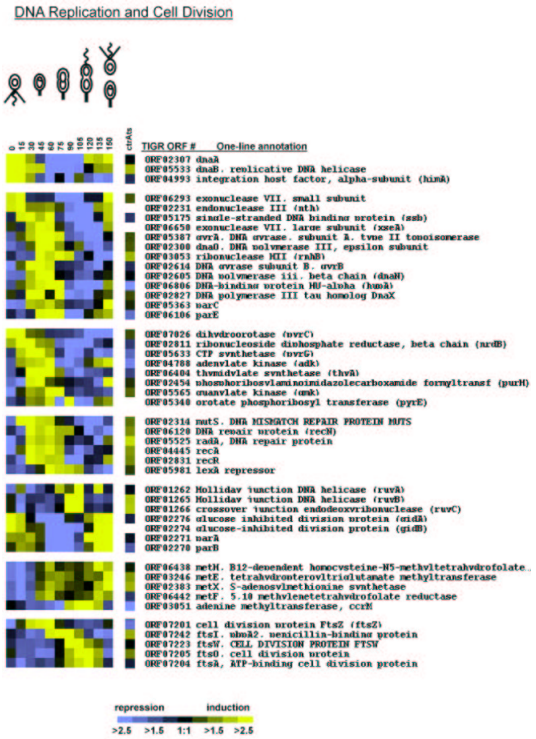
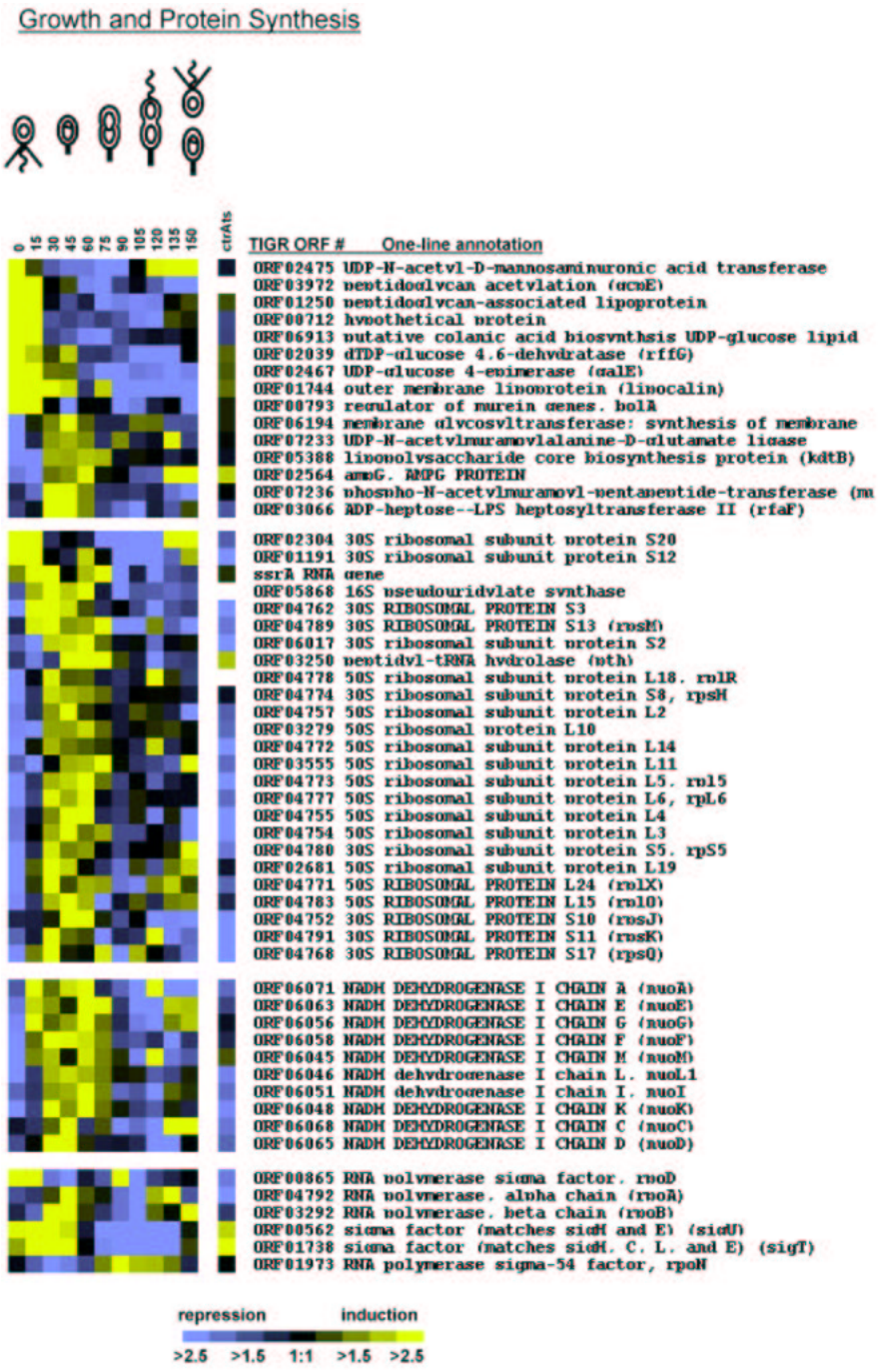
Supplemental Fig 2a

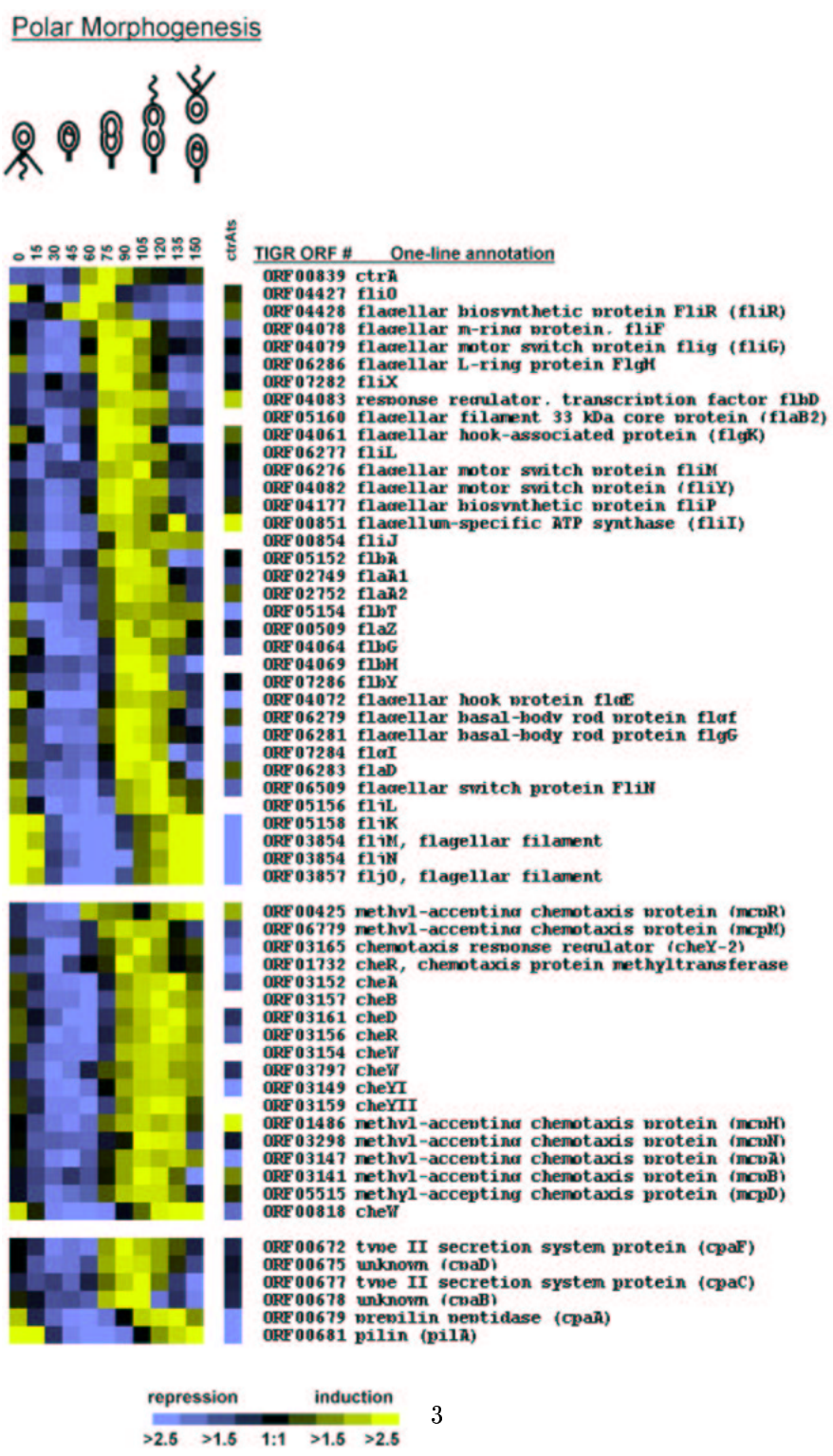
Expanded, annotated versions of Fig. 2, A to C, of the report, with expression profiles for functionally related sets of genes. For each cell cycle-regulated function, the expression profiles of associated genes are shown as color bars, coded as in Web figure 1. The class I-IV flagellar genes shown in (C) are expressed in temporal order consistent with their known ordering in flagella construction. The column labeled ctrAts shows the change in expression level for each gene in response to loss of CtrA function where blue, black, and yellow indicate decrease, no change, and increase, respectively. For each gene, the corresponding TIGR ORF number and one-line annotation information are listed to the right of its profile.



Supplemental Fig 2b

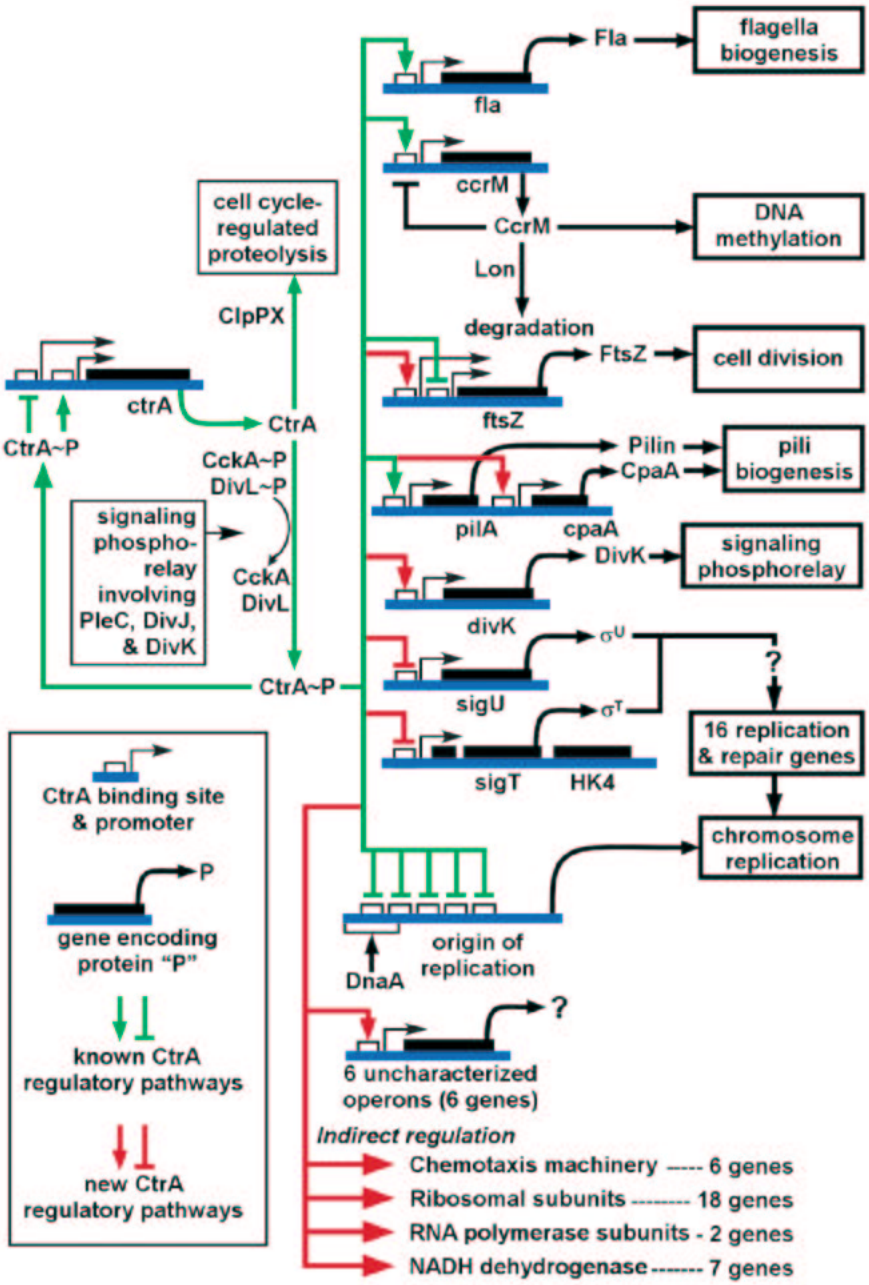


Supplemental Fig 2c



Supplemental Fig 3

An expanded, detailed version of Fig. 3 of the report. The CtrA regulatory network governing Caulobacter cell cycle progression with newly identified pathways shown in red. Phosphorylated CtrA, binding to a conserved motif, autoregulates its own transcription (26), and activates the transcription of sets of genes required for flagellar biogenesis, DNA methylation, pili biogenesis, cell division, and six other operons of unknown function. CtrA also acts as a repressor of the newly identified G1-S sigma factors sigT and sigU and, as previously suggested, of ftsZ (13). Binding of CtrA to sites in the origin of replication inhibits replication initiation (26). In addition, our results suggest there are at least 113 genes in a wide range of functional categories that appear to be indirectly regulated by CtrA; these include members of the chemotaxis machinery, ribosomal subunits, RNA polymerase subunits, and NADH dehydrogenase categories. CtrA is proteolyzed at the G1-S transition by the protease ClpPX (27). The histidine kinases CckA and DivL mediate CtrA phosphorylation. DivK, DivJ, and PleC are all thought to be involved in a cell cycle phosphorelay [reviewed in (28, 29)].

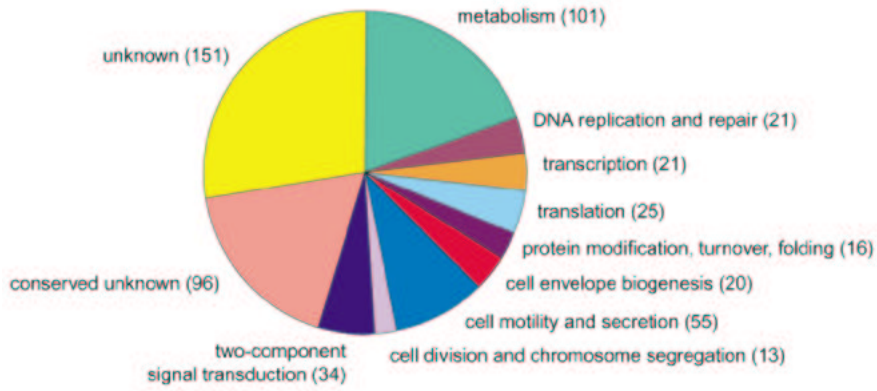

flagella biogenesis
Fla
fla
ccrM
CcrM
DNA methylation
Lon
degradation
cell cycle-regulated proteolysis
ClpPX
FtsZ
cell division
ftsZ
ctrA
CtrA
CtrA~P
Pilin
CpaA
pili biogenesis
pilA
cpaA
CckA~P
DivL~P
signaling phospho-relay involving PleC, DivJ, & DivK
CckA
DivL
DivK
signaling phosphorelay
divK
σU
?
sigU
σT
16 replication & repair genes
sigT
HK4
CtrA~P
CtrA binding site & promoter
P
gene encoding protein "P"
chromosome replication
origin of replication
DnaA
known CtrA regulatory pathways
?
6 uncharacterized operons (6 genes)
new CtrA regulatory pathways
Indirect regulation
Chemotaxis machinery ----- 6 genes
Ribosomal subunits -------- 18 genes
RNA polymerase subunits - 2 genes
NADH dehydrogenase ------- 7 genes

Supplemental Fig 4

Expanded, annotated version of Fig. 4 of the report. Regulatory genes expressed in a cell cycle-dependent pattern. Genes encoding histidine kinases, response regulators, and sigma factors that are transcribed as a function of the cell cycle are shown at the time of their peak expression levels during the cell cycle. Newly identified regulatory genes are denoted by their predicted type and numbered based on their order of expression in the cell cycle: black, response regulators with an identifiable output domain; green, single-domain response regulators; red, histidine kinases; blue, histidine kinases with a fused response regulator domain; and orange (, sigma factors. Underlined genes have been characterized previously. The genes in black boxes encode proteins that are known to be dynamically localized during the cell cycle. TIGR ORF numbers are listed for each gene shown.

Supplemental Fig 5

Distribution of cell cycle-regulated genes within functional categories. The number of genes in each category is in parentheses. Genes were assigned to a functional category based on the classification scheme defined by Clusters of Orthologous Genes (COGs) (30).



Methods

Wild-Type Time Course Methods

A culture of wild-type Caulobacter (CB15N) was grown in minimal media supplemented with glucose (M2G) at 28°C to an OD660 of 0.3 and then synchronized as described (4). After resuspension of isolated swarmer cells in fresh M2G, cells were harvested at 15-min intervals for the duration of the 150-min cell cycle.

RNA from cells harvested at each time point was reverse-transcribed into Cy3-labeled cDNA, and competitively hybridized on a microarray with a Cy5-labeled reference cDNA from a mid-log phase asynchronous culture. Thus, subsequent log ratios (Cy3:Cy5) greater or less than zero indicate higher or lower RNA abundance, respectively, relative to the averaged RNA levels in the unsynchronized reference population. Details of RNA preparation, cDNA probe preparation, and microarray hybridization are at (31).

Identification of Cell Cycle-Regulated Transcripts / Discrete Cosine Transform Analysis

To identify cell cycle-dependent transcripts, the discrete cosine transform (DCT) (see devbio1.stanford.edu/usr/hm/Mathematica/Discrete Transforms/index.html) was calculated for each of the 2966 expression profiles with valid data. Three parameters were calculated for each profile: (i) the maximum DCT coefficient, (ii) the percentage of signal power contained in the first four DCT coefficients, and (iii) the peak-to-trough ratio of the profile. A cell cycle-dependent profile is expected to have a large maximum DCT coefficient, a significant percentage of total signal power in the lower frequency (e.g., the first four) DCT coefficients, and a high peak-to-trough ratio. We selected profiles for which parameters i, ii, and iii (above) were (a) [geq] 0.0017, [geq] 0.8, and [geq] 1.8, or (b) [geq] 0.0017, [geq] 0.6, [geq] 3.0, respectively. These parameters were chosen to min-

imize the false positive rate and to ensure inclusion of previously known cell cycle-regulated genes as described below. This procedure selected 579 candidate profiles.

As the synchronization procedure involved a shift from 4°C to 28°C at the zero time point, some genes responding to this temperature shift were anticipated in the 579 candidate profiles. These potential heat shock genes were identified by two criteria: (i) their profiles showed maximal expression at the zero time point followed by an immediate drop to minimal expression for the remainder of the time course, and (ii) their expression levels in an asynchronous population subjected to the same temperature shift were similar to the levels seen in the synchronized, swarmer cell population at the zero time point (data not shown). Genes that are expressed at a higher level in the synchronized population relative to the unsynchronized population after the same temperature shift are thus swarmer-specific and not heat shock-responsive. The original group of 579 candidate cell cycle-regulated genes contained 26 genes that responded solely to the synchronization method and were discarded; this led to a final group of 553 apparently cell cycle-regulated genes.

To estimate the number of false positives included in this set of 553 genes, the entire data set was randomly shuffled (2356 genes with 11 time points each = 25,916 randomly shuffled data points) and subjected to the same DCT analysis procedure. Only 46 of the 2356 randomized expression profiles exceeded the chosen parameters; accordingly, the false positive rate is estimated to be less than 9

Analysis of ctrA Mutant Strains

Wild-type and ctrA401ts Caulobacter strains were grown at the permissive temperature of 28°C in peptone-yeast extract (PYE) media to an OD660 of 0.05 and then shifted to the restrictive temperature of 37°C. Cells were collected at 0 and 4 hours after the shift with total RNA harvested and treated as described in (31). RNA levels on microarrays were compared in three ways, where "mutant" represents the ctrA401ts strain and "control" is the isogenic, wild-type strain: (a) control at 0 hours vs. control at 4 hours, (b) control at 0 hours vs. mutant at 0 hours, (c) mutant at 0 hours vs. mutant at 4 hours. Each of the three comparisons was done twice, using independent wild-type and ctrA401ts colonies each time, with results averaged. The correlation coefficient for the duplicate experiments was greater than 95Strains LS3326 and LS3327 were grown overnight in PYE + 0.05

Promoter Analysis

A promoter database containing the 600 base pairs upstream of the predicted start codon of each predicted ORF was constructed. A profile of known CtrA binding sites was built using MEME (Multiple Expectation Maximization Motif Elicitation), available through the GCG software package (Genetics Computer Group, Madison, WI). This profile was then used with MotifSearch (GCG soft-

ware) to identify candidate CtrA binding sites in the promoter database. CtrA binding sites were selected by requiring a combined MotifSearch P value of $<$ 0.01 and a minimum 7 out of 9 match to the consensus CtrA binding site of TTAA-n7-TTAAC.