestion, ligation, polymerase chain reaction, and DNA sequencing.

## 2.2. Construction of in-Frame Gene Deletion and Molecular Cloning

A *ctrA* null allele (an in-frame deletion) was constructed using the gene knockout method [19]. The in-frame deletion allele (Δ*ctrA* allele) was constructed *in vitro* following two steps. A fragment containing 787-bp upstream of *ctrA* to 21-bp of the 5' *ctrA* coding sequence was amplified using *ctrA*-upstream F1 (a *XbaI* site was introduced at the 5' end of this primer) and *ctrA*-upstream R2 primers. Another fragment containing the 803-bp downstream of *ctrA* to 21-bp of the 3' *ctrA* coding sequence was amplified using *ctrA*-downstream F3 and *ctrA*-downstream R4 primers (a *XbaI* site was introduced at the 5' end of this primer). *ctrA*-upstream R2 and *ctrA*-downstream F3 primers were designed to be complementary with each other. These two PCR fragments were sequenced to confirm their sequences and then used in a second round of PCR as template DNA, using the *ctrA*-upstream F1 (forward primer) and the *ctrA*-upstream R2 (reverse primer). As a result, the complementary sequences at the ends of the PCR fragments produced a fusion *ctrA* deletion fragment (Δ*ctrA* allele), which contained only a 14 amino acid coding sequence corresponding to the first seven and the last seven codons of the *ctrA* gene.

The suicide vector, pLO1 [20], was linearized with *XbaI* and then ligated to the *XbaI*-digested Δ*ctrA* allele. pLO1 is an *E. coli* plasmid containing the $Km^r$ gene and the *Bacillus subtilis sacB* gene [21]. As a suicide vector, pLO1 cannot autonomously replicate in *R. sphaeroides*, but can replicate in *E. coli* [20]. *E. coli* DH5*α* cells were then transformed with the resulting recombinant plasmid, pLO1Δ*ctrA*. The deletion was confirmed by sequencing the entire insert of the recombinant plasmid using flanking sequencing primers, *ctrA*-sequencing F1 and *ctrA*-sequencing R2.

## 2.3. Construction of Δ*CtrA R. sphaeroides* Strain

The *E. coli* S17-1 cells were transformed with the pLO1Δ*ctrA* plasmid, since S17-1 was used for transferring the pLO1Δ*ctrA* plasmid to *R. sphaeroides* by conjugation. The plasmid was mobilized into *R. sphaeroides* using the biparental mating method [19] [22] [23]. *E. coli* S17-1 donor cells with recombinant plasmids (pLO1Δ*ctrA*) were grown in LB mediums with kanamycin to mid-log phase; then 200 μl cultures were spun down. Pellets were washed with LB medium for five times to get rid of kanamycin. On the other hand, *R. sphaeroides* recipient cells were also grown to mid-log phase and a one-milliliter culture was

then pelleted. Both pellets were then mixed in 50 µl LB medium and spotted onto solid LB medium for overnight incubation at 30˚C allowing plasmids to transfer to *R. sphaeroides* by conjugation.

After overnight incubation, cells on spotting plates were collected and washed with SIS medium 5 times to maximally avoid LB medium to prevent *E. coli* contamination in the next step. Ten independent heterogenotes of *R. sphaeroides*, which resulted from a single recombination event, were selected as $Km^r$ colonies on SIS agar plates with kanamycin. Plates were incubated for four days until pink colonies were visible. These $Km^r$ colonies were picked and grown in liquid culture without antibiotics for several times, and then plated out on LB agar plates containing 15% sucrose, which selected for the double crossover event leading to the replacement of the wild type *ctrA* gene by the deletion allele [24]. The *sacB* gene on the plasmid encodes levansucrase that is secreted in culture medium under induced conditions in the presence of sucrose, which causes cell-lysis or inhibition of cell growth [24]. Plates were incubated at 30˚C for several days until pink colonies appeared. Multiple colonies from each plate were patched on both SIS plates and SIS plus kanamycin replica plates. Plates were incubated for at least 5 days. Only $km^s$ colonies were selected from SIS replica plates and were grown under both with and without kanamycin to confirm their kanamycin sensitivities. Eventually, colony PCR was performed on 10 kanamycin-sensitive colonies to screen for the loss of the *ctrA* gene. Putative colonies with loss of *ctrA* were then grown in SIS liquid cultures, and genomic DNA was then extracted. Gene replacement was confirmed by amplifying the fragment containing the *ctrA* deletion by performing PCR using genomic DNA previously extracted from putative mutant strains and subsequent sequencing of the PCR product.

## 2.4. Gene Complementation

Also, *ctrA*-null strain was complemented with pSHSU0003 recombinant plasmid, pRK415 containing *ctrA* wild-type gene with native promoter. A 931 base pair fragment containing the wild type *ctrA* was amplified using ctrA2F and ctrA2R primers. *EcoRI* and *HndIII* sires were introduced in ctrA 2F and ctrA 2R primer, respectively. PCR product was digested with *EcoRI* and *HindIII* and then cloned into EcoRI/HindII sites of pRK415. The recombinant plasmid, pRK415ctrA was mobilized in to *R. sphaeroides* Δ*ctrA* strain for complementation study.

## 2.5. RNA Isolation and Reverse-Transcription PCR

Total RNA was isolated from *R. sphaeroides'* cells grown in lag, mid log, and stationary phases using a Qiagen RNeasy Midi Kit. All RNA samples were checked for concentration and DNA contamination by measuring absorbance at $OD_{260}$ and $OD_{280}$. A Power Sybr@ Green RNA-to-$C_T$™ 1-step Kit (Applied Biosystems) was used for qPCR. Three sets of internal primers as listed in **Table 2**

**Table 2.** Normalization of mRNA expression levels of *ctrA* during different growth stages in wild type and mutant strains.

| Strains | $OD_{600}$ | *rpoZ* $C_T$ | *ctrA* $C_T$ | Fold changes in gene expression ($2^{-\Delta\Delta C_T}$)[a] |
|---|---|---|---|---|
| Wild type | 0.5 | 12.71174908 | 11.84598255 | 1.109098462 |
| | | | 11.97644711 | 1.013202748 |
| | | | 12.16367912 | 0.889885039 |
| | | | | **Mean ± SE:** 1.004 ± 0.110 |
| | 0.6 | 10.15510941 | 9.628145218 | 0.876961117 |
| | | | 9.442531586 | 0.99736841 |
| | | | 9.98237133 | 0.686037179 |
| | | | | **Mean ± SE:** 0.853 ± 0.157 |
| | 1.2 | 11.48839092 | 10.82333279 | 0.965051082 |
| | | | 10.82483864 | 0.964044309 |
| | | | 10.97593403 | 0.868185944 |
| | | | | **Mean ± SE:** 0.932 ± 0.056 |
| Δ*ctrA* | 0.5 | 12.22843933 | 24.5294838 | 0.000120605 |
| | | | 24.67738152 | 0.000108854 |
| | | | 23.88406563 | 0.000188649 |
| | | | | **Mean ± SE:** 0.00014 ± 0.00004 |

[a]$\Delta\Delta C_T = \Delta C_T - \Delta C_{T.\,rpoZ}$. The average of expressions of genes derived from wild type cultures with $OD_{600}$ = 0.5 were set as the reference line for expression changes.

were used for measuring *ctrA* mRNA expression levels to ensure integrity of the full length mRNA. The *rpoZ* gene, which encodes the *ω* subunit of RNA polymerase, was used to normalize mRNA expression values for *ctrA*. The Applied Biosystem StepOne Real-time PCR System (Life Technologies, Gaithersburg, MD) was used to monitor amplifications to quantify mRNA expressions. The $2^{-\Delta\Delta C_T}$ method was used to analyze relative gene expression [25]. Statistical analysis was performed using Minitab 15 statistical software (Minitab, Minneapolis, MN).

## 2.6. Growth Curve, Motility Tests and Microscopy

Cell densities of both wild-type and *ctrA*-null mutant cells were measured by optical densities at 600 nm at different time points to plot growth curves for both wild type and mutant strains. Microscopic examination was performed by fixing cells onto the slides using 20 μL of poly-L-lysine (Sigma-Aldrich P4832), and then examined under 1000x magnification. Swimming motility was assayed on swim plates, which contained SIS medium with 100 μM succinate as the minimal carbon source and solidified with 0.25% (w/v) agar. Swarming motility was tested on swarm plates containing regular SIS medium solidified with 0.25%

(w/v) agar. *R. sphaeroides* cells (5 µl) were placed on swim plates or stabbed into semi-solid swarm plates and incubated at 30˚C aerobically in the dark for several days as previously described [26]. Swarming was also determined by stabbing cells with a straightened toothpick into test tubes with SIS medium solidified with 0.25% (w/v) agar as previously described [27]. Test tubes were incubated with constant illumination of medium light (15 watt/$m^2$) for at least 4 - 5 days. Alternatively, 3 µl samples from cultures grown to stationary phase were placed on the surface of swarm plates and incubated at 30˚C aerobically in the dark [28]. Ring formations were monitored after several days.

### 2.7. Phylogenetic Analysis

Protein sequences of CtrA homologs from 40 different *Proteobacteria* species were obtained from NCBI. Protein sequences were aligned using MUSCLE [29] and phylogenetic analyses were performed using PhyML [30], under the Jones-Taylor-Thornton (JTT) model [31] with an estimated transition/transversion ratio and estimated proportion of invariable sites. The number of bootstrap replications performed was 300. The workflow was carried out in Phylogeny.fr (Dereeper *et al.*, 2008). The phylogenetic tree output was treated as unrooted.

### 2.8. DNA Motif Search

The MEME program [32], a model based on statistical expectation maximization, was used to identify the CtrA binding sites within the promoter regions or upstream regions of genes identified and confirmed as being part of the CtrA regulons in *C. crescentus* [5]. These predicted consensus sequence that binds CtrA in *C. crescentus* has been identified as TTAA-$N_7$-TTAAC (gapped) and TTAACCAT (ungapped) [5], and these motifs were then compared with the experimentally validated CtrA binding motifs; the best match was subsequently used to scan the noncoding sequences in both chromosomes of *R. sphaeroides* for CtrA-binding motifs using the MAST program [33]. All DNA sequences and their corresponding protein sequences as well as upstream and downstream sequences for species examined in this study were obtained from the NCBI Database (http://www.ncbi.nlm.nih.gov/).

### 2.9. Cluster of Orthologous Groups (COGs) Analysis

The Cluster of Orthologous Groups [34] [35], a classification in the NCBI database provides a tool in examining gene roles. There are four major COG functions, which include 1) Information storage and Processing, 2) Cellular Processes, 3) Metabolism, 4) Poorly Characterized functions. These major groupings were further classified into 25 sub-groups (detailed in **Figure 1** legend). COG functions for *C. crescentus*, *R. capsulatus*, and *R. sphaeroides* *ctrA*-regulated genes were subsequently analyzed for differences and similarities in *ctrA* control across these species. The *ctrA*-regulated genes ascertained in *C.*

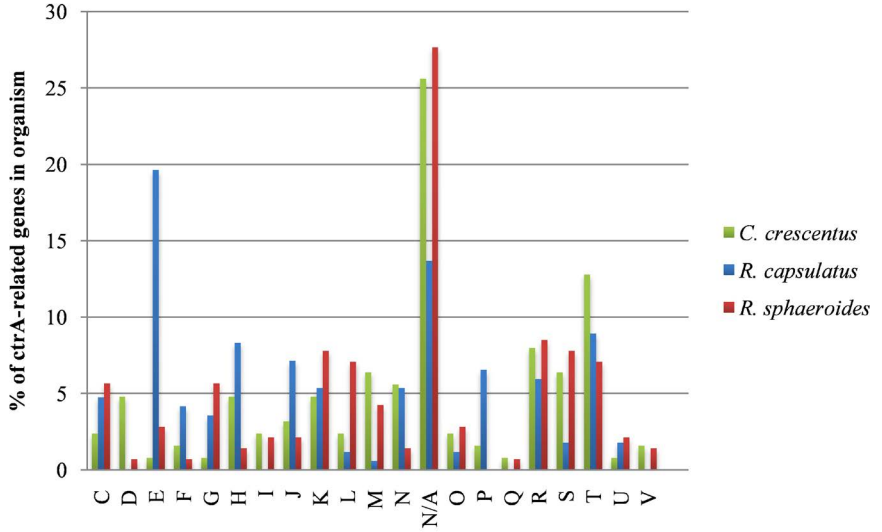



**Figure 1.** Cluster of orthologous group (COG) classification of genes in *Caulobacter crescentus*, *Rhodobacter capsulatus*, and *Rhodobacter sphaeroides* that are *ctrA*-associated. Data for *C. crescentus* and *R. capsulatus* is taken from previous studies [11] [27]. As shown by the distribution of the COGs, there is a wide variability of the clustering of the functions for the *ctrA*-associated genes across the three organisms. COG group values are as follows: A. RNA processing and modification; B. Chromatin structure and dynamics; C. Energy production and conversion; D. Cell division and chromosome partitioning; E. Amino acid transport and metabolism; F. Nucleotide transport and metabolism; G. Carbohydrate transport and metabolism; H. Coenzyme metabolism; I. Lipid metabolism; J. Translation, ribosomal structure and biogenesis; K. Transcription; L. DNA replication, recombination and repair; M. Cell envelope biogenesis, outer membrane; N. Cell motility and secretion; N/A. Not in COGs; O. Posttranslational modification, protein turnover, chaperones; P. Inorganic ion transport and metabolism; Q. Secondary metabolites biosynthesis, transport and catabolism; R. General function prediction only; S. Function unknown; T. Signal transduction mechanisms; U. Intracellular trafficking and secretion; V. Defense mechanisms; Z. Cytoskeleton.

*crescentus* and *R. capsulatus* were garnered from previous studies [5] [12], although these studies had different methodologies in comparison to the present one in determining the association of genes to *ctrA*.

### 2.10. Protein Homology Search

A protein homology search was performed using a gapped BLASTP [36] Genes with *ctrA* binding sites in *R. sphaeroides* were compared against the genomes of *R. capsulatus* and *C. crescentus*. The matches were then analyzed to see whether the *ctrA*-regulated genes of *R. sphaeroides* matched to the *ctrA*-regulated genes of those organisms.

## 3. Results and Discussion

### 3.1. Molecular Analysis of the *CtrA*-Null Mutant

The *ctrA*-null strain of *R. sphaeroides* was kanamycin-sensitive; this sensitivity of the mutant strains was checked repeatedly on both SIS plates and in SIS liquid

culture with kanamycin. Km$^s$ strains selected could be either *ctrA*-null mutants or wild type. To select for mutants, colony PCR was performed to distinguish mutants from wild type, as shown in **Figure 2**. When the primer *ctrA*-upstream F1 located on upstream of the *ctrA* gene and primer *ctrA*-qPCR-1 R2 located inside the deleted *ctrA* coding region was used to amplify the genomic DNA of mutants and wild type cells, it was predicted that wild type genomic DNA should amplify a 918-bp PCR product while the genomic DNA from mutant strains should not amplify any PCR product since *ctrA*-qPCR R2 primer binding site was missing in mutant strains. As shown in **Figure 2(a)**, no PCR product was detected for the mutant strain, while the wild type produced a ~1000 bp PCR product. This result was consistent with the prediction that wild type *ctrA* would be replaced by the deleted *ctrA* allele. Genomic DNA of the false positive strains derived from sucrose plates produced the same PCR product as wild type genomic DNA. Alternatively, using sequencing primers F1 and R2 that are located on 181-bp upstream and 164-bp downstream regions, respectively, a much smaller fragment (387 bp) was predicted to be amplified using mutant genomic DNA as template, compared with a 1059-bp PCR product amplified using wild type DNA as shown in **Figure 2(b)**. The sequencing of PCR products using this set of primers further confirmed the deletion of *ctrA* in this mutant strain. On the other hand, both fragments were amplified in the Km$^r$ strains since the wild-type *ctrA* and ctrA-null alleles were present in the chromosome. Sequencing the ctrA target from the null strain confirmed its integrity and this strain was used for further analysis.

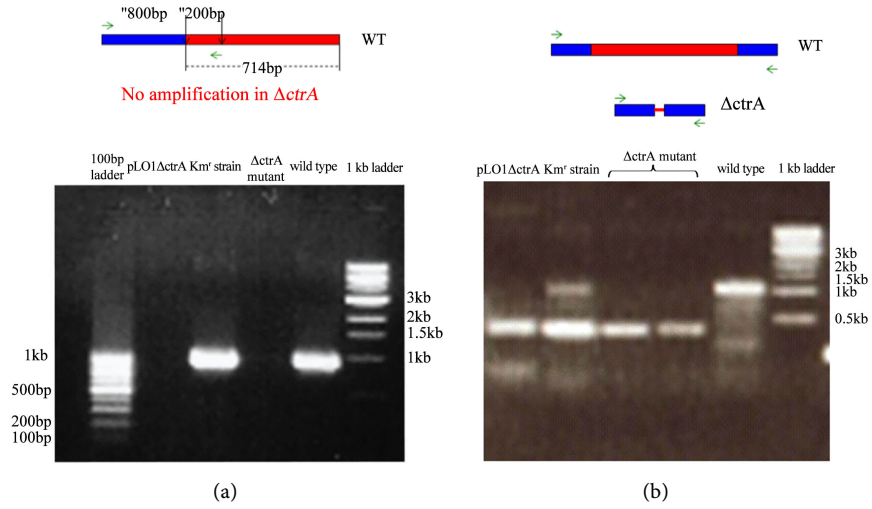



**Figure 2.** Analyses of PCR products of the *ctrA*-null mutant: The red segment refers to the *ctrA* gene, the blue segments refer to the gene flanking sequences, and the green arrows refer to the primers listed in **Table 1** designed from these locations. (a) Analysis of PCR products using primers (*ctrA*-2F and *ctrA*-2R) as shown in top panel. The reverse primer in this case is in the deletion segment; (b) Analysis of PCR products derived from reactions using primers (*ctrA*-upstream F1 and *ctrA*-downstream R4) as shown in top panel show that discrete PCR products can be amplified from wild type, *ctrA* deletion strains, and the cloned *ctrA* allele into pLO1 plasmid.

### 3.2. Loss of CtrA Does Not Affect Cell Growth

Growth curves for both wild type and *ctrA*-null mutant strains under both aerobic and photosynthetic conditions (10 watts/$m^2$) revealed no significant differences as shown in **Figure 3**. In addition, wild type and *ctrA*-null mutant strains of *R. sphaeroides* grown under these conditions displayed no differences in colony morphology or pigmentation characteristics examined under light microscopy as shown in **Figure 4**. Notably though, although both the wild type and the *ctrA*-null mutant strains displayed slow growth under the anaerobic-dark-DMSO condition, the *ctrA*-null mutant grew noticeably slower (approximately 1.5 times slower) and its doubling time was noticeably increased as well compared to the wild-type strain (data not shown).

These findings demonstrate that CtrA is not essential for cell growth and cell cycle regulation in *R. sphaeroides*. Although these findings seem similar to the observations previously reported in *R. capsulatus*, a bacterium closely related to *R. sphaeroides*, it seems that *ctrA* may play a more important role in growth under certain conditions (such as the anaerobic-dark-DMSO condition) in *R. sphaeroides* that was not seen in *R. capsulatus*. However, the role of *ctrA* still does not seem to be "essential" for cell cycle regulation in *R. sphaeroides.* These results support the hypothesis that symmetrically dividing species in the *α-Proteobacteria* group undergo cell division using a different yet unknown regulatory system from asymmetrically dividing species such as *C. crescentus* [37] [38].

### 3.3. Weak Motilities in Both Wild Type and the *CtrA*-Null Mutant Strains

The loss of *ctrA* has been shown to impair motility in *C. crescentus* as this gene positively regulates the expression of genes involved in flagellar assembly [2]. Additionally, *ctrA* has been shown to be possibly involved in motility and flagellar functions in *R. capsulatus* via motility assays [15] and via microarray expression analysis [12]. Since *R. sphaeroides* is closely related to *R. capsulatus*, it was originally hypothesized that *ctrA* may play an essential role in flagellar motility in *R. sphaeroides* as well.

*R. sphaeroides* has two distinct flagellar systems and its motility is controlled by approximately 30 proteins encoded by two different gene clusters (fla1 and fla2) on the primary chromosome. In general, polar flagellum in *R. sphaeroides* has been shown to be functional in liquid medium and motility can be observed in denser environments through the production of a large number of lateral flagella [27]. More specifically, the first flagellar system (fla1) of *R. sphaeroides* produces polar flagellum for swimming under both aerobic and photosynthetic growth conditions, while the expression of the second set of flagellar genes (fla2) seems to be positively regulated under anaerobic growth conditions and produces lateral flagella for swarming motility. The fla2 gene cluster is native to *R. sphaeroides* and does not code for the entire flagellum without the cooperation of some gene products from the fla1 gene cluster [27].

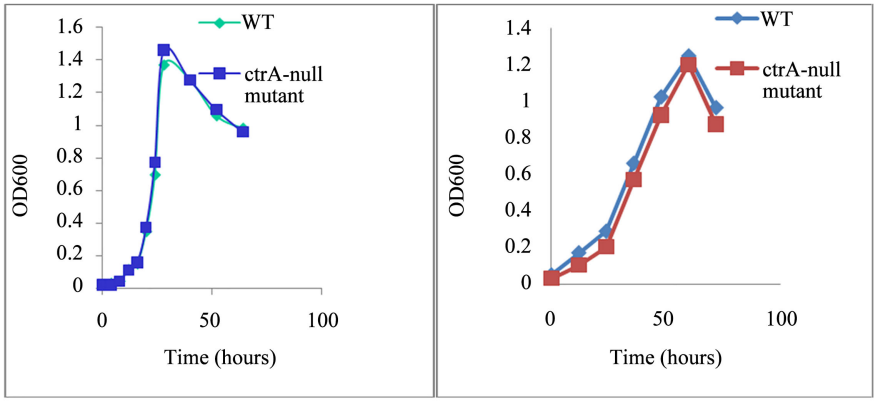



**Figure 3.** Growth curves for both wild type and *ctrA*-null mutant strains under aerobic and photosynthetic growth conditions are shown on the left and the right, respectively. The results reveal no significant differences in growth rates between the wild type and *ctrA*-null mutant strains under both these conditions.

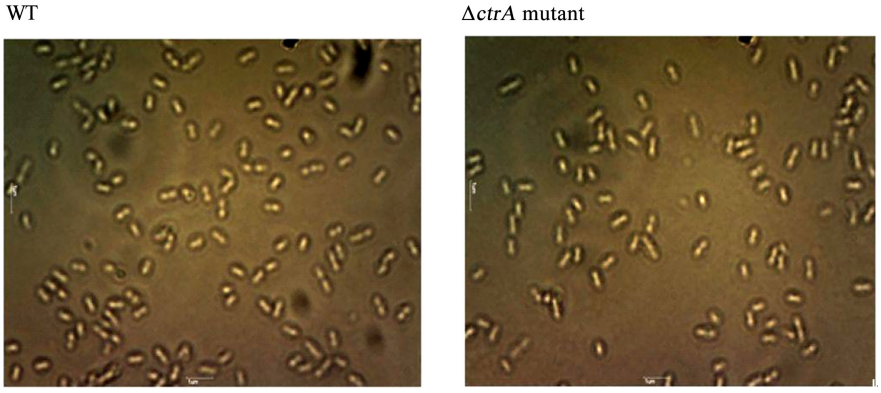
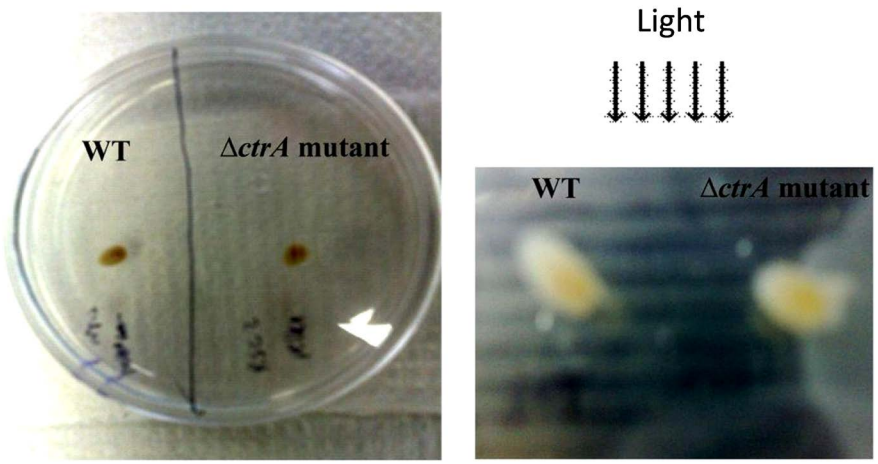



**Figure 4.** Cell morphologies of wild type (left) and *ctrA*-null mutant (right) of *R. sphaeroides* cells examined by light microscopy at 1000×.

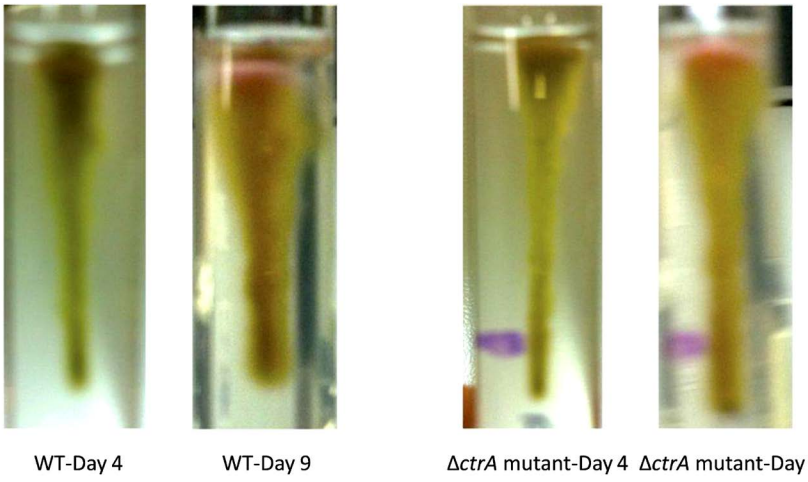
The results of this study revealed no significant change in swimming and swarming motility due to loss of *ctrA* in *R. sphaeroides*. As shown in **Figure 5(a)**, both wild type and mutant cells indicated very weak motilities during aerobic growth (dark) with a very small ring formation. No significant difference was observed. Furthermore, very weak phototaxis may have occurred in both strains when the illumination was provided from only one side while holding other side in the dark (**Figure 5(b)**). Results shown in soft agar test tubes were consistent with results on swarm plate as shown in **Figure 6**.

The wild type *R. sphaeroides* 2.4.1 has poor motility compared to another *R. sphaeroides* strain, WS8, which is usually used as a model strain for flagellar motility and chemotaxis studies [27] [39]. As such, even though no significant change of motility was observed due to loss of CtrA in *R. sphaeroides* 2.4.1, it cannot be definitively concluded that CtrA is not involved in flagellar motility. For instance, it is quite possible that the genes from the fla2 gene cluster compensate for the impairment of the gene regulation in fla1 genes that could be responsible for flagellum formation. Moreover, since the promoter of the fla1 operon




(a) (b)

**Figure 5.** Swimming assays of wild type and *ctrA*-null mutant strains: (a) Swimming patterns of wild type (left spot) and *ctrA*-null mutants (right spot) grown aerobically in the dark; (b) Swimming patterns of wild type (left spot) and *ctrA*-null mutants (right spot) grown with illumination from only one side as indicated by arrows. Pictures were taken 5 days after stabbing cells into soft agar plates.




(a) (b)

**Figure 6.** Motilities of wild type and *ctrA*-null mutant strains: (a) Motility of wild type cells (left panel) and mutant cells (right panel) grown photosynthetically after 4 and 9 days, respectively; (b) Motility of wild type cells (left panel) and mutant cells (right panel) grown in the dark for 10 days.

(between RSP_1312 and RSP_1313) contains a half-CtrA binding site, expressions of thesegenes could be modulated by a CtrA regulator. In short, the native fla2 genes could have allowed for normal or near-normal flagellum formation in *R. sphaeroides* 2.4.1 and therefore no defect in the motility phenotype was observed. Thus, CtrA could be involved in yet another unknown pathway in *R. sphaeroides* especially since the CtrA proteins have diverged across *α-Proteobacteria* species.

### 3.4. Analysis of *CtrA* Gene Expression

The mRNA expression levels of *ctrA* during different growth phases, normalized with the expression of a reference gene (*rpoZ*), are shown in Table 2. There were slight differences in *ctrA* expressions related to different growth stages; however, differences of relative expression levels were not statistically significant (one-way ANOVA, $p = 0.344$), which was similar to the results previously shown in *R. capsulatus* [12]. The expression pattern demonstrated that *ctrA* is expressed independently of cell densities in the batch culture. The effect of the deletion of *ctrA* was further confirmed on a transcriptional level. The mutant strain yielded 10,000-fold (virtually undetectable) lower level mRNA product compared to wild type cells. The decrease in relative mRNA expression levels was statistically significant (paired *t*-test, $p = 0.004$; Table 2). Moreover, the melt curves show a single sharp peak generated by all three sets of primers specific to three different regions of the *ctrA* gene. The methods indicate high target specificity where the primers are likely binding to single locations in the genome and producing uniform products in wild type cells and mutant cells. Complementation analysis revealed ~30% lower expression of *ctrA* gene in the complemented strains compared to the expression level in the wild type (data not shown). The initial over-expression of CtrA from plasmids may have possibly repressed its own expression indicative of the autorepression of the *ctrA* gene.

### 3.5. Phylogenetic Analysis Revealed Diversified Role of CtrA across *α-Proteobacteria*

Sequence alignments revealed that the CtrA protein is highly conserved among *α-Proteobacteria*. CtrA in *R. sphaeroides* shares 92.4% and 72.2% amino acid identities with CtrA in *R. capsulatus* and *C. crescentus*, respectively. All the CtrA homologs among the species have a conserved residue for a putative phosphorylation site (aspartic acid) at position 51. A phylogenetic tree of CtrA homologs for 40 *α-Proteobacteria* species is shown in Figure 7 and reveals that the CtrA homolog of *R. sphaeroides* is most closely related to that of *R. capsulatus*.

The non-essential function of CtrA has been observed in *R. capsulatus*, *Rugeria sp.* TM1040, and now in *R. sphaeroides*. Even though CtrA is highly conserved among species of the subgroup, the phylogenetic analysis demonstrated that these three species where CtrA exhibits non-essential functions are a very closely related clade as shown in Figure 7. Furthermore, these three species are separated from other species like *C. crescentus*, which exhibits asymmetric cell division. These results are in line with previous findings [37].

### 3.6. CtrA Promoter Analysis

A schematic diagram shown in Figure 8 demonstrates a variety of upstream flanking regions among *α-Proteobacteria* species. Species undergoing asymmetric cell division have significantly longer flanking upstream regions (197 bp to 550 bp), indicating a larger and more complex promoter region, while in species such as *R. sphaeroides*, *R. capsulatus* and *Rugeria sp.* TM1040 that divide sym-

metrically, relatively smaller promoter regions were found (95 bp to 125 bp). Furthermore, putative CtrA binding sites were found in the upstream flanking region of *ctrA* in *R. sphaeroides*, suggesting the possibility that CtrA is auto-regulated, similar to the findings in *C. crescentus* [8].

### 3.7. Putative CtrA Binding Sites Reveal Diverged Gene Regulatory Functions of CtrA

As previously described, the DNA consensus sequence that binds CtrA in *C. crescentus* has been identified as TTAA-$N_7$-TTAAC (gapped) and TTAACCAT

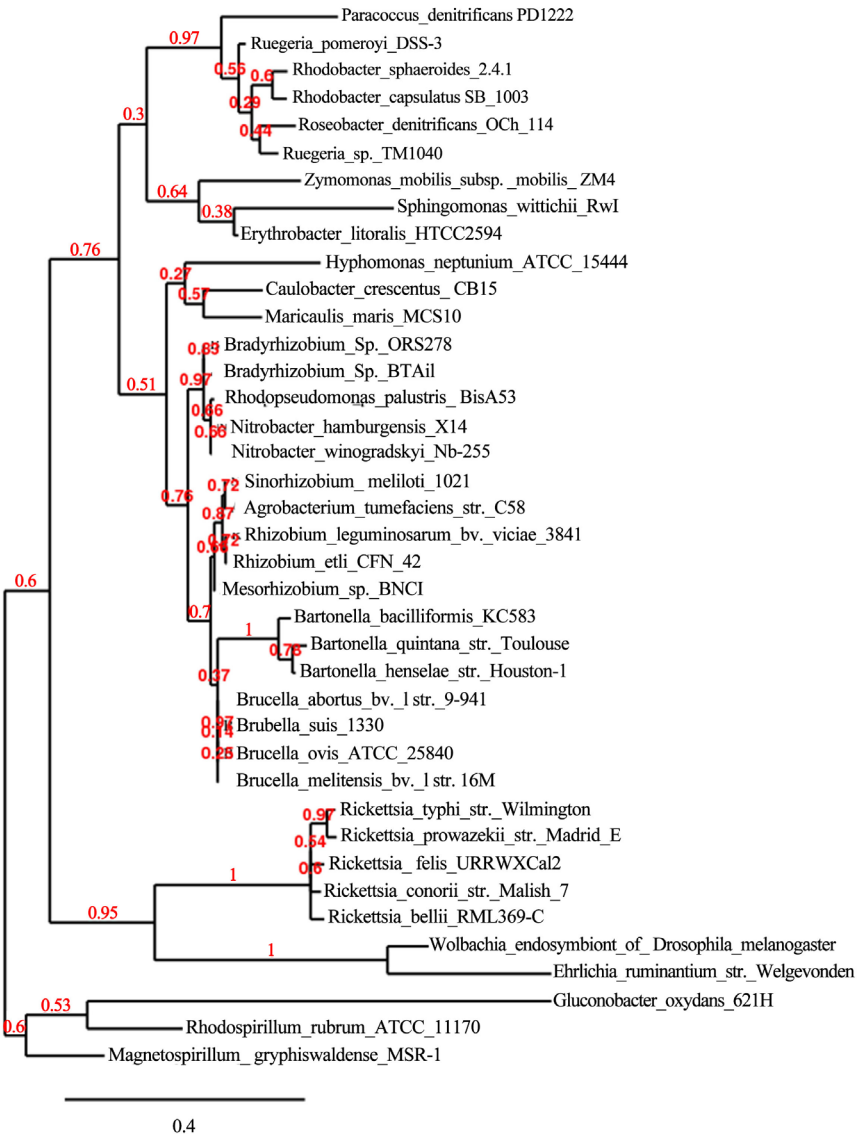



**Figure 7.** A phylogenetic tree of the *ctrA* gene that represents 40 *Proteobacteria* species: The scale bar at the bottom of the figure represents the substitutions per site while the numbers over the branches represent the bootstrap support values for those nodes. Bootstrap is given as a fraction of the number 1, with 1 being the strongest bootstrap support attainable for a given node. The bootstrap values were calculated using 300 replications.

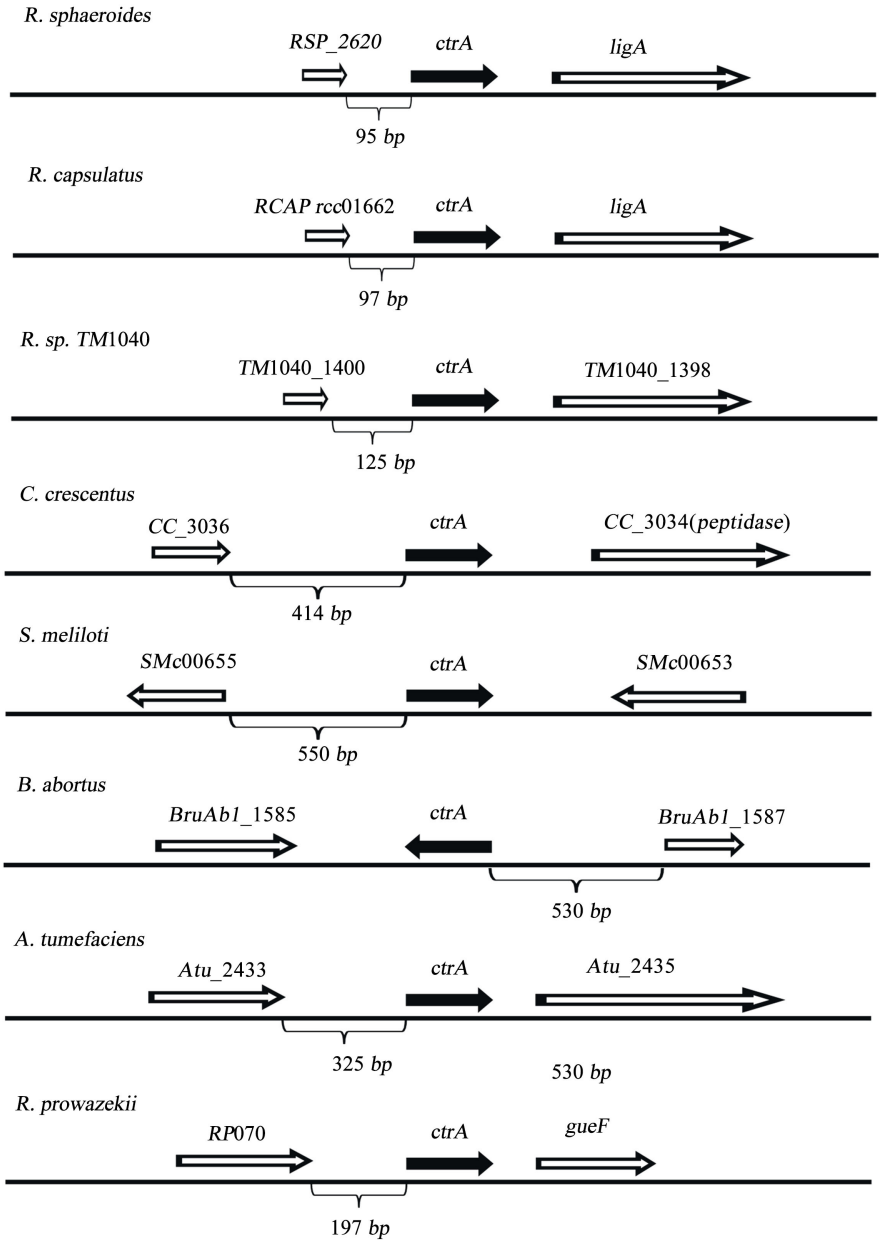



**Figure 8.** A schematic diagram depicts the alignment of *ctrA* and its upstream and downstream genes in eight different α-proteobacterial species. Species with asymmetric cell division, such as *A. tumefaciens*, *B. abortus*, *C. crescentus*, *R. prowazekii*, and *S. meliloti*, have longer upstream regions compared to species that divide symmetrically such as *R. sphaeroides*, *R. capsulatus*, and *Rugeria sp.* TM1040.

(ungapped) [5]. Using upstream sequences of 53 genes that were previously identified as being directly controlled by CtrA [5], the MEME program generated more restrictive DNA binding sequences for CtrA (as referred to above). Using the binding motif generated and validated by the MEME program, potential CtrA binding sites were identified by scanning the consensus binding motif against the *R. sphaeroides* genome and the results are shown in supplementary **Table 1**. A total of 127 genes distributed on both chromosomes of *R. sphaeroides*

were identified to have putative CtrA binding sites in their promoter regions. These genes are involved in diverse functions such as in energy production, iron utilization, transport function, and transcriptional regulation.

CtrA binding site search analysis in *R. sphaeroides* also revealed conservation of the CtrA regulatory circuit compared with *C. crescentus*. For example, putative binding sites were found in both *ctrA* and *cckA*, indicating the possibility of the autoregulation of CtrA and direct control of CckA. The major difference between CtrA of *R. capsulatus* and *C. crescentus* was the lack of CtrA control on genes involving cell division, DNA methylation, and energy production in *R. capsulatus* [12]. However, potential CtrA binding sites were found in promoter regions of genes displaying functions in the above pathways, such as NADH dehydrogenase gene clusters and SAM-dependent methyltransferase. Notably, not all intergenic regions with binding motifs bind with CtrA *in vivo* [5].

Protein homology searches revealed that only a small number of *ctrA*-regulated genes are common across *C. crescentus*, *R. capsulatus*, and *R. sphaeroides*. **Table 3** details the findings of these comparisons. More specifically, for *R. sphaeroides ctrA*-regulated genes, less than 20% of them had matches to *C. crescentus* and *R. capsulatus ctrA*-regulated genes. When filtered for an e-value < $10^{-5}$, only ~6% of the *R. sphaeroides* genes retained their matches to *R. capsulatus* and *C. crescentus ctrA*-regulated genes. The findings were similar in a comparison from *R. capsulatus* to *C. crescentus* as well. These results suggest a significant variability in *ctrA* gene control in these organisms and could partly explain the differences in the functions and essentiality of *ctrA* across these species.

### 3.8. Functions of *CtrA*-Regulated Genes across Species

**Figure 1** displays the COG group classification of *ctrA*-regulated genes in *C. crescentus*, *R. capsulatus*, and *R. sphaeroides* (data for which is available in Additional Files 2, 3, and 4). Putative CtrA-regulated genes in *R. capsulatus* [12] and *C. crescentus* [5] were taken from data collected in previous studies, although it must be noted that these studies had different methodologies in ascertaining how genes were regulated or controlled by *ctrA* in comparison to the present one. The number of *ctrA*-regulated genes utilized for analysis for *R. sphaeroides*, *C. crescentus*, and *R. capsulatus* are 127, 116, and 143 genes, respectively, and data concerning the functions of genes can be found in Additional Files 2, 3, and 4, respectively. As evidenced by **Figure 1**, there is a wide variability of the cluster of gene functions that are *ctrA*-regulated across the organisms but no specific pattern of control is apparent among the three, which is expected since the function of *ctrA* seems to differ among them. In *C. crescentus*, where *ctrA* is essential, there are three COG groups with a greater proportion of genes when compared to *R. capsulatus* or *R. sphaeroides*: D (cell division and chromosome partitioning), M (cell envelope biogenesis, outer membrane), and T (signal transduction mechanisms). Actually, *R. capsulatus* and *R. sphaeroides* possessed almost no genes in COG group D, which constitutes genes directly

**Table 3.** Number of matches for *ctrA*-associated genes between species.

| Query Organism (QO) | *R. sphaeroides* | *R. sphaeroides* | *R. capsulatus* |
|---|---|---|---|
| **# of *ctrA*-associated genes in QO** | 127 | 127 | 143 |
| **Comparison Organism (CO)** | *R. capsulatus* | *C. crescentus* | *C. crescentus* |
| **# of *ctrA*-associated genes in CO** | 143 | 116 | 116 |
| **Matches to *ctrA*-associated genes in CO** | 25 | 20 | 43 |
| **% of QO Total** | 19.7% | 15.7% | 30.1% |
| **Matches with E-Value < $10^{-5}$** | 8 | 8 | 25 |
| **% of QO Total** | 6.3% | 6.3% | 17.5% |

participating in the cell cycle. However, outside of these three COG groups, there was no other pattern that the ctrA in *C. crescentus* seemed to exert that illustrated a drastic difference in cellular regulation compared to *R. capsulatus* and *R. sphaeroides*. Thus, the essentiality of CtrA in *C. crescentus* may depend more on selective control that it exerts on a few genes and pathways that are not controlled by CtrA in *R. capsulatus* or *R. sphaeroides*. This is logical as all three seem to possess extensive and broad pathway associations across a variety of classifications.

## 4. Conclusion

In conclusion, the loss of CtrA in *R. sphaeroides* clearly demonstrates that CtrA was not essential for cell viability and was likely not required for major aerobic and photosynthetic growth. However, it is quite possible that CtrA may control unique functions yet unidentified in *R. sphaeroides*. As only less than 20% of *ctrA*-regulated genes in *R. sphaeroides* showed homology to *ctrA*-regulated genes in *C. crescentus* and *R. capsulatus*, it is conceivable that CtrA may have diversified its role in different evolutionary lineages. This is further confirmed by the finding that the distribution of the COG functions of *ctrA*-regulated genes differs widely across *C. crescentus*, *R. sphaeroides*, and *R. capsulatus*. However, *ctrA* of all three species seemed to exhibit broad pathway control overall. Therefore, this result also suggests that the cell cycle essentiality of CtrA in *C. crescentus* and other asymmetrically dividing species may depend more on selective control that it exerts on a few critical genes and pathways that are not controlled by CtrA in symmetrically dividing species such as *R. capsulatus* and *R. sphaeroides*.

## Acknowledgements

This work was supported by the Enhancement Grant for Research (EGR) from Sam Houston State University to Madhusudan Choudhary.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this pa-

per.

## References


[1] Huang, K.J. and Igo, M.M. (1996) Identification of the Bases in the OmpF Regulatory Region, Which Interact with the Transcription Factor OmpR. *Journal of Molecular Biology*, **262**, 615-628. https://doi.org/10.1006/jmbi.1996.0540

[2] Quon, K.C., Marczynski, G.T. and Shapiro, L. (1996) Cell Cycle Control by an Essential Bacterial Two-Component Signal Transduction Protein. *Cell*, **84**, 83-93. https://doi.org/10.1016/S0092-8674(00)80995-2

[3] Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., *et al.* (2005) Genome Streamlining in a Cosmopolitan Oceanic Bacterium. *Science*, **309**, 1242-1245. https://doi.org/10.1126/science.1114057

[4] Barnett, M.J., Hung, D.Y., Reisenauer, A., Shapiro, L. and Long, S.R. (2001) A Homolog of the CtrA Cell Cycle Regulator Is Present and Essential in *Sinorhizobium meliloti*. *Journal of Bacteriology*, **183**, 3204-3210. https://doi.org/10.1128/JB.183.10.3204-3210.2001

[5] 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. *Proceedings of the National Academy of Sciences of USA*, **99**, 4632-4637. https://doi.org/10.1073/pnas.062065699

[6] Laub, M.T., McAdams, H.H., Feldblyum, T., Fraser, C.M. and Shapiro, L. (2000) Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle. *Science*, **290**, 2144-2148. https://doi.org/10.1126/science.290.5499.2144

[7] Holtzendorff, J., Reinhardt, J. and Viollier, P.H. (2006) Cell Cycle Control by Oscillating Regulatory Proteins in *Caulobacter crescentus*. *BioEssays*, **28**, 355-361. https://doi.org/10.1002/bies.20384

[8] Domian, I.J., Reisenauer, A. and Shapiro, L. (1999) Feedback Control of a Master Bacterial Cell-Cycle Regulator. *Proceedings of the National Academy of Sciences of USA*, **96**, 6648-6653. https://doi.org/10.1073/pnas.96.12.6648

[9] Bellefontaine, A.F., Pierreux, C.E., Mertens, P., Vandenhaute, J., Letesson, J.J. and De Bolle, X. (2002) Plasticity of a Transcriptional Regulation Network among Alpha-Proteobacteria Is Supported by the Identification of CtrA Targets in *Brucella abortus*. *Molecular Microbiology*, **43**, 945-960. https://doi.org/10.1046/j.1365-2958.2002.02777.x

[10] McAdams, H.H., Srinivasan, B. and Arkin, A.P. (2004) The Evolution of Genetic Regulatory Systems in Bacteria. *Nature Reviews Genetics*, **5**, 169-178. https://doi.org/10.1038/nrg1292

[11] Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H. and Kurland, C.G. (1998) The Genome Sequence of *Rickettsia prowazekii* and the Origin of Mitochondria. *Nature*, **396**, 133-140. https://doi.org/10.1038/24094

[12] Mercer, R.G., Callister, S.J., Lipton, M.S., Pasa-Tolic, L., Strnad, H., Paces, V., Beatty, J.T. and Lang, A.S. (2010) Loss of the Response Regulator CtrA Causes Pleiotropic Effects on Gene Expression But Does Not Affect Growth Phase Regulation in *Rhodobacter capsulatus*. *Journal of Bacteriology*, **192**, 2701-2710. https://doi.org/10.1128/JB.00160-10

[13] Lang, A.S. and Beatty, J.T. (2000) Genetic Analysis of a Bacterial Genetic Exchange

Element: The Gene Transfer Agent of *Rhodobacter capsulatus*. *Proceedings of the National Academy of Sciences of USA*, **97**, 859-864. https://doi.org/10.1073/pnas.97.2.859

[14] Miller, T.R. and Belas, R. (2006) Motility Is Involved in *Silicibacter sp*. TM1040 Interaction with Dinoflagellates. *Environmental Microbiology*, **8**, 1648-1659. https://doi.org/10.1111/j.1462-2920.2006.01071.x

[15] Lang, A.S. and Beatty, J.T. (2002) A Bacterial Signal Transduction System Controls Genetic Exchange and Motility. *Journal of Bacteriology*, **184**, 913-918. https://doi.org/10.1128/jb.184.4.913-918.2002

[16] Kiley, P.J. and Kaplan, S. (1988) Molecular Genetics of Photosynthetic Membrane Biosynthesis in *Rhodobacter sphaeroides*. *Microbiological Reviews*, **52**, 50-69.

[17] Suwanto, A. and Kaplan, S. (1989) Physical and Genetic Mapping of the *Rhodobacter sphaeroides* 2.4.1 Genome: Presence of Two Unique Circular Chromosomes. *Journal of Bacteriology*, **171**, 5850-5859. https://doi.org/10.1128/jb.171.11.5850-5859.1989

[18] Sistrom, W.R. (1960) A Requirement for Sodium in the Growth of *Rhodopseudomonas spheroides*. *Journal of General Microbiology*, **22**, 778-785. https://doi.org/10.1099/00221287-22-3-778

[19] Oh, J.I. and Kaplan, S. (1999) The cbb3 Terminal Oxidase of *Rhodobacter sphaeroides* 2.4.1: Structural and Functional Implications for the Regulation of Spectral Complex Formation. *Biochem*, **38**, 2688-2696. https://doi.org/10.1021/bi9825100

[20] Lenz, O., Schwartz, E., Dernedde, J., Eitinger, M. and Friedrich, B. (1994) The *Alcaligenes eutrophus* H16 hoxX Gene Participates in Hydrogenase Regulation. *Journal of Bacteriology*, **176**, 4385-4393. https://doi.org/10.1128/jb.176.14.4385-4393.1994

[21] Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Treboul, G. and Gay, P. (1985) The DNA Sequence for the Secreted *Bacillus subtilis* Enzyme Levansucrase. *Molecular Genetics and Genomics*, **200**, 220-228. https://doi.org/10.1007/BF00425427

[22] Davis, J., Donohue, T.J. and Kaplan, S. (1988) Construction, Characterization, and Complementation of Puf Mutant of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, **170**, 320-329. https://doi.org/10.1128/jb.170.1.320-329.1988

[23] Tai, T.N., Moore, M.D. and Kaplan, S. (1988) Cloning and Characterization of the 5-Aminolevulinic Synthase Gene(s) from *Rhodobacter sphaeroides*. *Gene*, **70**, 139-151. https://doi.org/10.1016/0378-1119(88)90112-6

[24] Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T. and Kado, C.I. (1985) Positive Selection Procedure for Entrapment of Insertion Sequence Elements in Gram-Negative Bacteria. *Journal of Bacteriology*, **164**, 918-921.

[25] Livak, K.J. and Schmittgen, T.D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408. https://doi.org/10.1006/meth.2001.1262

[26] Sockett, R.E., Armitage, J.P. and Evans, M.C. (1987) Methylation-Independent and Methylation-Dependent Chemotaxis in *Rhodobacter sphaeroides* and *Rhodospirillum rubrum*. *Journal of Bacteriology*, **169**, 5808-5814. https://doi.org/10.1128/jb.169.12.5808-5814.1987

[27] Poggio, S., Abreu-Goodger, C., Fabela, S., Osorio, A., Dreyfus, G., Vinuesa, P. and Camarena, L. (2007) A Complete Set of Flagellar Genes Acquired by Horizontal Transfer Coexists with the Endogenous Flagellar System in *Rhodobacter sphaeroides*. *Journal of Bacteriology*, **189**, 3208-3216. https://doi.org/10.1128/JB.01681-06

[28] De la Mora, J., Ballado, T., Gonzalez-Pedrajo, B., Camarena, L. and Dreyfus, G.

(2007) The Flagellar Muramidase from the Photosynthetic Bacterium *Rhodobacter sphaeroides*. *Journal of Bacteriology*, **189**, 7998-8004. https://doi.org/10.1128/JB.01073-07

[29] Edgar, R.C. (2004) MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Research*, **32**, 1792-1797. https://doi.org/10.1093/nar/gkh340

[30] Guindon, S. and Gascuel, O. (2003) A Simple, Fast and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, **52**, 696-704. https://doi.org/10.1080/10635150390235520

[31] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) The Rapid Generation of Mutation Data Matrices from Protein Sequences. *Computer Applications in the Biosciences*, **8**, 275-282.

[32] Bailey, T.L. and Elkan, C. (1994) Fitting a Mixture Model by Expectation Maximization to Discover Motifs in Biopolymers. In: *Proceedings of the 2nd International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, Menlo Park, 28-36.

[33] Bailey, T.L. and Gribskov, M. (1998) Combining Evidence Using P-Values: Application to Sequence Homology Searches. *Bioinformatics*, **14**, 48-54. https://doi.org/10.1093/bioinformatics/14.1.48

[34] Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., *et al.* (2003) The COG Database: An Updated Version Includes Eukaryotes. *BMC Bioinformatics*, **4**, 41. https://doi.org/10.1186/1471-2105-4-41

[35] Tatusov, R.L., Koonin, E.V. and Lipman, D.J. (1997) A Genomic Perspective on Protein Families. *Science*, **278**, 631-637. https://doi.org/10.1126/science.278.5338.631

[36] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Research*, **25**, 3389-3402. https://doi.org/10.1093/nar/25.17.3389

[37] Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M. and Biondi, E.G. (2010) The Diversity and Evolution of Cell Cycle Regulation in Alpha-Proteobacteria: A Comparative Genomic Analysis. *BMC Systematic Biology*, **4**, 52. https://doi.org/10.1186/1752-0509-4-52

[38] Hallez, R., Bellefontaine, A.F., Letesson, J.J. and De Bolle, X. (2004) Morphological and Functional Asymmetry in Alpha-Proteobacteria. *Trends in Microbiology*, **12**, 361-365. https://doi.org/10.1016/j.tim.2004.06.002

[39] Sockett, R.E., Foster, J.C.A. and Armitage, J.P. (1990) Molecular Biology of the *Rhodobacter sphaeroides* Flagellum. In: Drews, G. and Dawes, E.A., Eds., *Molecular Biology of Membrane-Bound Complexes in Phototrophic Bacteria*, Springer, Berlin, No. 53, 473-479.