

Circadian Control of Global Gene Expression by the Cyanobacterial Master Regulator RpaA

Joseph S. Markson,^{1,3,5} Joseph R. Piechura,^{1,2,5} Anna M. Puszynska,^{1,2} and Erin K. O'Shea^{1,2,4,*}

¹Howard Hughes Medical Institute, Harvard University Faculty of Arts and Sciences Center for Systems Biology

²Department of Molecular and Cellular Biology

³Graduate Program in Biophysics

⁴Department of Chemistry and Chemical Biology
Harvard University, Cambridge, MA 02138, USA

⁵These authors contributed equally to this work

*Correspondence: erin.oshea@harvard.edu
<http://dx.doi.org/10.1016/j.cell.2013.11.005>

SUMMARY

The cyanobacterial circadian clock generates genome-wide transcriptional oscillations and regulates cell division, but the underlying mechanisms are not well understood. Here, we show that the response regulator RpaA serves as the master regulator of these clock outputs. Deletion of *rpaA* abrogates gene expression rhythms globally and arrests cells in a dawn-like expression state. Although *rpaA* deletion causes core oscillator failure by perturbing clock gene expression, rescuing oscillator function does not restore global expression rhythms. We show that phosphorylated RpaA regulates the expression of not only clock components, generating feedback on the core oscillator, but also a small set of circadian effectors that, in turn, orchestrate genome-wide transcriptional rhythms. Expression of constitutively active RpaA is sufficient to switch cells from a dawn-like to a dusk-like expression state as well as to block cell division. Hence, complex global circadian phenotypes can be generated by controlling the phosphorylation of a single transcription factor.

INTRODUCTION

The circadian clock of the cyanobacterium *Synechococcus elongatus* PCC7942 drives daily genome-wide oscillations in mRNA expression levels, controls genome compaction and supercoiling, and modulates cell division (Johnson et al., 2011). The clock contains a core oscillator consisting of the proteins KaiA, KaiB, and KaiC, which together generate circadian (i.e., ~24 hr) oscillations in KaiC phosphorylation (Markson and O'Shea, 2009). Remarkably, the KaiC phosphorylation oscillations observed *in vivo* can be reconstituted *in vitro* simply by mixing the three Kai proteins and ATP (Nakajima et al., 2005).

In vivo, this proteinaceous posttranslational oscillator (PTO) is embedded in a transcription-translation feedback loop (TTL) that regulates expression of the *kaiBC* operon, enhancing the precision of the clock by stabilizing its phase (Johnson et al., 2011; Qin et al., 2010; Teng et al., 2013; Zwicker et al., 2010).

Although much is known about the mechanism by which the PTO keeps time, less is understood about how the clock uses the time information encoded in the PTO to generate outputs like global gene expression rhythms and modulation of cell division. Most genes show circadian expression oscillations in constant light, displaying a variety of amplitudes, phases, and waveforms (Ito et al., 2009; Johnson et al., 2011; Vijayan et al., 2009). The distribution of phases is bimodal, with expression of one population peaking around subjective dusk (class 1) and the other peaking around subjective dawn (class 2). ("Subjective dusk" and "subjective dawn" refer to the times at which light-to-dark or dark-to-light transitions would occur in a 12 hr light-12 hr dark environmental cycle. The term "subjective" is used when, as here, environmental conditions are held constant in order to isolate clock-driven from environmentally driven processes.) Genome compaction and DNA supercoiling also oscillate in a circadian manner (Smith and Williams, 2006; Woelfle et al., 2007), and supercoiling oscillations contribute to the generation of gene expression oscillations (Vijayan et al., 2009). Finally, the clock gates cell division, prohibiting it during the subjective night (Dong et al., 2010; Mori et al., 1996; Yang et al., 2010).

Genetic and biochemical approaches have revealed key players in the output pathway that connects the PTO to these output responses. The response regulator RpaA has been implicated genetically in circadian gene expression control (Takai et al., 2006; Taniguchi et al., 2007). The phase of the PTO is transmitted to RpaA via the histidine kinases SasA and CikA (Gutu and O'Shea, 2013; Takai et al., 2006; Taniguchi et al., 2010): *in vivo*, SasA phosphorylates and CikA dephosphorylates RpaA via their respective kinase and phosphatase activities. The PTO generates temporal separation of SasA and CikA activities, producing circadian oscillations in RpaA phosphorylation levels (Gutu and O'Shea, 2013). The response regulator RpaB, a

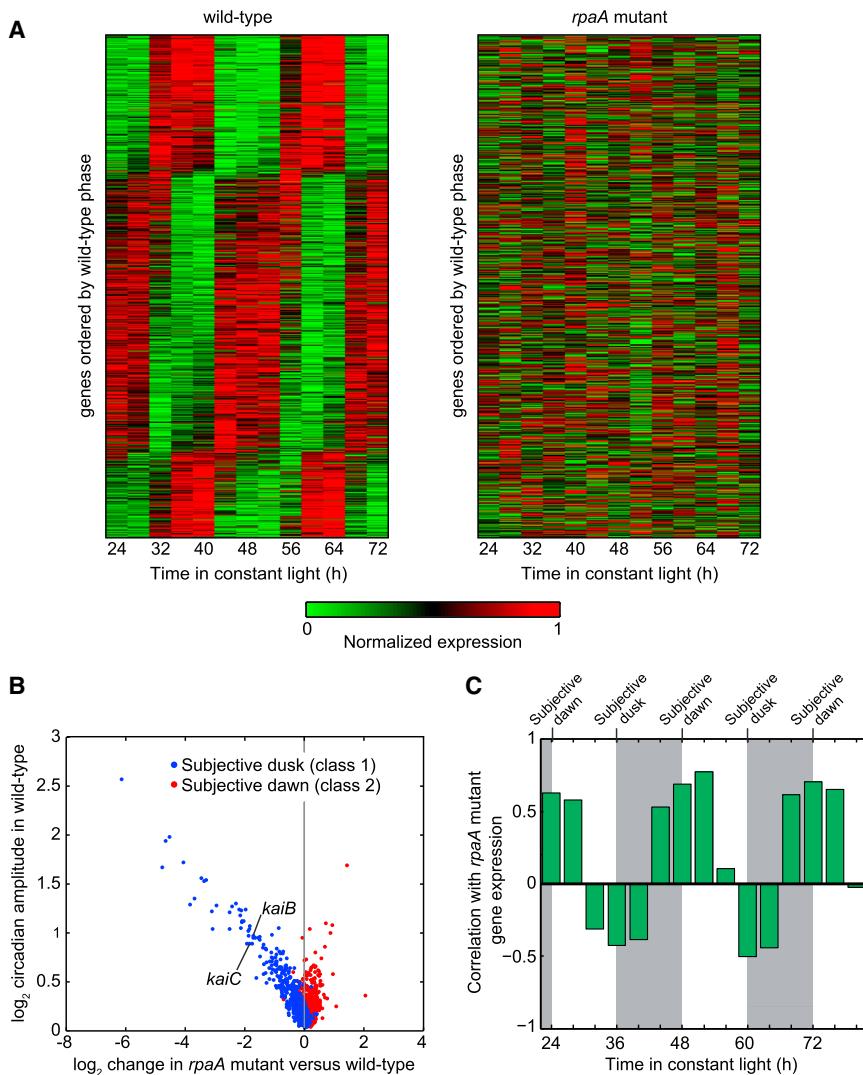


Figure 1. Gene Expression Is Globally Perturbed by Deletion of *rpaA*

(A) Global gene expression time course in the wild-type and *rpaA* mutant. (Left) Circadian gene expression in the wild-type strain (data from Vijayan et al. [2009]). Expression time courses of genes reproducibly oscillating with a circadian period ($n = 856$; see Extended Experimental Procedures) were normalized to the interval [0 1] and sorted by phase. (Right) Gene expression in the *rpaA* mutant. Expression time course of the same set of genes as in wild-type, displayed in the same order.

(B) Comparison of gene expression change in the *rpaA* mutant with circadian amplitude in the wild-type strain. We computed the expression change in the *rpaA* mutant by comparing the average *rpaA* mutant expression over 1 day with the average wild-type expression over 1 day (see Experimental Procedures). Only genes that oscillate with circadian periodicity in the wild-type strain are shown ($n = 856$). *kaiB* and *kaiC* are indicated; *kaiA* is not classified as circadian and hence is not shown.

(C) Correlation of global gene expression in the *rpaA* mutant with each time point in the wild-type time course. Expression of all genes (both circadian and noncircadian) in the *rpaA* mutant was time averaged as in (B), and the correlation between this time-averaged expression and the expression in the wild-type strain at each time point over a 60 hr time span was computed. Wild-type data are from Vijayan et al. (2009). Subjective night is shaded. Subjective dawn occurs at 24, 48, and 72 hr, and subjective dusk falls at 36 and 60 hr. See also Figure S1 and Table S1.

paralog of RpaA, also may play a role in circadian clock output: it binds in vitro and in vivo to the promoters of several circadian genes, and this binding is antagonized by RpaA in vitro (Hanaoka et al., 2012).

Deletion of *rpaA* eliminates oscillations in the activity of the ten circadian promoters that were assayed by bioluminescence reporters (Takai et al., 2006), but the role of RpaA in regulating circadian expression genome wide is not known. RpaA is predicted to be a transcription factor (Takai et al., 2006), but recent studies failed to detect binding of RpaA to candidate target promoters (Hanaoka et al., 2012; Takai et al., 2006). Similarly, cell division gating by the clock requires RpaA (Dong et al., 2010), but the mechanism is not understood.

Here, we employ a multifaceted approach to elucidate the molecular and functional roles of RpaA in circadian clock output. We find that circadian gene expression oscillations are absent genome wide in an *rpaA* mutant, with cells being arrested in a subjective dawn-like transcriptional state. Through chromatin immunoprecipitation with high-throughput sequencing

(ChIP-seq) and in vitro DNase I footprinting, we show that phosphorylated RpaA binds directly to >100 locations in the genome, including the promoter of *kaiBC*. Finally, we show that overexpression of a

phosphomimetic mutant of RpaA is sufficient to drive cells from the subjective dawn to the subjective dusk gene expression state and also to close the cell division gate, demonstrating that RpaA is the global regulator of circadian output in this organism.

RESULTS

RpaA Is Required for Global Circadian Gene Expression, and Its Deletion Arrests Cells in a Subjective Dawn-like State

To determine whether RpaA is required for global gene expression rhythms, we measured gene expression over 48 hr by microarray in an *rpaA* mutant strain (Figure 1A). We found that circadian oscillations were abolished genome wide (Figure 1A), even for the genes that oscillate with the highest amplitude in the wild-type (Figure S1 available online). Hence, RpaA is required for the generation of global circadian gene expression rhythms.

To gain insight into the role of RpaA in circadian gene regulation, we searched for genes that showed the greatest magnitude of transcript level difference between the wild-type and *rpaA* mutant strains (Table S1). Consistent with previous results (Takai et al., 2006), *kaiBC* expression decreased ~3.5-fold, whereas *kaiA* expression was not affected substantially; this disparate effect on *kai* gene expression perturbs Kai protein stoichiometry, likely situating it in a regime that does support PTO function (Nakajima et al., 2010). Overall, we found that the expression of 67 genes decreased >2-fold, whereas the expression of 16 genes increased by at least that amount. Strongly downregulated genes included four sigma factors, two transcription factors, and several genes encoding proteins involved in energy production and metabolism (particularly carbohydrate metabolism). Circadian genes downregulated in the *rpaA* mutant were highly enriched for subjective dusk phasing, whereas upregulated genes were enriched for subjective dawn phasing (Figure 1B). Interestingly, the decrease in expression of subjective dusk genes in the *rpaA* mutant is directly proportional to the gene's circadian amplitude in the wild-type strain (Figure 1B). These observations suggest that RpaA is responsible for promoting subjective dusk gene expression and repressing subjective dawn gene expression. Consistent with this scenario, global gene expression in the *rpaA* mutant is most positively correlated with wild-type subjective-dawn expression and is most negatively correlated with wild-type subjective dusk expression (Figure 1C). Hence, deletion of *rpaA* arrests cells in a subjective dawn-like state.

Rescue of PTO Function in the *rpaA* Mutant Reveals that RpaA Is Directly Responsible for the Orchestration of Circadian Gene Expression

Interpretation of the role of RpaA is confounded by the loss of PTO function in the *rpaA* mutant (Takai et al., 2006): the absence of gene expression oscillations in the *rpaA* mutant could merely be a secondary effect of the loss of PTO function rather than an indication of a master regulator role for RpaA. Indeed, it is possible that the primary role of RpaA is simply to sustain Kai posttranslational oscillations by, for example, modulating *kaiBC* expression to maintain permissive Kai protein stoichiometry.

To distinguish between direct and indirect contributions of RpaA to circadian gene expression, we asked whether rescuing the PTO in an *rpaA* mutant background would restore global gene expression oscillations. We rescued the PTO by ectopically expressing *kaiBC* from the IPTG-inducible *Ptrc* promoter (Murrayama et al., 2008) in a $\Delta rpaA \Delta kaiBC$ background ($\Delta rpaA \Delta kaiBC$ *Ptrc::kaiBC*) and, as a control, in a $\Delta kaiBC$ background ($\Delta kaiBC$ *Ptrc::kaiBC*). We refer to these strains as the $\Delta rpaA$ clock rescue and the control clock rescue, respectively. In both strains, we obtained strong KaiC phosphorylation rhythms of similar amplitude in the presence of 6 μ M IPTG (Figures 2 and S2A). Note that ectopic expression of KaiA was not necessary because its levels are not substantially perturbed by deletion of *rpaA* (Takai et al., 2006).

We used microarrays to examine global gene expression dynamics in the two rescue strains (Figure 2 and Table S2). In the control clock rescue, circadian gene expression oscillations were restored robustly for both subjective dusk and subjective

dawn genes (Figures 2, S2B, and S2C). In contrast, no strong oscillations were observed in either class of genes in the $\Delta rpaA$ clock rescue (Figures 2, S2B, and S2C). We conclude that RpaA is required for the orchestration of robust circadian gene expression by the PTO, ruling out the scenario in which RpaA's role is limited to supporting oscillator function.

Phosphorylated RpaA Binds to the *kaiBC* Promoter and Upregulates *kaiBC* Expression

Although the rescue experiments demonstrate that RpaA plays a critical and central role in producing global gene expression rhythms, they do not indicate the means by which it does so. As previous studies have failed to identify direct RpaA binding to DNA, it has been proposed that RpaA acts indirectly via displacement of RpaB from circadian promoters (Hanaoka et al., 2012). However, these studies have assayed only a few regions for RpaA binding, and it is possible that RpaA acts via rhythmic binding in conditions or at locations that have not been examined.

To determine whether and where RpaA associates with genome, we generated an anti-RpaA antibody (Figure S3A) and used it to perform chromatin immunoprecipitation (ChIP) over a 36 hr circadian time course. We found that RpaA localized to the *kaiBC* promoter in an oscillatory manner, in phase with both RpaA phosphorylation (RpaA~P) and expression of the *kaiBC* transcript (Figures 3A and S3B). These results are not an artifact of off-target binding of the anti-RpaA antibody, as ChIP-qPCR analysis of a strain expressing epitope-tagged RpaA (HA-RpaA) using an anti-HA antibody demonstrated that it localized to the *kaiBC* promoter with a circadian period and in phase with HA-RpaA phosphorylation (Figure S3C).

To determine whether localization of RpaA to the *kaiBC* promoter (*PkaiBC*) results from direct association with DNA, we assayed for physical interaction of RpaA with promoter DNA in vitro using DNase I footprinting. We observed a clear footprint, strictly dependent on RpaA phosphorylation, between 29 and 51 bp upstream of the transcription start site (Figure 3B). The strict phosphorylation dependence may explain the inability to detect RpaA binding to the *kaiBC* promoter in previous *in vitro* studies (Hanaoka et al., 2012; Takai et al., 2006): these studies used either unphosphorylated RpaA (Takai et al., 2006) or RpaA putatively phosphorylated by treatment with acetyl phosphate (Hanaoka et al., 2012), which we found does not actually result in RpaA phosphorylation (Figure S3D). Intriguingly, the footprint of RpaA~P in the *kaiBC* promoter coincides with a region in which mutations substantially reduce promoter activity *in vivo* (Kutsuna et al., 2005). We tested three of those mutations in the footprinting assay, finding that all of them reduced or eliminated RpaA~P binding (Figure S3E).

To test whether RpaA~P drives expression of the *kaiBC* transcript *in vivo*, we expressed phosphorylation site mutants mimicking phosphorylated or unphosphorylated RpaA in an *rpaA* mutant background and used a *PkaiBC*-driven luciferase reporter to assess the effect of these mimetics on *PkaiBC* activity (Figure 3C). Aspartate 53 (D53) of RpaA is predicted by sequence homology with the well-studied *E. coli* OmpR protein to be the site of phosphorylation by SasA, and we therefore used a glutamate mutation at residue 53 (D53E) to mimic RpaA~P and an

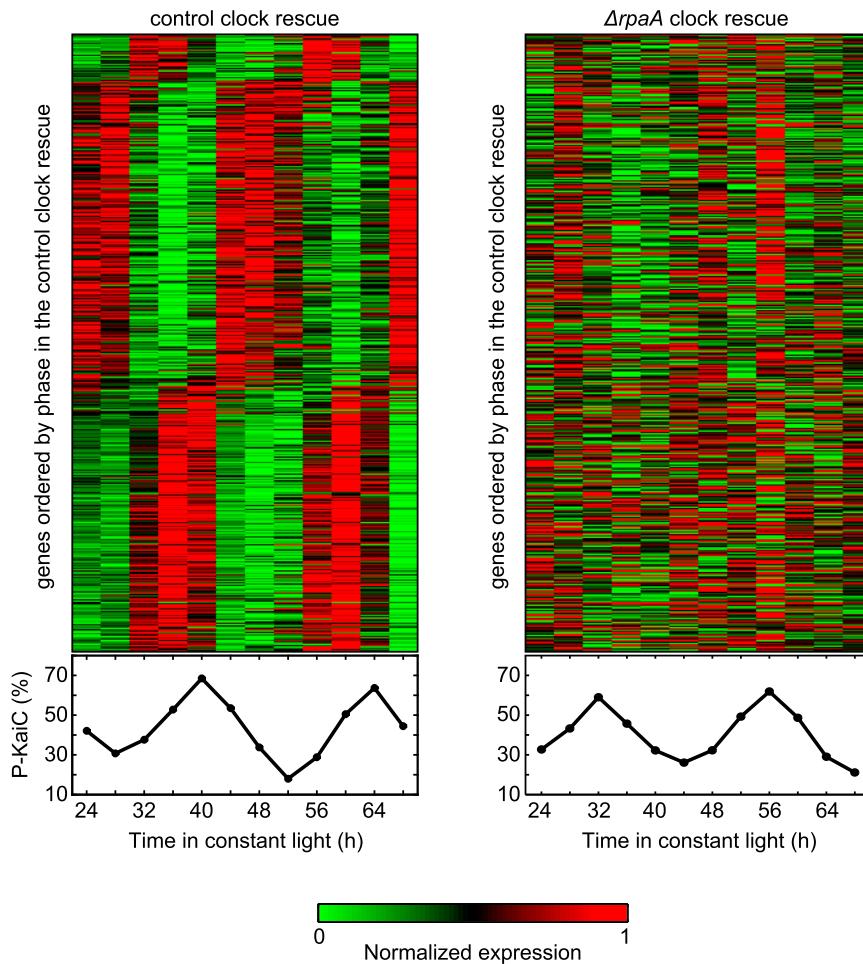


Figure 2. RpaA Is Required for Control of Global Gene Expression by the PTO

(Left) Gene expression time course upon rescue of Kai oscillator function in a $\Delta kaiBC$ background via ectopic expression of *kaiBC* ($\Delta kaiBC$ *Ptrc::kaiBC*, termed the “control clock rescue”). The heatmap shows the expression time course of the 471 genes that oscillate with a circadian period in the control clock rescue in each of two biological replicates, with genes normalized individually and ordered by phase. KaiC phosphorylation levels during the time course are shown below the heatmap.

(Right) Gene expression time course upon rescue of Kai oscillator function in the absence of RpaA via ectopic expression of *kaiBC* ($\Delta rpaA$ $\Delta kaiBC$ *Ptrc::kaiBC*, termed the “ $\Delta rpaA$ clock rescue”). The heatmap shows the expression time course of the same set of genes as in (A), displayed in the same order. KaiC phosphorylation levels during the time course are shown below the heatmap. See also Figure S2 and Table S2.

by high-throughput sequencing (ChIP-seq). We identified 110 binding sites (peaks), all located on the main chromosome (see [Experimental Procedures](#)) (Figure 4A and Table S3A). A well-defined binding site was identified upstream of the *kaiBC* locus, the occupancy of which varied with circadian time (Figure 4B). RpaA binds in a circadian manner and with a similar phase at all 110 binding sites (Figure 4C).

To ensure that the enrichment we observed did not reflect off-target binding

of the anti-RpaA antibody, we performed a ChIP-seq experiment using the HA-RpaA strain and the anti-HA antibody (Table S3B). The genome-wide HA-RpaA binding profile (Figure S4A) resembled that of wild-type RpaA, although enrichments were generally substantially lower (Figure S4B and Table S3B), consistent with the lower overall phosphorylation of HA-RpaA (Figure S4C). We found that 66 out of the 110 wild-type binding sites (60%) were also present in the HA-RpaA ChIP experiment at a minimum of 2-fold enrichment (Figure S4D and Table S3A).

To determine whether RpaA can bind directly to DNA at locations other than the *kaiBC* promoter, we performed *in vitro* DNase I footprinting on a strongly enriched region upstream of the gene encoding the sigma factor *RpoD6*. RpaA bound to the *rpoD6* promoter in a phosphorylation- and concentration-dependent manner (Figure S4E).

We identified an A/T-rich motif overrepresented within the RpaA-binding sites (Figure 4D and Table S4A). This motif is present in the footprints of RpaA~P on the *kaiBC* and *rpoD6* promoters (Figures 3B, 4D, and S4E), and more than half (55%) of the RpaA-binding sites contain one or more instances of this motif with p values less than 0.001 (Tables S3A and S4B). We conclude that the peaks observed in the ChIP-seq data primarily result from direct binding of RpaA to DNA.

alanine mutation (D53A) to mimic unphosphorylated RpaA. All of the RpaA variants were expressed in the presence of IPTG (Figure S3F). We observed no expression of the luciferase reporter in a control strain containing the *Ptrc* promoter without RpaA. Leaky expression of wild-type RpaA from the *Ptrc* promoter (Figure S3F) rescued circadian activity of the *kaiBC* reporter (Figure 3C). Consistent with a previous report (Taniguchi et al., 2007), increasing levels of wild-type RpaA with IPTG induction progressively repressed promoter activity, likely due to an increase in the ratio of unphosphorylated RpaA to RpaA~P (Figure S3G). In contrast, *kaiBC* promoter activity was absent at all doses of RpaA(D53A). Expression of RpaA(D53E), however, restored activity from the promoter in a dose-dependent manner (Figure 3C), consistent with positive regulation of promoter activity by RpaA~P. No aspartate phosphorylation was present in RpaA(D53A) or RpaA(D53E) (Figure S3G). Collectively, our observations suggest that rhythmic association of RpaA~P with the *kaiBC* promoter drives circadian expression of this operon. Therefore, the circadian TTL is directly mediated by RpaA.

ChIP-Seq Reveals the Landscape of RpaA Binding

To identify RpaA-binding sites genome wide, we used our anti-RpaA antibody to perform circadian time course ChIP analyzed

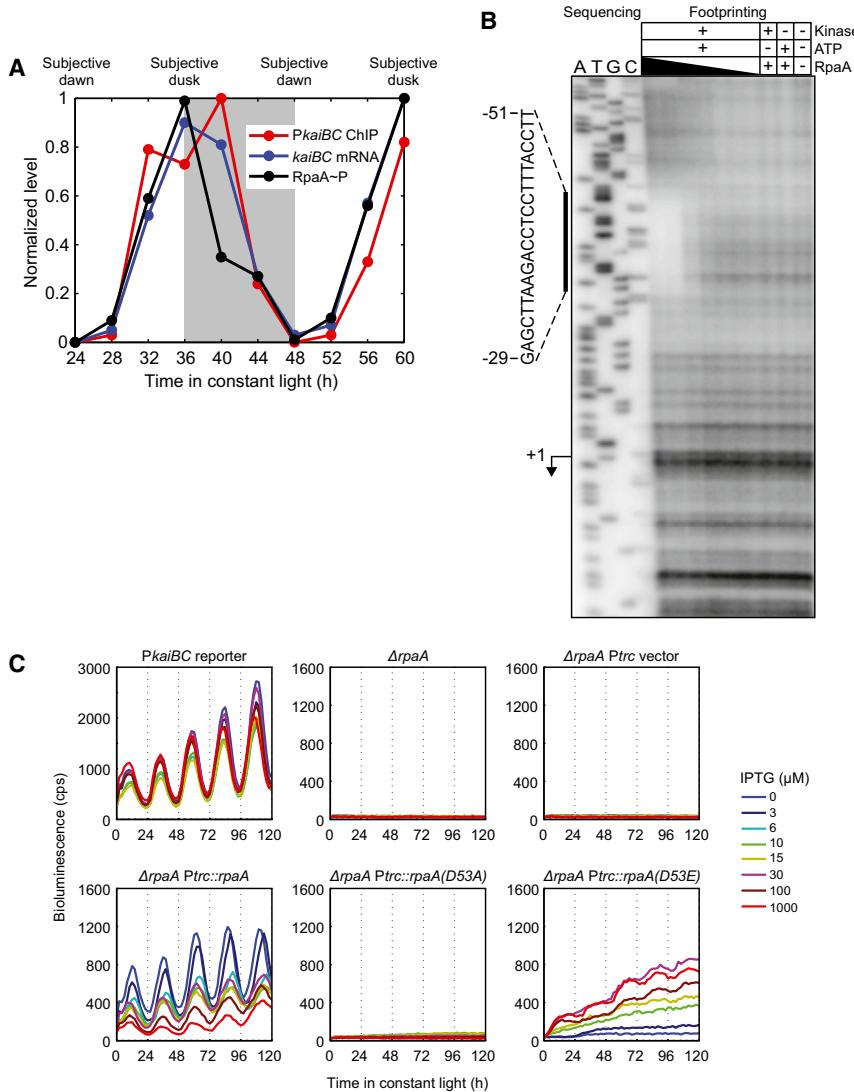


Figure 3. RpaA Binds to the *kaiBC* Promoter In Vivo and In Vitro and Promotes *kaiBC* Expression in a Phosphorylation-Dependent Manner

(A) Correlation between RpaA phosphorylation, RpaA enrichment at the *kaiBC* promoter (*PkaiBC*), and abundance of the *kaiBC* transcript. Subjective night is shaded. RpaA phosphorylation was measured by Phos-tag western blot (see Figure S3B for gel image), association with *PkaiBC* by ChIP-qPCR, and *kaiBC* expression by RT-qPCR. Note that, although the apex of the ChIP-qPCR enrichment is at 40 hr (4 hr after subjective dusk) in this experiment, the precise phase of RpaA binding (as well as that of global gene expression oscillations) varies between experiments, with RpaA binding typically peaking at or a few hours before subjective dusk.

(B) In vitro DNase I footprinting of RpaA on the *kaiBC* promoter as a function of recombinant RpaA phosphorylation and concentration. Sanger sequencing reactions used to identify the location of the footprint are shown on the left; footprinting reactions are shown on the right. The region protected from digestion by RpaA~P is indicated by the vertical bar. The *kaiBC* transcription start site (Kutsuna et al., 2005) is indicated with an arrow. RpaA pretreatment and concentration are indicated above each footprinting lane. RpaA was at least 50% phosphorylated in the presence of both kinase and ATP but was unphosphorylated otherwise (Figure S3D). RpaA was added to a final concentration of 6.0, 3.0, 0.6, 0.3, 0.06, or 0.03 μM as indicated by the thickness of the wedge.

(C) Activity of the *kaiBC* promoter was assayed using a *PkaiBC::luxAB* luciferase reporter in various genetic backgrounds: wild-type, *ΔrpaA*, and *ΔrpaA* expressing wild-type RpaA, unphosphorylatable RpaA (D53A), an RpaA phosphomimetic (D53E), or nothing (vector) from the IPTG-inducible *Ptrc* promoter. IPTG was added at the indicated concentration prior to entrainment with two 12 hr dark pulses.

See also Figure S3.

The RpaA Regulon Contains Genes Mediating a Variety of Cellular Processes

We systematically identified targets of RpaA genome wide by searching for genic (annotated mRNA, tRNA, or rRNA) transcripts and high-confidence noncoding transcripts (Vijayan et al., 2011) with 5' ends near the RpaA-binding sites (see Extended Experimental Procedures). We found 134 such target transcripts; together, these comprise the RpaA regulon. Ninety-three of these transcripts collectively encode 170 genes (many being coexpressed in operons) (Figure 5A), while 41 of the target transcripts are high-confidence noncoding RNAs (Tables S1 and S5). Three peaks were located too far from any transcripts to be associated with a target.

Expression of most of the genic RpaA ChIP targets oscillates with circadian periodicity, and these targets are strongly enriched for subjective dusk expression (Figure 5B). Expression of a small number of targets peaks at subjective dawn, including, quite interestingly, the canonical subjective dawn gene *purF*

(Paddock et al., 2013). In the *rpaA* mutant, subjective dusk RpaA ChIP targets decrease in expression, whereas subjective dawn targets increase in expression (Figure 5C), suggesting that RpaA functions as an activator of subjective dusk targets and a repressor of subjective dawn targets. This helps to explain the dawn phase arrest of the *rpaA* mutant (Figures 1B and 1C). We also found that most strongly downregulated genes in the *rpaA* mutant are RpaA ChIP targets, whereas more weakly downregulated genes typically are not targets (Figure 5D). This implies that RpaA acts as a master regulator by directly modulating the expression of a subset of high-amplitude circadian genes, whose products, in turn, effect the fine global pattern of circadian gene expression.

The RpaA regulon (Figure 5A and Tables S1 and S5) contains targets involved in transcription regulation, including four sigma factors. Interestingly, RpaA also targets the *himA* gene (synpcc7942_2248) encoding the nucleoid protein HU, which is downregulated 1.7-fold in the *rpaA* mutant. Because both DNA

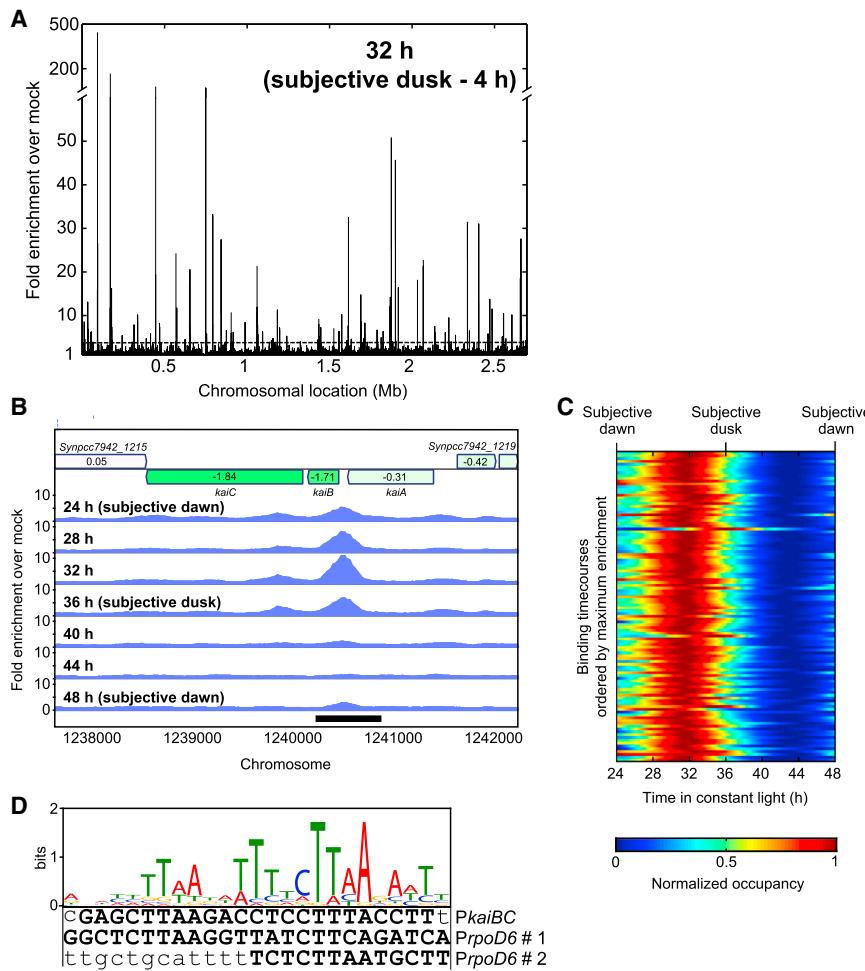


Figure 4. Identification of RpaA-Binding Sites by ChIP-Seq

(A) Genome-wide binding profile of RpaA by ChIP-seq. The enrichment of read density in the RpaA ChIP-Seq (anti-RpaA antibody on the wild-type strain) at 4 hr prior to subjective dusk (32 hr), relative to the mock ChIP-seq (anti-RpaA antibody on the *rpaA* mutant), is plotted as a function of position on the chromosome. The dotted line indicates the 3-fold enrichment cutoff for identification of RpaA-binding sites.

(B) Genome browser view of ChIP-seq enrichment profiles in the vicinity of the *kaiBC* locus over 1 day in the wild-type strain. Genes are shown at the top, with the \log_2 of their expression change in the *rpaA* mutant represented by shading (green is decreased expression, and red is increased expression) and indicated by the text inside the gene.

(C) Time course of RpaA enrichment at the 110 RpaA-binding sites. ChIP-seq was performed every 4 hr for 24 hr, and enrichment relative to the mock IP was calculated at the location of maximum wild-type ChIP-seq read density within each binding site. Enrichment at intermediate time points was computed by interpolation with cubic splines. Each row in the heatmap represents the binding time course for one binding site; the rows are sorted by the maximum enrichment observed during the time course, which ranged from 411-fold (top) to 3.1-fold (bottom). The dynamic range (maximum enrichment divided by minimum enrichment for each binding site) varied from 38-fold to 1.4-fold.

(D) A 25 basepair motif is overrepresented near RpaA-binding sites (E value, 1.7×10^{-36} , Table S4A) and is found within the RpaA~P footprint in the *kaiBC* promoter (Figure 3B) and in both footprints in the *pD6* promoter (Figure S4E). Bases protected by RpaA~P binding in the DNase I footprinting assays are capitalized and boldfaced. See also Figure S4 and Tables S3 and S4.

compaction (Smith and Williams, 2006) and *himA* expression are regulated as a function of circadian time (Vijayan et al., 2009), *himA* could link oscillatory RpaA activity to circadian genome compaction, providing another route for RpaA to influence gene expression globally.

Several RpaA targets are enzymes of the glycolysis, glycogen, and pentose phosphate metabolic pathways, suggesting a direct link between the circadian clock and energy production and storage. RpaA also directly targets the translation initiation factor IF-3 (*infC*), the protein chaperone trigger factor (*tig*), and a ClpXP protease system (*clpX* and *clpP2*), implying that the clock may modulate translation and protein homeostasis directly, connections that have not been reported previously in this organism. In addition, the *rpaA* promoter itself is an RpaA ChIP target, suggesting the presence of autoregulatory feedback. We note that some 37% of the RpaA ChIP targets have no known function; among these targets, there may be unforeseen control nodes with roles in global gene expression regulation.

Cells mutated for *rpaA* display increased efficiency of energy transfer from the light-harvesting phycobilisomes to photo-

system II relative to photosystem I, and for this reason, the gene was named regulator of phycobilisome association A (Ashby and Mullineaux, 1999). We find that RpaA directly targets the *rpaC* gene, which encodes an integral membrane protein implicated in controlling the stability of the photosystem II-phycobilisome interaction (Joshua and Mullineaux, 2005). *rpaC* is circadianly expressed and is downregulated ~2-fold in the *rpaA* mutant, consistent with positive regulation by RpaA.

Finally, we found that the RpaA regulon is enriched for genes whose expression increases in darkness in a *kaiABC*-dependent manner (Hosokawa et al., 2011) ($p = 4.4 \times 10^{-6}$, Fisher's exact test), suggesting that RpaA is involved in clock modulation of the gene expression response to darkness.

Active RpaA Is Sufficient to Switch Cells between the Two Major Gene Expression States Produced by the Circadian Clock

To test directly whether RpaA serves as the master regulator governing circadian gene expression, we asked whether

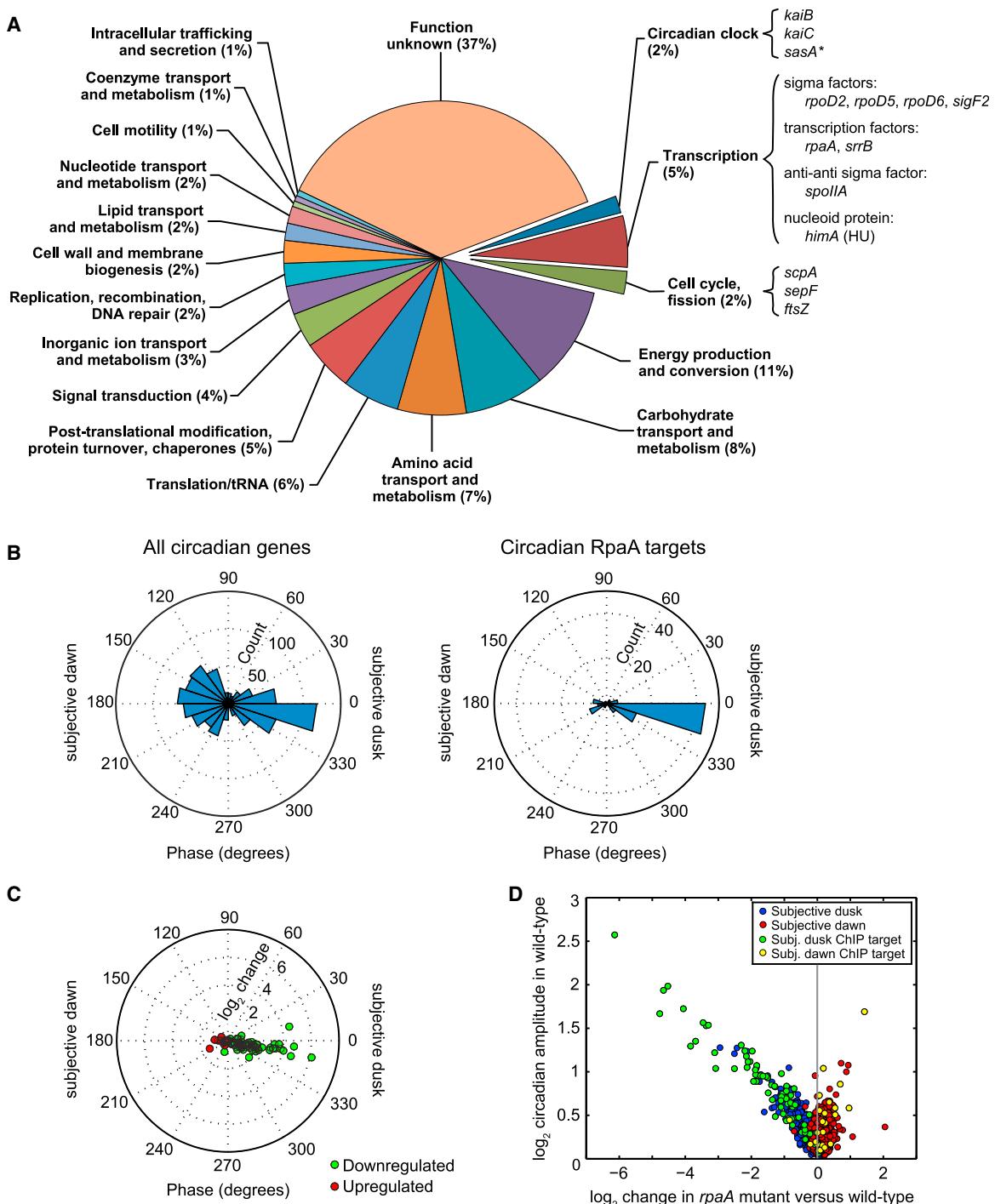


Figure 5. The RpaA Regulon

(A) Functional characterization of protein and tRNA ChIP targets of RpaA. We identified 134 transcripts closest to the 110 binding sites (see [Extended Experimental Procedures](#)). Of those transcripts, 93 encode proteins or tRNAs (corresponding to 170 genes, some of which are coexpressed in operons), whereas the other 41 are classified as noncoding RNAs ([Vijayan et al., 2011](#)). Because the function of the noncoding RNAs is not known, we restrict our functional analysis to the 170 protein-coding or tRNA genes ("RpaA ChIP target genes"). RpaA ChIP target genes were categorized as described in the [Extended Experimental Procedures](#). Some genes of particular interest are highlighted. The asterisk (*) indicates that the gene's classification as an RpaA ChIP target is artificial because of assignment to an incorrectly demarcated operon containing a bona fide target ([Vijayan et al., 2011](#)).

(B) Comparison of the distribution of phases of all circadian genes (left, n = 856, from [Figure 1A](#)) and of the ChIP target genes whose expression oscillates with circadian periodicity (right, n = 95).

(legend continued on next page)

the RpaA~P phosphomimetic RpaA(D53E) is sufficient to induce global changes in expression similar to those that occur over the course of a circadian cycle. We hypothesized that overexpression of RpaA(D53E) in an *rpaA* mutant would switch cells from the subjective dawn state in which the mutant resides to the subjective dusk state that coincides with the time of maximal RpaA activity in the wild-type strain. To isolate transcriptional changes resulting directly from activity of RpaA(D53E) away from potentially confounding Kai oscillator-dependent processes, we introduced the *Ptrc::rpaA(D53E)* construct into a strain lacking *kaiBC* as well as *rpaA*, producing a *ΔrpaA ΔkaiBC Ptrc::rpaA(D53E)* strain that we refer to as “OX-D53E.”

We used high-throughput RNA sequencing (RNA-seq) to compare gene expression changes caused by induction of RpaA(D53E) expression with IPTG (Figure S5A) to those experienced during the course of a day in constant light in the wild-type strain (Table S6). First, we calculated the correlation between expression of circadian genes during a time course of OX-D53E induction and during a wild-type circadian time course (Figure 6A). Consistent with the gene expression state of the *rpaA* mutant (Figure 1C), preinduction OX-D53E is most correlated with wild-type at 24 hr (subjective dawn) and is most anticorrelated with wild-type at 36 hr (subjective dusk). The correlations reverse over 12 hr of RpaA(D53E) induction with IPTG (Figure 6A): OX-D53E becomes most similar to the wild-type 36 hr time point and most anticorrelated with the wild-type 24 hr time point. Importantly, these shifts are not observed when IPTG is added to a control strain (OX-mock) in which no gene is inserted downstream of the *Ptrc* promoter or when IPTG is not added to the OX-D53E strain (Figures S5B and S5C).

The flip of the dawn-to-dusk gene expression switch is illustrated by plotting each circadian gene’s expression change upon IPTG induction in OX-D53E against its change between subjective dawn and dusk in the wild-type (Figure 6B). After induction, subjective dusk gene expression increases in proportion to its change from subjective dawn to dusk in the wild-type, whereas subjective dawn gene expression decreases in proportion to its change in the wild-type (correlation = 0.8). Furthermore, expression of 85% of circadian genes (725 out of 856) differed from baseline by >1.5-fold in at least one time point after IPTG induction. The strong correlation between the dawn-to-dusk expression change in the wild-type and the expression change upon RpaA(D53E) induction in OX-D53E shows that active RpaA suffices to switch cells between dawn and dusk expression states.

Consistent with our analysis of the ChIP-seq data (Figures 5B–5D), constitutively active RpaA strongly induced expression of ChIP dusk targets and had weaker repressive activity toward a minority of its subjective ChIP dawn targets (Figure 6B). Interestingly, RpaA ChIP targets comprise only a subset of significantly affected genes; the remaining genes must be activated or

repressed by one of RpaA’s direct targets. Collectively, these results demonstrate that RpaA is the master regulator of circadian gene expression, acting as the most upstream node in a network of circadian regulators that together orchestrate global gene expression rhythms.

The RpaA Regulon Produces Complex Gene Expression Dynamics that Parallel Those Observed during a Circadian Period

We examined whether the dynamics of the subjective dawn-to-dusk transition induced by active RpaA mirror those observed in the wild-type strain. We used K-means clustering to identify patterns in expression of circadian genes in the wild-type strain and of those same genes in the OX-D53E strain (Figures 6C and S5D). In the wild-type strain, clustering separated genes according to their time of maximum expression: dawn, morning, afternoon, evening, dusk, and night (Vijayan et al., 2009). We named the clusters obtained for the OX-D53E strain according to their dynamics following induction: repressed, latently repressed, transiently activated, activated, latently activated, and transiently repressed. Some genes in the repressed and latently repressed categories were slightly induced at 30 min after IPTG addition but were subsequently repressed as RpaA(D53E) levels further increased (Figure S5A). This suggests that low levels of RpaA promote the expression of these genes, whereas higher levels have a much stronger repressive effect on their expression. Notably, the timescale of response to induction varies among the clusters, with some clusters responding in concert with RpaA(D53E) accumulation (activated and repressed) and others more slowly (latently activated and latently repressed). The transiently activated and transiently repressed clusters display markedly nonmonotonic responses to RpaA (D53E) induction. Hence, the network of regulators downstream of RpaA encodes a variety of responses to continuous accumulation of active RpaA. Note that RpaA ChIP targets never comprise more than 27% of the genes in a cluster. Therefore, the majority of genes in each cluster are controlled by the RpaA regulon rather than by RpaA itself.

If the gene expression dynamics observed in the OX-D53E induction time course resemble those in the wild-type circadian time course, there should be an overlap between the clusters in each strain. Indeed, although each wild-type cluster is represented in each OX-D53E cluster and vice versa, we find a nearly one-to-one mapping in statistically significant overlap between the two sets of clusters (Figure 6D). The dawn cluster in the wild-type strain maps to the repressed cluster in OX-D53E; morning maps to latently repressed; afternoon maps to transiently activated; evening maps to activated; dusk maps to latently activated; and night maps to repressed ($p \leq 0.001$ for each pair). Remarkably, this mapping arises despite the difference in the timing of RpaA activation in the two strains. In the wild-type strain, RpaA~P abundance varies sinusoidally,

(C) Change in expression of circadian ChIP target genes downregulated (green, $n = 72$) or upregulated (red, $n = 23$) in the *rpaA* mutant plotted as a function of their phase in the wild-type strain.

(D) Comparison of gene expression change in the *rpaA* mutant with circadian amplitude in the wild-type strain (from Figure 1B). Only genes that oscillate with circadian periodicity in the wild-type strain are shown ($n = 856$). Circadianly expressed RpaA ChIP target genes ($n = 95$) are highlighted. See also Tables S1 and S6.

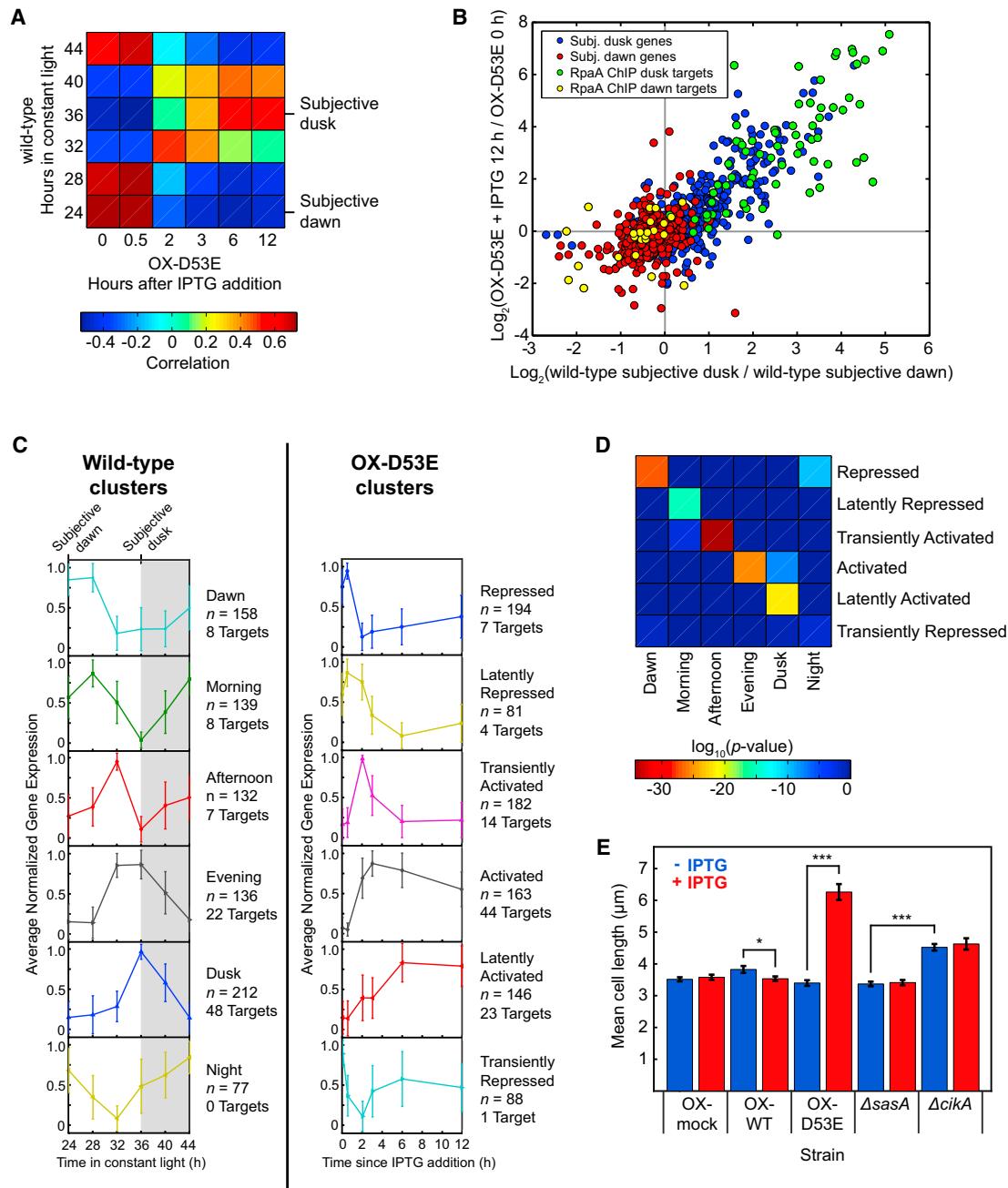


Figure 6. RpaA Orchestrates Global Circadian Gene Expression and Controls the Cell Division Gate

(A) Correlation between the expression of circadian genes ($n = 856$) in the wild-type strain over the course of 1 day and the expression of those genes in the $\Delta rpaA$ strain before ($t = 0$ hr) and after induction with IPTG. Gene expression was measured by RNA-seq.

(B) Correlation between the change in expression of circadian genes ($n = 856$) caused by induction of RpaA(D53E) in the OX-D53E strain (y axis) with the change in expression between subjective dusk and dawn in the wild-type strain (x axis). RpaA ChIP target genes are highlighted ($n = 95$; 71 subjective dusk and 24 subjective dawn). Gene expression was measured by RNA-seq.

(C) K-means identification of gene expression clusters in the wild-type and OX-D53E strains. Gene expression was measured by RNA-seq. With $K = 6$, wild-type circadian genes ($n = 854$; $kaiB$ and $kaiC$ were omitted because they are absent in the OX-D53E strain) were separated into six clusters with distinct expression phases (left), consistent with previous microarray observations (Vijayan et al., 2009). Time courses of the same set of genes in the OX-D53E strain were also clustered using $K = 6$ (right). The traces show the average normalized time course of genes within each cluster; error bars show SD. The numbers of all genes (n) and RpaA ChIP target genes in each cluster are indicated. Noncoding RNAs were not included in this analysis.

(D) Mapping between clusters in the wild-type (x axis) and OX-D53E (y axis) strains. Each element of the heatmap shows the \log_{10} of the statistical significance (Fisher's exact test) of the overlap between the corresponding clusters on each axis.

(legend continued on next page)

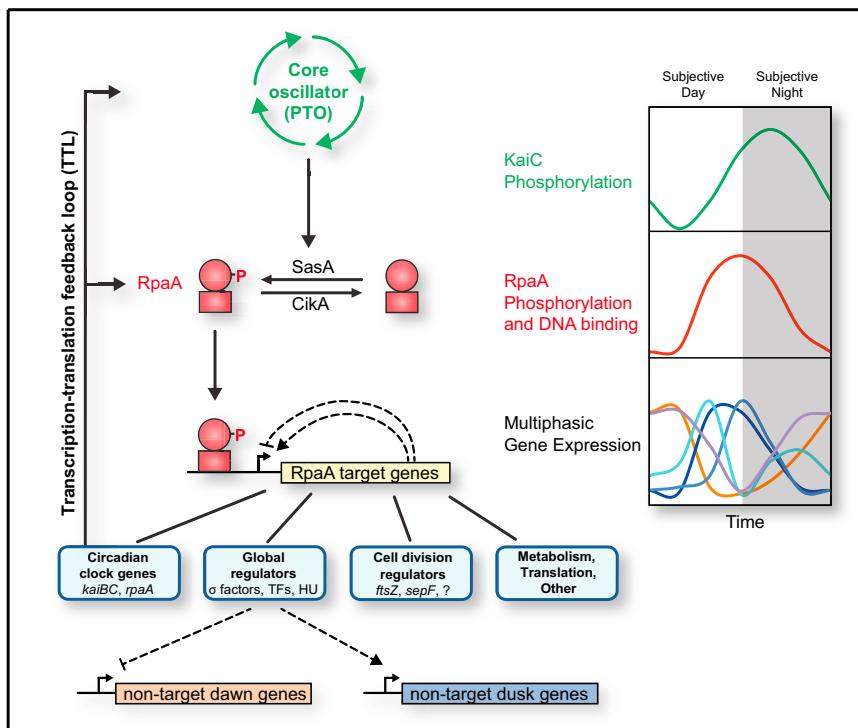


Figure 7. Model for RpaA and the Cyanobacterial Circadian Program

Model for control of clock output by RpaA. Time encoded in the PTO is transduced into RpaA phosphorylation via SasA and CikA, producing oscillations in RpaA~P (red) that phase lead those of phosphorylated KaiC (green) by ~4 hr (Gutu and O’Shea, 2013). RpaA~P binds to DNA and controls the expression of the RpaA regulon, which consists of global transcriptional regulators, cell division regulators, certain clock genes (*kaiBC* and *rpaA*), and genes involved in metabolism and translation. The global regulators are at the top of a transcriptional cascade that orchestrates multiphasic circadian gene expression, repressing subjective dawn genes while activating subjective dusk genes. Fine patterns within the dusk and dawn categories could be generated by a network of interactions among RpaA ChIP targets (hypothetical positive and negative feedbacks are shown as dotted lines). RpaA control of *kaiBC* and *rpaA* expression forms the clock TTL.

increasing over a ~12 hr time span and then decreasing over the same time span (Figure 3A). In contrast, RpaA(D53E) accumulates rapidly, plateauing after 3 hr of induction and remaining high for the remainder of the time course (Figure S5A). Nonetheless, genes respond to active RpaA in a largely stereotyped manner in both strains, suggesting that the dynamics of global circadian gene expression are hardwired into the RpaA regulon.

Active RpaA Closes the Cell Division Gate

During the subjective night, cells elongate but do not divide, a phenomenon referred to as cell division gating (Dong et al., 2010; Mori et al., 1996; Yang et al., 2010). When the gate is closed, cells form elongating rods, allowing the status of the gate in a given strain to be inferred from the cell length distribution (Dong et al., 2010). The observation that deletions of *sasA* and *cikA* have opposite effects on cell-cycle gating (Dong et al., 2010) suggests that RpaA~P could be responsible for closing the gate, as SasA and CikA have opposite effects on RpaA phosphorylation (Gutu and O’Shea, 2013).

To test this hypothesis, we examined the effect of overexpression of wild-type and phosphomimetic RpaA on cell length. We characterized cell length in strains ectopically expressing wild-type RpaA (OX-WT), RpaA(D53E) (OX-D53E), or an empty multicloning site (OX-mock) from the *Ptrc* promoter with and without IPTG treatment (Figures 6E, S5E, and S5F). Consistent with our

hypothesis, cells showed pronounced elongation in the OX-D53E strain after induction. Conversely, induction of OX-WT shortens the median cell length, consistent with the reduction in RpaA~P levels (Figure S5F) and the repression of *kaiBC* gene expression in this condition (Figure 3C). Induction of OX-mock had no effect on cell length. These data are consistent with a causative role for RpaA~P in clock-mediated cell division gating.

Several RpaA ChIP targets are involved in cell division and thus might mediate gating. The most prominent of these are the bacterial tubulin homolog *ftsZ* and the *FtsZ* regulator *sepF* (Marbouy et al., 2009). Intriguingly, *FtsZ* is mislocalized in $\Delta cikA$ strains (Dong et al., 2010). However, mean *ftsZ* and *sepF* expression are unchanged by deletion of *rpaA* (Table S1), so the functional relevance of these targets is unclear. Other ChIP targets of interest are genes involved in the peptidoglycan biosynthetic pathway (*synpcc7042_0482*, *synpcc7042_1740/murB*, *synpcc7042_1741/murC*), the last of which has been shown to interact with several Fts cell division proteins (Munshi et al., 2013).

DISCUSSION

RpaA Is the Hub through which the Circadian Clock Controls Cellular Physiology

RpaA phosphorylation links the core Kai oscillator to two of the most striking physiological outputs of the clock: global transcriptome oscillations and gating of cell division (Figure 7). Time information encoded in the PTO is read out through the histidine

(E) Mean cell lengths in $\Delta rpaA \Delta kaiBC$ strains containing a *Ptrc* promoter driving expression of wild-type RpaA (OX-WT), RpaA(D53E) (OX-D53E), or an empty multicloning site (OX-mock) grown in the presence (red) or absence (blue) of the inducer IPTG (100 μ M). At least 80 cells were analyzed for each strain. Error bars represent SEM. * $p < 0.05$; ** $p < 10^{-13}$ (one-way ANOVA).

See also Figure S5 and Table S6.

kinases SasA and CikA, which antagonistically regulate RpaA phosphorylation to generate oscillations of RpaA~P that peak at or immediately preceding subjective dusk (Gutu and O'Shea, 2013) (Figures 3A and 7). The accumulation of RpaA~P switches the cell's gene expression program from subjective dawn to subjective dusk (Figures 6A and 6B) via transcriptional programs hardwired into the RpaA regulon (Figures 6C and 6D). RpaA~P accumulation also initiates closure of the cell division gate (Figure 6E). RpaA~P decreases during the subjective night (Gutu and O'Shea, 2013) (Figure 3A), allowing gene expression to revert to its default dawn-like state (Figures 1B and C) and also opening the cell division gate.

Previous reports have implicated circadianly regulated DNA supercoiling in driving gene expression oscillations (Vijayan et al., 2009; Woelfle et al., 2007). We sought to investigate the relationship between RpaA binding and supercoiling by measuring supercoiling in the *rpaA* mutant, $\Delta rpaA$ clock rescue, and OX-D53E strains but had difficulty obtaining reproducible results and thus do not report any here. Nonetheless, we can gain insight into to the relative importance of supercoiling and RpaA binding in generating global gene expression oscillations by comparing the magnitudes of expression changes in response to perturbations in supercoiling and RpaA activity. When supercoiling is rapidly relaxed by treatment with a pharmacological inhibitor of DNA gyrase (Vijayan et al., 2009), global gene expression changes in the same manner as it does upon induction of RpaA(D53E) (Figure 6B). However, the magnitude of change is substantially larger for induction of RpaA(D53E) than for relaxation of supercoiling (compare Figure 6B here to Figure 4 in Vijayan et al. [2009]). Moreover, induction of RpaA(D53E) almost quantitatively reproduces the magnitude of gene expression change observed over circadian time in the wild-type strain (Figure 6B). On this basis, we suggest that circadian oscillations in RpaA activity play a dominant role in driving global circadian gene expression oscillations. Perhaps one or more members of the RpaA regulon induces oscillation in supercoiling, which acts in concert with RpaA via a feedforward loop to actuate circadian gene expression. Alternatively, RpaA binding affinity could be positively influenced by supercoiling, in which case relaxation of supercoiling by pharmacological inhibition of gyrase would release RpaA~P from the chromosome, leading to observed effects on global gene expression state (Vijayan et al., 2009). Future studies will be required to establish the molecular connections between RpaA activity and supercoiling. Our observation that RpaA targets the *himA* gene encoding the nucleoid protein HU could provide a starting point for such studies.

Also meriting future investigation are possible roles for RpaA in the integration of environmental and time information. Environmental cues may influence RpaA phosphorylation directly, as the activity of its phosphatase CikA is regulated not only by the PTO, but also by the cellular redox state, which in turn reflects environmental conditions such as light availability (Ivleva et al., 2006; Kim et al., 2012). Hence, information about the environment and circadian time may be integrated at the level of RpaA phosphorylation. RpaB also may play a role: RpaB phosphorylation is regulated by changes in light intensity (Moronta-Barrios et al., 2012), and RpaB binds to the promoters of *kaiBC* and *rpoD6* (Hanaoka et al., 2012), both of which are

RpaA targets. In fact, RpaB binding to the *kaiBC* promoter is antagonized by RpaA (Hanaoka et al., 2012), consistent with the overlap of the RpaA footprint (Figure 3B) with the HLR1 motif bound by RpaB. A systematic exploration of the interaction of RpaA and RpaB at promoters and its effect on gene expression will be required.

A recent study found that activity of the canonical subjective dawn promoter *PpurF* is affected by the presence of the *kaiC* gene in an *rpaA* mutant background, suggesting the presence of an RpaA-independent output pathway (Paddock et al., 2013). The authors proposed a model for *PpurF* control in which PTO-modulated RpaA~P levels repress the activity of a "predominant" RpaA-independent output pathway, which in turn links serine-phosphorylated KaiC to activation of the promoter. This model predicts that *purF* transcript abundance will oscillate in our $\Delta rpaA$ clock rescue experiment, in which KaiC phosphorylation oscillates with high amplitude (Figures 2 and S2). Instead, we found *purF* levels to be constant, locked at the level of maximum expression in the control clock rescue (Figure S2C).

Although our data do not rule out the existence of an RpaA-independent clock output pathway, they do suggest that any such pathway is weak relative to the RpaA-dependent output pathway that we describe here. Moreover, our ability to switch the genome-wide transcriptional program from subjective dawn to dusk by expression of an RpaA phosphomimetic in a strain lacking both *kaiB* and *kaiC* (Figures 6 and S5) shows that RpaA phosphorylation is sufficient to control the circadian gene expression program. Notably, *purF* expression is switched from a high to a low state upon phosphomimetic induction (Figure S5D); this is straightforwardly explained by the fact that *purF* is a direct ChIP target of RpaA (Tables S1 and S5). We found that RpaA~P binds to *PpurF* in vitro and that point mutations affecting its expression phase (Vijayan and O'Shea, 2013) fall within the RpaA~P footprint; at least two of these mutations impair RpaA~P binding (Figure S4F).

Generation of Complex Gene Expression Patterns with a One-Dimensional Signal

The complex gene expression dynamics observed upon induction of OX-D53E (Figure 6C) demonstrates that a smooth, univariate signal like RpaA phosphorylation (Figure 3A; Gutu and O'Shea, 2013) can generate a diverse array of dynamic responses. We suggest that the smooth RpaA~P signal is converted into a mosaic of dynamic patterns through network motifs composed of the gene expression regulator targets of RpaA (e.g., sigma factors) and their own downstream targets.

Interestingly, the RpaA~P level does not uniquely specify the time of day, as intermediate RpaA~P levels are experienced during both the subjective day and subjective night (Figure 3A). The single-bit encoding of time in RpaA phosphorylation contrasts sharply with its encoding in the PTO, in which differential phosphorylation at two residues of KaiC specifies two bits of information (Nishiwaki et al., 2007; Rust et al., 2007). The four phosphoforms of KaiC appear in an ordered pattern during each circadian cycle in a manner that uniquely maps the time of day to a particular phosphoform distribution (Nishiwaki et al., 2007; Rust et al., 2007). It seems paradoxical that time encoded in the PTO would be read out through a single-bit channel

that cannot uniquely encode it, causing a loss of information. However, the network motifs that likely generate the complex circadian gene expression dynamics (Figures 6C and S5D) could make the functional effect of a given RpaA~P level history dependent (hysteretic). The current level of RpaA~P and its history together would suffice to fully specify time.

Relevance to Eukaryotic Circadian Clock Output Pathways

In directly controlling a large number of transcriptional regulators, RpaA resembles eukaryotic circadian effectors like white collar complex (WCC) in *Neurospora crassa* (Smith et al., 2010), CLOCK and CYCLE in *Drosophila melanogaster* (Abruzzi et al., 2011), CLOCK and BMAL1 in mice (Koike et al., 2012; Rey et al., 2011), and PRR5 in *Arabidopsis thaliana* (Nakamichi et al., 2012). Transcription factors in these eukaryotic clock output pathways have been proposed to initiate hierarchical transcriptional cascades that effect widespread circadian rhythms in gene expression (Edery, 2011), similar to what we propose occurs with RpaA. However, the output pathways in these eukaryotic clocks are inseparable from the core circadian oscillators: the core oscillators are built from the output pathways themselves, which form time-delayed TTLs that drive oscillations. Multiple interlocking positive and negative transcriptional feedback loops exist in these clocks, with transcription factors serving simultaneously as clock outputs and TTL components. In contrast, the cyanobacterial clock is fundamentally a PTO with a subsidiary and dispensable TTL (Qin et al., 2010; Teng et al., 2013), with the output pathway serving mainly as a conduit for transmitting time information from the PTO to the genome in order to control circadian gene expression and cell division gating. More work is required to determine whether these are indeed bona fide differences in the topology of these systems or whether the eukaryotic and prokaryotic clocks share more similarities than currently appreciated.

Synthetic Biology Applications

Our work has implications for synthetic biology. Identification of the binding motif and binding dynamics of RpaA (Figures 3A and 4) provide the molecular details that are required to rationally design connections between the PTO and synthetic transcriptional outputs, enabling reconstitution of the PTO with a transcription-based reporter in orthogonal organisms. Industrial applications arise, as cyanobacteria present an attractive platform for the engineered production of chemicals directly from the abundant fuel source of sunlight (Ducat et al., 2011). Synthetic pathways could be coupled to the circadian clock via RpaA; as the clock enhances the fitness of the wild-type organism in oscillating environmental conditions (Ouyang et al., 1998), so too might it enhance the productivity of synthetic pathways.

EXPERIMENTAL PROCEDURES

Gene Expression Analysis

Microarray analysis of gene expression time courses was conducted as described previously (Vijayan et al., 2009). To compare gene expression in the wild-type strain to that in the *rpaA* mutant (Figures 1B, 1C, and S1), we prepared time-averaged, pooled cDNA samples for each strain (constructed by combining equal mass quantities of cDNA from each time point), labeled

them with different dyes, and hybridized them against one another. Dye ratios were normalized by lowess regression. To reduce dye bias, we analyzed a second microarray in which the dyes were swapped and averaged the lowess-normalized values from the two arrays.

RNA-seq library preparation is described in the [Extended Experimental Procedures](#). Expression levels were median and z score normalized across samples. K-means clustering (with K = 6) was performed in Matlab with squared Euclidean distance as the distance metric.

ChIP-qPCR and ChIP-Seq

ChIP was performed as described previously (Hanaoka and Tanaka, 2008; Vijayan et al., 2011) with modifications described in the [Extended Experimental Procedures](#). ChIP was performed using affinity-purified anti-RpaA antibody (see [Extended Experimental Procedures](#)) or anti-HA agarose beads (Pierce). Mock ChIP samples were prepared using the anti-RpaA antibody with cross-linked *rpaA* mutant cells or the anti-HA antibody with crosslinked wild-type cells.

For ChIP-qPCR, enrichment was computed by normalizing the abundance of the promoter of interest to the abundance of the coding region of *synpcc7942_0612*, which shows no enrichment for RpaA occupancy by ChIP-seq. See [Extended Experimental Procedures](#) for more details.

Preparation and sequencing of ChIP-seq libraries and ChIP-seq data analysis are described in the [Extended Experimental Procedures](#).

ACCESSION NUMBERS

High-throughput data are available from Gene Expression Omnibus under accession number GSE50922.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2013.11.005>.

ACKNOWLEDGMENTS

We thank Vikram Vijayan for the 48 hr wild-type gene expression time course and for code, technical assistance, and discussion; Andrian Gutu for Phos-tag western blots; and Andrian Gutu, Xu Zhou, Abbas Rizvi, Brian Zid, Michael Rust (University of Chicago), Guogang Dong (UCSD), and Susan Golden (UCSD) for discussion and advice. We thank Richard Losick, Bodo Stern, and Andrian Gutu for comments on the manuscript and Takao Kondo (Nagoya University) and Susan Golden for plasmids and strains. This work was supported by the Howard Hughes Medical Institute.

Received: July 10, 2013

Revised: September 23, 2013

Accepted: October 28, 2013

Published: December 5, 2013

REFERENCES

- Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S., and Rosbash, M. (2011). Drosophila CLOCK target gene characterization: implications for circadian tissue-specific gene expression. *Genes Dev.* 25, 2374–2386.
- Ashby, M.K., and Mullineaux, C.W. (1999). Cyanobacterial ycf27 gene products regulate energy transfer from phycobilisomes to photosystems I and II. *FEMS Microbiol. Lett.* 181, 253–260.
- Dong, G., Yang, Q., Wang, Q., Kim, Y.I., Wood, T.L., Osteryoung, K.W., van Oudenaarden, A., and Golden, S.S. (2010). Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in *Synechococcus elongatus*. *Cell* 140, 529–539.
- Ducat, D.C., Way, J.C., and Silver, P.A. (2011). Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* 29, 95–103.

- Edery, I. (2011). A master CLOCK hard at work brings rhythm to the transcriptome. *Genes Dev.* 25, 2321–2326.
- Gutu, A., and O’Shea, E.K. (2013). Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Mol. Cell* 50, 288–294.
- Hanaoka, M., and Tanaka, K. (2008). Dynamics of RpaB-promoter interaction during high light stress, revealed by chromatin immunoprecipitation (ChIP) analysis in *Synechococcus elongatus* PCC 7942. *Plant J.* 56, 327–335.
- Hanaoka, M., Takai, N., Hosokawa, N., Fujiwara, M., Akimoto, Y., Kobori, N., Iwasaki, H., Kondo, T., and Tanaka, K. (2012). RpaB, another response regulator operating circadian clock-dependent transcriptional regulation in *Synechococcus elongatus* PCC 7942. *J. Biol. Chem.* 287, 26321–26327.
- Hosokawa, N., Hatakeyama, T.S., Kojima, T., Kikuchi, Y., Ito, H., and Iwasaki, H. (2011). Circadian transcriptional regulation by the posttranslational oscillator without de novo clock gene expression in *Synechococcus*. *Proc. Natl. Acad. Sci. USA* 108, 15396–15401.
- Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., Terauchi, K., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2009). Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. USA* 106, 14168–14173.
- Ivleva, N.B., Gao, T., LiWang, A.C., and Golden, S.S. (2006). Quinone sensing by the circadian input kinase of the cyanobacterial circadian clock. *Proc. Natl. Acad. Sci. USA* 103, 17468–17473.
- Johnson, C.H., Stewart, P.L., and Egli, M. (2011). The cyanobacterial circadian system: from biophysics to bioevolution. *Annu. Rev. Biophys.* 40, 143–167.
- Joshua, S., and Mullineaux, C.W. (2005). The rpaC gene product regulates phycobilisome-photosystem II interaction in cyanobacteria. *Biochim. Biophys. Acta* 1709, 58–68.
- Kim, Y.I., Vinyard, D.J., Ananyev, G.M., Dismukes, G.C., and Golden, S.S. (2012). Oxidized quinones signal onset of darkness directly to the cyanobacterial circadian oscillator. *Proc. Natl. Acad. Sci. USA* 109, 17765–17769.
- Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* 338, 349–354.
- Kutsuna, S., Nakahira, Y., Katayama, M., Ishiura, M., and Kondo, T. (2005). Transcriptional regulation of the circadian clock operon kaiBC by upstream regions in cyanobacteria. *Mol. Microbiol.* 57, 1474–1484.
- Marbouy, M., Saguez, C., Cassier-Chauvat, C., and Chauvat, F. (2009). Characterization of the FtsZ-interacting septal proteins SepF and Ftn6 in the spherical-celled cyanobacterium *Synechocystis* strain PCC 6803. *J. Bacteriol.* 191, 6178–6185.
- Markson, J.S., and O’Shea, E.K. (2009). The molecular clockwork of a protein-based circadian oscillator. *FEBS Lett.* 583, 3938–3947.
- Mori, T., Binder, B., and Johnson, C.H. (1996). Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. *Proc. Natl. Acad. Sci. USA* 93, 10183–10188.
- Moronta-Barrios, F., Espinosa, J., and Contreras, A. (2012). In vivo features of signal transduction by the essential response regulator RpaB from *Synechococcus elongatus* PCC 7942. *Microbiology* 158, 1229–1237.
- Munshi, T., Gupta, A., Evangelopoulos, D., Guzman, J.D., Gibbons, S., Keep, N.H., and Bhakta, S. (2013). Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in *Mycobacterium tuberculosis*. *PLoS ONE* 8, e60143.
- Murayama, Y., Oyama, T., and Kondo, T. (2008). Regulation of circadian clock gene expression by phosphorylation states of KaiC in cyanobacteria. *J. Bacteriol.* 190, 1691–1698.
- Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* 308, 414–415.
- Nakajima, M., Ito, H., and Kondo, T. (2010). In vitro regulation of circadian phosphorylation rhythm of cyanobacterial clock protein KaiC by KaiA and KaiB. *FEBS Lett.* 584, 898–902.
- Nakamichi, N., Kiba, T., Kamioka, M., Suzuki, T., Yamashino, T., Higashiyama, T., Sakakibara, H., and Mizuno, T. (2012). Transcriptional repressor PRR5 directly regulates clock-output pathways. *Proc. Natl. Acad. Sci. USA* 109, 17123–17128.
- Nishiwaki, T., Satomi, Y., Kitayama, Y., Terauchi, K., Kiyohara, R., Takao, T., and Kondo, T. (2007). A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria. *EMBO J.* 26, 4029–4037.
- Ouyang, Y., Andersson, C.R., Kondo, T., Golden, S.S., and Johnson, C.H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 95, 8660–8664.
- Paddock, M.L., Boyd, J.S., Adin, D.M., and Golden, S.S. (2013). Active output state of the *Synechococcus* Kai circadian oscillator. *Proc. Natl. Acad. Sci. USA* 110, E3849–E3857.
- Qin, X., Byrne, M., Xu, Y., Mori, T., and Johnson, C.H. (2010). Coupling of a core post-translational pacemaker to a slave transcription/translation feedback loop in a circadian system. *PLoS Biol.* 8, e1000394.
- Rey, G., Cesbron, F., Rougemont, J., Reinke, H., Brunner, M., and Naef, F. (2011). Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biol.* 9, e1000595.
- Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O’Shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science* 318, 809–812.
- Smith, R.M., and Williams, S.B. (2006). Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. USA* 103, 8564–8569.
- Smith, K.M., Sancar, G., Dekhang, R., Sullivan, C.M., Li, S., Tag, A.G., Sancar, C., Bredeweg, E.L., Priest, H.D., McCormick, R.F., et al. (2010). Transcription factors in light and circadian clock signaling networks revealed by genome-wide mapping of direct targets for neurospora white collar complex. *Eukaryot. Cell* 9, 1549–1556.
- Takai, N., Nakajima, M., Oyama, T., Kito, R., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2006). A KaiC-associating Sasa-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 103, 12109–12114.
- Taniguchi, Y., Katayama, M., Ito, R., Takai, N., Kondo, T., and Oyama, T. (2007). labA: a novel gene required for negative feedback regulation of the cyanobacterial circadian clock protein KaiC. *Genes Dev.* 21, 60–70.
- Taniguchi, Y., Takai, N., Katayama, M., Kondo, T., and Oyama, T. (2010). Three major output pathways from the KaiABC-based oscillator cooperate to generate robust circadian kaiBC expression in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 107, 3263–3268.
- Teng, S.W., Mukherji, S., Moffitt, J.R., de Buyl, S., and O’Shea, E.K. (2013). Robust circadian oscillations in growing cyanobacteria require transcriptional feedback. *Science* 340, 737–740.
- Vijayan, V., and O’Shea, E.K. (2013). Sequence determinants of circadian gene expression phase in cyanobacteria. *J. Bacteriol.* 195, 665–671.
- Vijayan, V., Zuzow, R., and O’Shea, E.K. (2009). Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 106, 22564–22568.
- Vijayan, V., Jain, I.H., and O’Shea, E.K. (2011). A high resolution map of a cyanobacterial transcriptome. *Genome Biol.* 12, R47.
- Woelfle, M.A., Xu, Y., Qin, X., and Johnson, C.H. (2007). Circadian rhythms of superhelical status of DNA in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 104, 18819–18824.
- Yang, Q., Pando, B.F., Dong, G., Golden, S.S., and van Oudenaarden, A. (2010). Circadian gating of the cell cycle revealed in single cyanobacterial cells. *Science* 327, 1522–1526.
- Zwicker, D., Lubensky, D.K., and ten Wolde, P.R. (2010). Robust circadian clocks from coupled protein-modification and transcription-translation cycles. *Proc. Natl. Acad. Sci. USA* 107, 22540–22545.

Supplemental Information

EXTENDED EXPERIMENTAL PROCEDURES

Cyanobacterial Strains

Strains were constructed using standard procedures for genomic integration by homologous recombination (Clerico et al., 2007) and are described in Table S7A. Unless specified otherwise, our “wild-type” strain was AMC408, which contains a *PpurF* reporter but is otherwise wild-type (Table S7A).

Strains AMC395 and AMC408 were gifts from Susan Golden (University of California, San Diego). Plasmids pAM1573, pAM1580, pAM2055, and pAM2991 were gifts from Susan Golden. Plasmids pDRpaA(Km') and pNS2KmPtrc-kaiBC were gifts from Takao Kondo (Nagoya University).

Strain EOC66 was constructed by transforming AMC408 with pDRpaA(Km') (Takai et al., 2006). EOC72 was constructed by transforming AMC408 with pNS2KmPtrc-kaiBC (Murayama et al., 2008) and a plasmid deleting the *kaiBC* locus between positions +1 and +1813 (relative to the translation start site of *kaiB*) with the chloramphenicol resistance cassette from pAM1573 (Andersson et al., 2000). EOC101 was constructed by transforming AMC408 with pNS2KmPtrc-kaiBC, a plasmid deleting the *kaiBC* locus between positions +1 and +1813 (relative to the translation start site of *kaiB*) with a gentamycin cassette (SmaI-PvuII fragment from pAM2055 (Katayama et al., 2003)), and a plasmid deleting the same region of *rpaA* targeted by pDRpaA(Km') but with the chloramphenicol cassette from pAM1573.

The *PkaiBC* reporter strain EOC113 was created by transforming AMC395 with the *luxAB*-encoding plasmid pAM1580 containing the *kaiBC* promoter (Chabot et al., 2007) inserted between the XbaI and SalI restriction sites. Strains EOC339, EOC341, EOC345, and EOC346 were constructed by transforming EOC113 with the chloramphenicol-marked *rpaA* deletion vector described above for EOC101 and with a gentamycin-marked neutral site 2.2 (NS 2.2) targeting vector carrying the *Ptrc* promoter driving expression of the appropriate gene. The targeting vector was constructed by replacing the chloramphenicol cassette in EB2065 (Jain et al., 2012) with a gentamycin cassette from pAM2055. Specifically, primers CATAACTAGTGGATCTGGTAACCCCCAGCGCGG and ATA CACTAGTGGATCTAACGCTTGCTTGC were used to perform inverse PCR on EB2065, introducing an SphI site into which was cloned the gentamycin cassette amplified from pAM2055 with primers CATAACTAGTGGACGCACACCGTGGAAACG and ATACACTAGTGGCGCTTGACAATTTC.

To obtain the *Ptrc* constructs to place into this targeting vector, we first cloned the *rpaA* gene into the EcoRI and BamHI sites in vector pAM2991 (Mackey et al., 2008); for the empty-vector (mock / *Ptrc*::-) control, pAM2991 was used without modification. The D53A and D53E mutations then were introduced using the QuikChange II XL kit (Agilent) with primers ACCTGATCATGCTGGCTC TAATGCTGCCGCG and CGCGGCAGCATTAGAGCCAGCATGATCAGGT (D53A) or CCTGATCATGCTGGAGCTATGCTGC CGCGG and CCGCGGCAGCATTAGCTCCAGCATGATCAGG (D53E). The *lacZ*-*Ptrc-rpaA*/empty-*rrnB* region of these pAM2991 variants was amplified with primers P-GCGACATCTCCTGCTCCA and P-GCGGGCTTAACCAAGC and cloned into the SmaI site of the NS2.2 (Gm') targeting vector.

Strains EOC370, EOC371, and EOC374 were constructed in a true wild-type background (*Synechococcus elongatus* PCC7942, ATCC catalog number 33912). The ATCC strain was transformed with pDRpaA(Km'), the chloramphenicol-marked *kaiBC* deletion vector described above, and the NS2.2 (Gm') targeting vectors containing the appropriate *Ptrc* construct described above.

Strains EOC116 (*ΔsasA*) and EOC118 (*ΔcikA*) were described in Gutu and O’Shea, 2013.

The HA-RpaA strain EOC67 was constructed by knocking a KanR-3xHA-RpaA construct into the *rpaA* locus in an AMC408 background, such that the endogenous *rpaA* gene was completely replaced with the 3xHA-*rpaA* construct. The kanamycin cassette was placed such that the region between the RpaA binding site (peak 6 in Table S3A) and the 3xHA tag was wild-type. The sequence of the genome of EOC67 around the knocked-in KanR-3xHA-RpaA construct is:

synpcc7942_0096-KanR cassette-3xHA(with linker)-rpaA (synpcc7942_0095)
GGATTCTACAAACGAACTTGAGCTTGCGATGAAGAAAACCTCAAGGCAGCCCTAGAAAAAGCCAAAGCCGAAGCCCAAGCCTGCGTCTCAATTAAATCTCCACGCTTCGATCGCTCCTGATTACTTGAGTTGATCAAGGCTCAGATCAAGAGCGATCGCTACTTC AATACAGCTTACTTTGGCCGCTAGTCGACCTGCATCCCTTAACCTACTTATTAAATAATTAGCTATTGAAAGAGATAAGA ATTGTTCAAAGCTATATTGTTAAATCGTCAATTCTGCATTTAAGGAATTGTTAAATTGATTGATTTTGATTTGAAATATTCTTG TATTCTTGTTAAACCCATTCTATAACGAAATAATTATACTTTGTTATCTTGATATTCTGATTTTCTACTTAATCTG ATAAGTGAGCTATTCACTTAGGTTAGGATGAAAAGCTAGAGGATCTCAATGAATATTGTTGACACGGCGTATAAGACATG TTACTGTTGAATAACAAGGACGGATCTGATCAAGAGACAGGATGAGGATCGTTGCATGATTGAACAAGATGGATTGCAC GCAGGTTCTCGGCCGCTTGGTGGAGAGGCTATTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCTGATGCCGCC GTGTTCCGGCTGTCAGCGCAGGGCGCCGGTTCTTGTCAAGACCGACCTGTCGGCTGCCCTGAATGAACCTGCAAGGAC GAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCTGCGCAGCTGTCGCTGACGTTGTCACTGAAGCGGGAAAG GGACTGGCTGCTATTGGGCGAAGTGCAGGGCAGGATCTCCTGTCATCTCACCTGCTCTGCCAGAAAGTATCCATCATG GCTGATGCAATCGGGCGGCTGCATACGCTTGTACCGGCTACCTGCCATTGACCAACCGAAACATCGCATTGCGA GCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGACGAGAGCATCAGGGCTGCCAGCCGAACCTG TTCGCCAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTGTCGTCGACCCATGGCGATGCCCTGCTGCCGAATATCATG GTGGAAAATGGCCGCTTCTGATTGACTGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCT ACCCGTGAATTGCTGAAGAGCTGGCGGAATGGCTGACCGCTCCTCGTGCCTTACGGTATGCCGCTCCCATTG

CAGCGCATGCCCTCTATGCCCTTGTGACGAGTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGC
 CCAACCTGCCATCACGAGATTCGATTCCACCGCCCTCTATGAAAGGTTGGCTCGGAATCGTTCCGGACGCCGG
CTGGATGATCCTCCAGCGGGGATCTCATGCTGGAGTTCTCGCCCACCGGGATCCCTAACAGCAGTCAATCCTAAGAA
 ACTGCAGAAGCAGGGATTGAGTAATTGATGGTTGACTGCAGTCCTTATCAGTTGCTCTCAACAATTATCCAACCGCAATTACGGA
 CAGAGCACTGGACTAGCACCCTGCTAAAACCGCAAAGCTCGGACAAATATAAGACAACCTAAAGATTTCTACGCCCTGCGAC
 CTCACCAGAACAGAGCCACTATCATTTAAAGAAACATAAATTGATCCGCTGTAAGCCGCTGCCAGCACAAAGCTTACCGCGTT
 TTGCTAAAGCCTAGTCGCTGGTTGCTCCCGGAATGTTACCCCTAGTACCCCTACGATGTGCCGATTACGCTGGCAGCGCAGCATGAAACCCGATCCTCGT
TATGACGTCCCGATGCCGTAGCTATCCATATGACGTGCCGATTACGCTGGCAGCGCAGCATGAAACCCGATCCTCGT
GATCGATGATGACTCAGCCATCTTGAGCTGGTCGCCGTCATCTGGAGATGTCGGCTATGACGTACGCAAAGCTGAGG.

Cell Culture

For the experiments described in Figures 1, 3A, 4, 5, S1, S3A, S4A, S4B, and S4D, cells were grown in a turbidostat as described previously (Vijayan et al., 2009). Cultures were entrained by exposure to 12 hr of darkness, followed by 12 hr of light, followed by another 12-h of darkness, after which they were grown in constant, continuous light.

For Figures 2, 6, S2, S3B, S3C, S4C, and S5A-S5D, cultures were grown in tissue culture flasks illuminated with $100 \mu\text{E m}^{-2} \text{s}^{-1}$ ($\mu\text{moles photons m}^{-2} \text{s}^{-1}$) of cool fluorescent light and bubbled continuously with 1% CO₂ in air, with the OD₇₅₀ maintained near 0.3 by diluting the cultures every four hours with fresh medium. Medium was supplemented with 10 mM HEPES-KOH pH 8.0 maintain the pH. For the rescue and RpaA overexpression experiments (Figures 2 and S2, and 6 and S5A-D, respectively), cultures were grown initially in the absence of IPTG, treated with two 12-h dark pulses separated by 12 hr of light, and released to constant light ($100 \mu\text{E m}^{-2} \text{s}^{-1}$) concomitant with addition of IPTG to a final concentration of 6 μM (Figures 2 and S2) or 100 μM IPTG (Figure 6 and S5A-D). Culture conditions for cell length measurements are described in the Cell Length Measurements section below.

Identification of Reproducible Circadian Genes

In order to focus our analysis on genes that show reproducible circadian oscillations, we considered as circadian only those genes identified as circadian in both the 60 hr time course presented in Vijayan et al., 2009, and in a previously unpublished biological replicate time course 28 hr in length (V. Vijayan, personal communication). We also required that genes be expressed in the same class (class 1 or class 2) in both time courses. Specifically, we considered a gene to be reproducibly circadian if it had: (a) a Cosine period (Kuchō et al., 2005) of 22–26 hr in the 60-h time course described in Vijayan et al., 2009; (b) a Cosine period of 20–28 hr in the 28-h time course; and (c) had the same phase-based class (1 or 2) in both time courses. The period window is wider for the 28 hr time course because it was shorter and therefore the period was less precisely determined. 856 genes satisfy these criteria.

Anti-RpaA Antibody Production and Affinity Purification

Antibody against full-length recombinant RpaA protein was produced in rabbits and affinity purified prior to use. Recombinant RpaA was produced as described (Takai et al., 2006), and was used to inoculate two rabbits. Cocalico Biologicals performed the live animal work. To purify the antibody, an affinity column was prepared using Affi-Gel matrix (1:1 mixture of Affi-Gel 10 and Affi-Gel 15, Bio-Rad) to which recombinant RpaA was immobilized. Affinity purification was carried out at room temperature. Antisera from the two rabbits was combined in a 1:1 ratio by volume and applied to a column. The flow-through was passed back through the column. The column was then washed with 20 column volumes (CVs) of 10 mM TrisCl pH 7.5 followed by 20 CVs of 10 mM TrisCl pH 7.5 with 500 mM NaCl. Immobilized antibody was eluted with 10 CVs of 100 mM glycine pH 2.5 in 1 ml fractions and immediately neutralized with TrisCl pH 8.0. A base elution was then performed using 100 mM triethanolamine pH 11.5, neutralized with TrisCl pH 8.0. Protein abundance in the eluates was estimated by BCA assay. Protein was found predominantly in the acid eluate, with little to none present in the base eluate. When the base eluate showed no evidence of protein, it was discarded. Fractions containing protein were combined and dialyzed in 10 KDa MWCO Slide-A-Lyzer units (Pierce) against PBS containing 0.02% NaN₃. Antibody was then concentrated using a 10 KDa MWCO Amicon spin filter. Antibody concentration was determined by absorbance at 280 nm or by BCA assay.

The specificity of the antibody was assayed by comparing reactivity against wild-type and *rpaA* mutant ($\Delta rpaA$) lysate by Western blot (Figure S3A). Only a single band was observed for the wild-type lysate, while no reactivity was observed against the *rpaA* mutant lysate.

Western Blot Analysis

Cells were collected on nitrocellulose or cellulose acetate filters using vacuum filtration. Filters were flash-frozen in liquid nitrogen and stored at -80°C until lysis. To prepare lysates, cells were eluted from the filters using ice-cold lysis buffer (7.5 M urea, 20 mM HEPES pH 8.0, 1 mM DTT, and 1x Roche Complete protease inhibitor tablet, with or without 1 mM EDTA). Resuspensions were transferred to 500 μl screw-cap tubes containing 0.1 mm glass beads. The cells were then lysed by bead-beating at 4°C for a total of 5 min with periodic cooling on ice. The lysate was centrifuged for 5 min at $\geq 16,000 \times g$ at 4°C , after which the supernatant was transferred to a clean microcentrifuge tube. The protein concentration of each sample was then measured by BCA assay (Pierce) using bovine serum albumin (BSA, Bio-Rad) diluted into lysis buffer as the standard. For each Western blot, an equal mass quantity of each lysate was loaded onto an SDS-PAGE gel. SDS-PAGE gel compositions, electrophoresis conditions, and Western blotting procedures were performed as described in Gutu and O'Shea 2013, except that affinity-purified anti-RpaA antibody (0.17 $\mu\text{g/ml}$) was used in place of anti-RpaA serum.

Chromatin Immunoprecipitation

For each chromatin immunoprecipitation (ChIP) reaction, approximately 18 OD_{750 nm} units of culture were crosslinked for 15 min with 1% formaldehyde followed by quenching for 5 min with 125 mM glycine. Cells were collected by centrifugation for 10 min at 6000 x g at 4°C and then washed twice with 30 ml of ice-cold phosphate-buffered saline (PBS), centrifuging for 10 min at ~3000 x g at 4°C after each wash. Samples were then resuspended in 1 ml ice-cold PBS and pelleted in a microcentrifuge tube for 3 min at 3,200 x g at 4°C. The supernatant was discarded and the pellet was flash frozen in liquid nitrogen and stored at -80°C.

Samples were thawed on ice and resuspended in 500–600 µl of ice-cold lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 1x Roche Complete EDTA-free Protease Inhibitor Cocktail). Cells were lysed by beadbeating at 4°C in 2 ml screw-top tubes with 0.1 mm glass beads for 10 cycles of 30 s each separated by at least 30 s of cooling on ice. Lysate was separated from the beads by piercing the bottom of each tube with a small-diameter needle, placing the tube into a clean 1.5 ml microcentrifuge tube, and centrifuging for several minutes at < 500 x g at 4°C to transfer the lysate to the 1.5-ml tube. Chromatin was then sheared to a length of ~300 bp by sonication on ice in a Misonix sonicator 3000 for 9–11 cycles of 15 s separated by at least 90 s of cooling. Cell debris was removed by centrifuging twice at 14,000 x g for 15 min each at 4°C. Protein concentration in the lysates was determined by BCA assay using BSA as a standard.

For a given ChIP time course, equal mass quantities (typically 1–1.5 mg) of lysate from each time point were prepared in 500 µl of lysis buffer each. For anti-RpaA ChIP, equal amounts (typically 10–15 µg) of affinity-purified antibody were added to each tube; for anti-HA ChIP, 40 µl (bed volume) of anti-HA agarose beads (Pierce) equilibrated in lysis buffer were added to each tube. Samples were incubated overnight in the dark at 4°C with continuous rotation. The following morning, for anti-RpaA ChIP only, 50–70 µl of rProtein A Sepharose Fast Flow beads (GE Healthcare; 33% slurry in lysis buffer) was added to each sample, and the samples were then incubated for 2 hr at 4°C with rotation. For both anti-RpaA and anti-HA ChIP, beads were isolated by centrifugation for 1 min at 1000 x g at room temperature. Beads were then washed twice with 1 ml lysis buffer, once with 1 ml of buffer B (50 mM HEPES pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate), once with 1 ml of wash buffer (10 mM TrisCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.1% sodium deoxycholate) and finally with 1 ml of TE pH 7.5 (10 mM TrisCl pH 7.5, 1 mM EDTA); each wash was conducted for 5 min at room temperature on a tube rotator followed by isolation of beads by centrifugation for 1 min at 1000 x g at room temperature. Protein-DNA complexes were then eluted with 250 µl elution buffer (50 mM TrisCl pH 8.0, 10 mM EDTA, 1% SDS) for 1 hr at 65°C.

For both the eluate and a matched sample of lysate not subjected to immunoprecipitation (“input DNA”), crosslinks were reversed for 6–18 hr at 65°C. Next, 250 µl of TE was added to each sample to dilute the SDS, followed by addition of 100 µg of proteinase K and 80 µg of glycogen. Samples were incubated for 2 hr at 37°C to digest proteins. Samples were then supplemented with 55 µl of 4 M LiCl and extracted with 1 ml of phenol/chloroform/isoamyl alcohol followed by extraction with 1 ml chloroform. DNA in the aqueous phase was precipitated with ethanol and then washed once with 80% ethanol. Pellets were air-dried and then resuspended in 50 µl of TE containing 20 ng/µl of DNase-free RNase (Fermentas) and incubated for 1 hr at 37°C to digest RNA. Samples were then supplemented with 150 µl of TE and 22.2 µl of 3 M sodium acetate pH 5.2, extracted with phenol/chloroform/isoamyl alcohol, and precipitated with ethanol; alternatively, DNA was purified using a QIAGEN PCR purification kit. DNA concentration in ChIP samples was estimated either by PicoGreen assay (Invitrogen) or by qPCR based on the abundance of the coding region of *synpcc7942_0612*, a region showing no enrichment for RpaA binding in ChIP-Seq experiments; for both methods, input DNA quantified by absorption spectrophotometry was used as a standard.

Typical immunoprecipitation efficiencies (fraction of RpaA depleted from the lysate) were greater than 50%.

qPCR for ChIP and Gene Expression

qPCR was carried out using Taq polymerase, SybrGreen dye, and the promoter-specific primers described below.

Assay	Target	Primer sequences	Reference
ChIP-qPCR	<i>PkaiBC</i>	F: TTTACGAGGGCTCATACGC R: CCCACGAGAAACCTGAAAAG	This work
ChIP-qPCR	<i>synpcc7942_0612</i> CDS	F: AGAACGCAACGACAGCGTAGG R: GTCGCAGCTCGGATTTTG	This work
RT-qPCR	<i>kaiBC</i>	F: TACATTCTCAAGCTCTACG R: CGTCGCTAGGATTTATCC	(Vijayan and O’Shea, 2013)
RT-qPCR	<i>hsIO</i>	F: CAGACCAACTGATTCGAGCG R: GGAGGCCAGGAGCAGTC	(Vijayan and O’Shea, 2013)

RT-qPCR was performed as described previously (Vijayan and O’Shea, 2013), with the abundance of the *kaiBC* transcript normalized to that of *hsIO* transcript, whose abundance is constant in time (Vijayan and O’Shea, 2013; Vijayan et al., 2009).

For ChIP-qPCR, DNA abundances were calculated using a standard curve constructed from input DNA, which was prepared in the same manner as ChIP DNA but using sonicated lysate not subjected to immunoprecipitation.

Enrichment was computed by normalizing the abundance of the promoter of interest to the abundance of the coding region of *synpcc7942_0612*, which shows no enrichment for RpaA occupancy by ChIP-seq.

ChIP-Seq Library Preparation and Sequencing

Except for “replicate 2” samples in GSE50922, libraries for Illumina sequencing of ChIP DNA were prepared following a protocol developed by Ethan Ford (http://ethanomics.files.wordpress.com/2012/09/chip_truseq.pdf), with modifications. Specifically, 0.15–3 ng of ChIP DNA was used for each sample. DNA ends were blunted by treatment with 1.4 units of T4 DNA polymerase (NEB), 0.45 units of Klenow fragment (NEB), 4.5 units of T4 polynucleotide kinase (NEB), and 0.4 mM dNTPs (NEB) in 1x T4 DNA ligase buffer (NEB) in a total of 50 μ l volume for 30 min at 20°C. DNA was purified using 50 μ l of AMPure XP beads (Beckman) and 50 μ l of a solution containing 20% PEG8000 (Sigma) and 1.25 M NaCl. DNA was eluted in 16.5 μ l of TE/10 (10 mM TrisCl pH 8.0, 0.1 mM EDTA). DNA was then A-tailed at the 3' ends by treating the eluate with 2.5 units of Klenow fragment lacking 3'-5' exonuclease activity (Klenow fragment 3' → 5' exo-, NEB) and 0.2 mM dATP (GE Healthcare) in 1x NEB Buffer 2 in a total volume of 20 μ l for 30 min at 37°C. TruSeq adapters (Illumina or Eurofins MWG Operon) were ligated onto the A-tailed DNA by addition of 25 μ l of 2X Quick Ligase Buffer (NEB), 0.5 μ l of ~250 nM TruSeq adaptor (1:30 dilution of Illumina stock), 3 μ l of nuclease-free H₂O, and 1.5 μ l of Quick Ligase (NEB) followed by incubation for 20 min at 21°C. The ligation reaction was stopped by addition of 5 μ l of 0.5 M EDTA pH 8.0 (Ambion). Next, DNA was purified using 55 μ l of AMPure XP beads without additional PEG or salt. DNA was eluted in 15.5 μ l of TE/10. The Y-shaped adapters were then linearized with 5 cycles of PCR (initial denaturation of 30 s at 98°C followed by 5 cycles of [10 s at 98°C, 30 s at 60°C, 30 s at 72°C] followed by 5 min at 72°C) using Phusion polymerase (Thermo) in HF Buffer and 1 μ l of TruSeq primers (25 μ M) in a total volume of 31 μ L. Linearized DNA was purified using 30 μ l of AMPure XP beads without additional PEG or salt. DNA was eluted in 30 μ l of TE/10. Fragments between 300 and 500 bp were size-selected using agarose gel purification and the QIAquick gel extraction kit (QIAGEN) or using a Pippin Prep (SAGE Science). Purified DNA was further amplified with 13–14 cycles of PCR, as described above, in a total volume of 62.5 μ l. Following PCR, DNA was purified using 51 μ l of AMPure XP beads without additional PEG or salt. DNA was eluted in 12 μ l of TE/10. Libraries were assessed using a DNA High Sensitivity chip on an Agilent Bioanalyzer 2100. Samples were sequenced on a HiSeq instrument or Genome Analyzer II (Illumina) by the core facility at the Harvard FAS Center for Systems Biology. Reads were aligned to the *S. elongatus* genome using Bowtie ([Langmead et al., 2009](#)), counting only those aligning uniquely to one location with up to three mismatches. For computational simplicity, the chromosome (NC_007604.1) and two plasmids (NC_004073.2 and NC_004990.1) were treated as linear.

For “replicate 2” samples in GSE50922, libraries for sequencing were prepared using paired-end adapters and PCR primers from Illumina following the manufacturer’s protocol for ChIP-Seq sample preparation (part number 11257047 revision A) with modifications. All enzymes except for fast ligase were obtained from NEB instead of Illumina. First, 5–10 ng of ChIP or input DNA was blunted and phosphorylated with Klenow DNA polymerase, T4 DNA polymerase, and T4 polynucleotide kinase. 3' dA overhangs were added using exo⁻ Klenow fragment. Adapters (diluted 1:10 from the Illumina stock) were ligated for 15 min at room temperature using Ligafast DNA ligase (Promega). Adapter-ligated fragments were size-selected on a 2% agarose/TAE gel containing 400 μ g/ml ethidium bromide; a gel slice corresponding to the 200–300 bp region (based on 50 bp ladder, Promega) was excised. DNA was purified from the gel using the QIAGEN gel extraction kit according to the manufacturer’s protocol, except that the gel slice was dissolved at room temperature. Adapter-ligated fragments were then amplified by PCR using PE Primers 1.0 and 2.0 (Illumina). The resulting libraries were purified with the QIAGEN MinElute kit and eluted in 15 μ l of buffer EB. The concentration and fragment length distribution of the libraries was then determined using the DNA 1000 Bioanalyzer assay (Agilent). Only those reads in which the first 32 bases aligned perfectly and uniquely to the genome were counted. For computational simplicity, the chromosome and two plasmids were treated as linear.

The number of aligned reads and the most common insert size for each sample are listed in Table S7B.

RpaA-Binding Motif

Enriched motifs were identified with MEME ([Bailey and Elkan, 1994](#)) using sequences within 125 bp of each RpaA peak and a background consisting of a fourth-order Markov model of the entire genome. We searched for motifs between 6 bp and 26 bp in width, the latter value being the width of the largest DNase I footprint we observed (*PrpoD6* #1, [Figures 4D and S4E](#)). Only one statistically significant motif ([Figure 4D](#), E -value 1.7×10^{-36}) was located; the position-specific probability matrix (PSPM) for this motif is listed in Table S4A. We identified instances of this motif in the query sequences using FIMO ([Grant et al., 2011](#)) (Table S4B).

DNase I Footprinting

The *PkaiBC*, *PrpoD6*, and *PpurF* promoter sequences were amplified from *S. elongatus* genomic DNA with primers GACGGATCCTTTACGAGGGCTACGC and GACGAATTCGTGAGATGTATCGACGGTCTATCC for *PkaiBC*, GACGGGATCCCATCTCTTGTTCGTCGCTGAG and GACGAATTCTCCCTTACTTTCGACACA for *PrpoD6*, and GACGAATTCCATGCCCTTGCAGAGGCTC and GACGGATCCCGGATCGAACGTCGTTTG for *PpurF*. These primers added BamHI and EcoRI restriction sites (underlined) to the ends of the amplified promoters. Promoter sequences were ligated into BamHI/EcoRI-digested pUC18 to create pUC18-*PkaiBC*, pUC18-*PrpoD6*, and pUC18-*PpurF*. To generate mutations in *PkaiBC*, we utilized the QuikChange II XL site-directed mutagenesis kit (Agilent) with primers TCTATCCCACGGAGAAACCTGAAAAGGTTAGGAGGTCTTAAGC and GCTTAAGACCTCTAAACCTTTGAGGTTCTCGGGGGATAGA for m15, TCCCACCGGAGAAACCTGAAAAGGTAATCGAGGGTCTAAGCT and GAGGTTAAAGACCTCTGATTACCTTTCAGGTTTCTCGTGGGA for m16, and CTGAAAAGGTAAAGGGAGGTCAAAAGCTCGGCTCAATTTCTT and AGAGAATTGAGCCGAGCTTTGACCTCCTTACCTTTCAG for m20, using pUC18-*PkaiBC* as a template. We mutagenized *PpurF* similarly using primers GGCTAAATTAACATGTCTTTACCCTAGCGGTTAGTCTTTAGC and GCTAAAGACTAACCGC

TAGGGTAAAGACATTGTTAATTCGATTAGCC for m1-5, GCTAATGCAAATTAACAGTGACTTAACCCTAGCGGTTAGTCTTA and TAAAGACTAACCGCTAGGGTTAAGTCAGTCACTGTTAAATTGCATTAGC for m1-7, and AGCGGCTAATGCAAATTAGCAATGACT TAACCCTAGCG and CGCTAGGGTTAAGTCATTGCTAAATTGCATTAGCCGCT for m1-8, using pUC18-PpurF as a template.

DNA probes for DNase I footprinting were prepared as follows. First, a forward primer was 5' end-labeled with ^{32}P by incubation with T4 polynucleotide kinase (Promega) and γ - ^{32}P ATP (Perkin Elmer). For the *Pka/BC* probes, this primer was TTTTAC GAGGGCTCATACGC, for the *PrpoD6* probe this primer was ATCTCTGTTCGCTGAG, and for the *PpurF* probes this primer was CATGCCCTTGAGAGCTC. This primer was subsequently used to prepare the footprinting probe via PCR amplification from the appropriate plasmid template from above with a reverse primer. For the *Pka/BC* probes, the reverse primer was GTGAGATG TATCGACGGTCTATCC, for the *PrpoD6* probes, the reverse primer was TCCCTCTTACATTTGACACA, and for the *PpurF* probes, the reverse primer was GCGATCGAACGTCGTTG.

Binding reactions were carried out in a buffer containing 150 mM KCl, 5 mM MgCl₂, 20 mM HEPES-KOH pH 8.0, 10% glycerol (w/v), 1 mM DTT, and 30 ng/ μL poly dI-dC (Affymetrix). Binding reactions contained the appropriate DNA probe, varying concentrations of RpaA (purified as described in (Takai et al., 2006)), 1.5 μM CikA (purified as described previously (Gutu and O'Shea, 2013)), and 1 μM ATP, as indicated. Binding reactions were incubated for 1 hr at 30°C.

To initiate DNase I digestion, an equal volume of DNase I (0.025 U/ μL , New England Biolabs), in a buffer containing 25 mM TrisCl pH 8.0, 10 mM MgCl₂, 5 mM CaCl₂, 1 mM EDTA, and 10% glycerol was added to the reaction and incubated at room temperature for 2 min. Digestion was stopped by adding an equal volume of stop solution containing 200 mM NaCl, 30 mM EDTA, 1% SDS (w/v), and 100 $\mu\text{g}/\text{mL}$ yeast tRNA (Sigma).

DNA was then purified using phenol-chloroform extraction and ethanol precipitation and resuspended in a buffer containing 30 mM NaOH, 60% v/v formamide, and 0.1% w/v bromophenol blue. Sanger sequencing reactions were prepared with the Sequenase DNA sequencing kit (USB) using the appropriate plasmid as template. The reactions were run on a 6% acrylamide-urea gel prepared with the Ureagel system (National Diagnostics). Radiolabeled DNA was visualized with phosphorimaging.

For Phos-tag electrophoresis, reactions were carried out as described above but with *kaiBC* DNA probe prepared with unlabeled ATP. After one hour of incubation at 37°C, 2.5 μl of reaction mixture was diluted into 1x Phos-tag loading buffer and resolved on a 7% acrylamide gel containing 75 μM MnCl₂ and 50 μM Phos-tag AAL-107 reagent (Wako Chemicals) at 4°C. After electrophoresis, the gel was stained with Sypro Ruby (Invitrogen) according to manufacturer's instructions and imaged on a Typhoon Scanner (GE Healthcare).

ChIP-Seq Data Analysis

Data were analyzed using a custom-coded, modified form of the PeakSeq algorithm (Rozowsky et al., 2009) that narrows the regions identified as peaks by requiring that each 50-bp window within a putative peak be enriched ($p \leq 0.05$) relative to both the mock ChIP and the input DNA. (For computational simplicity, the chromosome and two plasmids were treated as linear.) The fold enrichment for each trimmed peak was calculated by finding the maximum ChIP-to-mock ratio within 50 bp of the location of the peak maximum in the raw ChIP-Seq signal (not in the enrichment ratio). For ChIP using the anti-RpaA antibody, we required peaks to be present in each of two biological replicates with enrichment ≥ 3 -fold in one replicate and ≥ 2.22 -fold in the other replicate, which had 26% lower enrichment overall. Multiple hypothesis-corrected q -values were less than 10^{-78} (smallest q -value between the two replicates).

To call targets of a given RpaA peak, we identified all transcripts with 5' ends lying within 500 bp of the location of maximum ChIP-Seq signal. (For this analysis, we considered only annotated mRNA, tRNA, or rRNA transcripts and high-confidence non-coding transcripts (Vijayan et al., 2011)). Within this set of nearby transcripts, we called as the target that transcript with its 5' end nearest the ChIP-Seq signal maximum. To account for possible inaccuracies in the 5' end identifications, we also assigned as targets any other transcripts with 5' ends whose distance to the ChIP-Seq signal maximum was within 50 bp of the distance between the 5' end of the closest transcript and the ChIP-Seq signal maximum.

RpaA ChIP target genes were categorized by their Cyanobase-assigned functional category (Nakao et al., 2010) using manually-updated assignments (Table S1).

RNA-Seq

Ribosomal RNA was depleted from 6 μg of total RNA (purified as described in Vijayan et al., 2009) using the MICROBExpress Bacterial mRNA enrichment kit (Applied Biosystems) according to manufacturer's instructions. Strand-specific RNA-sequencing libraries were prepared from 375 ng of rRNA-depleted RNA using the TruSeq Stranded mRNA Sample Prep Kit (Illumina). Samples were multiplexed and sequenced on an Illumina HiSeq machine by the core facility at the Harvard FAS Center for Systems Biology.

Sequencing reads were aligned to the *S. elongatus* chromosome as described above for ChIP-Seq. The number of aligned reads is listed in Table S7C.

To quantify gene expression, we counted the number of coding-strand sequencing reads with 5' ends between the start and stop positions of the coding region of each gene. To normalize gene expression values between samples, we utilized median normalization as described elsewhere (Anders and Huber, 2010). First, we calculated a pseudo-reference for each gene by determining the geometric mean of the expression counts for that gene across all samples, and calculated the ratio of the expression values with the appropriate pseudo-reference value. We determined the median value of these ratios within each sample and took this as a size factor that estimates the sequencing depth of each sample. To normalize, we divided all gene expression values within a sample

by the appropriate size factor. For Figure S5D, gene expression values were further normalized by dividing the gene expression values by the length of the appropriate open reading frame.

To calculate correlations for Figures 6A and S5B, we determined the average expression of each gene in wild-type cells over the course of a day. Then, we expressed gene expression at each time point as the log of the ratio of the expression at that time point to the average wild-type expression. If a circadian gene showed an expression value of 0 (meaning that no reads were present) in any time point from this experiment, we did not consider it when calculating correlations ($n = 24$ genes). We also ignored expression of *kaiB* and *kaiC* in this analysis, as these genes were disrupted in OX-D53E. We calculated Pearson's correlation coefficient between samples for the remaining high-confidence circadian genes with these log-transformed values.

For K-means clustering, we first normalized all gene expression values for reproducibly circadian genes (Table S1), except for *kaiB* and *kaiC*, in the OX-D53E time course or the wild-type time course using z-score normalization, and then clustered these genes into 6 groups using K-means clustering (MATLAB) with squared Euclidean distance as the distance metric.

Bioluminescence Time Courses

Cultures were inoculated into 96-well plates containing 250 μ l BG-11M and the specified concentration IPTG in each well. After incubation in light ($\sim 60 \mu\text{E m}^{-2} \text{s}^{-1}$) for a minimum of 12 hr, the cultures were entrained with two 12-h dark pulses and then released to constant light. Bioluminescence was measured every 2 hr in constant light in a TopCount luminometer (PerkinElmer) as described previously (Vijayan and O'Shea, 2013). Replicate experiments produced qualitatively similar results.

Cell Length Measurements

Cultures were grown in medium containing appropriate antibiotics but no IPTG under constant light ($74 \mu\text{E m}^{-2} \text{s}^{-1}$) at 30°C. After reaching OD_{750 nm} of 0.4, cell cultures were diluted 1:10 into medium containing 100 μM IPTG and then grown for 4 days. Control strains received no IPTG. Cells were imaged by capturing their red autofluorescence using an AxioObserver Z1 inverted microscope (Zeiss) equipped with a Plan-Apochromat 100X/1.40 Oil Ph3 objective (Zeiss) and an Evolve EMCCD Camera (Photometric). Cell length analysis was performed in ImageJ (National Institutes of Health).

SUPPLEMENTAL REFERENCES

- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11, R106.
- Andersson, C.R., Tsioroumas, N.F., Shelton, J., Lebedeva, N.V., Yarrow, J., Min, H., and Golden, S.S. (2000). Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.* 305, 527–542.
- Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.
- Chabot, J.R., Pedraza, J.M., Luitel, P., and van Oudenaarden, A. (2007). Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature* 450, 1249–1252.
- Clerico, E.M., Ditty, J.L., and Golden, S.S. (2007). Specialized techniques for site-directed mutagenesis in cyanobacteria. *Methods Mol. Biol.* 362, 155–171.
- Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018.
- Jain, I.H., Vijayan, V., and O'Shea, E.K. (2012). Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. *Proc. Natl. Acad. Sci. USA* 109, 13638–13643.
- Katayama, M., Kondo, T., Xiong, J., and Golden, S.S. (2003). *ldpA* encodes an iron-sulfur protein involved in light-dependent modulation of the circadian period in the cyanobacterium *Synechococcus elongatus* PCC 7942. *J. Bacteriol.* 185, 1415–1422.
- Kucho, K., Okamoto, K., Tsuchiya, Y., Nomura, S., Nango, M., Kanehisa, M., and Ishiura, M. (2005). Global analysis of circadian expression in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* 187, 2190–2199.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Mackey, S.R., Choi, J.S., Kitayama, Y., Iwasaki, H., Dong, G., and Golden, S.S. (2008). Proteins found in a CikA interaction assay link the circadian clock, metabolism, and cell division in *Synechococcus elongatus*. *J. Bacteriol.* 190, 3738–3746.
- Min, H., Liu, Y., Johnson, C.H., and Golden, S.S. (2004). Phase determination of circadian gene expression in *Synechococcus elongatus* PCC 7942. *J. Biol. Rhythms* 19, 103–112.
- Nakao, M., Okamoto, S., Kohara, M., Fujishiro, T., Fujisawa, T., Sato, S., Tabata, S., Kaneko, T., and Nakamura, Y. (2010). CyanoBase: the cyanobacteria genome database update 2010. *Nucleic Acids Res.* 38 (Database issue), D379–D381.
- Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M.B. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat. Biotechnol.* 27, 66–75.

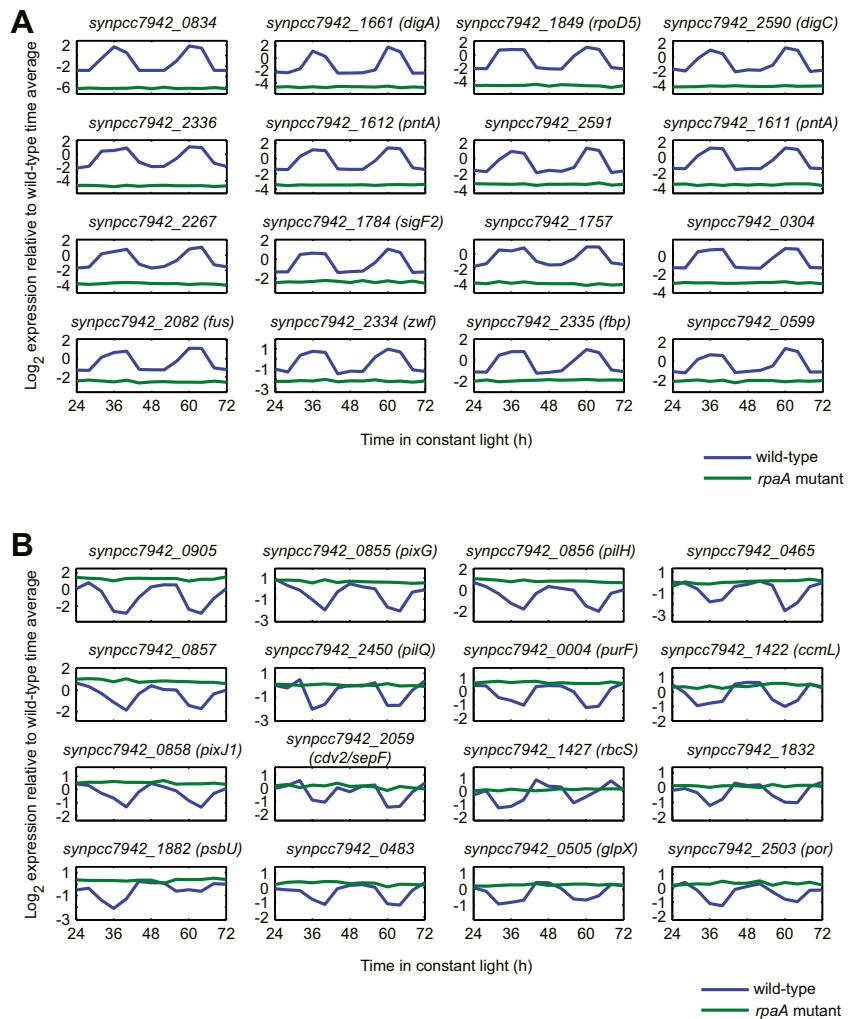


Figure S1. *rpaA* Mutant Gene Expression Time Course, Related to Figure 1

(A) Wild-type and *rpaA* mutant gene expression time courses for the highest-amplitude wild-type subjective dusk (class 1) genes. The expression time courses of the sixteen highest amplitude subjective dusk genes in the wild-type strain are shown in blue; expression at each time point is reported relative to the wild-type time average (Vijayan et al., 2009). The time courses of those same genes in the *rpaA* mutant strain are shown in green, also relative to the wild-type time average (see Experimental Procedures). Note that the overall expression level is significantly lower in the *rpaA* mutant than in the wild-type strain for all 16 genes. Gene expression was measured by microarray.

(B) Wild-type and *rpaA* mutant gene expression time courses for the highest-amplitude wild-type subjective dawn (class 2) genes. This plot is constructed in the same manner as is Figure S1A (above) but using the highest-amplitude wild-type subjective dawn genes. Note that the overall expression level generally is higher in the *rpaA* mutant than in the wild-type strain. Gene expression was measured by microarray.

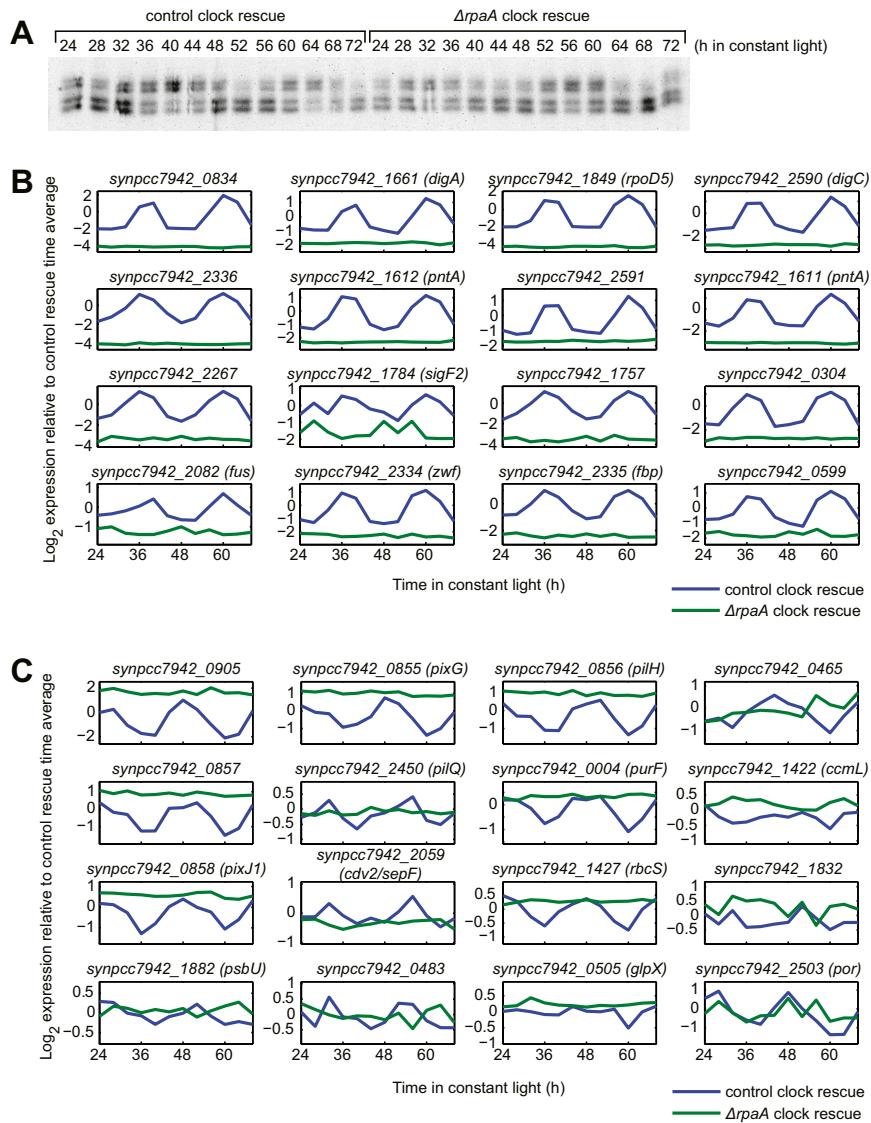


Figure S2. Rescue Experiments, Related to Figure 2

(A) Western blot of KaiC phosphorylation for the control and $\Delta rpaA$ clock rescue time courses shown in Figure 2. Quantification of these blots is shown below the heatmaps in Figure 2. Samples for Western blotting were acquired contemporaneously with those used for microarray analysis. Samples were lysed in urea lysis buffer, and equal amounts of total protein from each lysate were loaded onto a 4%–20% SDS-PAGE gradient gel for Western blot analysis. See Extended Experimental Procedures for details.

(B) Microarray-based gene expression time courses in the rescue experiments (Figure 2) for the subjective dusk (class 1) genes with the highest amplitude in the wild-type strain (same set as in Figure S1A). Expression at each time point in both strains is reported relative to the control clock rescue time average. The time average comparison was prepared in the same manner as in Figure S1 (see Experimental Procedures), except that the dye-swap experiment was not performed.

(C) Microarray-based gene expression time courses in the rescue experiments (Figure 2) for the subjective dawn (class 2) genes with the highest amplitude in the wild-type strain (same set as in Figure S1B). This plot is constructed in the same manner as is Figure S2A (above) but using the highest-amplitude subjective dawn genes in the wild-type strain.

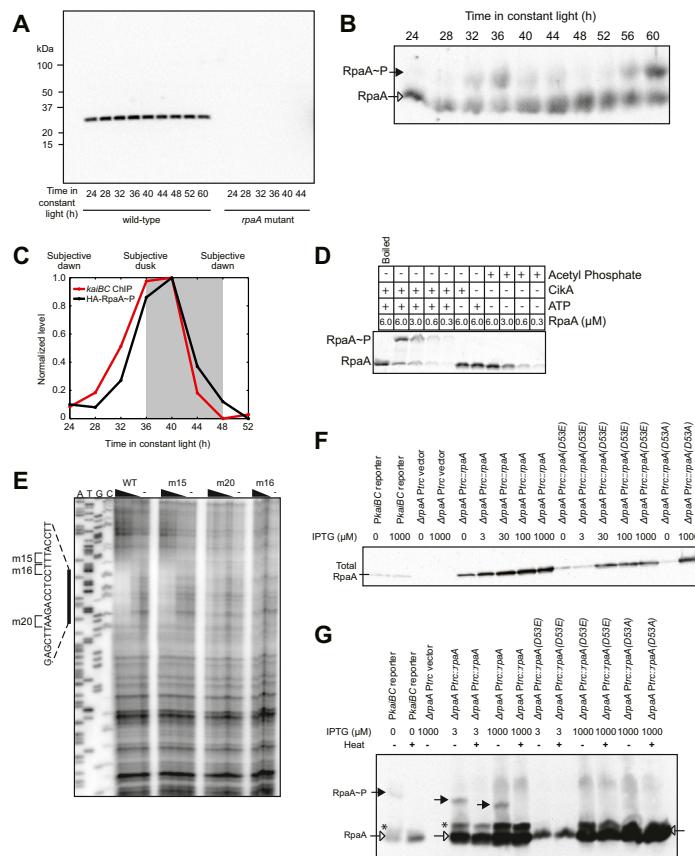


Figure S3. RpaA Binds to the *kaiBC* Promoter In Vivo and In Vitro and Promotes *kaiBC* Expression in a Phosphorylation-Dependent Manner, Related to Figure 3

(A) Analysis of affinity-purified anti-RpaA antibody specificity by Western blot. Samples from wild-type and *rpaA* mutant circadian time courses were lysed in urea lysis buffer. Equal masses of total protein from each lysate were separated on a 4%–20% gradient SDS-PAGE gel. Proteins were transferred from the gel to nitrocellulose, which was then blocked in milk and probed with the same affinity-purified anti-RpaA antibody used for ChIP-qPCR and ChIP-Seq experiments (0.17 µg/ml in TBST buffer containing 2.5% milk). Locations of Kaleidoscope Prestained molecular weight markers (Bio-Rad) are shown to the left of the gel.

(B) Western blot analysis of the phosphorylation state of RpaA. Closed, solid black arrows, phosphorylated RpaA (RpaA~P); open arrows, unphosphorylated RpaA. Equal masses of total protein from each lysate were separated on a Phos-tag gel and analyzed by Western blotting (Gutu and O’Shea, 2013).

(C) Correlation between HA-RpaA phosphorylation and HA-RpaA enrichment at the *kaiBC* promoter. RpaA phosphorylation was measured by Phos-tag Western blot (Gutu and O’Shea, 2013), while association with P*kaiBC* was measured by ChIP-qPCR performed with anti-HA antibody (see Extended Experimental Procedures). Subjective night is shaded in gray.

(D) Phosphorylation of RpaA used in DNase I footprinting reactions. RpaA was incubated DNase I footprinting reaction buffer for 1 hr at 30°C with *kaiBC* footprinting probe (end-labeled with non-radioactive ATP) in the presence of 1.5 µM recombinant CikA, 1 mM ATP, and/or 10 mM lithium potassium acetyl phosphate (Sigma Aldrich), as indicated. After incubation, 2.5 µl of each of these reactions was resolved on a Phos-tag SDS-PAGE gel and visualized by staining with Sypro Ruby as described in the Extended Experimental Procedures. The identity of the band corresponding to phosphorylated RpaA (RpaA~P) was revealed by boiling a reaction that was expected to contain RpaA~P for several minutes at 95°C to hydrolyze phosphoryl groups from RpaA (reaction is marked with ‘Boiled’). RpaA phosphorylation was observed only when CikA and ATP were present.

(E) DNase I footprinting of RpaA on *kaiBC* promoter mutants that produce diminished expression compared to the wild-type promoter (Kutsuna et al., 2005). *kaiBC* promoter mutants m15, m16, and m20 contain T to A, A to T, and/or G to C mutations at the positions indicated in the sequence next to the vertical bar, which marks the region of the wild-type promoter protected from digestion by high levels of RpaA~P. Sanger sequencing reactions used to identify the location of the footprint are shown on the left; footprinting reactions are shown on the right. All reactions contained 1.5 µM CikA and 1 mM ATP. Reactions marked with a ‘-’ did not contain RpaA, while reactions marked with a wedge contained 0.6, 3, or 6 µM RpaA as indicated by the thickness of the wedge. For m16, the reaction containing 0.6 µM was not run on the gel.

(F) Western blot analysis of the abundance of ectopically-expressed RpaA in the strains employed in Figure 3C. Strains were induced with indicated concentrations of IPTG, grown for 24 hr, entrained with two 12-h dark pulses, and then released to constant light. Samples were collected after 36 hr in constant light. Samples were lysed in denaturing buffer, run on a standard 4%–20% gradient SDS-PAGE gel, and analyzed by Western blotting. Under this gel condition, RpaA and RpaA~P co-migrate.

(G) Western blot analysis of the phosphorylation state of RpaA in the complementation strains employed in Figure 3C using Phos-tag Western blotting. Samples are the same as those described in Figure S3F. Where indicated (“heat”), samples were heated at 95°C for four minutes. Closed, solid black arrows, phosphorylated RpaA (RpaA~P); open arrows, unphosphorylated RpaA. Asterisk (*), heat-stable band of unknown identity, observed in all RpaA-expressing strains.

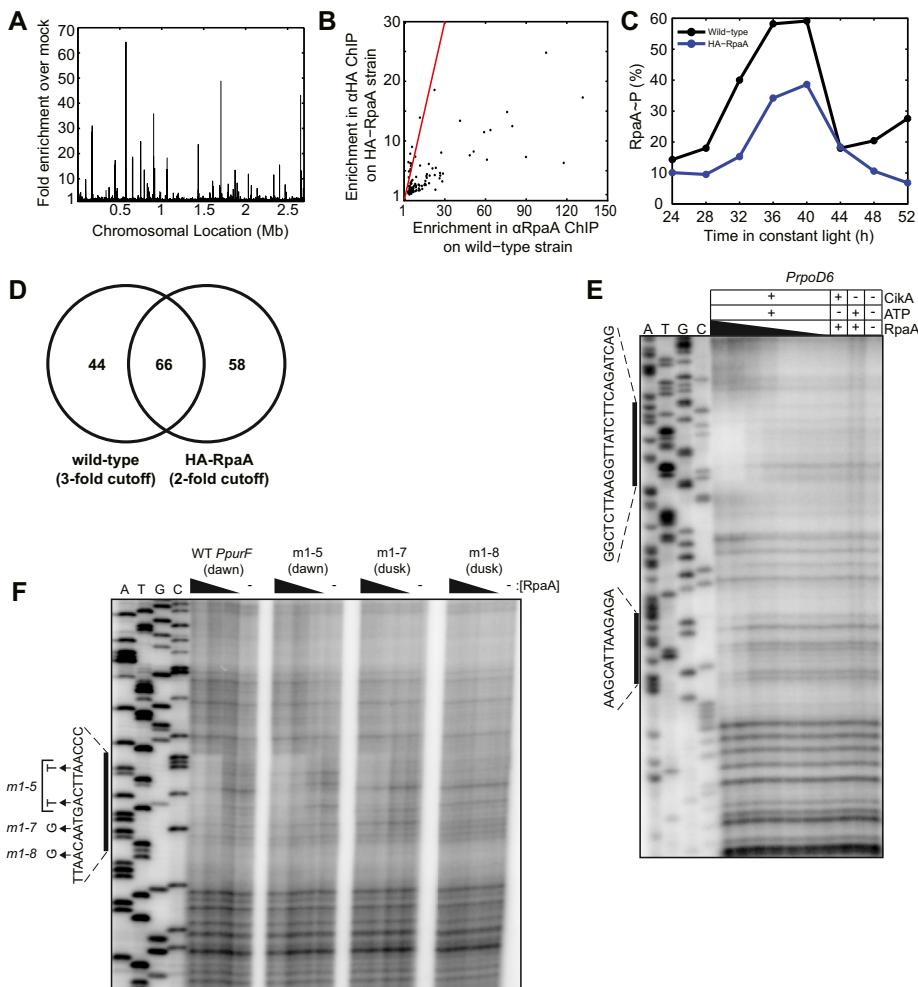


Figure S4. Identification of RpaA-Binding Sites by ChIP-Seq, Related to Figure 4

(A) Genome-wide binding profile of HA-RpaA measured by ChIP-Seq. The enrichment of read density in the HA-RpaA ChIP-Seq (anti-HA antibody with the HA-RpaA strain) at peak binding, relative to the mock ChIP-Seq (anti-HA antibody with the wild-type strain), is plotted as a function of position on the chromosome. Peak RpaA binding with the HA-RpaA strain occurred at subjective dusk (36 h), within the range observed for wild-type RpaA (32 - 40 hr, Figures 3A and 4).

(B) Comparison of ChIP-Seq enrichments in the wild-type and HA-RpaA strains at the time points of maximum RpaA binding in each strain. For each RpaA binding site identified in the wild-type strain (Figure 4; see Extended Experimental Procedures), we computed the maximal enrichment of read density in the RpaA ChIP-Seq (anti-RpaA antibody with the wild-type strain or anti-HA antibody with the HA-RpaA strain) relative to its respective mock ChIP-Seq (anti-RpaA antibody with the *rpaA* mutant strain or anti-HA antibody with the wild-type strain). The anti-RpaA ChIP-Seq experiment was carried out on a sample from the 32 hr time point (the time of maximum RpaA binding in the wild-type strain) while the anti-HA ChIP-Seq experiments was conducted on a sample from the 36 hr time point (the time of maximum HA-RpaA binding in the HA-RpaA strain). One data point, corresponding to the most strongly enriched binding site in both the anti-RpaA (411-fold) and anti-HA (31-fold) ChIP-Seqs, falls outside the bounds of the plot and is not shown. The red line denotes $y = x$, i.e., equal enrichments for wild-type and HA-RpaA.

(C) Comparison of phosphorylation levels of wild-type and HA-RpaA over 1.5 days. Wild-type and HA-RpaA strains were grown side-by-side, entrained with two 12-h dark pulses, and released to continuous light ($\sim 100 \mu\text{E m}^{-2} \text{s}^{-1}$) as described in Extended Experimental Procedures. RpaA phosphorylation was measured by Phos-tag Western blotting (Gutu and O'Shea, 2013).

(D) Venn diagram showing the overlap between RpaA binding sites identified in the wild-type strain using the anti-RpaA antibody (3-fold enrichment cutoff, Table S3A and Figure 4A) and those identified in the HA-RpaA strain using the anti-HA antibody (2-fold enrichment cutoff, Table S3B and Figure S4A). A lower threshold was used for calling binding sites in the HA-RpaA strain because of the weaker RpaA binding observed in that strain (Figure S4B); overlap with the wild-type strain increases for lower thresholds but with a concomitant increase in the number of non-overlapping false positives. The highest (least significant) Q-value for a peak obtained with the anti-RpaA antibody was 2.4×10^{-79} (Table S3A) and for the anti-HA antibody was 2.5×10^{-63} (Table S3B).

(E) In vitro DNase I footprinting of RpaA on the *PpoD6* promoter (*PpoD6*) as a function of recombinant RpaA phosphorylation and concentration. Two regions protected from digestion by high levels of RpaA~P are indicated by the vertical bars on the left. Sanger sequencing reactions used to identify the location of the footprint are shown on the left; footprinting reactions are shown on the right. RpaA pre-treatment and concentration are indicated above each footprinting lane. RpaA was added to a final concentration of 6.0, 3.0, 0.6, 0.3, 0.06, or 0.03 μM as indicated by the thickness of the wedge.

(F) Footprinting of RpaA~P on the *purF* promoter (*PpurF*). Footprinting was performed on the wild-type (WT) promoter and on mutants displaying altered phases of gene expression (Vijayan and O'Shea, 2013). Labeled as in Figure S3E with the phase of each promoter (subjective dawn or subjective dusk) indicated.

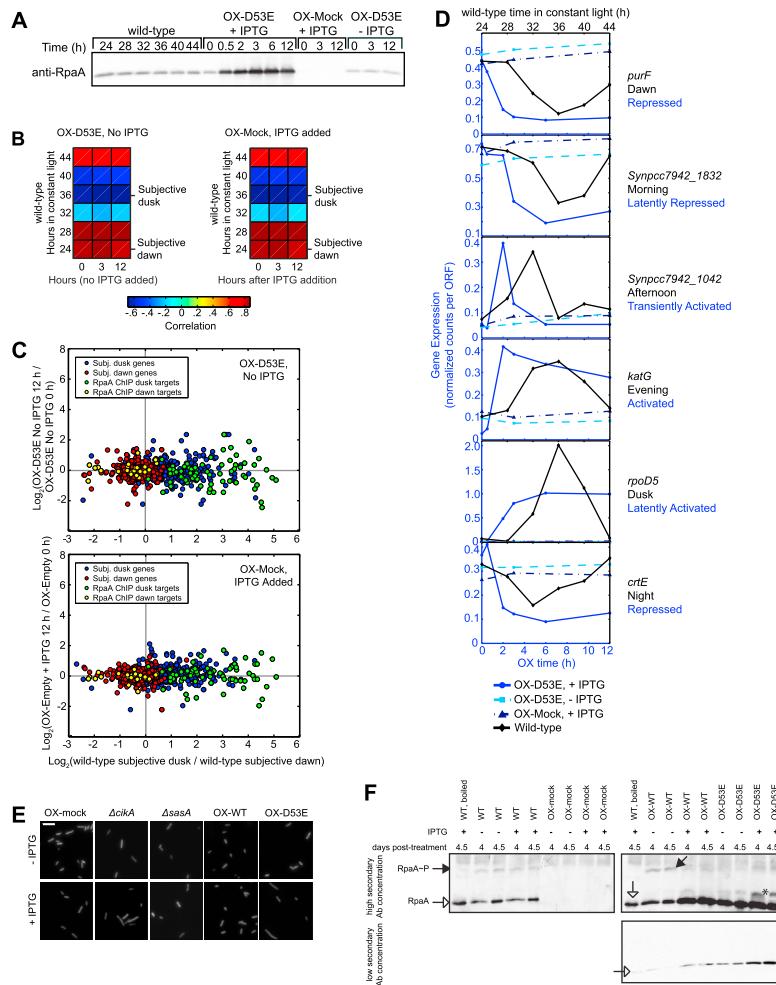


Figure S5. RpaA Orchestrates Global Circadian Gene Expression, Related to Figure 6

(A) RpaA protein abundance in samples used for RNA sequencing for Figure 6 and related experiments. We prepared whole protein lysates from samples acquired contemporaneously with those used for RNA-seq (Figure 6A-D and S5B-D) and probed for RpaA by Western blotting using the anti-RpaA antibody used for ChIP-seq. As a control, we added IPTG to a *ΔcikA* *ΔkaiBC* strain in which the *Ptrc* construct was inserted into the genome, but without *rpaA* inserted downstream of *Ptrc* (OX-Mock). We also carried out the same sampling procedure on OX-D53E with no IPTG added (OX-D53E - IPTG).

(B) Correlation between the expression of circadian genes ($n = 856$) in the wild-type strain over the course of one day and the expression of those genes in OX-D53E over time in the absence of IPTG addition (left), or those genes in OX-Mock before (time 0 h) and after IPTG addition (right). Gene expression was measured by RNA sequencing.

(C) Correlation between the change in expression of circadian genes ($n = 856$) caused by growing OX-D53E in the absence of IPTG for 12 hr (top, y axis) or incubating OX-Mock with IPTG for 12 hr (bottom, y axis) with the change in expression between subjective dusk and dawn in the wild-type strain (x axis).

(D) Gene expression time courses of representative genes assigned to corresponding clusters in the wild-type and OX-D53E experiments (Figure 6D). Normalized expression is shown on the y axis as described in the Extended Experimental Procedures. Time courses are shown in shades of blue for OX-D53E with IPTG (+ IPTG), OX-D53E without IPTG (- IPTG), or OX-Mock with IPTG (+ IPTG); blue traces correspond to the blue (bottom) x axis. Time courses of the same genes in entrained wild-type cells are shown in black; these traces correspond to the black (top) x axis. Beneath each gene name is shown the name of the wild-type cluster (black) and the OX-D53E cluster (blue) to which the gene was assigned by K-means clustering.

(E) Representative micrographs of cells from the cell length experiment analyzed in Figure 6E. Cells were imaged using their intrinsic red autofluorescence. Scale bar = 10 μ m.

(F) Western blot analysis of the phosphorylation state of RpaA in the strains employed in Figure 6E. Indicated strains were grown for four days in constant light (74 μ M IPTG). Western blots were performed with two different secondary antibody concentrations in order to capture the full range of RpaA abundances present. The top row shows blots probed with the higher secondary antibody concentration, while the bottom row shows a blot probed with the lower secondary antibody concentration; samples in the first column were analyzed with the higher secondary antibody concentration only. Closed, solid black arrows, phosphorylated RpaA (RpaA~P); open arrows, unphosphorylated RpaA. Asterisk (*), heat-stable band of unknown identity, observed in all RpaA-expressing strains. Note that several of the unphosphorylated RpaA bands in the overexpression strains are saturated in the top right image. When the same samples are analyzed with a lower secondary antibody concentration (bottom right image), the IPTG-dependent increase in unphosphorylated RpaA levels becomes more apparent. Also, the band of unknown identity (*) is not visible in this image, demonstrating that it is present at very low levels relative to unphosphorylated RpaA.