

Architecture and inherent robustness of a bacterial cell-cycle control system

Xiling Shen*, Justine Collier[†], David Dill[‡], Lucy Shapiro^{†§}, Mark Horowitz*, and Harley H. McAdams[†]

*Department of Electrical Engineering, 161 Packard Building, Stanford University, Stanford, CA 94305; [†]Department of Developmental Biology, Beckman B300, Stanford University, Stanford, CA 94305; and [‡]Department of Computer Science, Gates Building 344, Stanford University, Stanford, CA 94305

Contributed by Lucy Shapiro, May 30, 2008 (sent for review May 16, 2008)

A closed-loop control system drives progression of the coupled stalked and swarmer cell cycles of the bacterium *Caulobacter crescentus* in a near-mechanical step-like fashion. The cell-cycle control has a cyclical genetic circuit composed of four regulatory proteins with tight coupling to processive chromosome replication and cell division subsystems. We report a hybrid simulation of the coupled cell-cycle control system, including asymmetric cell division and responses to external starvation signals, that replicates mRNA and protein concentration patterns and is consistent with observed mutant phenotypes. An asynchronous sequential digital circuit model equivalent to the validated simulation model was created. Formal model-checking analysis of the digital circuit showed that the cell-cycle control is robust to intrinsic stochastic variations in reaction rates and nutrient supply, and that it reliably stops and restarts to accommodate nutrient starvation. Model checking also showed that mechanisms involving methylation-state changes in regulatory promoter regions during DNA replication increase the robustness of the cell-cycle control. The hybrid cell-cycle simulation implementation is inherently extensible and provides a promising approach for development of whole-cell behavioral models that can replicate the observed functionality of the cell and its responses to changing environmental conditions.

Caulobacter | model | regulatory circuit | hybrid system | symbolic model checking

The identification of regulatory pathways controlling the cell cycle of the bacterium *Caulobacter crescentus* has progressed to the point that an experimentally based system-level characterization of the cell-cycle control circuit is available. Aspects of the cyclical circuit that drives progression of the *C. crescentus* cell cycle are described in recent papers (1–5). Recent findings have identified novel mechanisms providing signals that tightly integrate the biochemical and genetic components of the core cell-cycle circuit with the 3D topology of the cell (4, 6, 7) and DNA replication (8).

Many of the regulatory mechanisms and pathways in the *C. crescentus* cell-cycle circuit operate very rapidly so that they approximate discrete switching elements, whereas others, e.g., the transcriptional regulatory networks, operate relatively slowly. This suggested use of hybrid control analysis methods (9–13) to simulate this cell-cycle control system and formal analysis methods to analyze its properties. Hybrid control systems, by definition, include both continuous and discrete regulatory mechanisms (9). We constructed a hybrid control system simulation model and derived an equivalent asynchronous discrete circuit model from the simulation model. The analysis showed that the cell-cycle control is robust and that DNA methylation-based cell-cycle regulatory mechanisms (8) enhance robustness of the cell cycle.

C. crescentus always divides asymmetrically, producing morphologically distinct daughters (Fig. 1A), so it is a model system not only for bacterial cell-cycle control but also for asymmetric cell division. The daughter cells have identical genotypes but different morphology and cell fates. The swarmer daughter cell has a single polar flagellum, polar chemotaxis receptors, and

polar pili, and it cannot initiate DNA replication until after a period of motility when the swarmer differentiates into a stalked cell. Swarmer-to-stalked cell differentiation involves loss of the flagellum and polar chemotaxis receptors, retraction of the pili, construction of a polar stalk, and initiation of DNA replication. In contrast, the stalked daughter cell initiates chromosome replication after cytoplasmic compartmentalization during cytokinesis. This compartmentalization event occurs ≈ 20 min before completion of cell division after chromosome decatenation (Fig. 1A) (6). Compartmentalization initiates divergent genetic programs in each compartment, so that the next cell cycle effectively begins in each compartment well before cell division (4, 6, 7).

The *C. crescentus* cell-cycle control system is hierarchical, distributed, asynchronous, and parallel. The control circuit comprises phosphosignaling pathways, genetic regulatory logic, dynamical control of regulatory protein localization, activation, and stability, and two processive subsystems, chromosome replication and cytokinesis. Parallelism arises from activation of multiple modular cell-cycle functions that operate at the same time. Four master regulator proteins (DnaA, GcrA, CtrA, and CcrM), organized as a cyclical genetic circuit, serially activate many subsystems to implement the cell cycle (Fig. 1B–D and supporting information (SI) Fig. S1) (3). Fig. 1B shows the changing concentrations of these regulatory proteins that activate or repress controlled subsystems. Fig. 1C shows principal signal pathways from the core circuit to the DNA replication and cell constriction subsystems. Figs. 1D and Fig. S1 show feedback signal pathways from these two controlled subsystems to the core circuit to tie the rate of progression of the core circuit to the progression of the cell cycle. The cell-cycle control system is synchronized tightly with the progression of chromosome replication and cytokinesis by both genetic and nongenetic mechanisms (8, 14). The rates of progress of the parallel independent reaction cascades are inherently unpredictable, but some pathways must be completed in a particular order. Synchronization is accomplished by several mechanisms: (i) tying subsystem activation time to cyclical regulatory protein concentrations and activation levels, (ii) tying gene expression levels to intermediate events in the processive subsystems [e.g., promoter hemimethylation during DNA replication (8)], and (iii) checkpoint signaling (e.g., delaying completion of cell constriction until after DNA decatenation).

In the next section, we describe the hybrid control simulation model and its validation by comparison with experimental observations. Then we describe use of model checking to analyze

Author contributions: L.S., M.H., and H.H.M. designed research; X.S., J.C., D.D., and H.H.M. performed research; X.S., J.C., D.D., and H.H.M. analyzed data; and X.S., J.C., D.D., and H.H.M. wrote the paper.

The authors declare no conflict of interest.

[§]To whom correspondence should be addressed. E-mail: shapiro@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0805258105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA

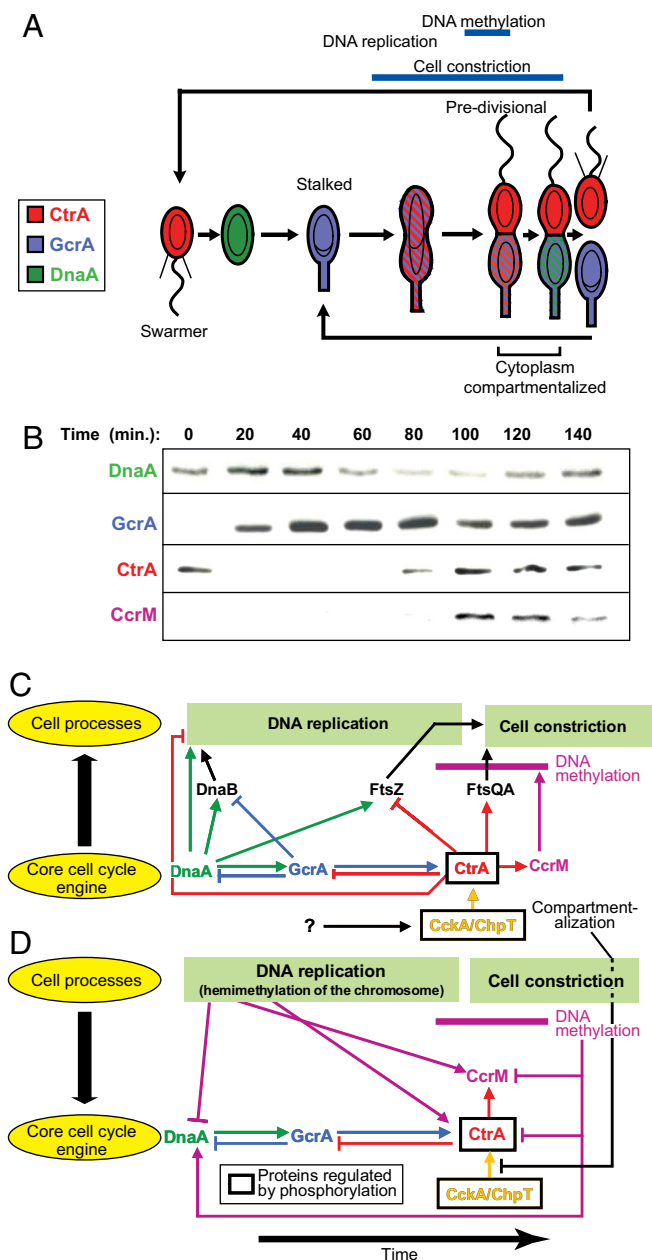


Fig. 1. Genetic circuit that drives cell-cycle progression. (A) Schematic of the *C. crescentus* cell cycle showing changes in master regulatory protein concentrations that control activation of numerous modular functions that implement the cell cycle. Predivisional cells are compartmentalized ≈ 20 min before cell separation (6). (B) Western blots showing concentrations of the master regulatory proteins during the cell cycle (8, 26). (C) Schematic of protein interactions that create the cyclical core engine and the regulatory connections from the core engine to the processive DNA replication and cell constriction functions. Events are positioned to indicate their approximate timing. (D) Signals returning from the controlled subsystems synchronize the core engine with the state of cell-cycle progression. The interval of DNA methylation is indicated in C and D by the horizontal purple bar. All of the pathways indicated in C and D are included in the cell-cycle control model.

robustness of the cell-cycle control and control circuit features that contribute to robustness.

Cell-Cycle Hybrid Control System Simulation Model. Several descriptions of the aspects of *C. crescentus* control system circuitry are available (1, 3, 5). A stalked cell-cycle subcircuit model has been

reported (15), but with no analysis of robustness or modeling of the control of CtrA phosphorylation. We developed a scalable simulation of control of the coupled swarmer and stalked cell cycles with emphasis on compatibility with formal analysis. We wanted to predict the progress of the regulatory machinery into either compartment of the predivisional cell to enable comparative analysis of regulatory events within the respective compartments. These events that establish asymmetry are difficult to observe experimentally.

Sources for the signaling in the cell-cycle control system, the regulatory logic in promoter regions, and protein stability models are in Table S1 in *SI Appendix*. The protein components of the genetic circuit include the four master regulator proteins (DnaA, GcrA, CtrA, and CcrM) that comprise the core cyclical circuit and DnaB, FtsZ, and FtsQA (Fig. S2B). DnaB, FtsZ, and FtsQA are components of pathways that connect the core engine with DNA replication and cytokinesis (Fig. 1 C and D). This is a parsimonious model of the cell-cycle control circuitry. For example, DnaB is only one of the proteins in the replication complex whose synthesis is activated by DnaA (16). FtsQA represents two proteins, FtsQ and FtsA (whose genes are in an operon), required for initiation of cell constriction. The simulation includes phenomenological models of the progress of chromosome replication and cell constriction. These two subsystems determine the timing of the changes in the methylation state of the *dnaA*, *ctrA*, and *ccrM* promoter regions and of cell compartmentalization (8, 17).

Cytoplasmic compartmentalization is essentially instantaneous (6). The promoter methylation state changes when the replication fork traverses promoters that involve this mechanism (8) are also instantaneous, as is the remethylation reaction. Positive autoregulation of the *ctrA* P₂ promoter by CtrA~P leads to rapid synthesis of CtrA (Fig. S1). Methylation state dependence of the *ctrA* P₁ promoter tends to delay *ctrA* activation until the gene is replicated, providing additional assurance that CtrA is synthesized rapidly and at the right time. CtrA~P is cleared from the cell by simultaneous activation of proteolysis and inhibition of its activation so that the elimination of CtrA~P is also rapid. The rapid creation and destruction of CtrA~P shortens average switching times of downstream genes and other processes by activated CtrA~P because of accelerated traversal of the “threshold” regions for these reactions (18).

Each of these events triggers discrete changes in the network of gene transcription, phosphosignaling, and protein-level reactions that control the cell cycle. Additional description of pathways is in *SI Appendix*. In modeling the system, we used mechanistic models where the reaction mechanisms are known, for example, the promoter control mechanisms of the master regulator genes in the core engine (Fig. S1). In the case of the CckA/ChpT phosphosignaling pathway (Fig. 1 C and D and Fig. S1), mechanisms of the pathway are not completely identified, but the function of the pathway and the timing of its operation within the cell cycle are well characterized (14). Similarly, the time from onset to completion of DNA replication and the time from initial cell constriction to compartmentalization are known (19). In these cases, we used phenomenological, but functionally accurate, models with correct cell-cycle timing. Reactions that occur essentially instantaneously were modeled as discrete switching events. The resulting hybrid control simulation enabled analysis of the *C. crescentus* cell-cycle control as an integrated system.

The simulation model is constructed by using the Matlab tools, Simulink and Stateflow, that are widely used for control system analysis (see *SI Appendix*). Innovative aspects of this simulation model are: (i) a top-down hierarchical simulation architecture that mimics the design of the cell’s control system (Figs. S2–S4) designed for future incorporation of additional regulatory subsystems (e.g., control of polar organelle development, a mech-

strate robustness, that is, “invariance of phenotypes in the face of perturbation” (21). Commonly, robustness of biological circuits is analyzed by assessing the range of parameter variation within which the circuit functions satisfactorily by checking performance with random choices of parameter values. We performed a limited parameter sensitivity analysis on a subset of the parameters in the simulation, and the results suggested that the operation of the cell-cycle control is robust. Then we used the validated simulation model to identify an equivalent asynchronous sequential digital circuit model and used formal model checking software for further analysis.

Fig. 3A illustrates the rationale underlying our approach to identifying the discrete abstraction to the hybrid simulation model and application of the model checking methodology for robustness analysis. The figure shows eight events in an interval centered on the time of daughter cell separation that will occur in a stalked cell cycle (Fig. 1A and Fig. S1) and affect operation of the control circuit. The events occur at times labeled by the T_i s. T_1 , for example, is the time of cytoplasmic compartmentalization that interrupts the CckA/ChpT phosphosignal and precipitates elimination of CtrA~P. The specific timing of events and the intervals between the events will depend on the kinetic parameters of the system, whether *in vivo* or *in silico*. (Because of stochastic variation in reaction rates from cell to cell in a population, actual reaction rates in different cells will vary, and consequently cells will show dispersion in the rate of progression through the cell cycle.) Thus, the rate of production of any protein X will depend on the kinetic parameters of its transcription and translation, and the time for X to activate a dependent process will depend on the kinetics of its binding at a target site that determine the concentration range for action. The rate of production of X will also depend on availability of substrates (e.g., amino acids) that will vary with environmental nutrient levels. There is a minimum average cell generation time under optimal high nutrient conditions. Under either carbon or nitrogen starvation, the cell cycle comes to a controlled stop (see below), suggesting there is also a maximum to the sustainable generation time. Thus, a defining element of the robustness of the *C. crescentus* cell-cycle control is its ability to achieve reliably correct ordering of cell-cycle functions between these extremes of cell growth rate. Our method for exploring this is to consider variation of the event times (i.e., the T_i s in Fig. 3A) directly, rather than individual variation of the many kinetic parameters that determine the times. This approach to robustness analysis is more computationally efficient and complete than using a Monte Carlo method to explore the parameter space.

For example, consider the particular situation shown in Fig. 3A, in which the condition for activation of DNA replication involving the CtrA, DnaA, and DnaB proteins is satisfied at T_3 . After a delay, the *dnaA* gene is duplicated at T_4 . Consequently, its promoter region becomes hemimethylated and the rate of *dnaA* transcription is greatly reduced (8). Although there are essential ordering relationships between some timing events, for example: $T_2 > T_1$, $T_4 > T_3$, $T_6 > T_5$, $T_6 > T_4$, and $T_8 > T_7$, the lengths of the various interevent intervals are determined by the specific reaction rates that occur by chance in each cell under the prevailing environmental conditions.

To facilitate exploration of the effects of timing variations, we simplify the protein concentration profiles as shown in Fig. 3A so that they are either “low” (under the thresholds of downstream sites of action) or “high” (above the thresholds of downstream sites of action), because other than time and direction of traversing the threshold region, details of the temporal profile of these variables are not significant to the operation of the circuit. This yields a discrete logic abstraction of the signaling and the time points that identify the timing of the transitions between the two states (e.g., T_5 and T_6 in Fig. 3A). Other events can also be characterized as binary signals or finite

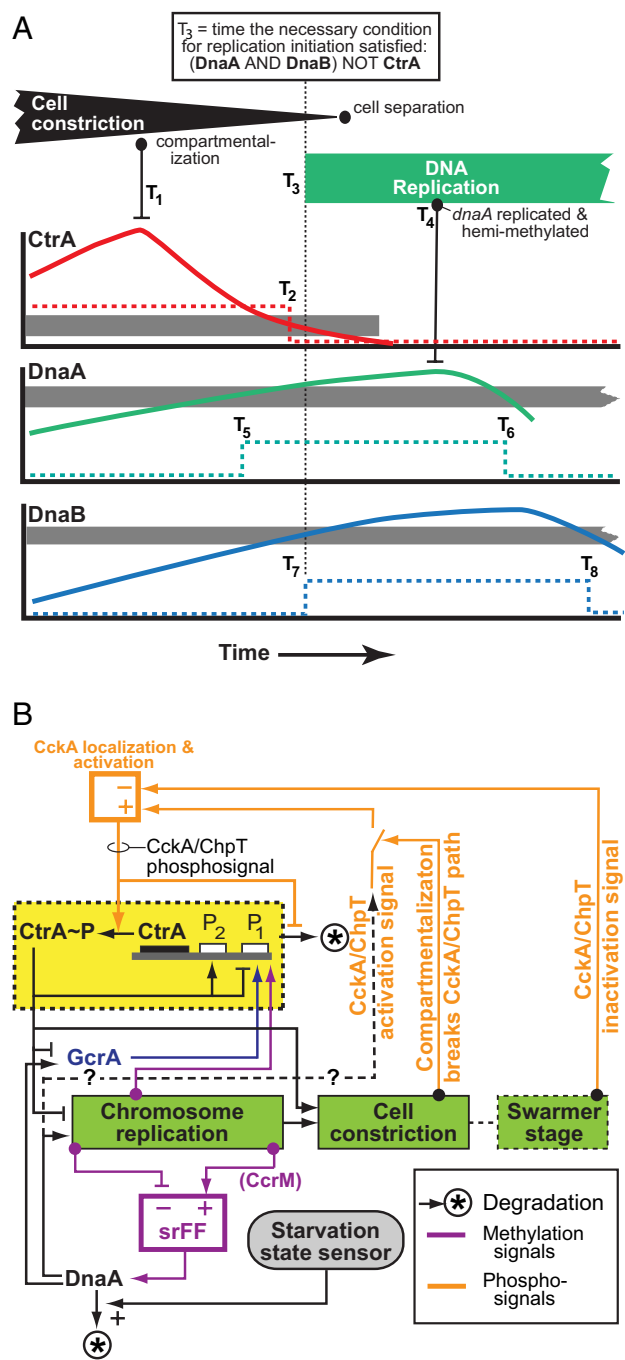


Fig. 3. Logic circuit of *Caulobacter* cell cycle control system. (A) Stalked cell-cycle events affecting control circuit operation in the interval around cell separation. Graphs show protein concentration patterns. Gray bands: “threshold” ranges relating to initiation of DNA replication. [CtrA] above the band represses initiation; below, it does not. The dotted CtrA line illustrates the binary signal abstraction used in the robustness analysis. Thus, T_2 approximates the time when [CtrA] transitions from repressing to not-repressing initiation. Other T_i s are times of other events. The robustness analysis examined the effect of all patterns of T_i orderings on cell-cycle control. (B) Logical diagram of cell-cycle control system operations. Green rectangles: Serially ordered processive subsystems. Yellow box: cyclical synthesis and destruction of CtrA~P that is a central element of the circuit. Orange pathways: CckA/ChpT phosphosignal pathways differentially controlled in swarmer and stalked cell compartments that play an essential role in both enabling resynthesis of CtrA~P when the signal is active and triggering CtrA~P destruction when it is inactive. Purple: DNA methylation-state based control of DnaA synthesis that creates a logical component functionally similar to an electrical set-reset flipflop (srFF) circuit (8). Carbon or nitrogen starvation halts cell-cycle progression by accelerating DnaA proteolysis (27).

state machines with discrete stages. For example, in Fig. 3A, the cell changes from not compartmentalized to compartmentalized at T_1 , and the *dnaA* promoter changes from methylated to hemimethylated at T_4 . This simplification of the cell-cycle control circuit by using discrete signal levels with discrete transition timing produces an abstraction of the biological circuit equivalent to an asynchronous sequential digital circuit, a network of logic elements and simple-state machines with variable delays. Within this abstraction, it is necessary only to examine each distinctive ordering of events to determine whether the circuit will function correctly. Following electrical circuit analysis procedures for similar problems, we use a software model checking tool to search the immense space of all feasible orderings of the T_i s for cases where the circuit might fail. This approach using a symbolic model checker (22) is used to check for correct operation of concurrent systems such as electrical circuits and network protocols; we use it to check the cell-cycle control circuit, which is also a concurrent system. The model checker we used, NuSMV (22, 23), takes as input the logic description of the *C. crescentus* cell-cycle model (created from the validated simulation model as described above) and a specification of the limitations on ordering of transitions that must apply for a viable cell. The program checks all feasible orderings of the transition times and notifies the user of any event timing “hazards” where the specification would be violated.

A challenge in any method for assessing biological robustness is defining requirements for successful operation of the modeled system. For the *C. crescentus* cell-cycle control, the key functionality of the control circuit is to ensure that actions that affect completion of the cell cycle, cell growth, chromosome replication, asymmetric cell division, and so forth, repeatedly occur in the proper order. This includes ordering of expression of the seven proteins in the model (Fig. S1), switching of the CckA/ChpT phosphosignaling pathway, and activation of DNA replication followed by cell constriction, compartmentalization into nascent swarmer and stalked cell compartments, and cell division. Thus, we required that for all feasible orderings of the event T_i s, the model system would successfully produce an unbounded succession of Stalked→DNA-replication→Compartmentalization→(Swarmer or Stalked) states as occurs in the tree of descendent daughter cells of an initial swarmer cell. This criterion is written in computational tree logic to cover all possible paths in the state space (24). For the subset of cell-cycle subsystems we model and for the asynchronous logic abstraction of the model, satisfaction of this requirement is evidence that the cell-cycle control is robust. A hazard detected by the model checking can be a biological possibility that has acceptably low probability of occurrence or it may represent a feature of the biological system that prevents the hazard from occurring, but is missing in the simulation model.

Model-Checking Results. The model-checking program exhaustively checked the enormous number of alternative possible T_i combinations, and only two potential hazards were identified. This suggests that the overall design of the cell-cycle control circuit has been optimized by evolutionary selection to operate over a wide range of nutrient conditions and to be resistant to stochastic variations in time to complete various subsystem operations or signaling pathways. Further, it is evidence that conclusions relating to operation of the cell-cycle control based on the simulation model do not depend on the exact parameter values for the simulation. The two hazards identified both relate to repressive feedback signals (GcrA repression of DnaB expression and CtrA repression of FtsZ expression; see Fig. S1). In both of these cases, the hazard identified is that anomalously slow synthesis of the protein coupled with anomalously fast synthesis of the repressive feedback signal might block cell-cycle progression. Both cases appear to have relatively low probability

of occurring, and there also may be undiscovered aspects of the regulation of both DnaB and FtsZ that eliminate the hazard.

We also examined the incremental contribution of the methylation-based control of the *ctrA* P_1 promoter to robustness of the cell-cycle control. When we simulated an *in silico* mutant where full methylation does not repress the *ctrA* P_1 promoter, the cell cycle operated correctly; however, applying the model checking program to this *in silico* mutant identified an additional hazard: in this mutant, CtrA could be synthesized too early, leading to premature repression of FtsZ and thus failure of cytokinesis. This suggests that, although this methylation-based mechanism is not essential, it contributes to fitness of the organism by assuring correct operation of the cell cycle in the fraction of cells where stochastic variation produces anomalously early expression of CtrA.

The DNA methylation-based control of *dnaA* transcription also contributes to robustness of cell-cycle control by reducing its expression during DNA replication, so that the likelihood of reinitiation of replication is reduced. However, the methylation-based repression of *dnaA* is not complete, so that some expression remains (8). Model checking analysis of an *in silico* mutant without this remaining basal expression identified another hazard: in a situation where DNA replication is unusually slow, GcrA could be degraded before the *ctrA* P_1 promoter becomes hemimethylated. Then, CtrA would not reaccumulate in predivisional cells to activate the synthesis of FtsQA, and cell constriction would not occur. The low-level synthesis of DnaA by the *dnaA* basal expression, however, can restart GcrA synthesis and, in turn, CtrA synthesis to rescue the cell.

Halting and Restarting the Cell Cycle. The DnaA protein is strategically located in the circuit to simultaneously activate DNA replication and the GcrA-CtrA-CcrM pathway at the beginning of each cell cycle (Fig. 1C and Fig. S1) (16, 25, 26). Elimination of DnaA halts the cell cycle (25), and control of DnaA stability is a mechanism used to halt the *C. crescentus* cell cycle in response to stress (27). When *C. crescentus* cells are starved for carbon or nitrogen, DnaA is rapidly proteolyzed, and relief of the starvation rapidly restabilizes DnaA (27). We used model checking to assess whether this DnaA stability-dependent method of halting and restarting the cell cycle is robust by adding an additional switch in the model to modulate DnaA stability. This analysis showed that the cell-cycle stops and restarts correctly regardless of when DnaA stability changes during the cell cycle.

Discussion

A small number of spatially and temporally controlled regulatory proteins create a core cell-cycle engine that drives progression of the *C. crescentus* cell cycle. The cell-cycle control system is hierarchically organized, and it activates functional modules just when needed. The cascaded cell-cycle reaction pathways proceed in a near mechanical step-like fashion. The engine and modular cell-cycle subfunctions are synchronized by mechanisms that connect the timing of gene expression and the phosphorylation of CtrA to progression of cell-cycle functions. There are checkpoint signals between modules that further assure proper ordering of functions. Top-level regulatory proteins are redundantly regulated at the levels of transcription, stability, and activity. We considered operation of a bacterial cell cycle from a hybrid control system engineering perspective to ask whether the cell-cycle control is robust, and whether stopping and restarting the cell cycle in response to nutritional stress is robust. Our analysis shows the control system design will operate correctly over a wide range of cell generation times as necessary to adapt to nutrient availability, and, in periods of starvation, to completely stop cell-cycle progression. Furthermore, we showed through analysis by model checking that the recently discovered “methylation ratchet” (8) is a mechanism that has evolved to

