A cryptic transcription factor regulates *Caulobacter* adhesin development

Maeve McLaughlin[1], David M. Hershey[2], Leila Reyes Ruiz[3], Aretha Fiebig[1], Sean Crosson[1]

[1]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, USA

[2]Department of Bacteriology, University of Wisconsin, Madison, WI, USA.

[3]Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC, USA.

Abstract

Alphaproteobacteria commonly produce an adhesin that is anchored to the exterior of the envelope at one cell pole. In *Caulobacter crescentus*, this adhesin enables permanent attachment to solid surfaces and is known as the holdfast. An ensemble of two-component signal transduction (TCS) proteins control *C. crescentus* holdfast biogenesis by indirectly regulating expression of HfiA, a potent inhibitor of holdfast synthesis. A genetic selection to discover direct *hfiA* regulators that function downstream of this adhesion TCS system identified a hypothetical gene that we have named *rtrC*. Though the primary structure of RtrC bears no resemblance to any defined protein family, RtrC directly binds and regulates dozens of sites on the *C. crescentus* chromosome via a pseudo-palindromic motif. Among these binding sites is the *hfiA* promoter, where RtrC functions to directly repress transcription and thereby activate holdfast development. RtrC, the DNA-binding response regulator SpdR, and the transcription factor RtrB together form an OR-gated type I coherent feedforward loop (C1-FFL) that regulates *hfiA* transcription. C1-FFL motifs are known to buffer gene expression against transient loss of regulating signals, which often occurs in fluctuating natural environments. We conclude that the formerly hypothetical gene, *rtrC*, encodes a transcription factor that functions downstream of the *C. crescentus* TCS adhesion control system to regulate development of the holdfast adhesin.

Introduction

The ability of microbial cells to adhere to surfaces and form biofilms is often a key determinant of fitness in both clinical and non-clinical contexts [1-3]. Colonization of substrates can support energy production [4], protect cells from toxic compounds [5, 6], and shield cells from grazing protist predators [7]. However, competition for resources in a multicellular biofilm can also slow growth; thus, there are evolutionary tradeoffs between surface attached and planktonic lifestyles [8]. Given that the fitness benefit of surface attachment varies as a function of environmental conditions, it follows that the cellular decision to adhere to a substrate is highly regulated.

Gram-negative bacteria of the genus *Caulobacter* are common in aquatic and soil ecosystems [9] and are dominant members of mixed biofilm communities in freshwater [10]. *Caulobacter* spp. often produce a secreted polar adhesin known as the holdfast, which enables high-affinity attachment to surfaces [11] and robust biofilm formation [12]. In the model *Caulobacter* species, *C. crescentus*, holdfast development is regulated at many levels. The transcription of holdfast synthesis genes exhibits periodic changes across the cell cycle [13, 14]. Holdfast biogenesis is also influenced by mechanical cues [15, 16], while the second messenger cyclic-di-GMP affects both synthesis [16] and physical properties of [17] the holdfast. Additionally, an elaborate regulatory pathway comprised of multiple two-component signaling (TCS) proteins and one-component regulators controls holdfast development and surface attachment [18]. We have previously shown that a *C. crescentus* strain expressing a non-phosphorylatable allele of the *lovK* sensor histidine kinase ($lovK_{H180A}$) overproduces holdfast and, consequently, has an enhanced adhesion phenotype in a static biofilm assay. The $lovK_{H180A}$ adhesion phenotype requires the *spdS-spdR* two-component system and the hybrid histidine kinase *skaH* [18]. Two XRE-family transcription factors, RtrA and RtrB, function downstream of the TCS regulators to promote holdfast synthesis by directly repressing transcription of the holdfast inhibitor, *hfiA* (Figure 1A). Though *rtrA* and *rtrB* clearly contribute to holdfast regulation downstream of the

adhesion TCS proteins, we hypothesized that there were additional regulators of *C. crescentus* holdfast biosynthesis in this pathway. Our hypothesis is based on the observation that deletion of both *rtrA* and *rtrB* does not completely abrogate holdfast synthesis when the TCS pathway is constitutively activated [18] (Figure 2A).

To search for these postulated downstream regulators, we used a high-throughput genetic approach to select for mutations that attenuate the hyper-holdfast phenotype of a $lovK_{H180A}$ mutant. Our selection uncovered a gene encoding a hypothetical protein that we have named RtrC, which functions as both a transcriptional activator and repressor in *C. crescentus*. RtrC binds a pseudo-palindromic DNA motif *in vivo* and *in vitro* and activates holdfast synthesis downstream of the *lovK-spdSR-skaH* TCS ensemble by directly repressing transcription of the holdfast inhibitor, *hfiA*. RtrC can also directly control the transcription of dozens of other genes in *C. crescentus* via this pseudo-palindromic site, including genes that impact flagellar motility, cyclic-di-GMP signaling, and aerobic respiration.

Results

A hypothetical protein functions downstream of the TCS regulators, LovK and SpdR, to activate holdfast synthesis

An ensemble of two-component signal transduction (TCS) proteins in *C. crescentus*, including LovK and SpdR, can control holdfast synthesis by regulating transcription of *hfiA*. A previous study identified two XRE-family transcription factors, RtrA and RtrB, that function downstream of this TCS system to repress *hfiA* and thereby activate holdfast synthesis [18] (Figure 1A). However, deleting *rtrA*, *rtrB,* or both [18] has only modest effects on holdfast synthesis when the TCS system is constitutively activated (Figure 2B). We therefore reasoned that there are additional downstream regulators in this pathway that can activate *C. crescentus* holdfast synthesis. To identify such genes, we constructed a randomly barcoded transposon mutant library in a *lovK* mutant background ($lovK_{H180A}$) in which holdfast synthesis is constitutively

activated. This barcoded library was cultivated and serially passaged in the presence of cheesecloth, a process that titrates adhesive cells from liquid medium as recently described [19]. Non-adhesive mutants become enriched in the media supernatant surrounding the cheesecloth, which is reflected as a positive fitness score when the total barcoded population is quantified (Figure 1B). Using this approach, we aimed to identify transposon insertions that ablated the hyper-holdfast phenotype of a mutant in which the adhesion pathway is constitutively active.

We expected that performing this genetic selection in a hyper-holdfast *lovK*$_{H180A}$ background would not only uncover previously identified loss-of-adhesion mutants [19] but would also identify new regulators that function to activate holdfast synthesis downstream of LovK. As expected, strains harboring transposon insertions in all the known adhesion TCS genes (e.g. *lovK*, *spdS*, *spdR* and *skaH*) had increased abundance in the supernatant (i.e. decreased adhesion to cheesecloth, and positive fitness scores) when grown in the presence of cheesecloth. Insertions in select polar development regulators, and in holdfast synthesis and anchoring genes also resulted in the expected positive fitness scores (Figures 1C and Table S1). Strains with insertions in gene locus *CCNA_00551*, which encodes a predicted 146-residue hypothetical protein, had strongly positive fitness scores after cheesecloth selection. In fact, strains with insertions in *CCNA_00551* were enriched in the supernatant to a greater extent than TCS adhesion mutants or *rtrA* and *rtrB* mutants (Figure 1C and Table S1). Consistent with these Tn-seq data, in-frame deletion of either *spdR* or *CCNA_00551* from the chromosome abrogated the hyper-holdfast phenotype of *lovK*$_{H180A}$ (Figure 2A). Expression of *CCNA_00551* is directly activated by the DNA-binding response regulator, SpdR [18, 20], which implicated *CCNA_00551* in the adhesion TCS pathway. Following the convention of previously named adhesion factors that function downstream of SpdR [18], we henceforth refer to *CCNA_00551* as *rtrC*.

SpdR functions downstream of LovK [18] (Figure 1A) and expression of a phosphomimetic allele of SpdR (SpdR$_{D64E}$) provides an alternative genetic approach to constitutively activate the *C. crescentus* adhesion TCS system. We predicted that deletion of *rtrC* would also abrogate the

hyperadhesive phenotype of a *spdR*$_{D64E}$ strain. Consistent with this prediction and with the Tn-seq data, we observed that the fraction of cells with visibly stained holdfasts was reduced in a *spdR*$_{D64E}$ Δ*rtrC* strain compared to the *spdR*$_{D64E}$ parent (Figure 2B). There was no significant difference in the percentage of cells with visibly stained holdfasts between *spdR*$_{D64E}$ Δ*rtrA* Δ*rtrB* Δ*rtrC* and *spdR*$_{D64E}$ Δ*rtrC* (Figure 2B). This provides evidence that RtrC is the primary downstream determinant of hyperadhesion when the TCS adhesion pathway is constitutively active. Indeed, overexpression of *rtrC* alone enhanced the fraction of cells with stained holdfasts more than overexpression of either *rtrA* or *rtrB* (Figure 2C).

RtrC is a predicted transcription factor

A search of protein domain family databases in InterPro [21] and the Conserved Domain Database [22] failed to identify conserved domains in RtrC. However, a primary and secondary structure profile matching approach [23] indicated that RtrC resembled classic transcription factors. To explore this possibility, we implemented AlphaFold [24] to predict the tertiary structure of RtrC. This approach predicted a fold that contained five α-helices (α1 – α5) and two β-strands (β1 – β2) that form an antiparallel β hairpin (Figure 3A). We compared this structure to the Protein Data Bank (PDB) using Dali [25], which revealed that the predicted structure of RtrC was most similar to MepR (PDB: 3ECO), a MarR-family transcriptional regulator from *Staphylococcus aureus* containing a winged helix-turn-helix motif [26]. Based on the structural alignments and 3D superposition with MepR, α1 and α5 of RtrC likely form a dimerization domain, while α2, α3, α4, β1, and β2 form a winged helix-turn-helix (Figure 3A and 3B). Considering these structural predictions, we hypothesized that *rtrC* encoded a transcription factor that functions downstream of the *C. crescentus* TCS adhesion regulatory system.

RtrC is a potent repressor of the holdfast inhibitor, hfiA

The transcription factors RtrA and RtrB are known to activate holdfast synthesis and adhesion by repressing transcription of the holdfast inhibitor, *hfiA* [18]. Given the correlated phenotypes of *rtrA*, *rtrB*, and *rtrC* mutants and the prediction that RtrC is a transcription factor (Figures 2 & 3), we hypothesized that RtrC functioned as a transcriptional repressor of *hfiA*. To test this model, we measured changes in expression from a fluorescent *hfiA* transcriptional reporter upon overexpression of *rtrC*. As expected, overexpression of *rtrA* and *rtrB* reduced signal from the P_{hfiA} fluorescent reporter by 80% and 30%, respectively. Overexpression of *rtrC* resulted in a 95% reduction in *hfiA* expression (Figure 2D).

RtrC binds to a pseudo-palindromic DNA motif in vivo and in vitro

We next sought to directly test the predicted DNA-binding function of RtrC. We performed chromatin immunoprecipitation sequencing (ChIP-seq) using a 3xFLAG-tagged *rtrC* allele and identified 113 statistically significant peaks across the genome (Table S3). Peaks were highly enriched near experimentally defined transcription start sites (TSS) [27, 28] when compared to a set of randomly generated peaks (Figure 4B); this TSS-proximal enrichment pattern is characteristic of proteins that directly bind DNA to regulate gene expression. To identify putative binding motifs in the ChIP-seq peaks, we analyzed the peak sequences using the XSTREME algorithm within the MEME Suite [29]. This revealed a pseudo-palindromic motif in 112 of the 113 *rtrC* peaks (E-value: $2.3e^{-12}$) that likely corresponded to an RtrC binding site (Figure 4C).

To test if RtrC bound to this predicted binding site, we performed electrophoretic mobility shift assays (EMSA) with purified RtrC. Increasing concentrations of RtrC shifted a labeled DNA probe, containing a 27 bp sequence from the *hfiA* promoter centered on the predicted RtrC binding motif (Figure 4A and 4D). RtrC bound to this pseudo-palindrome in the *hfiA* promoter with high affinity (k_d of 45 $\pm$ 9 nM) (Figure S1). Addition of excess unlabeled specific DNA probe competed with labeled probe bound to RtrC, while unlabeled non-specific probe did not compete for RtrC

binding (Figure 4D). These data provide evidence that RtrC directly represses *hfiA* transcription by specifically binding to a pseudo-palindromic motif in the *hfiA* promoter.

RtrC is a transcriptional activator and repressor

To further characterize the function of RtrC as a transcriptional regulator, we used RNA sequencing (RNA-seq) to measure changes in transcript levels upon *rtrC* overexpression (*rtrC*$^{++}$). By combining RNA-seq and ChIP-seq datasets, we identified genes that are directly controlled by RtrC. Direct targets were defined as genes that *a)* were differentially regulated in *rtrC*$^{++}$ relative to an empty vector control, *b)* contained an RtrC-enriched peak by ChIP-seq, and *c)* contained an RtrC binding motif in their promoter region [30]. Of the directly regulated genes, 63% were activated and 37% were repressed by RtrC (Table S2). Consistent with transcriptional reporter analysis (Figure 2D), *hfiA* transcript levels were ~5-fold lower in *rtrC*$^{++}$ compared to the vector control (Table S2). To confirm the RNA-seq results, we constructed several fluorescent transcriptional reporters for genes identified as direct targets of RtrC. Consistent with the RNA-seq data, *rtrC* overexpression significantly increased reporter signal for *CCNA_00629* (2.6-fold) and *CCNA_00538* (2.0-fold) and decreased reporter signal for *CCNA_00388* (6.7-fold) compared to an empty vector control (Figure 5B). RtrC bound to the *rtrC* promoter *in vivo* as demonstrated by ChIP-seq, and signal from a *rtrC* transcriptional reporter was 19-fold lower when *rtrC* was overexpressed (Figure 5B). From this, we conclude that RtrC is a negative autoregulator.

We also measured signal from transcriptional reporters for several genes that contained RtrC motifs in their promoters but did not meet the statistical threshold for differential regulation by *rtrC* overexpression in the RNA-seq dataset. There was no significant change in transcription from 11 of these 16 reporters in response to *rtrC* overexpression, which was consistent with RNA-seq measurements (Figure S2). *rtrC* overexpression significantly enhanced transcription from the remaining 5 reporters: *CCNA_03585* (2.0-fold), *CCNA_02901* (1.6-fold), *dgrB* (5.8-fold), *CCNA_01140* (1.6-fold), and *CCNA_02976* (1.4-fold) (Figure 5B). Together, these ChIP-seq,

RNA-seq and reporter data provide evidence that RtrC can function as both a direct transcriptional activator and repressor.

RtrC motif position within regulated promoters correlates with transcriptional activity

We hypothesized that the activity of RtrC as an activator or repressor depends on its binding position within a promoter relative to the transcription start site (TSS). To assess whether position correlated with regulatory activity, we analyzed the location of RtrC binding motifs within the promoters of genes that were up- or downregulated based on RNA-seq and transcriptional reporter data. Promoters directly repressed by RtrC typically had predicted motifs that overlapped the -10/-35 region of the promoter. In contrast, genes activated by RtrC had binding motifs that were located upstream of the -10/-35 region (Figure 5C). These data provide evidence that the regulatory activity of RtrC is related to the position of the RtrC binding site in a promoter. The results of this analysis are consistent with a well-described trend in which DNA-binding regulators that function as repressors bind at or near the transcription start site, while activators typically bind upstream of the -10/-35 region to promote transcription [31].

SpdR, RtrB, and RtrC form a Type I coherent feedforward loop

Transcript levels of *rtrB* were 12-fold higher in the *rtrC*$^{++}$ background relative to a vector control, placing it among the most highly activated direct targets of RtrC (Table S2). This result, along with previous findings that SpdR activates transcription of both *rtrB* and *rtrC* [18, 20], suggested that these three proteins could form a coherent type I feedforward loop (FFL) (Figure 6A). The regulatory role of this predicted coherent type I FFL depends on whether *C. crescentus* uses AND-gated logic, in which both SpdR and RtrC are required to activate *rtrB* expression, or OR-gated logic, in which either SpdR or RtrC can activate *rtrB* expression [32]. To test FFL gating, we deleted *spdR* and *rtrC* from the chromosome and measured fluorescence from a *rtrB* transcriptional reporter upon expression of *spdR*$_{D64E}$ and/or *rtrC* from inducible promoters.

206 Expression of either *rtrC* or *spdR*$_{D64E}$ alone increased transcription from the *rtrB* reporter by ~5-
207 fold, while expression of both *rtrC* and *spdR*$_{D64E}$ increased transcription by ~6-fold (Figure 6B).
208 *spdR* deletion significantly reduced transcription from a P$_{rtrB}$ reporter in stationary phase (Figure
209 S3). As expected, deletion of *rtrC* alone did not affect transcription from P$_{rtrB}$ as *spdR* is still present
210 on the chromosome (Figure S3). We conclude that either SpdR or RtrC can activate *rtrB*
211 expression and are therefore competent to form an OR-gated coherent type I FFL in *C.*
212 *crescentus*.
213
214 *RtrC deletion does not affect holdfast biogenesis in stationary phase*
215 SpdR affects gene regulation during stationary phase [20, 33], and was previously
216 reported to bind *rtrC* promoter DNA [20]. These published results led us to assess the effect of
217 *rtrC* gene deletion on holdfast synthesis in log and stationary phase in complex medium. The
218 fraction of cells with holdfasts in complex medium during early log phase was not significantly
219 different in strains with in-frame deletions of *spdR*, *rtrA*, *rtrB*, *rtrC,* or in a strain missing all three
220 *rtr* regulators (Figure S4). Stained holdfasts are greatly reduced in stationary phase, though this
221 effect does not require *spdR*. Diminished holdfast counts in stationary phase are more
222 pronounced in strains lacking *rtrB* but are unaffected by deletion of *rtrC* (Figure S4). A
223 constellation of changes in cell physiology in stationary phase, including a slowed cell cycle, LovK-
224 SpdSR-SkaH signaling and changes in levels of nucleotide signals such as (p)ppGpp and cyclic-
225 di-GMP, likely contribute to reduction in holdfast in stationary phase in complex medium. The
226 spectrum of environmental conditions that may affect holdfast development via *rtrC* remain
227 undefined, though published transcriptomic data provide evidence that carbon limitation [34], cell
228 cycle [35], and stringent response signaling [36] also significantly affect *rtrC* transcription.
229
230 *RtrC overexpression reduces cell motility in soft agar*

RtrC directly regulates several genes with predicted roles in motility (Figure 5 and Table S2). We therefore tested whether RtrC expression can affect cell motility by measuring cell spreading through soft agar. We observed a consistent reduction in swarm diameter in the strain overexpressing *rtrC* compared to the empty vector control, though the effect is small (Figure S5). The transcription of the cyclic-di-GMP receptor *dgrB* is activated by *rtrC* overexpression, and we postulated that the effect of *rtrC*$^{++}$ on motility may be affected by *dgrB*. However, deletion of *dgrB* had no effect on the motility phenotype of *rtrC*$^{++}$.

Discussion

We designed a forward genetic selection to search for novel holdfast regulators and identified RtrC. This formerly hypothetical protein functions downstream of an ensemble of TCS regulatory genes to activate surface adhesion. RtrC binds and regulates multiple sites on the *C. crescentus* chromosome, including the *hfiA* promoter where it represses *hfiA* transcription and thereby activates holdfast synthesis.

RtrC structure and regulatory activity

A comparison of the predicted three-dimensional structure of RtrC to experimental structures available in the PDB suggested structural similarity to MepR and several other MarR family transcriptional regulators. Members of this transcription factor family often bind as dimers to pseudo-palindromic DNA sequences [37-39]. MarR family transcriptional regulators are known to function as both activators and repressors, depending on the position of binding within regulated promoters. Similarly, we observed that the activity of RtrC as an activator or repressor was correlated with the position of the RtrC motif within the promoter; this positional effect on transcriptional regulation is a well-described phenomenon [40]. The sequence of RtrC is not broadly conserved. It is largely restricted to the Caulobacterales and Rhodospirillales where it is

annotated as a hypothetical protein. However, our data provide evidence that RtrC (like MarR) is a classic transcription factor.

A new layer of hfiA regulation

Holdfast-dependent surface attachment in *C. crescentus* is permanent. Accordingly, holdfast synthesis is highly regulated. The small protein, HfiA, is central to holdfast control. It represses holdfast biogenesis by directly interacting with the glycosyltransferase HfsJ, an enzyme required for synthesis of holdfast polysaccharide [41]. *hfiA* expression is influenced by multiple cell cycle regulators, TCS sensory/signaling systems, a transcriptional regulator of stalk biogenesis, and c-di-GMP [18, 41-43]. We have shown that RtrC functions downstream of the stationary phase response regulator, SpdR, to directly bind the *hfiA* promoter and represses its transcription. SpdR can therefore regulate expression of at least three different direct repressors of *hfiA* transcription – *rtrA*, *rtrB*, and *rtrC* (Figure 7).

Why, then, does the *spdR* response regulator have so many outlets to directly modulate *hfiA* transcription? We do not know whether the activities of RtrA, RtrB, or RtrC as transcription factors are allosterically regulated by small molecules, chemical modifications, or protein-protein interactions. If these transcription factors are subject to allosteric regulation, it may be the case that this suite of proteins serves to integrate multiple environmental or cellular signals. In such a model, primary signals that regulate the transcriptional activity of SpdR may enhance expression of RtrA, RtrB, or RtrC, which could then influence the scale of adhesion to substrates in response to secondary physical or chemical cues. Another possible explanation for multiple adhesion transcription factors downstream of SpdR is redundancy. Transcription factor redundancy may buffer the network against transient changes in signaling and gene regulation, ensuring that the decision to synthesize a holdfast (or not) is less subject to environmental fluctuations.

More on the RtrC regulon

The *cox* genes encode an aa_3-type cytochrome oxidase. *C. crescentus* encodes at least four different aerobic terminal oxidase complexes [44], and the data presented herein provide evidence that the *spdR-rtrC* axis activates transcription of the *cox* aa_3-type terminal oxidase genes (Table S2). In *Pseudomonas aeruginosa*, an aa_3-type oxidase is reported to provide a survival advantage for cells under starvation conditions [45]. The physiological relevance of *cox* regulation by RtrC in *C. crescentus* is the subject of ongoing investigation.

RtrC directly regulates expression of several genes involved in c-di-GMP signaling including the c-di-GMP receptor *dgrB*, a PAS-containing EAL phosphodiesterase (*CCNA_01140*), and a GGDEF-EAL protein (*CCNA_00089*). Deletion of *CCNA_00089* is reported to increase surface attachment [46]. Given that *rtrC* overexpression represses *CCNA_00089* expression (Table S2), it is possible that *rtrC* influences adhesion through *CCNA_00089* in addition to *hfiA*. DgrB directly binds c-di-GMP and represses cellular motility [47]. However, deletion of *dgrB* did not affect *C. crescentus* motility in soft agar in either *rtrC* overexpression or vector control strains (Figure S5). RtrC also directly activated expression of genes with predicted roles in chemotaxis, including two methyl-accepting chemotaxis proteins (*CCNA_00538* and *CCNA_02901*) and a *cheY* receiver domain protein (*CCNA_03585*). Additionally, RtrC activates transcription of an alternative chemotaxis cluster (*CCNA_00628* and *CCNA_00629-CCNA_00634*), which has been reported to influence *hfiA* transcription and *C. crescentus* surface adherence [42].

Feedback control in the adhesion pathway

Our data provide evidence that *spdR* and *rtrC* form a type I coherent feedforward loop (C1-FFL) with the adhesion factor *rtrB*. Experimental and theoretical studies of C1-FFLs indicate that they function as sign-sensitive delay elements [32, 48]. AND-gated C1-FFLs exhibit a delay in the ON step of output expression, which can allow circuits to function as persistence detectors [32, 49]. Conversely, OR-gated C1-FFLs delay the OFF step of output expression, which can buffer the circuit against the transient loss of activating signals [32, 48]. Expression of either *spdR*

or *rtrC* was sufficient to activate transcription from a P_{rtrB} reporter, indicating that the *spdR-rtrC-rtrB* C1-FFL is competent to function as an OR-gated system. Though the exact environmental signals that regulate the adhesion TCS pathway remain undefined, the architecture of the SpdR-RtrC-RtrB circuit suggests that RtrC can reinforce *rtrB* expression in particular environments where the levels of activating signals for SpdR are fluctuating or noisy.

The DNA-binding response regulator, SpdR, is regulated in a growth phase and media-dependent fashion [20, 33], and systems homologous to *C. crescentus* SpdS-SpdR are reported to respond to cellular redox state and to flux through the electron transport chain (ETC) via modulation of disulfide bond formation [50], modification of a reactive cysteine [51], or by binding of oxidized quinones [52, 53]. *C. crescentus* SpdS contains both the reactive cysteine and quinone-interacting residues observed in related bacteria, suggesting that SpdS may be regulated in a similar manner. The activity of SpdR as a transcriptional regulator is also affected by the sensor kinases LovK and SkaH [18]. Thus, multiple environmental signals apparently feed into SpdR-dependent gene regulation.

This study expands our understanding of a transcriptional network functioning downstream of a consortium of TCS proteins – LovK, SkaH, SpdS, and SpdR – that affect surface adherence in *Caulobacter*. The DNA-binding response regulator SpdR regulates expression of three transcription factor genes (*rtrA, rtrB,* and *rtrC*) that directly repress the holdfast inhibitor, *hfiA*. Of these three transcription factors, RtrC is the most potent regulator of *hfiA*. However, it is clear from the ChIP-seq and transcriptomic data presented in this study that the regulatory function of RtrC likely extends well beyond *hfiA* and holdfast synthesis (Figure 7). Efforts focused on comparative analysis of the SpdR, RtrA, RtrB, and RtrC regulons will provide a more complete understanding of the regulatory logic that underpins the highly complex process of holdfast adhesin development in *Caulobacter*.

Materials and Methods

Strain growth conditions

Escherichia coli was grown in Lysogeny broth (LB) or LB agar (1.5% w/v) at 37°C [54]. Medium was supplemented with the following antibiotics when necessary: kanamycin 50 µg ml^{-1}, chloramphenicol 20 µg ml^{-1}, oxytetracycline 12 µg ml^{-1}, and carbenicillin 100 µg ml^{-1}. *Caulobacter crescentus* was grown in peptone-yeast extract (PYE) broth (0.2% (w/v) peptone, 0.1% (w/v) yeast extract, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$), PYE agar (1.5% w/v), or M2 defined medium supplemented with xylose (0.15% w/v) as the carbon source (M2X) [55] at 30°C. Solid medium was supplemented with the following antibiotics where necessary: kanamycin 25 µg ml^{-1}, chloramphenicol 1 µg ml^{-1}, and oxytetracycline 2 µg ml^{-1}. Liquid medium was supplemented with the following antibiotics where necessary: chloramphenicol 1 µg ml^{-1}, and oxytetracycline 2 µg ml^{-1}.

Tn-Himar mutant library construction and mapping

Construction and mapping of the barcoded Tn-himar library was performed following protocols originally described by Wetmore and colleagues [56]. A 25 ml culture of the *E. coli* APA_752 barcoded transposon donor pool (obtained from Adam Deutschbauer Lab) was grown to mid-log phase in LB broth supplemented with kanamycin and 300 µM diaminopimelic acid (DAP). A second 25 ml culture of *C. crescentus lovK*$_{H180A}$ was grown to mid-log phase in PYE. Cells from both cultures were harvested by centrifugation, washed twice with PYE containing 300 µM DAP, mixed and spotted together for conjugation on a PYE agar plate containing 300 µM DAP. After incubating the plate overnight at room temperature, the cells were scraped from the plate, resuspended in PYE medium, spread onto 20, 150 mm PYE agar plates containing kanamycin and incubated at 30°C for three days. Colonies from each plate were scraped into PYE medium and used to inoculate a 25 ml PYE culture containing 5 µg ml^{-1} kanamycin. The

358 culture was grown for three doublings, glycerol was added to 20% final concentration, and 1 ml
359 aliquots were frozen at -80°C.

360 To map the sites of transposon insertion, we again followed the protocols of Wetmore et
361 al. [56]. Briefly, genomic DNA was purified from three 1 ml aliquots of each library. The DNA was
362 sheared and ~300 bp fragments were selected before end repair. A Y-adapter (Mod2_TS_Univ,
363 Mod2_TruSeq) was ligated and used as a template for transposon junction amplification with the
364 primers Nspacer_BarSeq_pHIMAR and either P7_mod_TS_index1 or P7_mod_TS_index2. 150-
365 bp single end reads were collected on an Illumina HiSeq 2500 in rapid run mode, and the genomic
366 insertion positions were mapped and correlated to a unique barcode using BLAT [57] and
367 MapTnSeq.pl to generate a mapping file with DesignRandomPool.pl. Using this protocol, we
368 identified 232903 unique barcoded insertions at 60940 different locations on the chromosome.
369 The median number of barcoded strains per protein-encoding gene (that tolerated Tn insertion)
370 was 34; the mean was 49.6. Median number of sequencing reads per hit protein-encoding gene
371 was 4064; mean was 6183.5. All code used for this mapping and analysis is available at
372 https://bitbucket.org/berkeleylab/feba/.

373

374 *Adhesion profiling of the lovK$_{H180A}$ Tn-Himar mutant library*

375 Adhesion profiling followed the protocol originally outlined in Hershey et al. [19]. 1 ml aliquots of
376 the barcoded transposon library were cultured, collected by centrifugation, and resuspended in 1
377 ml of M2X medium. 300 µl of this barcoded mutant pool was inoculated into a well of a 12-well
378 microtiter plate containing 1.5 ml M2X defined medium with 6-8 ~1 x 1 cm layers of cheesecloth.
379 These microtiter plates were incubated for 24 hours at 30°C with shaking at 155 rpm after which
380 150 µl of the culture was passaged by inoculating into a well with 1.65 ml fresh M2X containing
381 cheesecloth. Cells from an additional 500 µl of medium from each well was harvested by

centrifugation and stored at -20°C for barcode sequencing (BarSeq) analysis. Each passaging experiment was performed in triplicate, and passaging was performed sequentially for a total of five rounds of selection. Identical cultures grown in a plate without cheesecloth were used as a nonselective reference condition.

Cell pellets were used as PCR templates to amplify the barcodes in each sample using indexed primers [56]. Amplified products were purified and pooled for multiplexed sequencing. 50 bp single end reads were collected on an Illumina HiSeq4000. The Perl and R scripts MultiCodes.pl, combineBarSeq.pl and FEBA.R were used to determine fitness scores for each gene by comparing the $\log_2$ ratios of barcode counts in each sample over the counts from a nonselective growth in M2X without cheesecloth. To evaluate mutant phenotypes in each selection, the replicates were used to calculate a mean fitness score for each gene after each passage. Mean fitness (a proxy for adhesion to cheesecloth) was assessed across passages for each gene.

Plasmid and strain construction

Plasmids were cloned using standard molecular biology techniques and the primers listed in Table S4. To construct pPTM051, *CCNA_03380* (-21 to +15 bp relative to the start of the gene) was fused to *mNeonGreen* and cloned into pMT805 lacking the xylose-inducible promoter [58]. To construct pPTM056, site directed mutagenesis was used to introduce a silent mutation in the chloramphenicol acetyltransferase gene of pPTM051 to remove an EcoRI site. A cumate-inducible, integrating plasmid was constructed by fusing a backbone with a chloramphenicol resistance marker derived from pMT681 [58], the xylose integration site derived from pMT595 [58], and the cumate-responsive repressor and promoter derived from pQF through Gibson Assembly [59]. To construct pPTM057, the xylose integration site, cumate repressor, and cumate-inducible promoter of pPTM052 were fused to a backbone with a kanamycin resistance marker

407 derived from pMT426 [58]. For reporter plasmids, inserts were cloned into the replicating plasmid
408 pPTM056. For overexpression constructs, inserts were cloned into pPTM057 or pMT604 that
409 integrate at the xylose locus and contain either a cumate- (P_{Q5}) or xylose-inducible (P_{xyl}) promoter,
410 respectively [58]. For 3xFLAG-tagged RtrC overexpression, inserts were cloned into the
411 replicating plasmid pQF [59]. Deletion inserts were constructed by overlap PCR with regions up-
412 and downstream of the target gene and cloned into the pNPTS138 plasmid. Clones were
413 confirmed with Sanger sequencing.
414 Plasmids were transformed into *C. crescentus* by either electroporation or triparental mating [55].
415 Transformants generated by electroporation were selected on PYE agar supplemented with the
416 appropriate antibiotic. Strains constructed by triparental mating were selected on PYE agar
417 supplemented with the appropriate antibiotic and nalidixic acid to counterselect against *E. coli*.
418 Gene deletions and allele replacements were constructed using a standard two-step
419 recombination/counter-selection method, using *sacB* as the counterselection marker. Briefly,
420 pNPTS138-derived plasmids were transformed into *C. crescentus* and primary integrants were
421 selected on PYE/kanamycin plates. Primary integrants were incubated overnight in PYE broth
422 without selection. Cultures were plated on PYE agar plates supplemented with 3% (w/v) sucrose
423 to select for recombinants that had lost the plasmid. Mutants were confirmed by PCR amplification
424 of the gene of interest from sucrose resistant, kanamycin sensitive clones.
425
426 *Holdfast imaging and quantification*
427 Strains were inoculated in triplicate in M2X or PYE, containing 50 µM cumate when appropriate,
428 and grown overnight at 30°C. Strains were subcultured in M2X or PYE, containing 50 µM cumate
429 when appropriate, and grown for 3-8 hours at 30°C. Cultures were diluted to 0.0000057 – 0.00045
430 OD_{660} and incubated at 30°C until reaching 0.05 – 0.1 OD_{660},. For stationary phase cells, cultures
431 were diluted to 0.05 OD_{660} and incubated at 30°C for 24 hours. Alexa594-conjugated wheat germ
432 agglutinin (WGA) (ThermoFisher) was added to the cultures with a final concentration of 2.5 µg

433 ml^{-1}. Cultures were shaken at 30°C for 10 min at 200 rpm. Then, 1.5 ml early log phase culture or
434 0.75 ml stationary phase culture was centrifuged at 12,000 x g for 2 min and supernatant was
435 removed. Pellets from early log phase in M2X and PYE were resuspended in 35 µl M2X or 100
436 µl H_2O, respectively. Pellets from stationary phase in PYE were resuspended in 400 µl H_2O. Cells
437 were spotted on 1% (w/v) agarose pads in H_2O and imaged with a Leica DMI6000 B microscope.
438 WGA staining was visualized with Leica TXR ET (No. 11504207, EX: 540-580, DC: 595, EM: 607-
439 683) filter. Cells with and without holdfasts were enumerated using the image analysis suite, FIJI.
440 Statistical analysis was carried out in GraphPad 9.3.1.
441
442 *Structure prediction and comparison*
443 The structure of CCNA_00551 was predicted with AlphaFold [24] through Google Colab using the
444 ChimeraX interface [60]. The predicted structure from AlphaFold was submitted to the Dali server
445 [25] for structural comparison to the Protein Data Bank (PDB).
446
447 *Chromatin immunoprecipitation sequencing (ChIP-seq)*
448 Strains were incubated in triplicate at 30°C overnight in 10 ml PYE supplemented with 2 µg/ml
449 oxytetracycline. Then, 5 ml overnight culture was diluted into 46 ml PYE supplemented with 2
450 µg/ml oxytetracycline, grown at 30°C for 2 hours. Cumate was added to a final concentration of
451 50 µM and cultures were grown at 30°C for 6 hours. Cultures were crosslinked with 1% (w/v)
452 formaldehyde for 10 min, then crosslinking was quenched by addition of 125 mM glycine for 5
453 min. Cells were centrifuged at 7196 x g for 5 min at 4°C, supernatant was removed, and pellets
454 were washed in 25 ml 1x cold PBS pH 7.5 three times. Pellets were resuspended in 1 ml [10 mM
455 Tris pH 8 at 4°C, 1 mM EDTA, protease inhibitor tablet, 1 mg ml^{-1} lysozyme] and incubated at
456 37°C for 30 min. Sodium dodecyl sulfate (SDS) was added to a final concentration of 0.1% (w/v)
457 and DNA was sheared to 300-500 bp fragments by sonication for 10 cycles (20 sec on/off). Debris
458 was centrifuged at 15,000 x g for 10 min at 4°C, supernatant was transferred, and Triton X-100

was added to a final concentration of 1% (v/v). Samples were pre-cleared through incubation with 30 µl SureBeads™ Protein A magnetic beads for 30 min at room temp. Supernatant was transferred and 5% lysate was removed for use as input DNA.

For pulldown, 100 µl Pierce™ anti-FLAG magnetic agarose beads (25% slurry) were equilibrated overnight at 4°C in binding buffer [10 mM Tris pH 8 at 4°C, 1 mM EDTA, 0.1% (w/v) SDS, 1% (v/v) Triton X-100] supplemented with 1% (w/v) bovine serum albumin (BSA). Pre-equilibrated beads were washed four times in binding buffer, then incubated with the remaining lysate for 3 hours at room temperature. Beads were washed with low-salt buffer [50 mM HEPES pH 7.5, 1% (v/v) Triton X-100, 150 mM NaCl], high-salt buffer [50 mM HEPES pH 7.5, 1% (v/v) Triton X-100, 500 mM NaCl], and LiCl buffer [10 mM Tris pH 8 at 4°C, 1 mM EDTA, 1% (w/v) Triton X-100, 0.5% (v/v) IGEPAL® CA-630, 150 mM LiCl]. To elute protein-DNA complexes, beads were incubated for 30 min at room temperature with 100 µl elution buffer [10 mM Tris pH 8 at 4°C, 1 mM EDTA, 1% (w/v) SDS, 100 ng μl^{-1} 3xFLAG peptide] twice. Elutions were supplemented with NaCl and RNase A to a final concentration of 300 mM and 100 µg ml^{-1}, respectively, and incubated at 37°C for 30 min. Then, samples were supplemented with Proteinase K to a final concentration of 200 µg ml^{-1} and incubate overnight at 65°C to reverse crosslinks. Input and elutions were purified with the Zymo ChIP DNA Clean & Concentrator™ kit and libraries were prepared and sequenced at the Microbial Genome Sequencing Center (Pittsburgh, PA). Raw chromatin immunoprecipitation sequencing data are available in the NCBI GEO database under series accession GSE201499.

ChIP-seq analysis

Paired-end reads were mapped to the *C. crescentus* NA1000 reference genome (GenBank accession number CP001340) with Bowtie2 on Galaxy. Peak calling was performed with the Genrich tool (https://github.com/jsh58/Genrich) on Galaxy; peaks are presented in Table S3. Briefly, PCR duplicates were removed from mapped reads, replicates were pooled, input reads

were used as the control dataset, and peak were called using the default peak calling option [Maximum q-value: 0.05, Minimum area under the curve (AUC): 20, Minimum peak length: 0, Maximum distance between significant sites: 100]. An average AUC $\geq$ 2500 was used as the cutoff for significant peaks. Distance between called peaks and the nearest transcription start sites (TSS) (modified from [30]) was analyzed using ChIPpeakAnno [61]. For genes/operons that did not have an annotated TSS, the +1 residue of the gene (or start of the operon) was designated as the TSS.

RtrC motif discovery

For motif discovery, sequences of enriched ChIP-seq peaks were submitted to the XSTREME module of MEME suite [29]. For the XSTREME parameters, shuffled input files were used as the control sequences for the background model, checked for motifs between 6 and 30 bp in length that had zero or one occurrence per sequence.

RNA preparation, sequencing, and analysis

Strains were incubated in quadruplicate at 30°C overnight in M2X broth supplemented with 50 µM cumate. Overnight replicate cultures were diluted into fresh M2X/50 µM cumate to 0.025 OD_{660} and incubated at 30°C for 8 hours. Cultures were diluted into 10 ml M2X/50 µM cumate to 0.001 – 0.003 OD_{660} and incubated at 30°C until reaching 0.3 – 0.4 OD_{660}. Upon reaching the desired OD_{660}, 6 ml culture was pelleted at 15,000 x g for 1 minute, supernatant was removed, and pellets were resuspended in 1 ml TRIzol and stored at -80°C. Samples were heated at 65°C for 10 min. Then, 200 µl chloroform was added, samples were vortexed, and incubated at room temperature for 5 min. Samples were centrifuged at 17,000 x g for 15 min at 4°C, then the upper aqueous phase was transferred to a fresh tube, an equal volume of 100% isopropanol was added, and samples were stored at -80°C overnight. Samples were centrifuged at 17,000 x g for 30 min at 4°C, then supernatant was removed. Samples were washed with cold 70% ethanol. Samples

were centrifuged at 17,000 x g for 5 min at 4°C, supernatant was removed, and the pellet was allowed to dry. Pellets were resuspended in 100 µl RNase-free water and incubated at 60°C for 10 min. Samples were treated with TURBO™ DNase and cleaned up with RNeasy Mini Kit (Qiagen).

Library preparation and sequencing was performed at the Microbial Genome Sequencing center with the Illumina Stranded RNA library preparation and RiboZero Plus rRNA depletion (Pittsburgh, PA). Reads were mapped to the *C. crescentus* NA1000 reference genome (GenBank accession number CP001340) using CLC Genomics Workbench 20 (Qiagen). Differential gene expression was determined with the CLC Genomics Workbench RNA-seq Analysis Tool ($|\text{fold-change}| \geq 1.5$ and FDR p-value ≤ 0.001). Raw RNA sequencing data are available in the NCBI GEO database under series accession GSE201499.

RNA-seq and ChIP-seq overlap analysis

The Bioconductor package was used to identify overlap between RtrC-regulated genes defined by RNA-seq and RtrC binding sites defined by ChIP-seq [61]. Promoters for genes were designated as the sequence -400 to +100 around the TSS [30]. Overlap between promoters and RtrC motifs identified from XSTREME was analyzed using ChIPpeakAnno within Bioconductor. Genes were defined as direct targets of RtrC if their transcript levels were differentially regulated in the RNA-seq analysis and had an RtrC motif within a promoter for their operon. To analyze RtrC motif distribution in directly regulated promoters, promoters were grouped based on the effect of RtrC on gene expression (i.e. upregulated vs. downregulated). The number of predicted RtrC motifs at each position with these promoters was calculated and plotted. If an operon contained more than one promoter, then each promoter for that operon that contained an RtrC motif was analyzed.

Analysis of transcription using fluorescent fusions

Strains were incubated in triplicate at 30°C overnight in PYE broth supplemented with 1 µg ml^{-1} chloramphenicol and 50 µM cumate. Overnight cultures were diluted to 0.05 OD_{660} in PYE/chloramphenicol/cumate broth and incubated at 30°C for 24 hours. For Figure S3, log phase (0.144 – 0.296 OD_{660}) cultures were diluted to 0.025 OD_{660} and incubated at 30°C for 48 hours. Then, 200 µl culture was transferred to a black Costar® 96 well plate with clear bottom (Corning). Absorbance at 660 nm and fluorescence (excitation = 497 $\pm$ 10 nm; emission = 523 $\pm$ 10 nm) were measured in a Tecan Spark 20M plate reader. Fluorescence was normalized to absorbance. For the *rtrB* reporter, strains were incubated in triplicate at 30°C overnight in PYE broth supplemented with 1 µg ml^{-1} chloramphenicol. Overnight cultures were diluted to 0.025 OD_{660} in PYE broth supplemented with 1 µg ml^{-1} chloramphenicol, 50 µM cumate, and 0.15% (w/v) xylose. Cultures were incubated at 30°C for 24 hours, then 100 µl overnight was diluted with 100 µl PYE and transferred to a black Costar® 96 well plate with clear bottom (Corning). Fluorescence and absorbance were measured as indicated above in a Tecan Spark 20M plate reader. Fluorescence was normalized to absorbance. Statistical analysis was carried out in GraphPad 9.3.1.

Protein purification

For heterologous expression of RtrC, plasmids were transformed into the BL21 Rosetta (DE3)/pLysS background. Strains were inoculated into 20 ml LB broth supplemented with 100 µg ml^{-1} carbenicillin and incubated overnight at 37°C. Overnight cultures were diluted into 1000 ml LB supplemented with carbenicillin and grown for 3 – 4 hours at 37°C. Protein expression was induced by 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubation at 37°C for 3 – 4.5 hours. Cells were pelleted at 11,000 x g for 7 min at 4°C, pellets were resuspended in 25 ml lysis buffer [25 mM Tris pH 8 at 4°C, 500 mM NaCl, 10 mM imidazole], and stored at -80°C. Samples were thawed, supplemented with PMSF and benzonase to a final concentration of 1 mM and 50 Units ml^{-1}, respectively. Samples were sonicated with a Branson Digital Sonifier at 20% output in 20" intervals until sufficiently lysed and clarified by centrifugation at 39,000 x g for 15

563 min at 4°C. Clarified lysates were batch incubated with 4 ml Ni-NTA Superflow Resin (50% slurry)
564 that had been pre-equilibrated in lysis buffer for 60 min at 4°C. Column was then washed with 25
565 ml lysis buffer, high salt buffer [25 mM Tris pH 8 at 4°C, 1 M NaCl, 30 mM imidazole], and low
566 salt buffer [25 mM Tris pH 8 at 4°C, 500 mM NaCl, 30 mM imidazole]. For elution, column was
567 batch incubated with 25 ml elution buffer [25 mM Tris pH 8 at 4°C, 500 mM NaCl, 300 mM
568 imidazole] for 60 min at 4°C.
569 Elution was supplemented with ULP1 protease to cleave the His_6-SUMO tag and dialyzed in 1 L
570 dialysis buffer [25 mM Tris pH 8 at 4°C, 500 mM NaCl] overnight at 4°C. Dialyzed sample was
571 batch incubated with 4 ml Ni-NTA Superflow Resin (50% slurry) that had been pre-equilibrated in
572 dialysis buffer for 60 min at 4°C. Flowthrough that contained untagged RtrC was collected and
573 concentrated on an Amicon® Ultra-15 concentrator (3 kDa cutoff) at 4,000 x g at 4°C. Samples
574 were stored at 4°C until needed.
575
576 *Electrophoretic mobility shift assay (EMSA)*
577 To prepare labeled DNA fragments, an Alexa488-labeled universal forward primer and an *hfiA*
578 specific reverse primer listed in Table S4 were annealed in a thermocycler in as follows: 94°C for
579 5 min, then ramp down to 18°C at 0.1°C s^{-1}. Overhangs were filled in with DNA polymerase I,
580 Large Klenow fragment at 25°C for 60 min. DNA fragments were then treated with Mung Bean
581 Nuclease for 120 min at 30°C to remove any remaining overhangs. DNA fragments were purified
582 with the GeneJet PCR purification kit, eluted in 10 mM TE/NaCl [Tris pH 8 at 4°C, 1 mM EDTA,
583 50 mM NaCl], and diluted to 0.5 µM in TE/NaCl. Unlabeled DNA fragments were prepared by
584 annealing primers listed in Table S4 with protocol listed above. For non-specific chase, the
585 sequence of the *hfiA* specific probe was shuffled.
586 RtrC was incubated with 6.25 nM labeled DNA in binding buffer at 20°C for 30 min in the dark and
587 subsequently cooled to 4°C on ice. For EMSA to analyze binding curves, DNA binding buffer
588 consisted of 32.5 mM Tris pH 8 at 4°C, 200 mM NaCl, 1 mM EDTA, 30% (v/v) glycerol, 1 mM

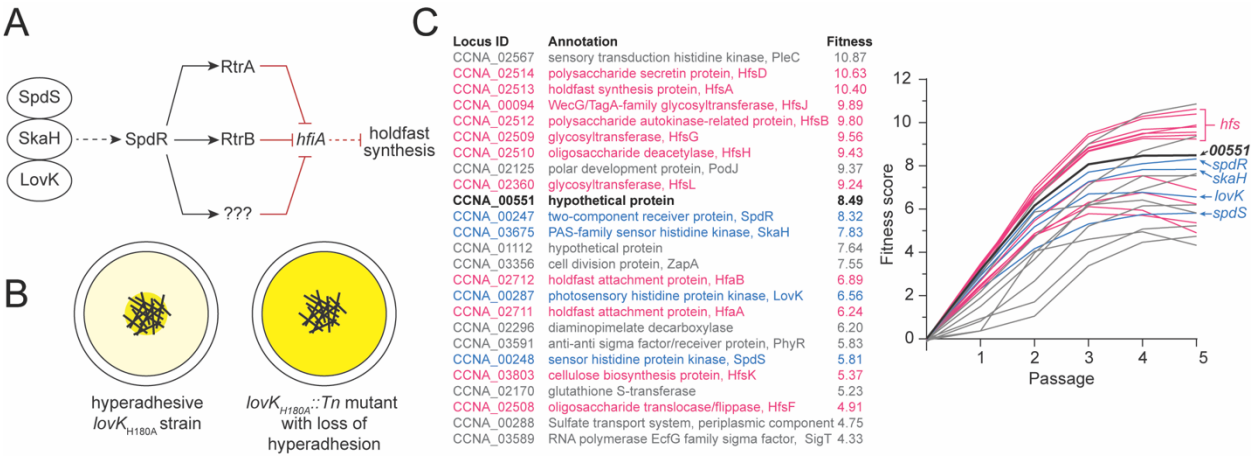
589 DTT, 10 µg ml^{-1} BSA, and 50 ng $µl^{-1}$ poly(dI-dC). For EMSAs with unlabeled chases, DNA binding
590 buffer consisted of 30 mM Tris pH 8 at 4°C, 150 mM NaCl, 1 mM EDTA, 30% (v/v) glycerol, 1 mM
591 DTT, 10 µg ml^{-1} BSA, and 50 ng $µl^{-1}$ poly(dI-dC). For non-specific and specific chases, reactions
592 were supplemented with 2.5 µM unlabeled DNA. Then, 15 µl reaction was loaded on to a
593 degassed polyacrylamide gel [10% (v/v/) acrylamide (37.5:1 acrylamide:bis-acrylamide), 0.5x
594 Tris-Borate-EDTA buffer (TBE: 45 mM Tris, 45 mM borate, 1 mM EDTA)] and run at 100 V for 40
595 min at 4°C in 0.5x TBE buffer. Gels were imaged on a ChemiDoc™ MP imaging system [Light:
596 blue epi illumination, Filter: 530/28, Exposure: 30 sec] and bands were quantified with FIJI. For
597 calculating k_d, percent bound probe at each protein concentration was calculated as (1 – [intensity
598 free probe at x nM protein]/[intensity of free probe at 0 nM protein]). Binding curve was derived
599 from One site – specific binding analysis using GraphPad 9.3.1.
600
601 *Soft agar swarm assay*
602 Strains were incubated in quadruplicate at 30°C overnight in PYE broth supplemented with 50 µM
603 cumate. Overnight cultures were diluted to 0.05 OD_{660} in PYE/50 µM cumate, then incubated at
604 30°C for 24 hours. Cultures were diluted to 0.5 OD_{660} in PYE broth, 0.75 µl diluted culture was
605 pipetted into PYE plate supplemented with 50 µM cumate, incubated at 30°C for 3 days. Plates
606 were imaged on a ChemiDoc™ MP imaging system and swarm size was measured with FIJI.
607 Statistical analysis was carried out in GraphPad 9.3.1.
608

References

1. Dufrene, Y.F. and A. Persat, *Mechanomicrobiology: how bacteria sense and respond to forces.* Nat Rev Microbiol, 2020. **18**(4): p. 227-240.
2. Figueiredo, A.M.S., et al., *The role of biofilms in persistent infections and factors involved in ica-independent biofilm development and gene regulation in Staphylococcus aureus.* Crit Rev Microbiol, 2017. **43**(5): p. 602-620.
3. Moons, P., C.W. Michiels, and A. Aertsen, *Bacterial interactions in biofilms.* Crit Rev Microbiol, 2009. **35**(3): p. 157-68.
4. Reguera, G. and K. Kashefi, *The electrifying physiology of Geobacter bacteria, 30 years on.* Adv Microb Physiol, 2019. **74**: p. 1-96.
5. Deschaine, B.M., et al., *Biofilm formation and toxin production provide a fitness advantage in mixed colonies of environmental yeast isolates.* Ecol Evol, 2018. **8**(11): p. 5541-5550.
6. Harrison, J.J., H. Ceri, and R.J. Turner, *Multimetal resistance and tolerance in microbial biofilms.* Nat Rev Microbiol, 2007. **5**(12): p. 928-38.
7. Seiler, C., et al., *Grazing resistance of bacterial biofilms: a matter of predators' feeding trait.* FEMS Microbiol Ecol, 2017. **93**(9).
8. Lowery, N.V., et al., *Division of Labor, Bet Hedging, and the Evolution of Mixed Biofilm Investment Strategies.* mBio, 2017. **8**(4).
9. Wilhelm, R.C., *Following the terrestrial tracks of Caulobacter - redefining the ecology of a reputed aquatic oligotroph.* ISME J, 2018. **12**(12): p. 3025-3037.
10. Niederdorfer, R., et al., *Ecological strategies and metabolic trade-offs of complex environmental biofilms.* NPJ Biofilms Microbiomes, 2017. **3**: p. 21.
11. Tsang, P.H., et al., *Adhesion of single bacterial cells in the micronewton range.* Proc Natl Acad Sci U S A, 2006. **103**(15): p. 5764-8.
12. Entcheva-Dimitrov, P. and A.M. Spormann, *Dynamics and control of biofilms of the oligotrophic bacterium Caulobacter crescentus.* J Bacteriol, 2004. **186**(24): p. 8254-66.
13. McGrath, P.T., et al., *High-throughput identification of transcription start sites, conserved promoter motifs and predicted regulons.* Nat Biotechnol, 2007. **25**(5): p. 584-92.
14. Schrader, J.M., et al., *Dynamic translation regulation in Caulobacter cell cycle control.* Proc Natl Acad Sci U S A, 2016. **113**(44): p. E6859-E6867.
15. Ellison, C.K., et al., *Obstruction of pilus retraction stimulates bacterial surface sensing.* Science, 2017. **358**(6362): p. 535-538.
16. Hug, I., et al., *Second messenger-mediated tactile response by a bacterial rotary motor.* Science, 2017. **358**(6362): p. 531-534.
17. Sprecher, K.S., et al., *Cohesive Properties of the Caulobacter crescentus Holdfast Adhesin Are Regulated by a Novel c-di-GMP Effector Protein.* mBio, 2017. **8**(2).
18. Reyes Ruiz, L.M., A. Fiebig, and S. Crosson, *Regulation of bacterial surface attachment by a network of sensory transduction proteins.* PLoS Genet, 2019. **15**(5): p. e1008022.
19. Hershey, D.M., A. Fiebig, and S. Crosson, *A Genome-Wide Analysis of Adhesion in Caulobacter crescentus Identifies New Regulatory and Biosynthetic Components for Holdfast Assembly.* mBio, 2019. **10**(1).
20. da Silva, C.A., et al., *Transcriptomic analysis of the stationary phase response regulator SpdR in Caulobacter crescentus.* BMC Microbiol, 2016. **16**: p. 66.
21. Blum, M., et al., *The InterPro protein families and domains database: 20 years on.* Nucleic Acids Res, 2021. **49**(D1): p. D344-D354.
22. Marchler-Bauer, A., et al., *CDD: conserved domains and protein three-dimensional structure.* Nucleic Acids Res, 2013. **41**(Database issue): p. D348-52.
23. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.

24. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
25. Holm, L., *Using Dali for Protein Structure Comparison.* Methods Mol Biol, 2020. **2112**: p. 29-42.
26. Kumaraswami, M., et al., *Structural and biochemical characterization of MepR, a multidrug binding transcription regulator of the Staphylococcus aureus multidrug efflux pump MepA.* Nucleic Acids Res, 2009. **37**(4): p. 1211-24.
27. Schrader, J.M., et al., *The coding and noncoding architecture of the Caulobacter crescentus genome.* PLoS Genet, 2014. **10**(7): p. e1004463.
28. Zhou, B., et al., *The global regulatory architecture of transcription during the Caulobacter cell cycle.* PLoS Genet, 2015. **11**(1): p. e1004831.
29. Grant, C.E. and T.L. Bailey, *XSTREME: Comprehensive motif analysis of biological sequence datasets.* bioRxiv, 2021.
30. Bharmal, M.H., J.R. Aretakis, and J.M. Schrader, *An Improved Caulobacter crescentus Operon Annotation Based on Transcriptome Data.* Microbiol Resour Announc, 2020. **9**(44).
31. Babu, M.M. and S.A. Teichmann, *Functional determinants of transcription factors in Escherichia coli: protein families and binding sites.* Trends Genet, 2003. **19**(2): p. 75-9.
32. Mangan, S. and U. Alon, *Structure and function of the feed-forward loop network motif.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 11980-5.
33. da Silva, C.A., et al., *SpdR, a response regulator required for stationary-phase induction of Caulobacter crescentus cspD.* J Bacteriol, 2010. **192**(22): p. 5991-6000.
34. Britos, L., et al., *Regulatory response to carbon starvation in Caulobacter crescentus.* PLoS One, 2011. **6**(4): p. e18179.
35. Biondi, E.G., et al., *Regulation of the bacterial cell cycle by an integrated genetic circuit.* Nature, 2006. **444**(7121): p. 899-904.
36. Boutte, C.C. and S. Crosson, *The complex logic of stringent response regulation in Caulobacter crescentus: starvation signalling in an oligotrophic environment.* Mol Microbiol, 2011. **80**(3): p. 695-714.
37. Kaatz, G.W., C.E. DeMarco, and S.M. Seo, *MepR, a repressor of the Staphylococcus aureus MATE family multidrug efflux pump MepA, is a substrate-responsive regulatory protein.* Antimicrob Agents Chemother, 2006. **50**(4): p. 1276-81.
38. Hong, M., et al., *Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family.* Mol Cell, 2005. **20**(1): p. 131-41.
39. Dolan, K.T., E.M. Duguid, and C. He, *Crystal structures of SlyA protein, a master virulence regulator of Salmonella, in free and DNA-bound states.* J Biol Chem, 2011. **286**(25): p. 22178-85.
40. Bervoets, I. and D. Charlier, *Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology.* FEMS Microbiol Rev, 2019. **43**(3): p. 304-339.
41. Fiebig, A., et al., *A cell cycle and nutritional checkpoint controlling bacterial surface adhesion.* PLoS Genet, 2014. **10**(1): p. e1004101.
42. Berne, C., et al., *Feedback regulation of Caulobacter crescentus holdfast synthesis by flagellum assembly via the holdfast inhibitor HfiA.* Mol Microbiol, 2018. **110**(2): p. 219-238.
43. Lori, C., et al., *A Single-Domain Response Regulator Functions as an Integrating Hub To Coordinate General Stress Response and Development in Alphaproteobacteria.* mBio, 2018. **9**(3).
44. Crosson, S., et al., *Conserved modular design of an oxygen sensory/signaling network with species-specific output.* Proc Natl Acad Sci U S A, 2005. **102**(22): p. 8018-23.

45. Osamura, T., et al., *Specific expression and function of the A-type cytochrome c oxidase under starvation conditions in Pseudomonas aeruginosa.* PLoS One, 2017. **12**(5): p. e0177957.
46. Abel, S., et al., *Bi-modal distribution of the second messenger c-di-GMP controls cell fate and asymmetry during the caulobacter cell cycle.* PLoS Genet, 2013. **9**(9): p. e1003744.
47. Christen, M., et al., *DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in Caulobacter crescentus.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 4112-7.
48. Kalir, S., S. Mangan, and U. Alon, *A coherent feed-forward loop with a SUM input function prolongs flagella expression in Escherichia coli.* Mol Syst Biol, 2005. **1**: p. 2005 0006.
49. Mangan, S., A. Zaslaver, and U. Alon, *The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks.* J Mol Biol, 2003. **334**(2): p. 197-204.
50. Swem, L.R., et al., *Signal transduction by the global regulator RegB is mediated by a redox-active cysteine.* EMBO J, 2003. **22**(18): p. 4699-708.
51. Wu, J., et al., *RegB kinase activity is repressed by oxidative formation of cysteine sulfenic acid.* J Biol Chem, 2013. **288**(7): p. 4755-62.
52. Wu, J. and C.E. Bauer, *RegB kinase activity is controlled in part by monitoring the ratio of oxidized to reduced ubiquinones in the ubiquinone pool.* mBio, 2010. **1**(5).
53. Swem, L.R., et al., *Identification of a ubiquinone-binding site that affects autophosphorylation of the sensor kinase RegB.* J Biol Chem, 2006. **281**(10): p. 6768-75.
54. Maniatis, T., E.F. Fritsch, and J. Sambrook, *Molecular cloning : a laboratory manual*. 1982, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory. x, 545 p.
55. Ely, B., *Genetics of Caulobacter crescentus.* Methods Enzymol, 1991. **204**: p. 372-84.
56. Wetmore, K.M., et al., *Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons.* mBio, 2015. **6**(3): p. e00306-15.
57. Kent, W.J., *BLAT--the BLAST-like alignment tool.* Genome Res, 2002. **12**(4): p. 656-64.
58. Thanbichler, M., A.A. Iniesta, and L. Shapiro, *A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in Caulobacter crescentus.* Nucleic Acids Res, 2007. **35**(20): p. e137.
59. Kaczmarczyk, A., J.A. Vorholt, and A. Francez-Charlot, *Cumate-inducible gene expression system for sphingomonads and other Alphaproteobacteria.* Appl Environ Microbiol, 2013. **79**(21): p. 6795-802.
60. Pettersen, E.F., et al., *UCSF ChimeraX: Structure visualization for researchers, educators, and developers.* Protein Sci, 2021. **30**(1): p. 70-82.
61. Zhu, L.J., et al., *ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data.* BMC Bioinformatics, 2010. **11**: p. 237.



745

746 **Figure 1. Adhesion TCS pathway and adhesion profiling of *lovK*$_{H180A}$::Tn mutants. A)**
747 Schematic of the LovK-SkaH-SpdS-SpdR adhesion TCS system that regulates holdfast synthesis
748 as described by Reyes-Ruiz et al. [18]. Question marks indicate postulated additional regulator(s)
749 in the adhesion control pathway. Dashed lines indicate post-transcriptional regulation and solid
750 lines indicate transcriptional regulation. Black arrows indicate activation and red bar-ended lines
751 indicate repression. **B)** Genetic selection to identify insertions that disrupt the hyper-holdfast
752 phenotype of *lovK*$_{H180A}$. Tn-*himar* strains are cultivated and serially passaged for five days in the
753 presence of cheesecloth (black cross-hatched lines in the center of the well). Mutants that do not
754 permanently adhere to cheesecloth are increasingly enriched in the supernatant with each
755 passage. Darker yellow color indicates non-adhesive *lovK*$_{H180A}$::Tn strains that are enriched after
756 five days of serial passaging. **C)** (left) List of the 25 genes for which transposon insertion has the
757 largest disruptive effect on adhesion of the *lovK*$_{H180A}$ strain. (right) Enrichment in the supernatant
758 is reflected in an increasing calculated fitness score with each daily passage. Mutations that
759 disrupt *lovK*$_{H180A}$ adhesion to cheesecloth include the expected holdfast synthesis and
760 modification genes (pink), and genes encoding the LovK-SkaH-SpdS-SpdR regulatory system
761 (blue). The hypothetical gene *CCNA_00551* is listed in black; all remaining genes are colored
762 grey.

763

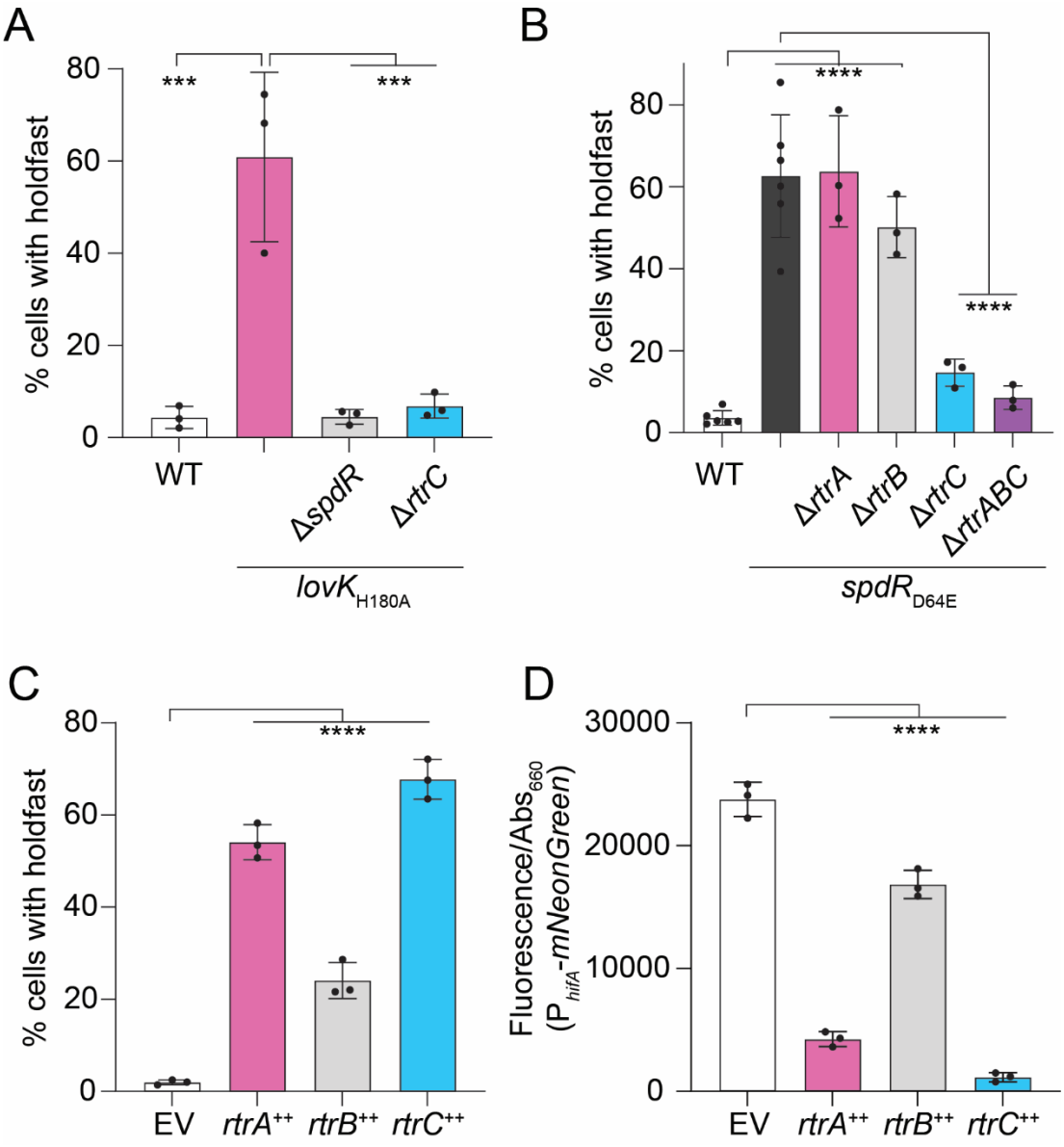


Figure 2. *CCNA_00551* (*rtrC*) regulates holdfast synthesis and *hfiA* expression. A) Percentage of cells with stained holdfast in wild type (WT) or *lovK*$_{H180A}$ strains bearing in-frame deletions (Δ) in *spdR* and *CCNA_00551* (*rtrC*). **B)** Percentage of cells with stained holdfast in WT, *spdR*$_{D64E}$, or *spdR*$_{D64E}$ strains bearing in-frame deletions in *rtrA*, *rtrB*, and *rtrC*. **C)** Percentage of cells with stained holdfasts in *rtrA*, *rtrB*, and *rtrC* overexpression (++) backgrounds. Stained holdfast were quantified by fluorescence microscopy. **D)** *hfiA* transcription in *rtrA*, *rtrB*, and *rtrC* overexpression (++) backgrounds as measured using a P$_{hfiA}$-*mNeonGreen* fluorescent reporter. Fluorescence was normalized to cell density; data show the mean fluorescence. Error bars are

773 standard deviation of three biological replicates, except WT and *spdR*$_{D64E}$ in panel B, which have
774 six biological replicates. Statistical significance was determined by one-way ANOVA followed by
775 **A-B)** Tukey's multiple comparisons test or **C-D)** Dunnett's multiple comparison (p-value $\leq$
776 0.001,***; p-value $\leq$ 0.0001,****).
777

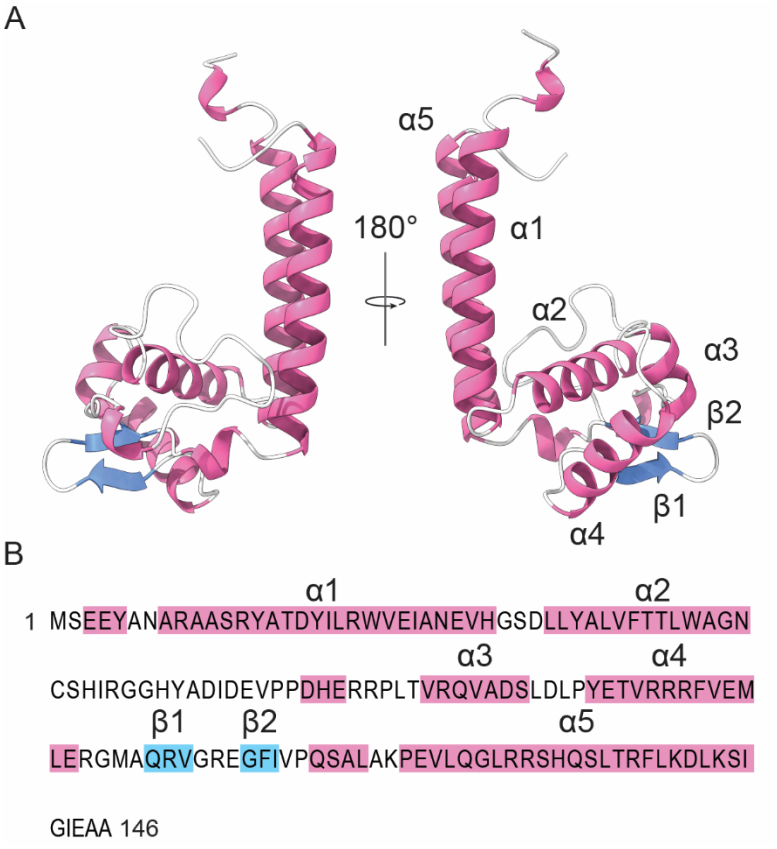


Figure 3. Structural analysis of RtrC A) Tertiary structure of RtrC predicted by AlphaFold [24]. Pink ribbons indicate alpha (α) helices and blue arrows indicate beta (β) strands. Labels (α1-5 and β1-2) correspond to the sequence highlighted in panel B. **B)** Primary structure of RtrC. Amino acids highlighted in pink are in predicted α-helices and residues highlighted in blue are in β-strands, as labeled above the sequence.

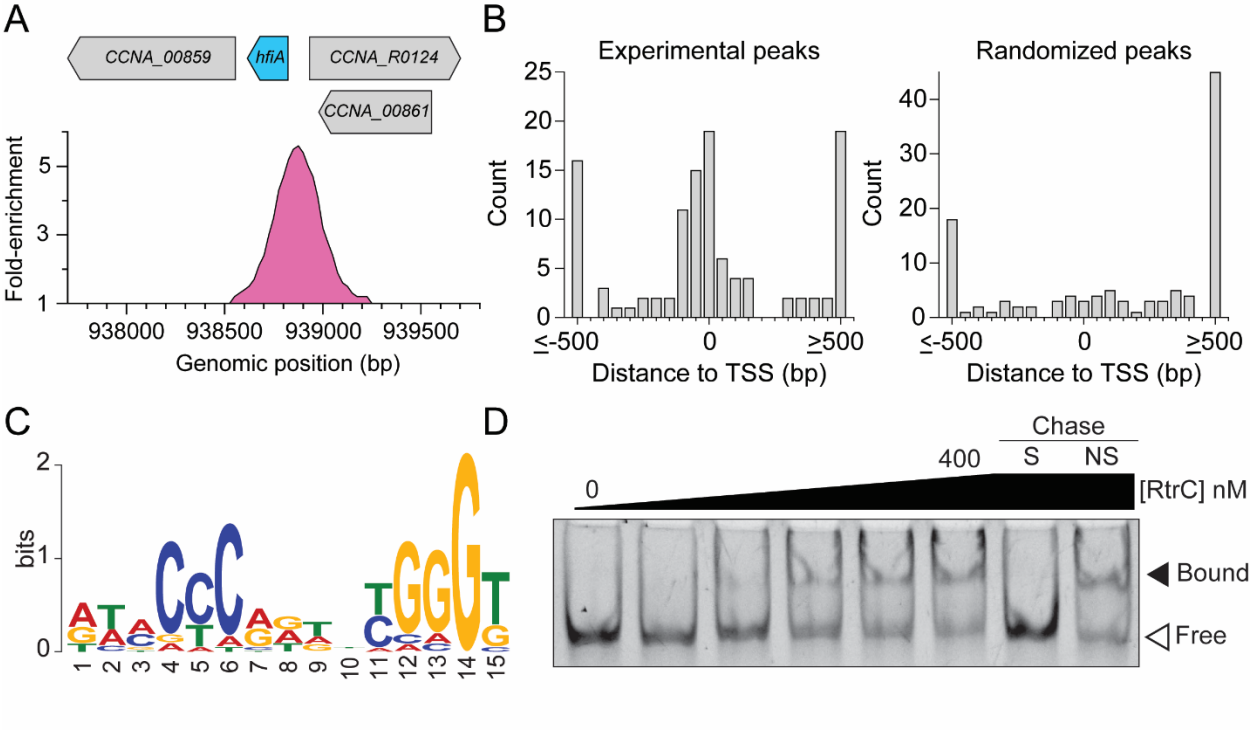


Figure 4. RtrC binds DNA *in vivo* and *in vitro*. A) RtrC binds the *hfiA* promoter *in vivo*. ChIP-seq profile from RtrC-3xFLAG pulldowns were plotted as fold-enrichment in read counts compared to the input control. Genomic position of the binding peak (in pink) on the *C. crescentus* chromosome and relative gene locations are marked. **B)** Distribution of RtrC peaks relative to experimentally defined transcription start sites (TSS). Distance from summit of RtrC ChIP-seq peak or randomized peaks to the nearest TSS (113 peaks) were analyzed and are plotted as a histogram. **C)** DNA sequence motif enriched in RtrC ChIP-seq peaks identified by XSTREME [29]. **D)** Electrophoretic mobility shift assay using purified RtrC and *hfiA* promoter sequence. Increasing concentrations of purified RtrC (0, 50, 100, 200, 300, and 400 nM) were incubated with 6.25 nM labeled *hfiA* probe. Specific chase (S) and non-specific chase (NS) contained 2.5 μM unlabeled *hfiA* probe and unlabeled shuffled *hfiA* probe, respectively. Blot is representative of two biological replicates.

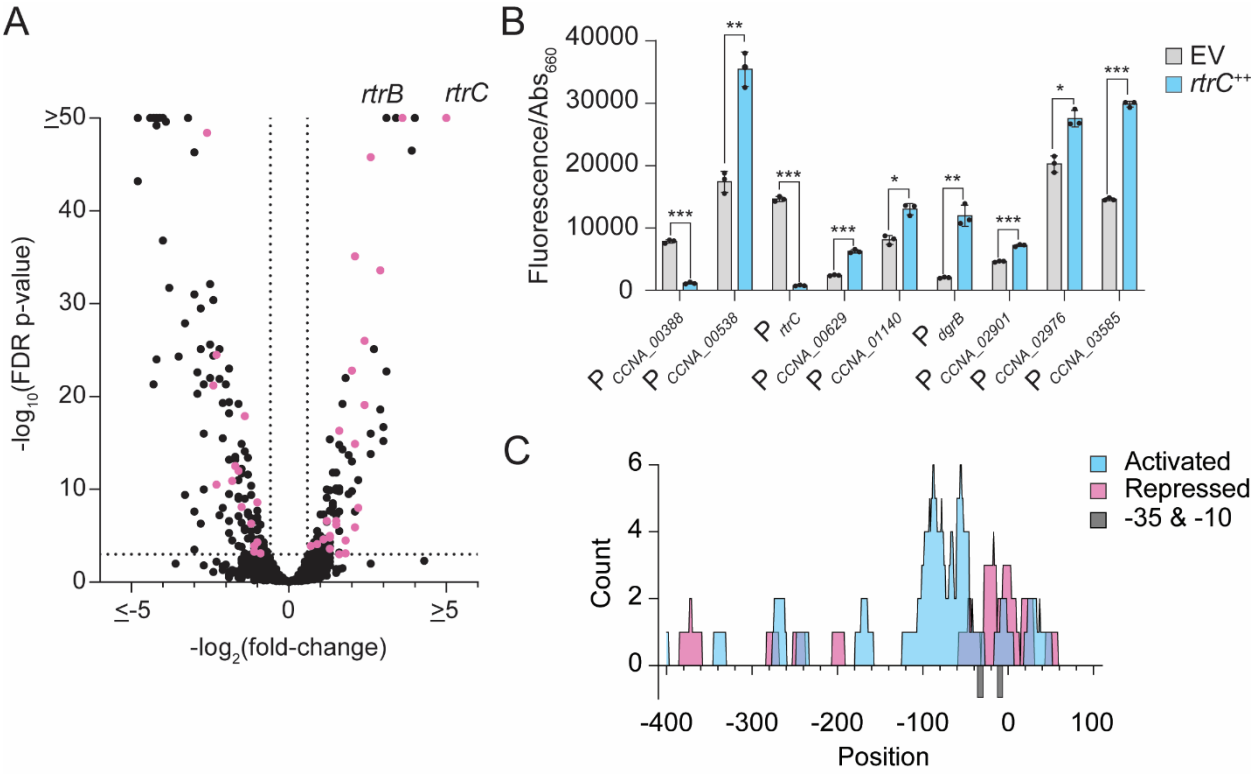


Figure 5. RtrC functions as a transcriptional activator and repressor. A) RNA-seq analysis of genes significantly regulated upon *rtrC* overexpression. Volcano plot showing $\log_2$(fold-change) in transcript levels in an *rtrC* overexpression strain (*rtrC*$^{++}$) versus empty vector (EV) are plotted against -$\log_{10}$(FDR p-value). Black dots indicate genes without RtrC motifs and pink dots indicate genes with RtrC motifs in their promoters. Data show the mean of three biological replicates. **B)** Transcription from predicted RtrC-regulated promoters measured by promoter fusions to mNeonGreen. Fluorescence measured in *rtrC*$^{++}$ or empty vector (EV) backgrounds was normalized to cell density (OD_{660}). Data show the mean signal; error bars are standard deviation of three biological replicates. Statistical significance was determined by multiple unpaired t tests, correcting for multiple comparisons using the Holm-Šídák method (p-value $\leq$ 0.05,*; p-value $\leq$ 0.01,**; p-value $\leq$ 0.001,***). **C)** Activity of RtrC as a transcriptional activator or repressor correlates with position of the pseudo-palindromic RtrC motif in the promoter. Distribution of RtrC motifs in promoters (-400 to +100 bp from the transcription start site; TSS) directly regulated by RtrC. Blue indicates motif position in promoters activated by RtrC (n=26) and pink indicates motif

814 position in promoters repressed by RtrC (n=16). Grey bars below the x-axis indicate the -35 and

815 -10 positions relative to the annotated TSS.

816

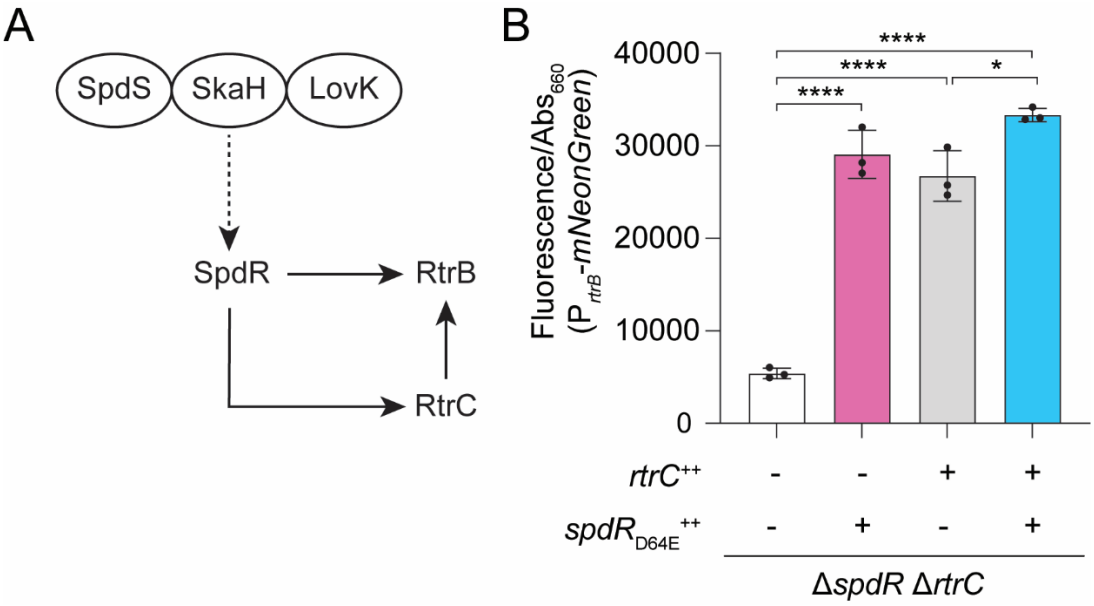


Figure 6. *spdR-rtrC-rtrB* form an OR-gated type I coherent feedforward loop. A) Schematic of the Type I coherent feedforward loop. The sensor histidine kinases SpdS, SkaH, and LovK function upstream and regulate the DNA-binding response regulator, SpdR [18]. SpdR can activate transcription of both *rtrC* and *rtrB*; RtrC activates transcription of *rtrB*. Dashed arrows indicate post-transcriptional activation and solid arrows indicate transcriptional activation. **B)** *rtrB* transcription measured using a P_{rtrB}-*mNeonGreen* transcriptional reporter; transcription measured upon *rtrC* and/or $spdR_{D64E}$ overexpression (++). Reporter strains were built in a background in which *rtrC* and *spdR* are deleted from the chromosome (Δ*spdR* Δ*rtrC*). Fluorescence signal was normalized to cell density. Data are the mean; errors bars represent standard deviation of three biological replicates. Statistical significance was determined by one-way ANOVA followed by Tukey's multiple comparisons test (p-value $\leq$ 0.05,*; p-value $\leq$ 0.0001,****).

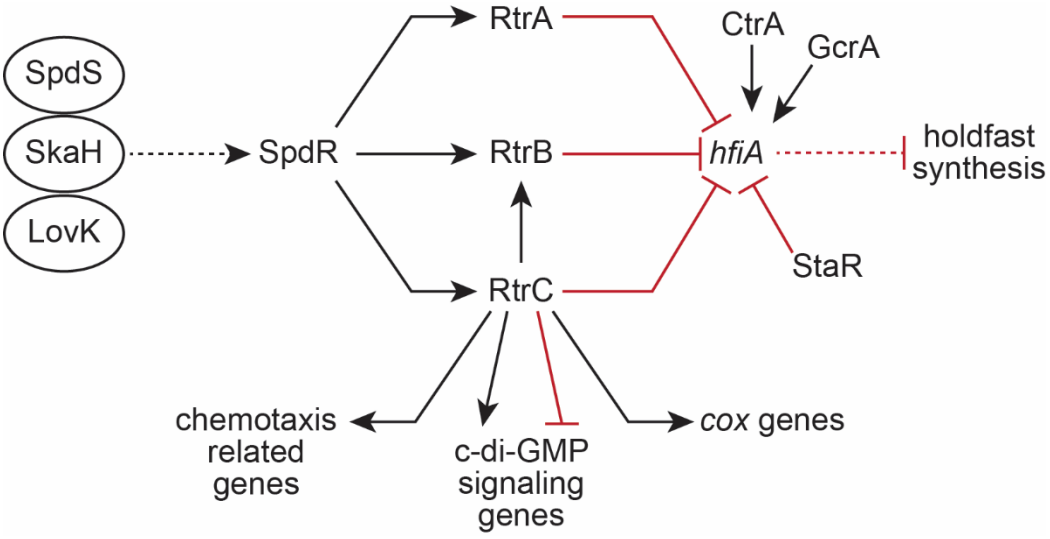
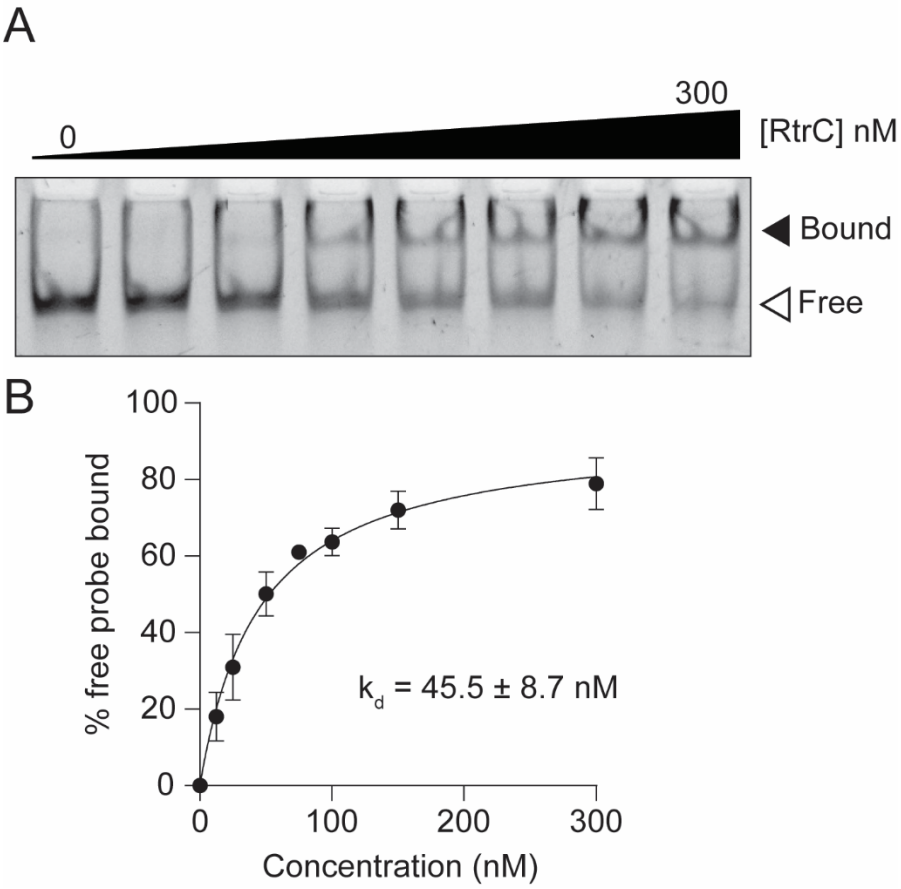


Figure 7. *C. crescentus* adhesion TCS regulatory network. The LovK, SkaH, and SpdS sensor histidine kinases function upstream of the DNA-binding response regulator, SpdR [18]. SpdR directly activates transcription of *rtrA, rtrB,* and *rtrC*. RtrA, RtrB, RtrC, and the XRE-family transcription factor, StaR, all directly repress *hfiA* transcription, while the cell cycle regulators CtrA and GcrA directly activate *hfiA* transcription. In addition to regulating *hfiA,* RtrC can function as a transcriptional activator and repressor for several groups of genes, including those with predicted roles in chemotaxis, c-di-GMP signaling, and aerobic respiration (*cox*). Dashed lines indicate post-transcriptional regulation, solid black arrows indicate transcriptional activation, and red bar-ended lines indicate transcriptional repression.



841

842 **Figure S1. RtrC binds DNA *in vitro*. A)** Electrophoretic mobility shift assay (EMSA) using purified
843 RtrC. Increasing concentrations of purified RtrC (0, 12.5, 25, 50, 75, 100, 150, and 300 nM) were
844 incubated with 6.25 nM labeled *hfiA* probe. Blot is representative of three biological replicates. **B)**
845 RtrC DNA binding curve derived from triplicate EMSA data. k_d was calculated based on
846 assumption of one site specific binding.

847

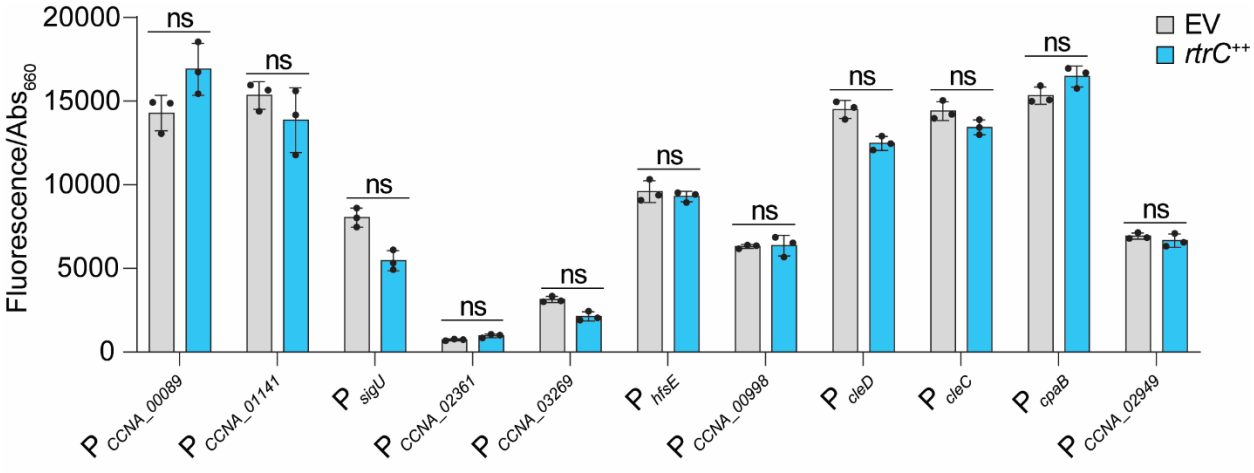


Figure S2. Genes that contain an RtrC motif in their promoters but that are not differentially regulated by *rtrC* overexpression. There are several genes that are not regulated by *rtrC* overexpression in the RNA-seq dataset (Table S2) despite the presence of an RtrC binding site in their promoter**.** To confirm this result, transcription from these genes was measured using P_{gene}-*mNeonGreen* transcriptional fusion reporters in a strain overexpressing *rtrC* (*rtrC*$^{++}$). Fluorescence was normalized to cell density (OD_{660}). Data show the mean; error bars represent standard deviation of three biological replicates. Statistical significance was determined by multiple unpaired t tests, correcting for multiple comparisons using the Holm-Šídák method (ns – not significant).

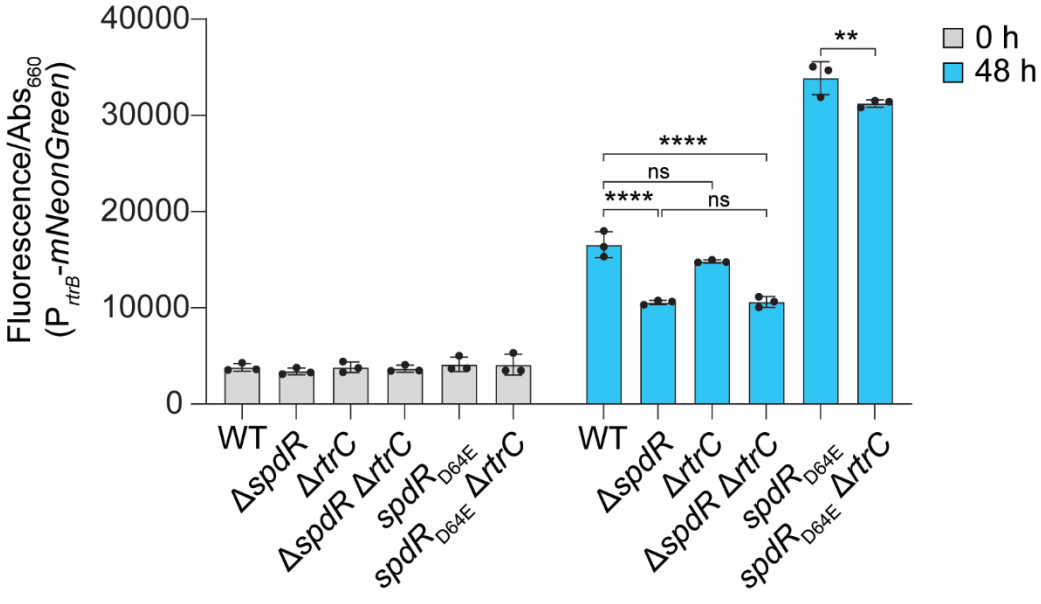


Figure S3. *rtrB* expression in logarithmic versus stationary phase. *rtrB* transcription can be activated by both SpdR and RtrC (see Figure 6). *rtrB* transcription was measured using a P_{rtrB}-*mNeonGreen* transcriptional fusion reporter in a wild type (WT) or *spdR*$_{D64E}$ background with in-frame deletions (Δ) in *spdR* and/or *rtrC* in early logarithmic phase cultures (0 h) and in stationary phase (48 h). Fluorescence measurements were normalized to cell density (OD_{660}). Data are the mean; errors bars represent standard deviation of three biological replicates. Statistical significance was determined by Two-way ANOVA followed by Tukey's multiple comparisons within each time point (p-value $\leq$ 0.01,**; p-value $\leq$ 0.0001,****; ns – not significant).

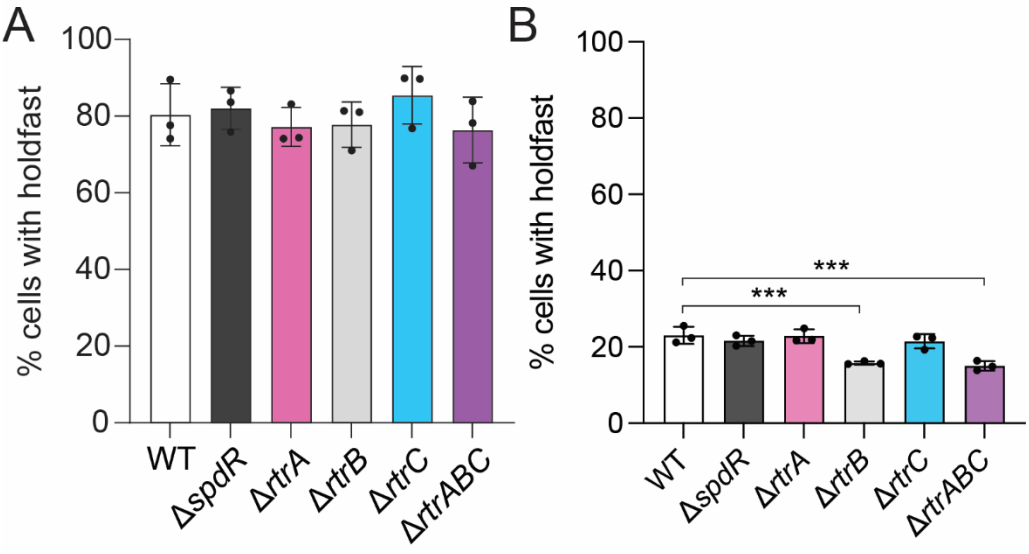


Figure S4. Regulation of holdfast synthesis in complex media as a function of growth phase. A-B) Percentage of cells with stained holdfast in wild type (WT) and in strains bearing in-frame deletions (Δ) in *spdR*, *rtrA*, *rtrB*, and *rtrC*. Holdfast counts were performed on cultures grown in complex medium (PYE) in **A)** early log phase or **B)** stationary phase (after 24 hours of growth). Data show the mean holdfast percentage; error bars are standard deviation of three biological replicates. Statistical significance was determined by one-way ANOVA followed by Dunnett's multiple comparison (p-value ≤ 0.001,***).

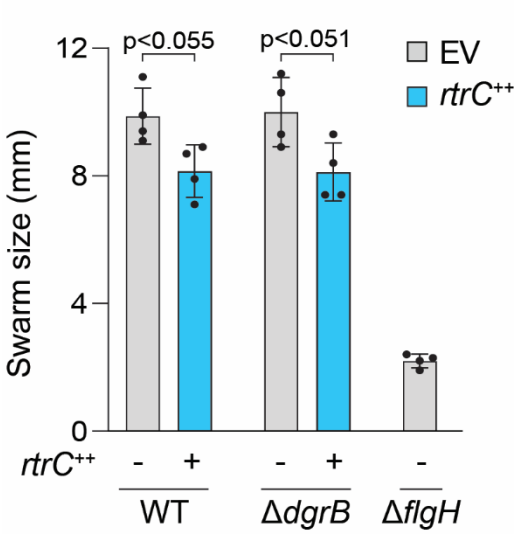


Figure S5. *rtrC* overexpression reduces motility in soft agar. Motility of strains overexpressing *rtrC* in soft agar. Wild type (WT) and strains bearing in-frame deletions (Δ) of *dgrB* and *flgH* are shown. Δ*flgH* does not assemble a flagellum and is shown as a zero motility control. *rtrC*$^{++}$ is expressed from pPTM059 (shown in blue); pPTM059 alone is the empty vector control (EV). Swim diameter was measured after 3 days of incubation. Expression of *rtrC* was induced from a cumate-regulated promoter with 50 μM cumate. Data presented are means and associated standard deviations from four biological replicates. Statistical analysis was multiple unpaired t tests, correcting for multiple comparisons using the Holm-Šídák method; p-values are marked above the bars that were compared.