This conditioned medium was then separated from the tissue and frozen.
19. Supplemental Web material may be found at www.sciencemag.org/cgi/content/full/290/5499/2140/DC1.
20. T. Perlmann, L. Jansson, *Genes Dev.* **9**, 769 (1995).
21. B. M. Forman, K. Umesono, J. Chen, R. M. Evans, *Cell* **81**, 541 (1995).
22. Conditioned medium was acidified with hydrochloric acid, mixed with an equal volume of hexane, and shaken vigorously for 5 min. This mixture was centrifuged, and the upper organic phase was recovered. Extraction of the aqueous phase was repeated once. The hexane extract was evaporated under nitrogen gas, and for activity assays, the residue was redissolved in 100 to 200 μl of ethanol. Hexane extract of conditioned medium was evaporated, redissolved in hexane, and injected onto a normal-phase HPLC column (Genesis Silica 4-μm particle size, 4.6 mm by 25 cm, Jones Chromatography Ltd., UK). Elution was performed using a linear gradient from hexane to hexane/dichloromethane/isopropanol (85:10:5, v/v), both containing 1% acetic acid, in 30 min at a flow rate of 0.5 ml/min. Starting 17 min after injection, 30 fractions containing 0.25 ml each, were collected. Active fractions were pooled, evaporated, and redissolved in 50 μl 80% methanol. We injected 30 μl onto a reversed-phase HPLC column (Genesis C18, 4-μm particle size, 3.0 mm by 25 cm). A mobile phase of methanol/isopropanol/water (80:10:10, v/v) containing 1% acetic acid at a flow rate of 0.3 ml/min was used. Active fractions along with the preceding and subsequent fractions were analyzed by nano-ES mass spectrometry.
23. Negative-ion nano-ES spectra were recorded on a AutoSpec-OATOFFPD (Micromass, Manchester, UK) hybrid double focusing magnetic sector–orthogonal acceleration (OA) time-of-flight (TOF) tandem mass spectrometer focal plane detector (FPD). Mass spectra were recorded as magnet scans over an *m/z* range of 70 to 1000 at an instrument resolution of 3000 (10% valley definition). Collision-induced dissociation (CID) spectra were recorded by selecting the precursor $[M\text{-}H]^-$ ions with the double-focusing sectors of the instrument, and focusing them into the fourth field–free region gas cell containing xenon collision gas. Undissociated precursor ions and fragment ions were pulsed into an OATOF mass analyzer.
24. All fatty acid ligands were purchased from Sigma, except C21:1*cis*12 and C23:1*cis*14, which were purchased from Nu-Chek-Prep (Elysian, MN).
25. A. Mata de Urquiza *et al.*, data not shown.
26. D. J. Peet, D. F. Doyle, D. R. Corey, D. J. Mangelsdorf, *Chem. Biol.* **5**, 13 (1998).
27. S. A. Oñate, S. Y. Tsai, M.-J. Tsai, B. W. O'Malley, *Science* **270**, 1354 (1995).
28. Bacterially expressed glutathione *S*-transferase (GST) or GST fusion proteins were bound to glutathione-Sepharose 4B beads (Pharmacia Biotech). SRC-1 cDNA was used to generate [^{35}S]methionine-labeled proteins (*49*). The labeled proteins were incubated with beads containing GST or GST fusion proteins in the presence of dimethyl sulfoxide (DMSO), DHA, or 1 μM estradiol. After an overnight incubation, free proteins were washed away. Bound proteins were eluted in loading buffer, separated by SDS–polyacrylamide gel electrophoresis, and visualized by fluorography.
29. W. Bourguet *et al.*, *Mol. Cell* **5**, 289 (2000).
30. S. Kitareewan *et al.*, *Mol. Biol. Cell* **7**, 1153 (1996).
31. M. Neuringer, G. J. Anderson, W. E. Connor, *Annu. Rev. Nutr.* **8**, 517 (1988).
32. N. Salem Jr., H.-Y. Kim, J. A. Yergey, in *Health Effects of Polyunsaturated Fatty Acids in Seafoods*, A. P. Simopoulos, R. R. Kifer, R. E. Martin, Eds. (Academic Press, New York, 1986), chap. 15, pp. 263–317.
33. M. C. Garcia, H.-Y. Kim, *Brain Res.* **768**, 43 (1997).
34. P. Homayoun *et al.*, *Neurochem. Res.* **25**, 269 (2000).
35. A. A. Spector, *Lipids* **34**, S1 (1999).
36. M. Makrides, M. Neumann, K. Simmer, J. Pater, R. Gibson, *Lancet* **345**, 1463 (1995).
37. P. Kastner *et al.*, *Cell* **78**, 987 (1994).
38. S. Gamoh *et al.*, *Neuroscience* **93**, 237 (1999).
39. R. Sheaff Greiner, T. Moriguchi, A. Hutton, B. M. Slotnick, N. Salem Jr., *Lipids* **34**, S239 (1999).
40. J. D. Fernstrom, *Lipids* **34**, 161 (1999).
41. M.-Y. Chiang *et al.*, *Neuron* **21**, 1353 (1998).
42. L. H. Storlien, A. J. Hulbert, P. L. Else, *Curr. Opin. Clin. Nutr. Metab. Care* **1**, 559 (1998).
43. D. P. Rose, J. M. Connolly, *Pharmacol. Ther.* **83**, 217 (1999).
44. J. J. Repa *et al.*, *Science* **289**, 1524 (2000).
45. R. Mukherjee *et al.*, *Nature* **386**, 407 (1997).
46. E. D. Bischoff, M. M. Gottardis, T. E. Moon, R. A. Heyman, W. W. Lamph, *Cancer Res.* **58**, 479 (1998).
47. T. M. Willson, P. J. Brown, D. D. Sternbach, B. R. Henke, *J. Med. Chem.* **43**, 527 (2000).
48. A. Chawla, E. Saez, R. M. Evans, *Cell* **103**, 1 (2000).
49. V. Cavaillés, S. Dauvois, P. S. Danielian, M. G. Parker, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10009 (1994).
50. We thank R. Heyman and R. Bissonnette for LG849, L. Foley for SR11237, D. Mangelsdorf for plasmids encoding RXRL436F and RXRF313I derivatives, L. Solomin for helpful comments on this manuscript, and members of the Perlmann lab for valuable discussion. Supported by the Human Frontiers Science Program, the Göran Gustafssons Foundation, and the Swedish Medical Research Council grant 03X-12551 (S.L., W.G., and J.S.), the Swedish Wennergren Foundation (M.S.), and the Swedish Cancer Society (M.S.).

26 May 2000; accepted 14 November 2000

Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle

Michael T. Laub,[1] Harley H. McAdams,[1] Tamara Feldblyum,[2] Claire M. Fraser,[2] Lucy Shapiro[1]*

This report presents full-genome evidence that bacterial cells use discrete transcription patterns to control cell cycle progression. Global transcription analysis of synchronized *Caulobacter crescentus* cells was used to identify 553 genes (19% of the genome) whose messenger RNA levels varied as a function of the cell cycle. We conclude that in bacteria, as in yeast, (i) genes involved in a given cell function are activated at the time of execution of that function, (ii) genes encoding proteins that function in complexes are coexpressed, and (iii) temporal cascades of gene expression control multiprotein structure biogenesis. A single regulatory factor, the CtrA member of the two-component signal transduction family, is directly or indirectly involved in the control of 26% of the cell cycle–regulated genes.

In the bacterium *Caulobacter crescentus*, a complex genetic network controls essential cell cycle functions, including the ordered biogenesis of structures at the cell poles and division plane (*1*). Because different functions occur at specific times in the *Caulobacter* cell cycle (Fig. 1A), regulation of the cell cycle must be examined as an integrated system. The availability of the full *Caulobacter* genome sequence and microarray gene expression assays now allow such an approach.

To determine the contribution of transcriptional control to bacterial cell cycle progression, we constructed DNA microarrays containing 2966 predicted open reading frames, representing about 90% of all *Caulobacter* genes (*2*). More than 19% of the *Caulobacter* genes exhibited discrete times of transcriptional activation and repression during a normal cell cycle. Prokaryotic biology has focused on external, environmental cues as the major mechanism for turning bacterial genes on and off, rather than internal cues. Surprisingly, the transcription of genes required for many cell cycle functions, such as DNA replication, chromosome segregation, and cell division, occurred just before or nearly coincident with the time of execution of that function, paralleling temporal patterns of gene expression observed in the yeast cell cycle (*3, 4*).

Swarmer cells from wild-type *C. crescentus* were isolated and allowed to proceed synchronously through their 150-min cell cycle (Fig. 1A). RNA was harvested from cell samples taken at 15-min intervals. RNA levels for each gene at each time point were compared to RNA levels in a mixed, unsynchronized reference population by means of microarrays (*5*). To determine which RNAs varied as a function of the cell cycle, we used a discrete cosine transform algorithm to identify expression profiles that varied in a cyclical manner (*5*). This analysis identified 553 genes whose RNA levels changed as a function of the cell cycle. The 72 genes with previously characterized cell cycle–regulated promoters were in this set of temporally regulated transcripts, and they exhibited peak times of expression consistent with earlier data (*5*).

A self-organizing map clustering technique was applied to the 553 cell cycle–

[1]Department of Developmental Biology, Beckman Center, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA.

*To whom correspondence should be addressed. E-mail: shapiro@cmgm.stanford.edu

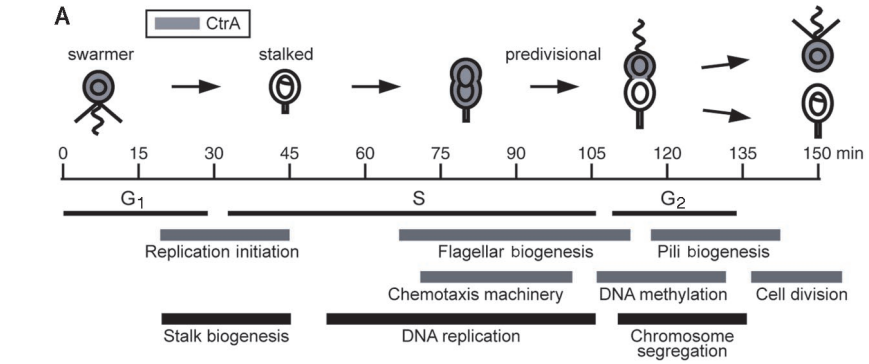
regulated expression profiles to identify groups of genes with similar expression patterns (*5*). Temporally regulated genes were maximally expressed at specific times distributed throughout the entire cell cycle (Fig. 1B). These genes were distributed along the single *Caulobacter* chromosome, with no evident correlation between time of expression and chromosomal position (*6*).

Fifty-five percent of the cell cycle–regulated genes had significant similarity to previously characterized genes in categories spanning all known biological processes (*5, 7*), showing that the *Caulobacter* genetic circuitry maintains tight temporal transcriptional control over a wide range of metabolic and morphological processes. The function of the other 45% of the cell cycle–regulated genes is unknown; 151 had no significant sequence similarity to other genes in public databases, whereas 96 had similarity to at least one other hypothetical protein.

For each cell cycle–regulated event, we observed a set of associated genes that were induced immediately before or coincident with the event (Fig. 2).[8] Further, genes encoding proteins that form multiprotein complexes were coexpressed and, when participating in the biogenesis of a complex structure, were expressed in a transcriptional cascade that, at least in two cases, reflected their order of assembly.

The expression profiles of genes associated with DNA replication and cell division are shown in Fig. 2A. Three genes encoding proteins that function in replication initiation were maximally expressed in swarmer cells, just preceding and overlapping replication initiation, and 13 genes encoding components of the replication machinery significantly increased transcription at the start of S phase. In addition, we found that homologs of genes required for the synthesis of nucleotides were induced at the G_1-S transition, perhaps ensuring sufficient nucleotide pools during S phase in part by transcriptional regulation of nucleotide synthesis genes. Additional genes with peak expression in early S phase included homologs of genes that participate in recombination and DNA repair. Temporal control of DNA methylation in *Caulobacter* appears to be executed by staged transcription of *S*-adenosylmethionine (SAM) metabolism genes and the SAM-dependent DNA methyltransferase *ccrM* gene. Four genes implicated in chromosome segregation were strongly induced during the interval after chromosome replication (*8, 9*). Three genes of the *ruvABC* operon that we predict to encode the resolvasome, an assemblage of enzymes that resolves Holliday junctions, were also expressed at this time. Given their well-characterized roles in resolving DNA structure in *Escherichia coli* and their expression in late predivisional cells in *Caulobacter*, we predict that the *ruvABC* genes play a role in the late stages of chromosome duplication in *Caulobacter*. Homologs of the *E. coli* cell division genes *ftsZ*, *ftsI*, *ftsW*, *ftsQ*, and *ftsA* were all expressed in a cell cycle–dependent fashion, with peak expression of these genes widely distributed in time. The distribution of transcription peaks of cell division genes suggests that assembly of the cell division machinery, which includes the tubulin-like FtsZ, is temporally ordered by regulation at the transcriptional level.

As swarmer cells differentiate into stalked cells, membrane synthesis rates (*10*) and overall translation rates (*11*) increase, suggesting that cell growth is regulated during the cell cycle. The S phase–specific induction of genes encoding ribosomes, RNA polymerase, and the NADH (reduced form of nicotinamide adenine dinucleotide) dehydrogenase complex of oxidative respiration (Fig. 2B) suggests transcriptional regulation of the general increase in metabolism observed in stalked cells relative to swarmer cells. Two sets of genes encoding enzymes that synthesize membranes and peptidoglycan were identified, one set with peak expression in swarmer cells and one with maximal expression in early stalked cells (Fig. 2B). These genes may play a role in cell envelope growth during the ensuing S and G_2 phases. Undoubtedly, some of these genes also participate in the synthesis of *Caulobacter*'s polar stalk, which occurs at the G_1-S transition. Previous rifampicin studies have shown that stalk biogenesis requires transcription during the first 45 min of the cell cycle (*12*), precisely the time of peak expression of the



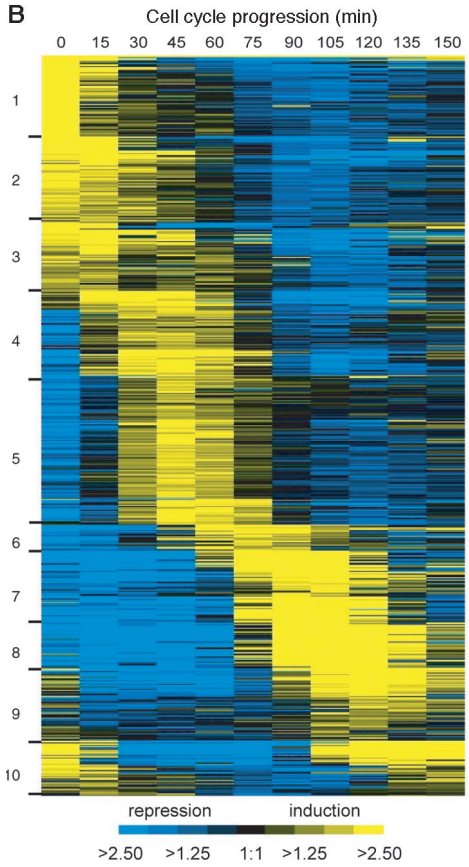


Fig. 1. (**A**) Temporally coordinated events of the *Caulobacter* cell cycle. Motile, piliated swarmer cells differentiate into stalked cells at the G_1-S transition by shedding their polar flagellum, growing a stalk at that site, losing the polar pili, and initiating DNA replication. Circles and "theta" structures in the cells represent quiescent and replicating chromosomes, respectively. CtrA is present in the shaded cells, where it represses DNA replication initiation and is cleared by proteolysis during the swarmer cell–stalked cell (G_1-S) transition. Cell division yields distinct progeny, a swarmer cell and a stalked cell. Bars below indicate timing of cell cycle functions (gray indicates a function controlled by CtrA). (**B**) Clustered expression profiles for the 553 identified cell cycle–regulated transcripts are organized by time of peak expression. Expression profiles for genes are in rows with temporal progression from left to right, as indicated at the top. Ratios are represented using the color scale at the bottom. Expression profiles were clustered using the self-organizing map analysis of the GeneCluster software and plotted using TreeView software. Each cluster is numbered; for an expanded, annotated view of these clusters, see (*5*).

membrane and peptidoglycan enzyme genes found in our expression assays (Fig. 2B).

Polar differentiation in *Caulobacter* includes construction of the polar flagellum and chemotaxis complex in predivisional cells and pili formation at the flagellated pole of swarmer cells. Cell cycle–dependent transcription is required for each of these events (*12*). At least 41 genes responsible for flagellar biogenesis are organized in a four-level transcriptional hierarchy in which the expression of each class of genes is required for expression of all subsequent classes (*13*). This temporal cascade is evident in our flagellar gene expression profiles (Fig. 2C). The flagellum, under the control of the chemotaxis machinery, moves the cell toward an attractant or away from a repellent. The majority of the 18 chemotaxis genes (Fig. 2C) were expressed in parallel with the construction of the flagellum so that the chemotaxis machinery is available when needed. Six genes constituting two adjacent operons encoding the membrane-embedded pilin secretory apparatus and the prepilin peptidase were expressed in late predivisional cells (Fig. 2C). The *cpaBCDF* genes, predicted to encode the secretion apparatus for pilin (*14*), were expressed first, followed immediately by expression of the prepilin peptidase gene *cpaA*. Finally, the pilin subunit, encoded by *pilA*, reached maximal expression 30 min later in early swarmer cells, coincident with the time in the cell cycle when pili are first observed by electron microscopy (*14*). Thus, pili biogenesis is apparently organized as a temporal transcriptional cascade, similar to the flagellar cascade.

The CtrA response regulator, a member of the two-component signal transduction family, controls several cell cycle functions (Fig. 1A). This critical regulator is periodically activated by phosphorylation and is cleared from stalked cells by temporally regulated proteolysis. A complex spatially and temporally regulated network of two-component signal transduction proteins (histidine kinases and response regulators) is thought to control the phosphorylation of CtrA (*1, 15*).

To identify all cell cycle–dependent genes that are directly or indirectly regulated by CtrA, we used microarrays to analyze gene expression levels in a mutant strain bearing a loss-of-function allele of *ctrA*, *ctrA401*^{ts} (*16*). Genes previously known to be activated by CtrA showed decreased expression levels in the *ctrA401*^{ts} strain. In all, we found 84 genes whose expression level (i) decreased in the *ctrA401*^{ts} strain, and (ii) increased in wild-type cells after the transcriptional activation of *ctrA*. Conversely, 60 genes negatively regulated by CtrA were identified on the basis of expression levels that were (i) increased upon loss of CtrA function, and (ii) maximally expressed in wild-type cells at the time when CtrA is proteolytically cleared from the cell (during the G_1-S transition) (Fig. 1A). Thus, 26% (144 of 553) of all cell cycle–dependent transcripts are significantly affected by the loss of functional CtrA (Fig. 2) (*2*). The *ctrA* loss-of-function mutant is known to accumulate cell mass, maintain viability, and regularly replicate the chromosome until at least 4 hours after the shift to the restrictive temperature (*16, 17*). The changes in expression levels in this strain can therefore be attributed to loss of CtrA activity, and not simply to a block of cell cycle progression.

We searched for CtrA binding sites in the promoter regions of all predicted genes (*5*). Of 1784 genes whose mRNA was determined not to vary as a function of the cell cycle, only two had an upstream consensus CtrA binding site. In contrast, of 553 genes with cell cycle–dependent mRNAs, 38 had consensus CtrA binding sites. Those genes with sites closely matching the consensus TTAA-n7-TTAAC (n = any nucleotide) are candidates for direct regulation by CtrA, whereas those without it are likely to be downstream, indirect targets of CtrA. Our criteria for identifying a gene as directly regulated by CtrA include (i) a predicted CtrA binding site in the regulatory region, (ii) altered expression in the *ctrA* loss-of-function mutant, and (iii) timing of the gene's expression pattern consistent with the time of CtrA availability during the cell cycle (Fig. 3).[1] These criteria correctly identified genes previously shown

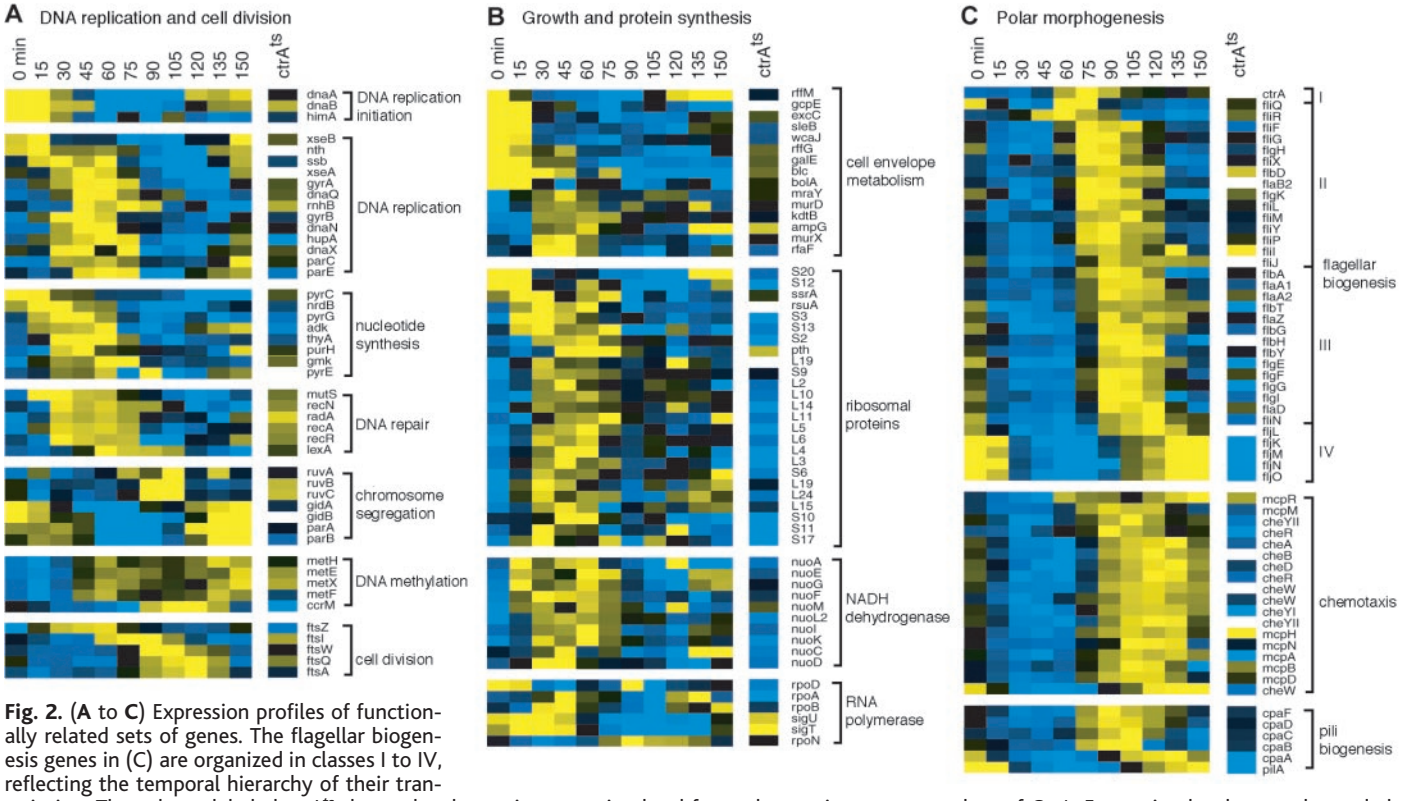


Fig. 2. (**A** to **C**) Expression profiles of functionally related sets of genes. The flagellar biogenesis genes in (C) are organized in classes I to IV, reflecting the temporal hierarchy of their transcription. The column labeled *ctrA*^{ts} shows the change in expression level for each gene in response to loss of CtrA. Expression levels are color-coded as in Fig. 1. For expanded, annotated profiles of each set of genes, see (*5*).

by in vitro footprinting and site-directed mutagenesis to be directly regulated by CtrA.

Among the genes meeting these criteria are three genes encoding regulatory factors that were expressed at the G_1-S transition: a histidine kinase (*HK4*) and two RNA polymerase sigma factors (*sigT* and *sigU*) (Fig. 4). These newly identified genes are likely candidates for regulation of early S-phase gene expression. The two sigma factors are members of the extracytoplasmic function (ECF) family (*18*). Changes in expression levels of *sigT* and *sigU* lead to aberrant cells (*19*). The *sigT* and *HK4* genes are in one operon and *sigU* in another. Expression of both the *sigT* and *sigU* operons was significantly enhanced in the *ctrA* loss-of-function experiment, and a consensus CtrA binding site was found in the predicted promoter regions of both operons (*5*), suggesting that CtrA normally acts to repress expression of these potential G_1-S regulators (*5*).

Expression of most of the chemotaxis genes was also significantly lower in the *ctrA401*^{ts} strain (Fig. 2C). The promoter region of the multigene *mcpA* operon was found to have a highly conserved half CtrA binding site (TTAAC) (*5*), suggesting that CtrA might directly activate this operon of chemotaxis genes. Transcription of the pilin subunit *pilA* was strongly reduced in the *ctrA401*^{ts} strain, consistent with a previous report (*14*). We also found that expression of the prepilin peptidase gene *cpaA* was markedly reduced in the *ctrA401*^{ts} strain and that *cpaA* has a CtrA-binding motif in its promoter region (*5*), expanding the potential role of CtrA in controlling pilus assembly.

Expression levels of genes encoding the RNA polymerase alpha subunit (*rpoA*) and major sigma subunit (*rpoD*), most ribosomal subunits, and NADH dehydrogenase subunits were reduced by a factor of more than 2 in the *ctrA401*^{ts} strain (Fig. 2B). However, none of these genes had upstream CtrA binding sites (*5*); this finding suggests indirect regulation by CtrA.

Changes in global transcription patterns were also assayed in a strain containing the inducible gain-of-function *ctrA* allele *ctrAD51EΔ3Ω*, which encodes a constitutively active, proteolysis-resistant form of CtrA (*5*). Induction of this allele blocks replication initiation, causing a strong G_1 arrest of *Caulobacter* cells (*20*). Normally, CtrA proteolysis at the G_1-S transition relieves the repression of replication initiation, but induction of the gain-of-function allele leads to continued binding of CtrA–P at the origin of replication, thereby blocking the formation of the replisome (*17, 20*).

Transcript levels of 125 genes increased significantly after induction of the *ctrAD51EΔ3Ω* gain-of-function allele (*2*). In wild-type cells, nearly 85% of these genes were expressed during the swarmer (G_1) phase when CtrA was present; expression levels of these genes normally decrease at the G_1-S transition when CtrA is cleared from the differentiating swarmer cell. Conversely, of the 153 genes that showed a significant decrease in expression in this gain-of-function strain, 80% were normally expressed in predivisional or late S and G_2 cells (*2*). Thus, blocking replication initiation through a *ctrA* gain-of-function allele reduced the expression of a wide range of genes normally induced during S and G_2, including 70% of those identified in the *ctrA401*^{ts} experiment as dependent on CtrA for expression. Persistence of active CtrA alone was thus not sufficient to induce expression of these CtrA-dependent genes after a block in the initiation of replication; this suggests a role for factors normally expressed later in the cell cycle.

The global analysis of bacterial cell cycle regulation initiated here has established the outline of the complex genetic circuitry that controls bacterial cell cycle progression and identified candidate genetic pathways for further exploration (Fig. 3). We found that the master regulator CtrA directly or indirectly controls at least 26% of cell cycle–regulated genes. Other currently unidentified factors must regulate the periodic expression of the 409 cell cycle–dependent genes that did not respond to loss of CtrA function. We predict that the regulation of cell cycle–dependent genes not controlled by CtrA will involve a hierarchical architecture with two or three additional master regulatory proteins acting in a coordinated fashion. Twenty-seven newly identified temporally controlled genes encoding two-component signal transduction proteins and sigma factors are candidates for these master regulatory roles (Fig. 4). Sixteen histidine kinases were found to be expressed

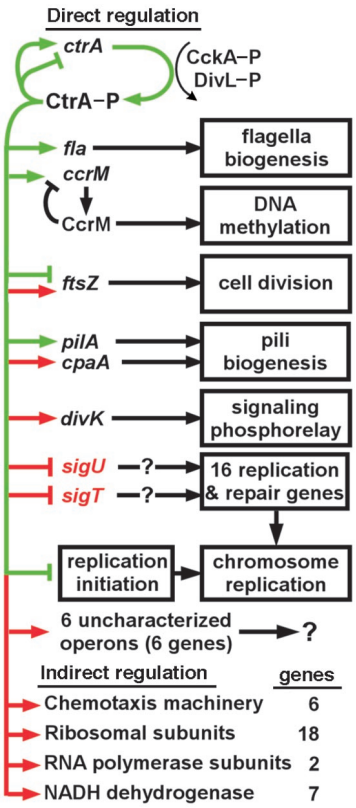


Fig. 3. The CtrA regulatory network governing *Caulobacter* cell cycle progression. Phosphorylated CtrA autoregulates its own transcription (*23*) and activates or represses the transcription of multiple sets of genes. CtrA bound to sites in the origin of replication inhibits replication initiation (*17*). In addition, there are at least 113 genes in several functional categories that appear to be indirectly regulated by CtrA. Green, previously known pathways; red, newly identified pathways. For an expanded network with added details, see (*5*).

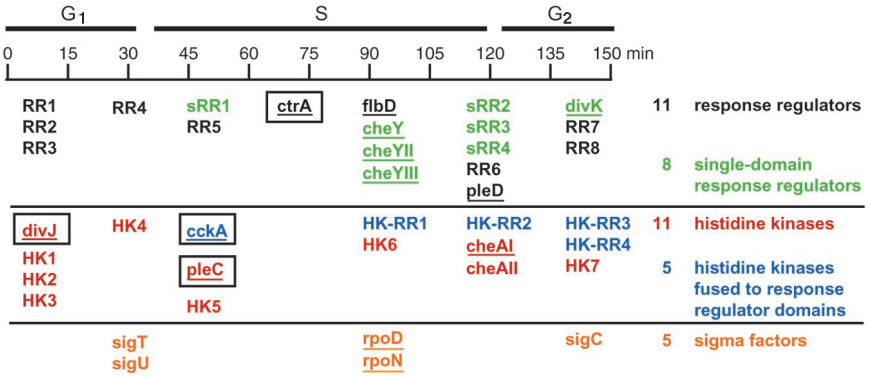


Fig. 4. Regulatory genes expressed in a cell cycle–dependent pattern. Genes encoding histidine kinases, response regulators, and sigma factors that are transcribed as a function of the cell cycle are shown at the time of their peak expression levels during the cell cycle. Newly identified regulatory genes are denoted by their predicted type and are numbered according to order of expression. 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; orange, sigma factors. Underlines indicate previously identified regulatory genes. Black boxes indicate genes encoding proteins known to be dynamically localized during the cell cycle. For additional details, see (*5*).

in a cell cycle–dependent manner. Only four of these (DivJ, CckA, PleC, and CheA) have been characterized, and three of the four (boxed in Fig. 4) are known to be dynamically localized to the cell poles at different times in the cell cycle (*21, 22*). The spatial distribution of CtrA also varies during the cell cycle (Fig. 1A) (*20*); together, these observations add another regulatory dimension. The operation of *Caulobacter*'s genetic circuitry controlling cell cycle progression and asymmetric cell division must ultimately be analyzed and modeled as an integrated, three-dimensional system of coupled chemical reactions incorporating genome-scale information on mRNA and protein levels, posttranslation modifications, and spatial distributions.

References and Notes

1. D. Hung, H. McAdams, L. Shapiro, in *Prokaryotic Development,* Y. V. Brun, L. J. Shimkets, Eds. (ASM Press, Washington, DC, 2000), pp. 361–378.
2. For complete data sets, gene lists, and details of microarray procedures, see http://caulobacter.stanford.edu/CellCycle.
3. P. T. Spellman *et al*., *Mol. Biol. Cell* **9**, 3273 (1998).
4. R. J. Cho *et al*., *Mol. Cell* **2**, 65 (1998).
5. For complete methods (including experimental procedures, details of the discrete cosine transform algorithm, clustering analyses, and promoter analysis), as well as annotated figures and a listing of all previously characterized genes (with references), see *Science* Online (www.sciencemag.org/cgi/content/full/290/5499/2144/DC1).
6. M. T. Laub, H. H. McAdams, T. Feldblyum, L. Shapiro, data not shown.
7. A complete functional categorization of the 553 cell cycle–regulated genes is available (*5*).
8. D. A. Mohl, J. W. Gober, *Cell* **88**, 675 (1997).
9. G. C. Draper, H. H. Ho, J. W. Gober, *Interactions Between GidA and the Chromosome Partitioning Apparatus of* Caulobacter crescentus (published in abstracts of the American Society for Microbiology General Meeting, Los Angeles, June 2000).
10. E. A. O'Neill, R. A. Bender, *J. Bacteriol.* **169**, 2618 (1987).
11. H. Iba, A. Fukuda, Y. Okada, *J. Bacteriol.* **135**, 647 (1978).
12. A. Newton, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 447 (1972).
13. J. W. Gober, J. C. England, in *Prokaryotic Development,* Y. V. Brun, L. J. Shimkets, Eds. (ASM Press, Washington, DC, 2000), pp. 319–339.
14. J. M. Skerker, L. Shapiro, *EMBO J.* **19**, 3223 (2000).
15. N. Ohta, T. W. Grebe, A. Newton, in *Prokaryotic Development,* Y. V. Brun, L. J. Shimkets, Eds. (ASM Press, Washington, DC, 2000), pp. 341–359.
16. K. C. Quon, G. T. Marczynski, L. Shapiro, *Cell* **84**, 83 (1996).
17. K. C. Quon, B. Yang, I. J. Domian, L. Shapiro, G. T. Marczynski, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 120 (1998).
18. D. Missiakas, S. Raina, *Mol. Microbiol.* **28**, 1059 (1998).
19. M. T. Laub, unpublished data.
20. I. J. Domian, K. C. Quon, L. Shapiro, *Cell* **90**, 415 (1997).
21. R. T. Wheeler, L. Shapiro, *Mol. Cell* **4**, 683 (1999).
22. C. Jacobs, I. J. Domian, J. R. Maddock, L. Shapiro, *Cell* **97**, 111 (1999).
23. I. J. Domian, A. Reisenauer, L. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6648 (1999).

24. *Caulobacter* sequence data were provided by The Institute for Genomic Research (W. O. Nierman *et al*., submitted). Sequencing of *C. crescentus* was accomplished with support from the U.S. Department of Energy. We thank K. Keiler and K. Ryan for helpful discussions, K. Ryan for strains LS3326 and LS3327, and J. Wang and S. Kim for assistance in microarray manufacturing. Supported by a Howard Hughes Medical Institute predoctoral fellowship and a Stanford Graduate Fellowship (M.T.L.), NIH grants GM32506 and GM51426 (L.S.), and Defense Advanced Research Projects Agency–Office of Naval Research grant N00014-99-1-05631 (H.H.M.).

17 August 2000; accepted 16 October 2000

The Bacterial Flagellar Cap as the Rotary Promoter of Flagellin Self-Assembly

Koji Yonekura,[1]* Saori Maki,[1]* David Gene Morgan,[2] David J. DeRosier,[3] Ferenc Vonderviszt,[4] Katsumi Imada,[1] Keiichi Namba[1,5]†

The growth of the bacterial flagellar filament occurs at its distal end by self-assembly of flagellin transported from the cytoplasm through the narrow central channel. The cap at the growing end is essential for its growth, remaining stably attached while permitting the flagellin insertion. In order to understand the assembly mechanism, we used electron microscopy to study the structures of the cap-filament complex and isolated cap dimer. Five leg-like anchor domains of the pentameric cap flexibly adjusted their conformations to keep just one flagellin binding site open, indicating a cap rotation mechanism to promote the flagellin self-assembly. This represents one of the most dynamic movements in protein structures.

Self-assembly is used in biological systems to construct large molecular complexes and cellular organelles. In most cases, the assembly processes are regulated by conformational adaptability between the assembled proteins and those that form the site of assembly. The assembly mechanism is not based on simple lock-and-key interactions of compactly folded molecules but involves dynamic conformational changes and partial folding. Tobacco mosaic virus (*1*, *2*) and the bacterial flagellum (*3*, *4*) are classical examples of such mechanisms. These two systems vividly illustrate how the conformational flexibility and adaptability of biological macromolecules regulate the assembly processes. However, the structures of the assembly sites have never been visualized, even for these well-defined structures. In order to study the assembly process of the bacterial flagellar filament, we have used electron microscopy and single particle image analysis.

In many bacteria, swimming results from rotation of helical flagella driven by rotary motors at their bases (*5–7*). In *Escherichia coli* and *Salmonella*, the motor structure called the flagellar basal body crosses both cytoplasmic and outer membrane and continues as an extracellular structure called the hook and the filament. The assembly process starts with FliF ring formation in the cytoplasmic membrane (*8*) (Fig. 1A). A dedicated export apparatus, homologous to the type III protein export system (*9*), is believed to be integrated at the cytoplasmic opening of the FliF ring channel and to export selectively a set of flagellar proteins into the channel in the flagellum by using the energy of ATP hydrolysis (*10*, *11*). The flagellar proteins travel through the channel of the growing structure to the distal end, where the assembly occurs (*12*, *13*). The filament is only about 200 Å in diameter but grows to a length of up to 15 μm by polymerization of as many as 30,000 flagellin subunits. The central channel through which those flagellin subunits are transported is only 30 Å wide. The entire flagellum is built by self-assembly of the component proteins.

Just before the filament elongation starts in the growth process of the bacterial flagellum, HAP2 (hook-associated protein 2, also called FliD) forms a cap on top of the hook-HAP1-HAP3 complex (*14*). Then, flagellin subunits passing through the channel polymerize just below the cap one after another to form the long helical filament (*15*) (Fig. 1A). The cap stays attached at the distal end during the filament growth, and the simplest role of

[1]Protonic NanoMachine Project, ERATO, JST, 3-4 Hikaridai, Seika, Kyoto 619-0237, Japan. [2]Department of Biophysics, Boston University School of Medicine, Boston, MA 02118, and Department of Cell Biology, Harvard Medical School, Boston, MA 02254, USA. [3]Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02254, USA. [4]Department of Physics, University of Veszprém, Egyetem Street 10, H-8201 Veszprém, Hungary. [5]Advanced Technology Research Laboratories, Matsushita Electric Industrial Co., Ltd., 3-4 Hikaridai, Seika, Kyoto 619-0237 Japan.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: keiichi@crl.mei.co.jp