Cell cycle progression in *Caulobacter* requires a nucleoid-associated protein with high AT sequence recognition

Dante P. Ricci[a,1], Michael D. Melfi[a,b,1], Keren Lasker[a], David L. Dill[c], Harley H. McAdams[a], and Lucy Shapiro[a,2]

[a]Department of Developmental Biology, Stanford University, Stanford, CA 94305; [b]Department of Chemistry, Stanford University, Stanford, CA 94305; and [c]Department of Computer Science, Stanford University, Stanford, CA 94305

Contributed by Lucy Shapiro, August 3, 2016 (sent for review June 3, 2016; reviewed by James W. Gober and William Margolin)

Faithful cell cycle progression in the dimorphic bacterium *Caulobacter crescentus* requires spatiotemporal regulation of gene expression and cell pole differentiation. We discovered an essential DNA-associated protein, GapR, that is required for *Caulobacter* growth and asymmetric division. GapR interacts with adenine and thymine (AT)-rich chromosomal loci, associates with the promoter regions of cell cycle-regulated genes, and shares hundreds of recognition sites in common with known master regulators of cell cycle-dependent gene expression. GapR target loci are especially enriched in binding sites for the transcription factors GcrA and CtrA and overlap with nearly all of the binding sites for MucR1, a regulator that controls the establishment of swarmer cell fate. Despite constitutive synthesis, GapR accumulates preferentially in the swarmer compartment of the predivisional cell. Homologs of GapR, which are ubiquitous among the α-proteobacteria and are encoded on multiple bacteriophage genomes, also accumulate in the predivisional cell swarmer compartment when expressed in *Caulobacter*. The *Escherichia coli* nucleoid-associated protein H-NS, like GapR, selectively associates with AT-rich DNA, yet it does not localize preferentially to the swarmer compartment when expressed exogenously in *Caulobacter*, suggesting that recognition of AT-rich DNA is not sufficient for the asymmetric accumulation of GapR. Further, GapR does not silence the expression of H-NS target genes when expressed in *E. coli*, suggesting that GapR and H-NS have distinct functions. We propose that *Caulobacter* has co-opted a nucleoid-associated protein with high AT recognition to serve as a mediator of cell cycle progression.

Caulobacter | nucleoid-associated protein | asymmetry | AT-rich | cell cycle

The organization of chromosomal DNA in prokaryotes, which lack structural equivalents of histones, is attributed to a diverse family of nucleoid-associated proteins (NAPs) (1). Although a growing body of evidence points to a physical chromosome that is ordered into regular, structured domains in bacteria (2–6), very little is known about the establishment and dynamics of these domains or the mechanism(s) by which NAPs integrate chromosome topology and gene regulation.

The topological problem of gene expression is compounded by the existence of cell type-specific genetic programs that underlie differentiation and specialization. The bacterium *Caulobacter crescentus* uses cell type-specific genetic programs to regulate growth, differentiation, and cell cycle progression (7). During each round of the cell cycle, *Caulobacter* divides asymmetrically to yield a stalked cell, which immediately initiates a new round of DNA replication and cell cycle progression, and a motile swarmer cell, which does not reinitiate the cell cycle until it differentiates into a stalked cell (Fig. 1). Each morphotype expresses a unique complement of genes that encode cell type-specific functions (e.g., motility, chemotaxis, encapsulation, or chromosome replication), the regulation of which can be attributed, at least in part, to an ensemble of interdependent, spatiotemporally restricted global transcriptional regulators that are organized into a highly robust control circuit (7–15). However, the control of more than 40% of cell cycle-regulated promoters cannot be accounted for by these master regulators (14), implying the existence of additional regulatory proteins that influence development through interaction(s) with the *Caulobacter* genome. Such factors could include NAPs, which can directly or indirectly modulate gene regulation (1). NAPs with broad roles in the regulation of chromosome dynamics and condensation (e.g., HU, IHF, and SMC) have been identified in *Caulobacter* (4, 16–18).

We report the discovery of GapR, a NAP that globally interacts with the GC-rich *Caulobacter* genome at high adenine and thymine (AT) loci. We demonstrate through ChIP-seq analysis that GapR associates with promoters of cell cycle-controlled genes bound by master transcriptional regulators, including more than 90% of sites bound by MucR1 (12), a recently identified transcription factor that directs the establishment of swarmer cell identity. The depletion or overexpression of GapR (growth-associated A/T-binding protein involved in regulation) causes pleiotropic defects in cell growth and division. We provide evidence for the asymmetric accumulation of GapR in the swarmer compartment of predivisional cells despite constitutive GapR synthesis across the cell cycle, implying the presence of a posttranslational mechanism for asymmetric distribution of this essential DNA-associated protein. Conversely, the histone-like nucleoid structuring protein from *Escherichia coli* (H-NS), which also exhibits affinity for regions of high AT (19), does not localize preferentially to the swarmer compartment

Significance

In all organisms, morphological and functional diversity is the product of cell type-specific genetic programs. Asymmetric cell division in *Caulobacter* yields daughter cells that differ functionally due to the differential read-out of their genomes. Here, we report the discovery of GapR, a conserved DNA-binding protein required for cell cycle progression. We show that GapR only associates with DNA sequences of high adenine and thymine (AT) content, colocating with cell cycle master regulators that control genes mediating swarmer cell development. GapR protein distributes asymmetrically, accumulating on the compacted chromosome of the daughter swarmer cell compartment prior to division. We argue that *Caulobacter* has co-opted a protein that associates with AT-rich DNA to provide spatial control during an asymmetric cell division.

Author contributions: D.P.R., M.D.M., K.L., D.L.D., H.H.M., and L.S. designed research; D.P.R., M.D.M., and K.L. performed research; D.P.R., M.D.M., K.L., and H.H.M. analyzed data; and D.P.R., M.D.M., K.L., and L.S. wrote the paper.

Reviewers: J.W.G., University of California, Los Angeles; and W.M., University of Texas Medical School.

The authors declare no conflict of interest.

Data deposition: Sequence data have been deposited to the Sequence Read Archive (SRA) (accession nos. SRR4076697–SRR4076704).

[1]D.P.R. and M.D.M. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: shapiro@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612579113/-/DCSupplemental.

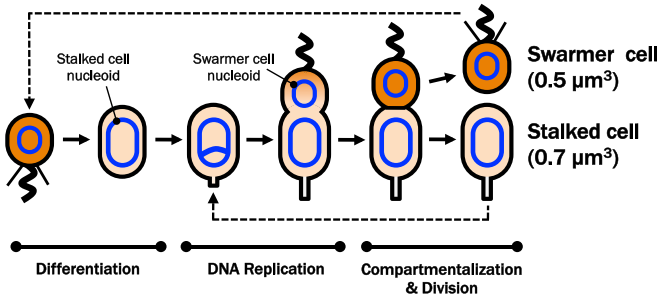

Fig. 1. The *C. crescentus* cell cycle. *Caulobacter* exists as one of two independent morphotypes, the swarmer cell and stalked cell, which are differently sized and which have distinct polar appendages. The swarmer cell is unable to initiate chromosome replication and does not grow or divide. Following differentiation, the stalked cell initiates DNA replication and segregation of DNA. As the cell cycle progresses, an asymmetric predivisional cell arises that elaborates a flagellum at the nascent swarmer cell pole, which forms opposite the stalked pole. Completion of chromosome segregation is followed by the compartmentalization of a small swarmer and large stalked cell cytoplasmic space and ultimately by complete cytokinesis. Each asymmetric division event yields two cells with identical genomes but significantly different cytoplasmic volumes into which those genomes must be packaged.

of predivisional *Caulobacter* cells. H-NS is also functionally distinct from GapR; whereas H-NS is dispensable and silences promoters of high AT content in *E. coli*, GapR associates with expressed genes and is essential for *Caulobacter* viability. The work presented here establishes a link between the recognition of chromosomal sites of high AT content and swarmer cell-specific functions through an essential NAP and suggests that GapR function may impact the establishment and/or maintenance of cell type-specific regulatory programs controlling *Caulobacter* cell cycle progression.

Results

Identification of Putative Essential DNA-Associated Proteins in *Caulobacter*. Top-level control of the *Caulobacter* cell cycle is achieved through the coordinated activity of transcription factors that temporally restrict the expression of genes with cell type-specific functions (20, 21). To identify additional factors that have an essential role in cell cycle control, we mined the list of essential genes identified by global Tn-seq analysis of the *Caulobacter* genome (22) and characterized those proteins predicted to have DNA-binding activity. We used a combination of primary-sequence, secondary-structure, and domain-based annotation algorithms (*SI Materials and Methods*) to generate functional predictions for the encoded products. This analysis yielded a candidate list of putative essential DNA-binding proteins not previously identified in *Caulobacter*. Here, we describe the characterization of one of these candidates, CCNA_03428 (hereafter referred to as GapR), a small (89-residue) protein comprising a single domain of unknown function.

GapR has little sequence similarity to characterized proteins. The nearest characterized homologs are the DNA-binding transcriptional regulator DsbA from bacteriophage T4, the prokaryotic MerR-family transcriptional regulators, the nucleic-acid-binding arm of valyl-tRNA synthetase, and the coiled-coil domain of the eukaryotic intermediate filament protein vimentin (Table S1). Although the functions of these weak homologs imply DNA-binding activity, the function of GapR cannot be predicted from primary sequence alone given the paucity of high-confidence homologs of known function.

GapR Is Essential for Normal Growth and Cell Division in *Caulobacter*. We attempted to delete the *gapR* gene from the chromosome to test the predicted essentiality of *gapR* and to determine the consequences of its dysregulation, but were unable to isolate a stable Δ*gapR* strain. This finding is consistent with the prediction that *gapR* is an essential gene in *Caulobacter* (22). To determine whether the inability to construct a Δ*gapR* strain is indeed due to the essentiality of GapR, we attempted to construct a *gapR* depletion strain using standard techniques but found that *gapR* expression from either the P_{xylX} promoter or the P_{vanA} promoter did not restore a WT phenotype to a Δ*gapR* strain, indicating that *gapR* transcription needs to be precisely controlled for optimal fitness (*SI Text*, *Viable* ΔgapR *Mutants of* Caulobacter *NA1000 Cannot Be Isolated*).

To circumvent this problem, we constructed a strain in which *gapR* is expressed at native levels but GapR protein abundance is dynamically controlled by inducible regulated proteolysis. This regulation was accomplished by exploiting the species-specific activity of the *E. coli* SspB adapter protein, which targets mutant derivatives of C-terminally *ssrA*-tagged proteins for degradation in a highly specific manner (23). By tagging the native *gapR* allele in frame with a sequence encoding an *E. coli* SspB-dependent *ssrA* tag (-AANDENYSENYADAS) and placing the *sspB* gene from *E. coli* under control of the xylose-inducible P_{xylX} promoter, we generated a *Caulobacter* strain in which GapR is conditionally degraded by growing cells in the presence of xylose (Fig. 2*A*). This system enables the efficient clearance of GapR (Fig. S1*A*) and offers improved depletion kinetics compared with conventional methods involving inducible heterologous promoters (Fig. S1*B*) while preserving native transcriptional and translational regulation. The GapR "proteolytic depletion" strain grows in the absence of xylose (i.e., when expression of $sspB^{Ecoli}$ is silenced), but does not grow on media containing xylose, suggesting that SspB-dependent proteolysis of GapR is sufficient to cause lethality (Fig. S1*C*). Because the expression of $sspB^{Ecoli}$ in *Caulobacter* has no effect on viability (Fig. S1*C*), we conclude that *gapR* is an essential gene.

Using our proteolytic depletion strain, we determined the physiological consequences of induced GapR elimination. A culture of the GapR proteolytic depletion strain was grown to midexponential phase, incubated for 4 h in the presence or absence of xylose, and examined microscopically. Although cells grown in the absence of xylose (in which GapR was present) displayed normal shapes and sizes, cells grown in GapR-depleting conditions (i.e., + xylose) exhibited a variety of growth defects. Specifically, some cells became filamentous and formed conjoined vermiform minicell daughters that never separated from the mother (18% of 141 cells; Fig. 2*B*). Fluorescence microscopy of GapR-depleted cells stained with DAPI confirmed that these minicells contain DNA (Fig. S2). These phenotypic abnormalities indicate that GapR is critical for normal growth and cell division.

To determine whether *gapR* overexpression compromises *Caulobacter* growth, we expressed *gapR* from low-copy-number (pRMCS) or high-copy-number (pBMCS) replicating vectors under control of the native promoter (P_{gapR}). $gapR^+$ merodiploid strains carrying low-copy P_{gapR}-*gapR* could be constructed, but these strains formed small, slow-growing colonies. Viable strains carrying high-copy pBMCS::P_{xylX}-*gapR* or pBMCS::P_{gapR}-*gapR* constructs could not be obtained, indicating that GapR causes lethality *Caulobacter* when constitutively expressed at high levels.

We examined cells in which *gapR* was overexpressed from a low-copy-number plasmid either from its native promoter or an inducible P_{xylX} promoter (Fig. 2*C*). We found pleiotropic defects in *Caulobacter* cell cycle progression that included filamentation, morphological defects, and the formation of polar minicells (8% of 546 cells display polar minicells). These data demonstrate the requirement for tight regulation of GapR concentration in *Caulobacter* and suggest that GapR influences an array of critical cellular processes related to growth, division, and cell cycle progression.

GapR Globally Binds the *Caulobacter* Chromosome at AT-Rich Loci. Given the putative DNA-binding activity implied by our bioinformatic analysis of GapR, we wished to determine whether GapR associates with DNA and to simultaneously identify all

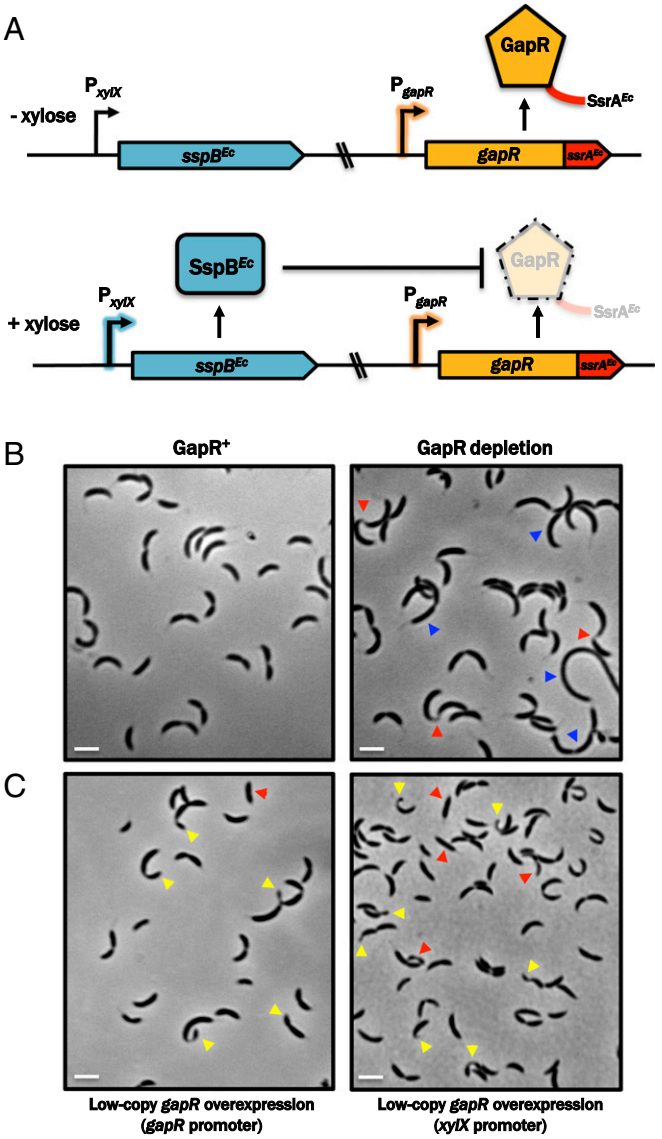

Fig. 2. Depletion and overexpression of GapR cause morphological defects. (*A*) Schematic of the *E. coli* SsrA/SspB-based inducible proteolysis system co-opted for the specific degradation of GapR in *Caulobacter*, which is tagged with an SspB-dependent *E. coli* SsrA tag ($SsrA^{Ec}$) and expressed at the native genomic locus to preserve native levels and regulation. In the absence of xylose, the *xylX* promoter is inactive and *sspB* is not expressed. Addition of xylose leads to production of *E. coli* SspB ($SspB^{Ec}$), which promotes ClpXP-dependent proteolysis of $SsrA^{Ec}$-tagged GapR. (*B*) The GapR proteolytic depletion strain described in *A* was grown to midexponential phase in M2G and then propagated for an additional 2 h in the absence (permissive condition; *Left*) or presence (depletion condition; *Right*) of 0.3% (wt/vol) xylose. GapR-depleted cells exhibit incomplete separation (red arrowheads) and filamentation (blue arrowheads). (Scale bar, 1 μm.) (*C*) *gapR* was overexpressed in WT *Caulobacter* NA1000 on low-copy plasmids constitutively from the native promoter (*Left*) or for 4 h from the xylose-inducible *xylX* promoter [0.3% (wt/vol) xylose; *Right*]. GapR overexpression leads to morphological defects (red arrowheads) and aberrant division events (yellow arrowheads). (Scale bar, 1 μm.)

potential sites of GapR occupancy in vivo through chromatin immunoprecipitation coupled to deep sequencing (ChIP-seq). To ensure selective and sensitive immunoprecipitation of GapR, we generated strains in which *gapR* is replaced at the native locus with an allele encoding an N- or C-terminally FLAG-tagged variant (FLAG-GapR or GapR-FLAG, respectively). Each of the resulting strains grew normally, indicating that the addition of an epitope tag to either terminus of GapR does not compromise its function. These strains, along with a WT control strain expressing untagged GapR, were grown to midexponential phase, incubated briefly with a crosslinking agent, and subjected to chromatin immunoprecipitation (ChIP) using an IP-grade α-FLAG antibody. GapR-bound DNA was purified and subjected to deep sequencing. As a control, we performed ChIP-seq on a strain expressing FLAG-tagged CtrA, encoded at its native locus (Dataset S1), and compared our CtrA ChIP-seq profile with that obtained previously using custom antisera specific for CtrA (12), which showed good concordance. The ChIP-seq profiles of the FLAG-GapR and GapR-FLAG strains were strongly correlated, with somewhat higher signal observed in the FLAG-GapR sample. ChIP-seq profiling revealed broad GapR occupancy of the *Caulobacter* chromosome (Fig. 3*A* and Dataset S2), with 599 peaks above background (q-value threshold = 0.0001; *SI Materials and Methods*). Although the *Caulobacter* genome is 91.6% genic, we predominantly observed GapR ChIP-seq peak summits in intergenic regions of the chromosome, particularly among the most significant peaks: 76.5% of the top 200 GapR peak summits were located within intergenic regions, demonstrating a clear propensity for GapR to bind between, rather than within, genes.

The *Caulobacter* genome exhibits a marked nucleotide bias with an overall GC content of ∼67% (24). However, intergenic regions of the otherwise GC-rich genome are punctuated frequently by local spikes in AT content that can reach 75% A/T. To determine whether GapR binding is correlated with a local bias in nucleotide content, we calculated the local GC content

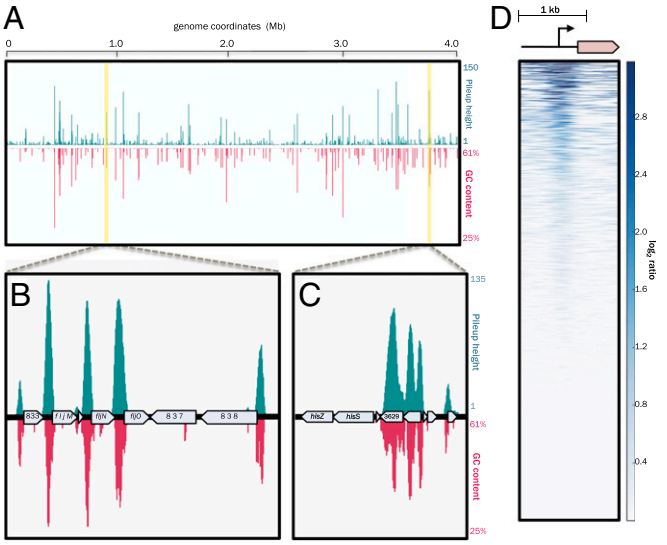

Fig. 3. GapR binds globally to AT-rich regions of the *Caulobacter* genome. (*A*) Genome-wide FLAG-GapR ChIP-seq profile (blue peaks) with read counts normalized to reads per kilobase per million (RPKM) plotted against local GC content (magenta peaks) as calculated using a 100-bp sliding window across the NA1000 genome. Regions highlighted in yellow are featured in greater detail to reveal intergenic (*B*) and intragenic (*C*) GapR ChIP-seq enrichment. Seventy-six percent of the top 200 GapR ChIP-seq peak summits lie within intergenic regions. (*D*) Enrichment of GapR peaks in promoter regions. Each row in the heatmap represents a *Caulobacter* ORF (stretched or compressed to 750 bp and oriented as shown in the gene cartoon above the plot such that the start codon occupies the same position in each row), as well as the 1,000 bp preceding the translation start site of each ORF (not stretched or compressed). Rows were subdivided into 10-bp bins, with each bin colored to reflect the degree of GapR ChIP-seq signal enrichment (bluer color = greater GapR enrichment). Finally, rows were sorted by maximum bin value ($\log_2$ of GapR ChIP-seq signal enrichment) from highest (*Top*) to lowest (*Bottom*). Heatmaps were generated using deepTools (*SI Materials and Methods*).

across the *Caulobacter* genome using sliding window analysis and measured the correlation between GapR occupancy and AT content. We indeed observed a striking positive correlation between GapR ChIP-seq signal strength and local AT content (Fig. 3 *A–C* and Fig. S3*A*), with an average GC content of only 45% among sequences comprising the top 100 GapR peaks.

To determine whether the AT-richness of GapR-associated sites reflects an AT-rich consensus motif or nonspecific affinity for high AT DNA, we mined the sequences under GapR peaks for a common motif that could represent a consensus sequence for GapR binding (25). We were unable to identify any sequence motif(s) common to more than 10% of the GapR-bound regions (Fig. S3*B*). This result, in light of the correlation between GapR ChIP-seq enrichment and local AT richness, suggests that the association of GapR with DNA is specified by local AT content rather than by an explicit consensus binding sequence. Indeed, DNA-associated proteins with generic affinity for AT-rich sequences have been described in various prokaryotic and eukaryotic organisms, including mammals (26–32).

GapR Associates with Promoters That Are Active During Normal Growth and Differentiation. Given its association with intergenic loci, we next looked for the intersection of GapR-bound loci with annotated genomic features (i.e., genes and promoters). We found that GapR primarily associates with the 5′ regulatory regions of genes, as indicated by the clustering of GapR peaks within a 200-bp window preceding annotated *Caulobacter* ORFs (Fig. 3*D*). To determine whether GapR associates with promoters that are normally active during growth, we measured overlap between the GapR ChIP-seq profile with that of the housekeeping sigma factor in *Caulobacter* (σ^{73}, i.e., RpoD) trapped on DNA in initiation complexes with target promoters as a consequence of rifampicin treatment (33). We observed a correlation between GapR and RpoD occupancy, with 80.4% of the top 500 GapR peaks overlapping an RpoD peak (Fig. 4*A*), indicating that GapR primarily associates with active promoters.

GapR ChIP-seq peaks were observed near an abundance of genes whose expression is known to vary over the course of the cell cycle. For example, GapR ChIP-seq signal is enriched at promoters for genes that encode structural proteins and regulators of the swarmer-specific polar appendages (e.g., *pilA*, *cpaA*, *flaFN*, *fljKLMNO*, *tipF*), a swarmer-specific protease (*perP*), a cell type-specific encapsulation system (*hvyA*, *pssZ*, CCNA_00162-168, CCNA_00465-472) (11), components of the divisome (*ftsZ*), and the replisome (*dnaA, dnaE*), as well as cell cycle signaling factors (e.g., *cckA*, *podJ, pleA*) and master regulators of the cell cycle (*ctrA*, *sciP, dnaA, gcrA*). Because the expression of each of these genes varies as a function of the cell cycle, we used transcriptomic data obtained from synchronous *Caulobacter* populations throughout the cell cycle to determine whether GapR-bound promoters commonly exhibit cell cycle-dependent activity. Specifically, we identified the set of *Caulobacter* genes defined as cell cycle-regulated by two independent groups (14, 34) and determined how frequently the transcription start sites (TSSs) for these genes are bound by GapR (based on ChIP-seq data, from mixed population). We found that GapR peaks overlap nearly one-third of all cell cycle-regulated promoters (31.6%, Pearson's χ^2 test, $P < 0.01$; Dataset S3), although cell cycle-regulated promoters are not, in general, more AT-rich than noncell cycle-regulated promoters (*SI Text*, *Cell Cycle-Regulated Promoters Are Not Especially AT-Rich*). These findings further suggesting a physiological connection between GapR and the regulation of cell cycle progression.

Many cell cycle-regulated genes are subject to combinatorial control by two or more master regulators of cell cycle progression (14). To determine whether GapR binds promoters that are also occupied by known master regulators, we compared the GapR genome occupancy profile with that of transcriptional regulators known to modulate target gene expression in a cell cycle-dependent manner. We observed overlap between GapR occupancy and ChIP-seq peaks corresponding to each of the known master regulators and found statistically significant pairwise overlap in occupancy between GapR and each of the regulators CtrA, MucR1, MucR2, and GcrA (Pearson's χ^2 test, $P < 2.2 \times 10^{-16}$; *SI Materials and Methods*, Fig. 4 *B* and *C*, and Fig. S4*A*). The intersection between GapR and MucR1 ChIP-seq peaks is particularly noteworthy, as 91% of all MucR1 peaks overlap a GapR peak (102/112; Fig. 4*B*).

CtrA and MucR1 together control the S → G1 transition by regulating the expression of many swarmer-specific genes (12). Although sites bound independently by either CtrA or MucR1 are distributed throughout the *Caulobacter* chromosome, those

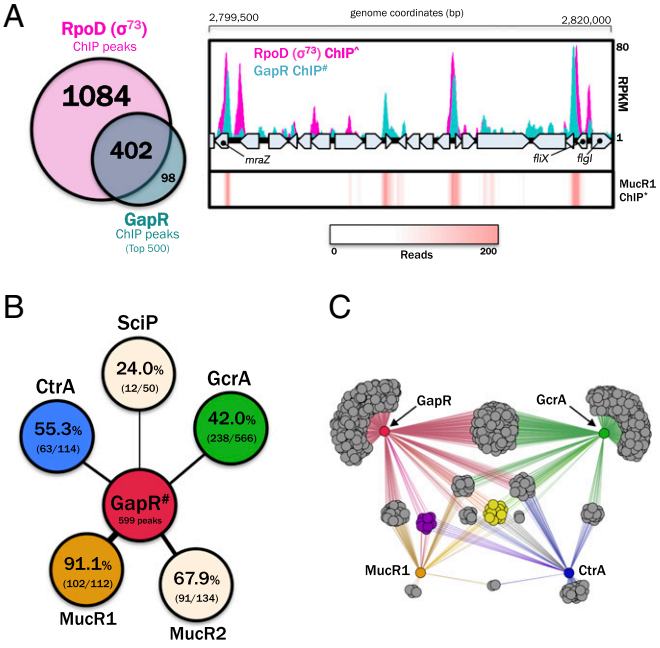

Fig. 4. GapR binds at active promoters controlled by master regulators of cell cycle progression. (*A*) Overlap between the top 500 FLAG-GapR ChIP-seq peaks and all RpoD ChIP-seq peaks (*Left*), with a specific genomic region featured (*Right*) to indicate the presence of shared binding sites between RpoD, GapR, and MucR1 (RpoD ChIP data obtained from ref. 33; MucR1 ChIP-seq data from ref. 12). MucR1 ChIP-seq signal is presented as a heatmap of piled reads. (*B*) ChIP-seq peak overlap between GapR and master transcriptional regulators bound to promoters of cell cycle-regulated genes. Percentages in the colored circles represent the proportion of ChIP-seq peaks for the indicated master regulator (e.g., CtrA and SciP) that intersect a GapR ChIP-seq peak (median peak width = 400 bp), with each fraction expressing the number of peaks that intersect a GapR binding site out of the total number of peaks identified for that TF. The length and thickness of the lines connecting the shapes reflect the degree of overlap in occupancy between GapR and the master TFs shown. MucR1/2 and SciP ChIP-seq data and peak calls were obtained from ref. 12; GcrA ChIP-seq data were obtained from ref. 33 and peaks were called using the same workflow used for GapR/CtrA (see *SI Materials and Methods* for details). (*C*) Network diagram of intergenic regions (gray, purple, and yellow nodes) associated with different proteins (red, green, blue, and orange nodes). Connections (lines) indicate the presence of a ChIP-seq peak connecting an intergenic region to one or more of the proteins. Intergenic regions are clustered with others that share the same combination of connections to given proteins. Three-way intersection of the three proteins MucR1, CtrA, and GapR is represented by two clusters of intergenic regions: the purple cluster (no GcrA peak associated) and the yellow cluster (GcrA peak also associated). The total number of observed three-way intersections in intergenic regions is greater than expected (under the assumption that ChIP-seq peaks of these three proteins are independently distributed over the chromosome; 21 observed vs. 11.2 expected; *SI Materials and Methods*). Genes associated with all of the chromosomal GapR-CtrA-MucR1 overlapping sites are listed in Table S2.

sites bound by both CtrA and MucR1 are almost entirely found in the origin-proximal half of the chromosome (binomial test, $P < 0.01$ for MucR1, $P < 0.05$ for CtrA; *SI Materials and Methods*). Further, GapR ChIP-seq peaks overlap a vast majority of these sites (24/26, 92.3%; Fig. S4*B* and Table S2). Overall, the overlap in occupancy of cell cycle-regulated promoters suggests that GapR is connected to the master regulatory circuit that drives cell cycle progression and cell type specification.

Minor Changes in Transcript Abundances Are Observed Following GapR Depletion or Overexpression. GapR exhibits widespread association with promoters and overlapping occupancy with known master transcriptional regulators; this is consistent with a role for this protein in regulating gene expression. To determine whether GapR is involved in transcriptional control, we selected individual genes whose promoters are occupied by GapR and quantitatively assessed transcript abundance following GapR overexpression or depletion. We found that overexpression (Fig. S5*A*) or depletion (Fig. S5*B*) of GapR leads to at least a twofold decrease or increase in expression, respectively, of candidate GapR target genes evaluated (*pilA,* CCNA_02831, CCNA_03711). To determine the global regulatory response to GapR perturbation, if any, we performed mRNA-seq following proteolytic clearance of GapR and conducted differential gene expression analysis compared with a GapR-replete strain to identify any GapR-dependent differences in transcription. In general, there are relatively few significant global changes in gene expression, with <1% of transcripts displaying differential expression following GapR depletion (Dataset S4). We identified 18 genes that that are differentially expressed on proteolytic depletion of GapR (Table S3); in 72% of these cases, the promoters of these genes overlap a GapR ChIP-seq peak. We also identified two genes that are differentially expressed up on mock depletion (in the *gapR*$^{+}$ mock depletion strain), but whose abundance changes in opposite directions between mock and true GapR depletion (Table S4). Additionally, we found that the vast majority of genes associated with GapR ChIP-seq peaks do not exhibit a greater than twofold change in expression following GapR depletion, suggesting that transcriptional regulation is an indirect or secondary function of GapR.

GapR Forms Compact Clusters Enriched in the Swarmer Compartment of the Predivisional Cell. The physical association between GapR occupancy and cell cycle-activated genes is somewhat surprising in light of time-resolved translational profiling experiments that show little variation in GapR synthesis over the *Caulobacter* cell cycle (~1.7-fold change) (35). To resolve this incongruity, we examined the subcellular distribution of GapR in vivo, reasoning that cell type-specific functions might be achieved through biased localization in lieu of "just-in-time" expression. We therefore generated fusions of GapR to the fluorescent protein mCherry and visualized the localization of the fusion protein in single cells (Fig. 5*A*). We observed in dividing *Caulobacter* cells a marked asymmetry in the intracellular distribution of mCherry-GapR, with the fluorescent signal found primarily in the swarmer compartment of at least half of the predivisional cells observed (Fig. 5*B*). This biased localization is not likely to be the result of swarmer compartment-specific expression of *gapR*, because GapR protein levels are roughly constant across the cell cycle in synchronous populations (Fig. S6).

Within single swarmer or stalked cells, we found that mCherry-GapR forms asymmetric foci in both cell types (Fig. 5*A*). In *Caulobacter*, and unlike in *E. coli*, the nucleoid fills the entire cytoplasmic space, as evidenced by the absence of DNA-free regions in cells stained with DNA-binding fluorescent dyes (16, 18); nevertheless, the fluorescent signal corresponding to GapR occupies only a fraction of the cytoplasm. The localized aggregation of GapR therefore suggests that GapR occupies regions of the

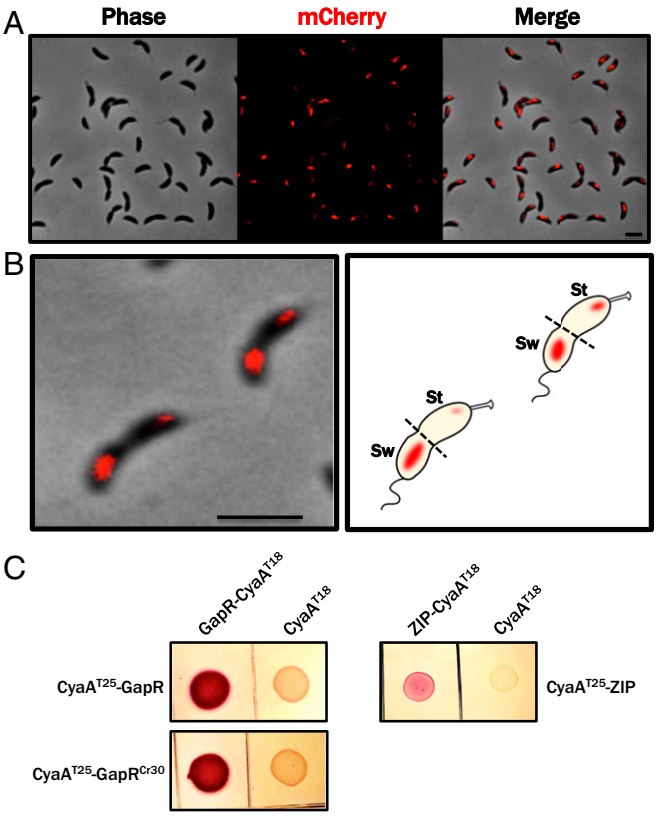

Fig. 5. GapR exhibits biased subcellular localization and self-associates in vivo. (*A*) WT *Caulobacter* NA1000 containing chromosomal *mCherry-gapR* under control of the P_{xylX} promoter was grown to midexponential phase in M2G, grown an additional 2 h in the presence of 0.3% xylose, and then imaged by phase contrast and epifluorescence microscopy. (Scale bar, 1 μm.) (*B*) Biased swarmer compartment localization of mCherry-GapR in predivisional cells, presented as a fluorescence/phase contract overlay (*Left*) and as a diagrammatic projection (*Right*). (Scale bar, 1 μm.) (*C*) Bacterial two-hybrid assay indicating direct GapR-GapR and GapR-GapR^{Cr30} interactions in vivo. Reconstitution of split adenylate cyclase (CyaA) activity, which implies a direct interaction between the domains fused to the T18 and T25 subunits of CyaA, is indicated by a Lac^{+} (red colony) phenotype on MacConkey agar. GapR or GapR^{Cr30} was fused in-frame to the C terminus of CyaA^{T25} and coexpressed with GapR-CyaA^{T18} or, as a negative control, CyaA^{T18} alone (*Right*). As a positive control, the self-associating leucine zipper domain of yeast GCN4 (ZIP) was fused to the T18 and T25 fragments of CyaA (*Left*).

Caulobacter chromosome that are particularly close to one another in 3D space.

GapR Self-Associates. The formation of H-NS clusters in *E. coli* requires both DNA binding and homo-oligomerization (36, 37). The clustering of GapR we observed in *Caulobacter* suggests that this NAP also multimerizes. Indeed, we observed that GapR is capable of self-association using a bacterial two-hybrid assay that reports on bimolecular complementation of *E. coli* adenylate cyclase (CyaA), an enzyme whose activity can be reconstituted when the two essential functional domains are, respectively, fused to each of two cytoplasmic proteins that directly interact (Fig. 5*C*, *Left*) (38).

GapR was fused in frame to each functional domain of adenylate cyclase (CyaA^{T18} and CyaA^{T25}), and complementary fusions were coexpressed in a strain of *E. coli* that cannot activate the cAMP-dependent reporter gene *lacZ* unless CyaA^{T18} and CyaA^{T25} are brought into close proximity through the direct interaction of their fusion partners (38). Coexpression of GapR-CyaA^{T18} with either GapR-CyaA^{T25} or CyaA^{T25}-GapR resulted in a

Lac$^+$ phenotype (Fig. 5*C*, *Right*), indicating that GapR can bind other GapR molecules. Although this assay does not reveal the stoichiometry of GapR multimers, we can infer that GapR is minimally capable of forming dimers in vivo; our results do not exclude the possibility that GapR forms higher-order oligomers in *Caulobacter*. Because auto-association of GapR was detected in an orthologous host (*E. coli*), we additionally conclude that these GapR-GapR interactions are direct and do not require additional *Caulobacter*-derived factors.

GapR Activity Is Conserved Among α-Proteobacteria. We identified 1,036 GapR homologs in sequenced genomes, 1,025 of which were found in the α-proteobacterial clade, indicating that GapR is essentially restricted to this lineage. Furthermore, GapR is nearly ubiquitous among the α-proteobacteria, with homologs identified in all but one of the free-living species in that clade. Indeed, GapR is a "signature" protein that is distinctive of the α-proteobacterial class (39–41). GapR homologs are also encoded within the genomes of several bacteriophages that infect α-proteobacterial hosts (including the *Caulobacter*-specific phage ΦCr30), suggesting that GapR is targeted or co-opted as a facet of the host-phage interaction.

To test the possibility that homologs of *Caulobacter* GapR exhibit a conserved function, we cloned GapR homologs from several divergent α-proteobacterial species (*Agrobacterium tumefaciens, Rhodobacter capsulatus,* and *Sinorhizobium meliloti*), as well as from the caulophage ΦCr30, and induced the expression of each in a WT *Caulobacter* strain. We found that strains overexpressing *gapR* homologs from each of these organisms led to the same morphological and cell division defects caused by overexpression of the native *gapR* gene (Fig. 6*A*), suggesting a conserved function and providing evidence for an ancestral, common role for GapR proteins across the α-proteobacteria and, curiously, their phages. In addition, the results of a bacterial two-hybrid assay indicate that the ΦCr30 homolog of GapR (GapR^{Cr30}) interacts with GapR (Fig. 5*C*), suggesting that the phage-encoded homolog may directly bind *Caulobacter* GapR during the course of ΦCr30 infection.

As the function and localization of a protein are often linked, we also visualized the localization that α-proteobacterial GapR homologs adopt in *Caulobacter*. We therefore fused the GapR homolog from *Rhodobacter capsulatus* (GapR^{Rhodo}) to mCherry, expressed this fusion in *Caulobacter*, and monitored the localization of the mCherry-GapR^{Rhodo} fusion protein in live cells. We found that GapR^{Rhodo} expressed in *Caulobacter* exhibited the same asymmetric distribution in predivisional cells as does *Caulobacter* GapR (Fig. 6*B*), providing additional evidence for a conserved, species-independent behavior of this novel family of NAPs.

In *Caulobacter*, precise protein localization is frequently achieved through protein-protein interactions that underlie hierarchical localization dependency networks. The fact that GapR^{Rhodo} exhibits a GapR-like localization pattern in *Caulobacter* is remarkable in light of the fact that these homologs are only 45% identical (Fig. S7), reducing the likelihood that specific protein-protein interactions involving GapR^{Rhodo} in *R. capsulatus* would be preserved in *Caulobacter* (42). Given the relatively low probability of conserved interactions between GapR^{Rhodo} and endogenous *Caulobacter* proteins, we asked whether clustering of GapR might instead be a consequence of generic binding to AT-rich regions of the genome, which have indeed been shown to colocalize in *E. coli* through association with the AT-associated NAP known as H-NS (37). If binding to AT-rich DNA is sufficient to establish the localization pattern observed for GapR, then an orthogonal AT-rich DNA-binding protein that bears no resemblance to any *Caulobacter* protein should nonetheless adopt a GapR-like localization pattern in *Caulobacter.* To test this, we expressed in *Caulobacter* a fluorescent protein fused to *E. coli* H-NS, a protein that lacks a homolog in *Caulobacter* and that is known to autonomously bind

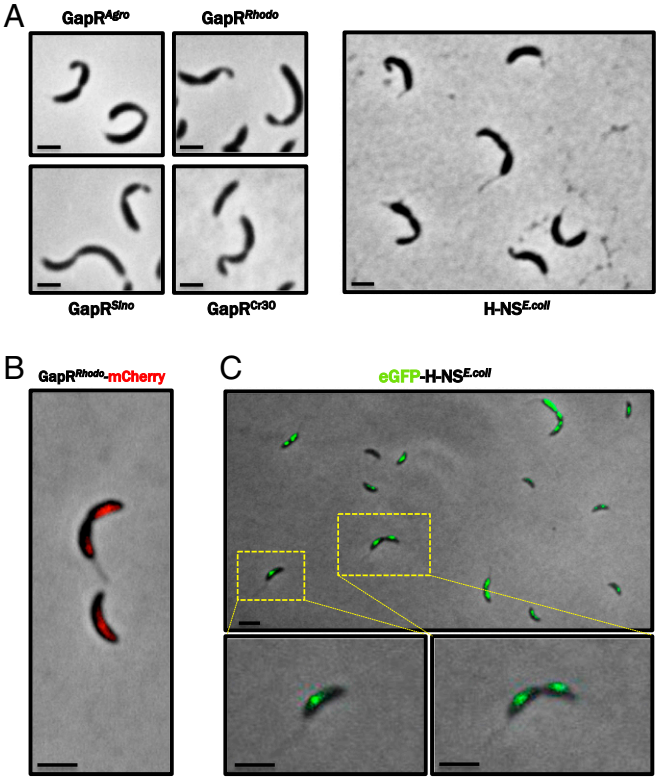

Fig. 6. GapR-like activity and localization of GapR homologs and *E. coli* H-NS in *Caulobacter*. (*A*) GapR homologs from divergent α-proteobacteria and *Caulobacter*-specific bacteriophage ΦCr30 (*Left*) or *E. coli hns* (*Right*) were introduced into *C. crescentus* NA1000 on a low-copy replicating plasmid driven by the vanillate-inducible *vanA* promoter, grown for 4 h in M2G after addition of 0.5 mM vanillate and imaged by phase contrast microscopy. Agro, *Agrobacterium tumefaciens*; Rhodo, *Rhodobacter capsulatus*; Sino, *Sinorhizobium meliloti*; Cr30, Caulophage ΦCr30. (Scale bar, 1 μm.) (*B*) The GapR homolog from *R. capsulatus* (Rcc02587) was tagged with mCherry, expressed in NA1000 from the chromosomal P_{xylX} promoter in the presence of 0.3% xylose and visualized by epifluorescence microscopy (shown overlaid with the phase contrast image), revealing an asymmetric distribution within predivisional cells. (Scale bar, 1 μm.) (*C*) eGFP-tagged H-NS from *E. coli* was expressed from the xylose promoter and its localization visualized as described above, showing a symmetric distribution within predivisional cells. (Scale bar, 1 μm.)

AT-rich DNA in vivo and in vitro. We observed that H-NS, like GapR, forms discrete subpolar clusters in *Caulobacter*, supporting the notion that GapR localizes through interactions with AT-rich loci on the *Caulobacter* chromosome. However, unlike GapR homologs, H-NS is apparently not asymmetrically distributed in *Caulobacter*, suggesting that swarmer-specific accumulation of GapR is a property specific to that protein family (Fig. 6*C*).

GapR Function Is Distinct from That of H-NS–Like Proteins. GapR and H-NS each bind AT-rich DNA and form clusters in *Caulobacter*, suggesting that these divergent DNA-binding proteins may perform a common function in their respective hosts. We reasoned that if GapR and H-NS are functionally equivalent, then (*i*) H-NS should exhibit GapR-like activity in *Caulobacter* and (*ii*) GapR should exhibit H-NS–like activity in *E. coli*.

To test these predictions, we first overexpressed H-NS in *Caulobacter* to determine whether the phenotypes observed on GapR overexpression can also be induced by H-NS. When *hns* was expressed in *Caulobacter* from the vanillate-inducible P_{vanA} promoter on a low-copy replicating plasmid, we observed the same defects in cell shape, size, and division that occur on *gapR*

overexpression, including the polar minicells that are typical of *gapR* mutant strains (Fig. 6*A*). This finding suggests that the constitutive binding of AT-rich DNA may be sufficient to cause the defects associated with GapR overexpression.

DNA binding-deficient variants of H-NS fail to form localized clusters in live *E. coli* (37), indicating that the association of H-NS with AT-rich DNA underlies its localization. To test whether GapR exhibits H-NS–like activity in *E. coli*, we first asked whether GapR adopts an H-NS–like subcellular localization by expressing fluorescently labeled GapR in a Δ*hns E. coli* mutant and determining its localization. Although the expression of mCherry alone led to a homogeneous, diffuse fluorescent signal that occupied the entire cytoplasmic space, mCherry-GapR formed two discrete clusters per cell during exponential growth (Fig. 7*A*), a localization pattern that has been previously observed for H-NS using both diffraction-limited and superresolution fluorescence microscopy (37). On simultaneous coexpression of H-NS and GapR fused to eGFP and mCherry (respectively) in *E. coli*, we were able to observe a precise overlap in the subcellular distributions of the two fluorescent proteins (Fig. 7*B*), confirming a shared localization pattern and indicating a common nucleoid association profile for H-NS and GapR in *E. coli*. This observation further implies that the localization of GapR, at least in *E. coli* cells, is a consequence of recognizing AT-rich DNA (37).

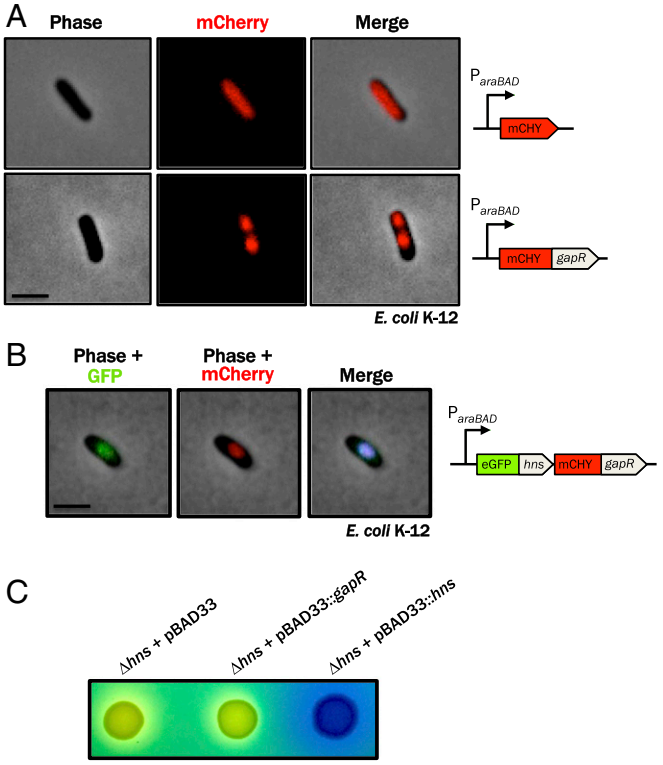

Fig. 7. GapR recapitulates the subcellular localization, but not the function, of H-NS in *E. coli*. (*A*) An *E. coli* Δ*hns* mutant expressing mCherry alone (*Top*) or mCherry-GapR (*Bottom*) under control of the arabinose-inducible *araBAD* promoter was grown for 2 h in rich media with induction (0.02% arabinose) and imaged by epifluorescence and phase contrast microscopy. (Scale bar, 1 μm.) (*B*) An *E. coli* Δ*hns* mutant containing a plasmid expressing both mCherry-GapR and eGFP-HNS from a single *araBAD* promoter was grown and imaged as in *A*. (*C*) *E. coli* Δ*hns* mutant strains transformed with pBAD33 containing *gapR*, pBAD33 containing *hns* (positive control), or empty vector (negative control) were spotted onto salicin agar (which reports on expression of the cryptic *bgl* operon) seeded with 0.02% arabinose. Yellow colony color, Bgl^+ (no complementation); blue colony color, Bgl^- (complementation). (Scale bar, 1 μm.)

H-NS is a global transcriptional silencer that represses the expression of ∼5% of *E. coli* genes. Accordingly, we next determined whether GapR can perform the silencing function of H-NS. One locus H-NS silences is the *bgl* operon, which encodes a functional phospho-β-glucosidase and a transport system for its substrate(s); Δ*hns* mutants exhibit a Bgl^+ phenotype and can use β-glucosides as a sole carbon source (43). We expressed *gapR* in a Δ*hns* mutant of *E. coli* under control of the arabinose inducible P_{araBAD} promoter and monitored the expression of the *bgl* operon using a colorimetric assay that reports on the metabolism of salicin, a β-glucoside. Although the *trans* expression of hns^+ in a Δ*hns* background restored the Bgl^- phenotype, and although *gapR* can be efficiently expressed in *E. coli* (Fig. 7 *A* and *B*), expression of *gapR* did not result in repression of the *bgl* operon, because Δ*hns* strains expressing *gapR* retained the Bgl^+ phenotype (Fig. 7*C*). Taken together, these findings show that, although GapR retains the AT-binding property of H-NS, it seems to lack the capacity to act as an autonomous transcriptional silencer. This observation distinguishes these two NAPs and suggests that although both GapR and H-NS associate with AT-rich regions of DNA, they are nonetheless functionally divergent.

Discussion

GapR Is an Essential DNA-Associated Proteins. This report describes the discovery of GapR, a NAP that globally binds the *Caulobacter* chromosome at AT-rich loci and plays an essential role in the growth and division of the dimorphic bacterium *Caulobacter crescentus*. GapR is required for viability, but is also toxic when expressed at high levels, indicating that normal growth also requires robust regulation of GapR abundance. Strains depleted for or overexpressing *gapR* exhibit compromised growth, and at the single-cell level, defects in cell size, cell shape, and division are observed. Because there is a positive correlation between GapR ChIP-seq enrichment and average local AT content, and because overexpression of the exogenous AT-binding protein H-NS in *Caulobacter* reproduces the GapR overexpression phenotype, it is conceivable that GapR overproduction leads to ectopic association with nonnative binding sites (once all native sites are occupied), which directly or indirectly leads to the disruption of one or more developmental processes. Lethal overproduction of AT-binding NAPs has been previously observed, with overexpression of H-NS in *E. coli* leading to extreme condensation of the nucleoid and fatal inhibition of macromolecular synthesis (44). Although the mechanism of cell cycle disruption by excess GapR in *Caulobacter* remains to be determined, we can conclude that its activity and/or abundance must be regulated so as to mitigate the disruptive effects of unchecked GapR accumulation in the cell.

Caulobacter GapR is unusual among NAPs in that it is essential for viability; although H-NS family members can influence many processes through silencing of as much as 5% of the bacterial genome (31), they are often dispensable in the γ-proteobacterial species in which they are found (29, 45, 46). Notable exceptions include the ortholog of H-NS in *Salmonella* and the semi-redundant pair of H-NS-like NAPs in *Pseudomonas aeruginosa*, MvaT/MvaU (32, 47). In both of these cases, the essential role of these NAPs is to silence the expression of horizontally acquired genes that are toxic when constitutively expressed. This phenomenon is unlikely to explain the essentiality of GapR, as most promoters bound by GapR are transcriptionally active (as evidenced by the presence of bound RpoD at >80% of these sites), and as we do not observe global de-repression of GapR-associated genes after GapR depletion (as is observed following deletion of *hns* in *E. coli*) (19). Our results therefore do not support an H-NS–like role for GapR in direct transcriptional silencing of target genes, suggesting that GapR represents a NAP family that is functionally distinct from H-NS–like proteins and whose divergent functions are critical for cell growth and division.

GapR Is a Signature α-Proteobacterial Protein. GapR comprises a single domain of unknown function (DUF2312) that is ubiquitous among the α-proteobacteria, with *gapR* coding sequences found in nearly all members of this clade (40, 41); this includes endosymbionts with minimal genomes (e.g., *Wolbachia*). This conservation indicates a common, ancient ancestor for all orthologous GapR sequences in the α-proteobacteria. We have shown here that GapR homologs from various α-proteobacterial species can recapitulate the toxic consequences of *gapR* overproduction in *Caulobacter*, indicating conservation of function activity between GapR orthologs.

Not only is GapR conserved throughout the α-proteobacteria, but several bacteriophages with α-proteobacterial hosts, including the *Caulobacter*-specific lytic phage ΦCr30, also encode GapR homologs. It is noteworthy that ΦCr30 carries a homolog of *Caulobacter* GapR, a protein that accumulates in the swarmer compartment of predivisional cells, as this phage exhibits a specific tropism for the swarmer cell due to an inability to infect stalked cells (11). The identification of an AT-associated GapR-like protein in ΦCr30 is also particularly intriguing when considered in the context of genomic nucleotide bias: whereas the *Caulobacter* genome is 67.2% GC, the ΦCr30 genome is instead markedly GC poor (38.2% GC) (48). When we express the $GapR^{Cr30}$ homolog in *Caulobacter,* we observe growth and division defects. We also observe that $GapR^{Cr30}$ directly binds GapR in a bacterial two-hybrid experiment. These observations together are consistent with the notion that $GapR^{Cr30}$ is a viral "mimic" protein (49) that acts by inhibiting, antagonizing, displacing, or otherwise subverting GapR and hijacking the critical biological process(es) in which it is engaged. Recently, the divergent swarmer-specific *Caulobacter* phage ΦCbK and several of its sequenced relatives were shown to encode a homolog of the cell cycle transcriptional regulator GcrA (50, 51), indicating that phage-encoded mimics could act by modulating the *Caulobacter* cell cycle. We expect that investigation into the role of $GapR^{Cr30}$ in the phage infection cycle will inform our understanding of the essential function of GapR in *Caulobacter* and may additionally reveal novel virulence strategies used by bacteriophages.

The GapR homolog from a phylogenetically distinct α-proteobacterium (*Rhodobacter capsulatus*) adopts a GapR-like localization pattern, including not only subcellular clustering, but also asymmetric distribution to the swarmer compartment in the predivisional stage when expressed in *Caulobacter*. This observation highlights the functional conservation between these homologs despite their phylogenetic distance. *E. coli* H-NS, by contrast, does not exhibit this type of asymmetry, implying a GapR-specific mechanism that drives the swarmer-specific accumulation of this NAP. We infer that AT binding is the primary driver of GapR subcellular clustering, however, because the *R. capsulatus* GapR homolog and *E. coli* H-NS form foci like GapR in single *Caulobacter* cells. Although it is possible that GapR is passively distributed to the swarmer cell compartment in predivisional cells through an interaction with AT-rich DNA, the lack of swarmer compartment-specific enrichment of H-NS expressed in *Caulobacter* may instead indicate that additional mechanisms (e.g., compartment-specific proteolysis, active translocation, posttranslational modification, or upstream polarity factors) contribute to this effect.

GapR Associates Primarily with the Swarmer Cell Nucleoid. Two separate lines of evidence support the possibility that the swarmer and stalked cell types of *Caulobacter* exhibit distinct nucleoid compaction or higher-order nucleoid structures. First, the nucleoids of swarmer and stalked cells exhibit markedly different sedimentation rates and contain distinct complements of associated protein (8, 52–54). Second, because the chromosome fills the cytoplasmic space in both daughter compartments (16), the swarmer nucleoid must be packaged into only ∼70% the volume afforded to that of the stalked cell nucleoid. It has been suggested that asymmetries in swarmer and stalked cell nucleoid structure are the consequence of one or more asymmetrically distributed NAPs that impart cell type-specific structure to the *Caulobacter* nucleoid (12). We are intrigued by the possibility that GapR, a protein that localizes to the swarmer nucleoid during cell division, represents a histone-like factor that contributes to the swarmer/stalked cell nucleoid asymmetry in *Caulobacter*.

GapR associates with hundreds of sites across the genome yet localizes to a discrete focus in vivo. This clustering suggests that distant genomic loci are colocalized in 3D space, as has been observed for H-NS–bound loci in *E. coli* (37). In addition to H-NS, multiple NAPs are known to promote long-range intrachromosomal interactions through DNA bending, looping, compaction, and the stabilization of discrete subdomains (4, 55–63). It may be that the self-association of GapR molecules links together noncontiguous loci that contain AT-rich sequences.

The findings presented here provide the foundation for the hypothesis that the accumulation of GapR in the swarmer cell compartment contributes to the control of swarmer cell fate. Conformation-capture experiments reveal a relationship between gene expression and chromosome structure in *Caulobacter* (4,6); if GapR plays a role in mediating nucleoid conformation of the swarmer, it could also establish or otherwise influence the program of gene expression that is specifically activated in the swarmer cell during the S → G1 transition.

GapR Is Linked to Spatiotemporal Regulation of Cell Cycle-Controlled Genes. Although its synthesis is constitutive across the cell cycle, we have shown that GapR associates with cell cycle-regulated genes and has hundreds of binding sites in common with the known master cell cycle regulators, indicating a role for this essential NAP in cell cycle control. By comparing the ChIP-seq footprint of GapR against those of other master regulators, we discovered that GapR shares in vivo target loci with the master cell cycle regulators SciP, CtrA, GcrA, and MucR1/2. Although GapR ChIP-seq peaks overlap hundreds of intergenic regions in *Caulobacter*, we have shown that the extent of GapR overlap with binding sites of each of the regulators CtrA, GcrA, MucR1, and MucR2 is statistically significant (*SI Materials and Methods*). GapR overlaps most extensively with MucR1, covering more than 90% of its binding sites. MucR1/2 are homologous transcriptional regulators that are critical in establishing the swarmer cell developmental state during the S → G1 transition, presumably by inhibiting expression of G1-specific CtrA-regulated genes in the S phase. Products of these swarmer-activated genes include the pilus subunit protein PilA, the buoyancy switch factor and encapsulation inhibitor HvyA, and, importantly, the G1-specific CtrA inhibitor SciP (12).

In *Caulobacter*, the chromosomal positions of genes dictate not only their timing of replication/segregation but also their spatial placement in the *Caulobacter* cell (64). Two separate observations support the possibility that the spatial positioning of a gene locus is important for its proper expression. First, chromosomal rearrangements that have occurred in *Caulobacter* species have largely preserved the distances between rearranged genes and the origin, suggesting that the longitudinal subcellular positions of genes are also generally conserved (65). Second, we observe that promoters bound by CtrA and MucR1, whose timing of expression and downstream gene function are largely swarmer specific, are almost completely restricted to the origin-proximal half of the chromosome and therefore to the stalked and flagellar pole-proximal regions of the cell (the majority of these genes are listed in Table S2). Proteins like MucR1/2 and CtrA that control expression of these swarmer genes may interact with spatially restricted proteins, such as GapR, to carry out their tasks.

The overlap of GapR with nearly all MucR1-occupied CtrA sites raises the possibility that GapR may be directly involved in

MucR1/2 function, i.e., modulating expression of CtrA-regulated genes and controlling the S → G1 switch. This connection between GapR and the activation of the swarmer cell program is even more striking in light of the subcellular distribution of GapR in predivisional cells, where the majority of GapR localizes asymmetrically to the swarmer cell compartment. Although MucR1/2 have been shown to repress CtrA-regulated genes, it remains unknown how such repression is restricted to S-phase (12). We hypothesize that the asymmetric distribution of GapR to the swarmer compartment provides spatial regulation to MucR1/2 activity. Because perturbation of GapR function leads only to minor changes in transcript abundances, we predict that this regulation is either indirect or that it fine-tunes essential events. We propose that GapR is a high AT-associated protein that has been co-opted to mediate the faithful asymmetric division of *Caulobacter* cells.

Materials and Methods

Bacterial Strains, Plasmids, and Growth Conditions. Strains and plasmids used in this study are detailed in Tables S5 and S6. Plasmid construction, strain engineering, and growth conditions are described in *SI Materials and Methods*.

ChIP-Seq. A full description of methods used for cell crosslinking and harvesting, chromatin immunoprecipitation, deep sequencing, and data processing is provided in *SI Materials and Methods*.

Subcellular Localization of Fluorescently Tagged Proteins. *C. crescentus* strains were cultured to log phase in peptone yeast extract medium (PYE) containing appropriate antibiotic. Where necessary, gene expression was induced by adding 0.3% xylose for 120 min during growth at 30 °C before imaging by phase contrast and epifluorescence microscopy. *E. coli* strains were grown to *A*600 of about 0.2 and induced with 0.2% arabinose for 90 min during growth at 37 °C before imaging.

Western Blot Analysis. Cultures were grown overnight in PYE and back-diluted 1:100 into fresh media. One-milliliter samples were collected from cultures grown in each condition at $OD_{600} = 1$. Harvested samples were normalized, pelleted (10,000 × *g*, 10 min), resuspended in SDS/PAGE sample buffer, lysed by incubating at 100 °C for 15 min, and subjected to electrophoresis through gradient (4–12%) SDS/PAGE gels. Rabbit polyclonal antisera that recognize the FLAG epitope (1:5,000 dilution; Sigma-Aldrich) or CtrA (1:10,000 dilution) were used for immunoblots. Protein bands were visualized using the Western Lightning ECL antibody detection kit (PerkinElmer) and Hyblot CL film (Denville Scientific).

ACKNOWLEDGMENTS. We thank Emma Chory, Gerald Crabtree, Elizabeth Zuo, and Michael Eckart for outstanding technical assistance, and members of the L.S. laboratory for critical review of this manuscript. This work was supported by Ruth L. Kirchstein NIH F32 Postdoctoral Fellowship F32GM109650 (to D.P.R.), National Science Foundation (NSF) Graduate Research Fellowship DGE-114747 (to M.D.M.), Gordon & Betty Moore Foundation Life Science Research Foundation Postdoctoral Fellowship GBMF2550.03 (to K.L.), NSF Inspire Award CCF344284 (to D.L.D. and H.H.M.), and NIH Grants R01 GM32506 and R35 GM11807101 (to L.S.).

1. Dillon SC, Dorman CJ (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* 8(3):185–195.
2. Cagliero C, Grand RS, Jones MB, Jin DJ, O'Sullivan JM (2013) Genome conformation capture reveals that the Escherichia coli chromosome is organized by replication and transcription. *Nucleic Acids Res* 41(12):6058–6071.
3. Dorman CJ (2013) Genome architecture and global gene regulation in bacteria: Making progress towards a unified model? *Nat Rev Microbiol* 11(5):349–355.
4. Le TB, Imakaev MV, Mirny LA, Laub MT (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* 342(6159):731–734.
5. Marbouty M, et al. (2015) Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. *Mol Cell* 59(4):588–602.
6. Umbarger MA, et al. (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Mol Cell* 44(2):252–264.
7. Kirkpatrick CL, Viollier PH (2012) Decoding Caulobacter development. *FEMS Microbiol Rev* 36(1):193–205.
8. Rizzo MF, Shapiro L, Gober J (1993) Asymmetric expression of the gyrase B gene from the replication-competent chromosome in the Caulobacter crescentus predivisional cell. *J Bacteriol* 175(21):6970–6981.
9. Domian IJ, Quon KC, Shapiro L (1997) Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell* 90(3):415–424.
10. Holtzendorff J, et al. (2004) Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science* 304(5673):983–987.
11. Ardissone S, et al. (2014) Cell cycle constraints on capsulation and bacteriophage susceptibility. *eLife* 3:3.
12. Fumeaux C, et al. (2014) Cell cycle transition from S-phase to G1 in Caulobacter is mediated by ancestral virulence regulators. *Nat Commun* 5:4081.
13. Gora KG, et al. (2010) A cell-type-specific protein-protein interaction modulates transcriptional activity of a master regulator in Caulobacter crescentus. *Mol Cell* 39(3): 455–467.
14. Zhou B, et al. (2015) The global regulatory architecture of transcription during the Caulobacter cell cycle. *PLoS Genet* 11(1):e1004831.
15. Collier J (2012) Regulation of chromosomal replication in Caulobacter crescentus. *Plasmid* 67(2):76–87.
16. Jensen RB, Shapiro L (1999) The Caulobacter crescentus smc gene is required for cell cycle progression and chromosome segregation. *Proc Natl Acad Sci USA* 96(19): 10661–10666.
17. Lee SF, Thompson MA, Schwartz MA, Shapiro L, Moerner WE (2011) Super-resolution imaging of the nucleoid-associated protein HU in Caulobacter crescentus. *Biophys J* 100(7):L31–L33.
18. Schwartz MA, Shapiro L (2011) An SMC ATPase mutant disrupts chromosome segregation in Caulobacter. *Mol Microbiol* 82(6):1359–1374.
19. Dorman CJ (2004) H-NS: A universal regulator for a dynamic genome. *Nat Rev Microbiol* 2(5):391–400.
20. Goley ED, Toro E, McAdams HH, Shapiro L (2009) Dynamic chromosome organization and protein localization coordinate the regulatory circuitry that drives the bacterial cell cycle. *Cold Spring Harb Symp Quant Biol* 74:55–64.
21. McAdams HH, Shapiro L (2009) System-level design of bacterial cell cycle control. *FEBS Lett* 583(24):3984–3991.
22. Christen B, et al. (2011) The essential genome of a bacterium. *Mol Syst Biol* 7:528.
23. Griffith KL, Grossman AD (2008) Inducible protein degradation in Bacillus subtilis using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. *Mol Microbiol* 70(4):1012–1025.
24. Nierman WC, et al. (2001) Complete genome sequence of Caulobacter crescentus. *Proc Natl Acad Sci USA* 98(7):4136–4141.
25. Bailey TL, et al. (2009) MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server Issue):W202–W208.
26. Cui F, Zhurkin VB (2009) Distinctive sequence patterns in metazoan and yeast nucleosomes: Implications for linker histone binding to AT-rich and methylated DNA. *Nucleic Acids Res* 37(9):2818–2829.
27. Gordon BR, et al. (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci USA* 108(26):10690–10695.
28. Timchenko T, Bailone A, Devoret R (1996) Btcd, a mouse protein that binds to curved DNA, can substitute in Escherichia coli for H-NS, a bacterial nucleoid protein. *EMBO J* 15(15):3986–3992.
29. Gordon BR, Imperial R, Wang L, Navarre WW, Liu J (2008) Lsr2 of Mycobacterium represents a novel class of H-NS-like proteins. *J Bacteriol* 190(21):7052–7059.
30. Ding P, et al. (2015) A Novel AT-Rich DNA Recognition Mechanism for Bacterial Xenogeneic Silencer MvaT. *PLoS Pathog* 11(6):e1004967.
31. Ali SS, Xia B, Liu J, Navarre WW (2012) Silencing of foreign DNA in bacteria. *Curr Opin Microbiol* 15(2):175–181.
32. Navarre WW, et al. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. *Science* 313(5784):236–238.
33. Haakonsen DL, Yuan AH, Laub MT (2015) The bacterial cell cycle regulator GcrA is a σ70 cofactor that drives gene expression from a subset of methylated promoters. *Genes Dev* 29(21):2272–2286.
34. Fang G, et al. (2013) Transcriptomic and phylogenetic analysis of a bacterial cell cycle reveals strong associations between gene co-expression and evolution. *BMC Genomics* 14:450.
35. Schrader JM, et al. (2014) The coding and noncoding architecture of the Caulobacter crescentus genome. *PLoS Genet* 10(7):e1004463.
36. Ueguchi C, Seto C, Suzuki T, Mizuno T (1997) Clarification of the dimerization domain and its functional significance for the Escherichia coli nucleoid protein H-NS. *J Mol Biol* 274(2):145–151.
37. Wang W, Li GW, Chen C, Xie XS, Zhuang X (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* 333(6048):1445–1449.
38. Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* 95(10): 5752–5756.
39. Gupta RS (2005) Protein signatures distinctive of alpha proteobacteria and its subgroups and a model for alpha-proteobacterial evolution. *Crit Rev Microbiol* 31(2): 101–135.
40. Gupta RS, Mok A (2007) Phylogenomics and signature proteins for the alpha proteobacteria and its main groups. *BMC Microbiol* 7:106.
41. Kainth P, Gupta RS (2005) Signature proteins that are distinctive of alpha proteobacteria. *BMC Genomics* 6:94.
42. Lewis AC, Jones NS, Porter MA, Deane CM (2012) What evidence is there for the homology of protein-protein interactions? *PLOS Comput Biol* 8(9):e1002645.
43. Prasad I, Schaefler S (1974) Regulation of the beta-glucoside system in Escherchia coli K-12. *J Bacteriol* 120(2):638–650.
44. Spurio R, et al. (1992) Lethal overproduction of the Escherichia coli nucleoid protein H-NS: Ultramicroscopic and molecular autopsy. *Mol Gen Genet* 231(2):201–211.

45. May G, Dersch P, Haardt M, Middendorf A, Bremer E (1990) The osmZ (bglY) gene encodes the DNA-binding protein H-NS (H1a), a component of the Escherichia coli K12 nucleoid. *Mol Gen Genet* 224(1):81–90.
46. Dorman CJ (2004) Virulence gene regulation in Shigella. *Ecosal Plus* 1(1).
47. Castang S, Dove SL (2012) Basis for the essentiality of H-NS family members in Pseudomonas aeruginosa. *J Bacteriol* 194(18):5101–5109.
48. Ely B, Gibbs W, Diez S, Ash K (2015) The Caulobacter crescentus transducing phage Cr30 is a unique member of the T4-like family of myophages. *Curr Microbiol* 70(6): 854–858.
49. Elde NC, Malik HS (2009) The evolutionary conundrum of pathogen mimicry. *Nat Rev Microbiol* 7(11):787–797.
50. Gill JJ, et al. (2012) The Caulobacter crescentus phage phiCbK: Genomics of a canonical phage. *BMC Genomics* 13:542.
51. Panis G, Lambert C, Viollier PH (2012) Complete genome sequence of Caulobacter crescentus bacteriophage φCbK. *J Virol* 86(18):10234–10235.
52. Evinger M, Agabian N (1977) Envelope-associated nucleoid from Caulobacter crescentus stalked and swarmer cells. *J Bacteriol* 132(1):294–301.
53. Evinger M, Agabian N (1979) Caulobacter crescentus nucleoid: Analysis of sedimentation behavior and protein composition during the cell cycle. *Proc Natl Acad Sci USA* 76(1):175–178.
54. Ward D, Newton A (1997) Requirement of topoisomerase IV parC and parE genes for cell cycle progression and developmental regulation in Caulobacter crescentus. *Mol Microbiol* 26(5):897–910.
55. Ghosh S, Mallick B, Nagaraja V (2014) Direct regulation of topoisomerase activity by a nucleoid-associated protein. *Nucleic Acids Res* 42(17):11156–11165.
56. Skoko D, et al. (2006) Mechanism of chromosome compaction and looping by the Escherichia coli nucleoid protein Fis. *J Mol Biol* 364(4):777–798.
57. Brackley CA, Taylor S, Papantonis A, Cook PR, Marenduzzo D (2013) Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. *Proc Natl Acad Sci USA* 110(38):E3605–E3611.
58. Tupper AE, et al. (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* 13(1):258–268.
59. Zimmerman SB (2006) Shape and compaction of Escherichia coli nucleoids. *J Struct Biol* 156(2):255–261.
60. Summers EL, et al. (2012) The structure of the oligomerization domain of Lsr2 from Mycobacterium tuberculosis reveals a mechanism for chromosome organization and protection. *PLoS One* 7(6):e38542.
61. Azam TA, Ishihama A (1999) Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* 274(46):33105–33113.
62. Swiercz JP, Nanji T, Gloyd M, Guarné A, Elliot MA (2013) A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res* 41(7):4171–4184.
63. Kleine Borgmann LA, Ries J, Ewers H, Ulbrich MH, Graumann PL (2013) The bacterial SMC complex displays two distinct modes of interaction with the chromosome. *Cell Reports* 3(5):1483–1492.
64. Viollier PH, Shapiro L (2004) Spatial complexity of mechanisms controlling a bacterial cell cycle. *Curr Opin Microbiol* 7(6):572–578.
65. Ash K, et al. (2014) A comparison of the Caulobacter NA1000 and K31 genomes reveals extensive genome rearrangements and differences in metabolic potential. *Open Biol* 4(10):140128.
66. Casadaban MJ (1976) Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. *J Mol Biol* 104(3):541–555.
67. Ried JL, Collmer A (1987) An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker exchange-eviction mutagenesis. *Gene* 57(2-3):239–246.
68. Kirkpatrick CL, Viollier PH (2014) Synthetic interaction between the TipN polarity factor and an AcrAB-family efflux pump implicates cell polarity in bacterial drug resistance. *Chem Biol* 21(5):657–665.
69. Ely B (1991) Genetics of Caulobacter crescentus. *Methods Enzymol* 204:372–384.
70. Thanbichler M, Iniesta AA, Shapiro L (2007) A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in Caulobacter crescentus. *Nucleic Acids Res* 35(20):e137.
71. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One* 3(11):e3647.
72. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177(14):4121–4130.
73. Battesti A, Bouveret E (2012) The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods* 58(4):325–334.
74. Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* 498:349–361.
75. Blankenberg D, et al. (2010) Galaxy: A web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* 19:19.10.1–19.10.21.
76. Giardine B, et al. (2005) Galaxy: A platform for interactive large-scale genome analysis. *Genome Res* 15(10):1451–1455.
77. Robinson JT, et al. (2011) Integrative genomics viewer. *Nat Biotechnol* 29(1):24–26.
78. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
79. Taylor J, Schenck I, Blankenberg D, Nekrutenko A (2007) Using galaxy to perform large-scale interactive data analyses. *Curr Protoc Bioinformatics* 19:10.5.1–10.5.25.
80. Zhang Y, et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9): R137.
81. Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. *Nucleic Acids Res* 43(W1):W39-49.
82. Machanick P, Bailey TL (2011) MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* 27(12):1696–1697.
83. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T (2014) deepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42(Web Server Issue): W187–W191.
84. Mazzon RR, Lang EA, Silva CA, Marques MV (2012) Cold shock genes cspA and cspB from Caulobacter crescentus are posttranscriptionally regulated and important for cold adaptation. *J Bacteriol* 194(23):6507–6517.
85. Jacomy M, Venturini T, Heymann S, Bastian M (2014) ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* 9(6):e98679.
86. Laub MT, Chen SL, Shapiro L, McAdams HH (2002) Genes directly controlled by CtrA, a master regulator of the Caulobacter cell cycle. *Proc Natl Acad Sci USA* 99(7):4632–4637.
87. Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L (2000) Global analysis of the genetic network controlling a bacterial cell cycle. *Science* 290(5499):2144–2148.
88. Quon KC, Marczynski GT, Shapiro L (1996) Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* 84(1):83–93.
89. Dorman CJ, Deighan P (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet Dev* 13(2):179–184.
90. Dorman CJ (2009) Global regulators and environmental adaptation in Gram-negative pathogens. *Clin Microbiol Infect* 15(Suppl 1):47–50.