LETTER

doi:10.1038/nature11942

Non-optimal codon usage is a mechanism to achieve circadian clock conditionality

Yao Xu[1], Peijun Ma[1], Premal Shah[2], Antonis Rokas[1], Yi Liu[3] & Carl Hirschie Johnson[1]

Circadian rhythms are oscillations in biological processes that function as a key adaptation to the daily rhythms of most environments. In the model cyanobacterial circadian clock system, the core oscillator proteins are encoded by the gene cluster *kaiABC*[1]. Genes with high expression and functional importance, such as the *kai* genes, are usually encoded by optimal codons, yet the codon-usage bias of the *kaiBC* genes is not optimized for translational efficiency. We discovered a relationship between codon usage and a general property of circadian rhythms called conditionality; namely, that endogenous rhythmicity is robustly expressed under some environmental conditions but not others[2]. Despite the generality of circadian conditionality, however, its molecular basis is unknown for any system. Here we show that in the cyanobacterium *Synechococcus elongate*, non-optimal codon usage was selected as a post-transcriptional mechanism to switch between circadian and non-circadian regulation of gene expression as an adaptive response to environmental conditions. When the *kaiBC* sequence was experimentally optimized to enhance expression of the KaiB and KaiC proteins, intrinsic rhythmicity was enhanced at cool temperatures that are experienced by this organism in its natural habitat. However, fitness at those temperatures was highest in cells in which the endogenous rhythms were suppressed at cool temperatures as compared with cells exhibiting high-amplitude rhythmicity. These results indicate natural selection against circadian systems in cyanobacteria that are intrinsically robust at cool temperatures. Modulation of circadian amplitude is therefore crucial to its adaptive significance[3]. Moreover, these results show the direct effects of codon usage on a complex phenotype and organismal fitness. Our work also challenges the long-standing view of directional selection towards optimal codons[4–7], and provides a key example of natural selection against optimal codons to achieve adaptive responses to environmental changes.

Most amino acids are encoded by several codons, and species vary in their preferences for specific codons for the same amino acid. This preference, or codon-usage bias, is thought to reflect a balance between mutational biases and selection for translational efficiency and accuracy[4,6–9]. Although the relative importance of various factors affecting codon usage is debated[10,11], the degree of codon-usage bias is known to increase with higher gene expression[12]. In particular, genes under stronger selection for translational efficiency and/or accuracy show greater preference for using codons that have complementary transfer RNAs with higher abundances. Despite the fact that codon-usage bias is a strong indicator of selection on genes[13] and that the specific nature of codon bias can be changed by environmental factors such as temperature and hypersalinity[14,15], its direct effect on a complex phenotype and organismal fitness remains largely unknown.

In the cyanobacterium *S. elongatus* PCC 7942, circadian rhythms confer a strong selective advantage in rhythmic environments[3,16]. Expression of the entire *S. elongatus* genome is controlled by the circadian pacemaker, as shown by rhythms of promoter activity[17], messenger RNA abundance[18–20], and the topology of the entire chromosome[19,21]. The core circadian clock is composed of three components, KaiA, KaiB and KaiC, that are expressed as monocistronic *kaiA* and dicistronic *kaiBC* transcripts[1]. Both transcripts are expressed at very high levels, falling within the top 5% of mRNA abundances in *S. elongatus*[18,19]. Whereas initial examination of codon usage in *kaiBC* suggested an unusual codon bias (Supplementary Table 1 and Fig. 1a), further analyses indicated that amino acid-specific codon usage in *kaiB* and *kaiC* is not significantly different from that of the average codon usage in the rest of the genome (Fig. 1b). To quantify the degree of selection on the *kai* gene cluster, we focused on its codon adaptation index (CAI)[13] and the 5′ mRNA folding energies of its transcripts. Given that *kaiBC* transcripts are highly abundant, it was surprising that the CAI for *kaiB* and *kaiC* is less than the average CAI value of all *S. elongatus* genes (Fig. 1b). Moreover, even though the folding energy of the *kaiB* transcript is less negative than that of most transcripts (Fig. 1c)—indicating that it has a relatively weaker secondary structure than most mRNAs in this organism and is therefore likely to initiate translation efficiently—the weak secondary structure of the *kaiB* transcript is insufficient to account for its high mRNA abundance given that the CAI is considerably below average. The *kaiC* portion of the *kaiBC* transcript has similar CAI (Fig. 1b) and folding energy (Fig. 1c) values to those of the *kaiB* portion.

To test whether the lower CAI of *kaiBC* might be adaptive in the core clock mechanism and/or in the output pathways controlled by the pacemaker in cyanobacteria, we generated two strains in which the endogenous *kaiBC* gene was replaced with modified versions of *kaiBC* where codon usage was 'optimized' to be similar to that of highly expressed genes[18,19], thereby increasing their CAI values from 0.60 to 0.95 for *kaiB* and 0.61 to 0.82 for *kaiC* (Fig. 1a and Supplementary Tables 2 and 3). Changing the codon usage of a gene in its 5′ region affects its folding energy and hence its rate of translation initiation[8,22–24]. Indeed, the optimized versions of *kaiB* and *kaiBC* also had significantly lower 5′ folding energy than the *kaiBC* transcript (Fig. 1c, d). Because the *kaiBC* gene is transcribed as a single discistronic mRNA with *kaiB* at its 5′ end[1], one optimized strain (optKaiB) replaced the entire wild-type $kaiB^{WT}$ gene with an optimized $kaiB^{opt}$ gene, whereas the other optimized strain (optKaiBC) replaced both the $kaiB^{WT}$ gene and the 5′ half of the $kaiC^{WT}$ gene with optimized versions ($kaiB^{opt}$ and $kaiC^{opt}$; Fig. 2a). Our initial hypothesis was that the 'non-optimal' codon bias of the endogenous *kaiBC* gene is essential for the expression of circadian rhythmicity, which would predict that the intrinsic rhythmicity of the optKaiB and optKaiBC strains would be poorer than that of wild type (as is true for the studies of the *Neurospora* FRQ protein[25]). Unexpectedly, the observed circadian rhythms of gene expression in the optKaiB and optKaiBC strains were as robust as those of the wild-type *S. elongatus* at the optimal growth temperature of 30 °C (Fig. 2b). However, to be adaptive, circadian clocks must be able to keep time accurately over the range of physiological temperatures for a given organism[26], so we tested real-time gene expression profiles of the optimized strains relative to wild type at temperatures from 18 °C to 38 °C. Again, to our surprise the optKaiB and

[1]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235, USA. [2]Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [3]Department of Physiology, The University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

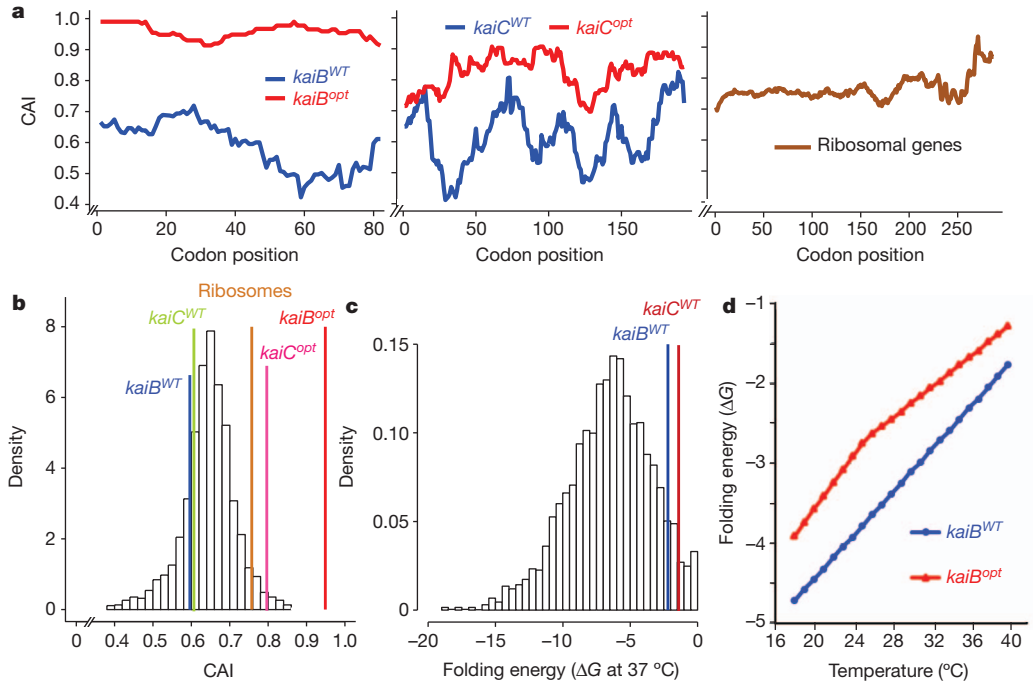

Figure 1 | Non-optimal codon usage of cyanobacterial clock genes. **a**, CAI values along the entire length of different genes at a sliding window of 20 codons. Left, $kaiB^{WT}$ versus $kaiB^{opt}$. Middle, KaiC-I domain (encoded by the 5′ half of the *kaiC* gene) of $kaiC^{WT}$ versus $kaiC^{opt}$. Right, average of all ribosomal protein genes. **b**, Comparison of CAI values of the $kaiB^{WT}$, $kaiB^{opt}$, $kaiC^{WT}$, $kaiC^{opt}$ and ribosomal genes in the CAI histogram distribution of the genome. **c**, Calculated 5′ folding energy of the mRNA for all genes in the *S. elongatus* genome. **d**, Comparison of calculated minimum free energy of folding (ΔG) over a range of temperatures between $kaiB^{WT}$ and $kaiB^{opt}$.

optKaiBC strains exhibited robustly rhythmic gene expression over a broad range of temperatures (Fig. 2b and Supplementary Table 4). By contrast, the rhythm of the wild-type strain damps within a few cycles at cool temperatures (18–23 °C; Fig. 2b and Supplementary Table 4). There are no significant differences in the free-running period of the circadian rhythms between wild-type and optKaiB/optKaiBC at 18–20 °C and 26–32 °C, and only small differences at other temperatures in the 18–38 °C range (Fig. 2c). Overall, the temperature compensation of the optimized strains was slightly poorer ($Q_{10} \approx 1.13$) over the range of 23–36 °C than that for wild type ($Q_{10} \approx 1.10$ from ref. 27, and $Q_{10} \approx 1.04$ from Fig. 2c and Supplementary Table 6), which might relate to the difference in mRNA folding energies (Fig. 1d).

We further confirmed that improved rhythmicity at lower temperatures of the optKaiB and optKaiBC strains (Fig. 2b) was due to higher protein production levels of KaiB and KaiC. First, we confirmed that KaiC levels are considerably higher in the optKaiBC strain than in the wild type (Fig. 3a and Supplementary Fig. 1), and showed that KaiB levels are similarly increased in both the optKaiB and optKaiBC strains (Fig. 3b and Supplementary Fig. 2). Interestingly, the amplitude of the KaiC phosphorylation rhythm is comparable between wild type and optKaiBC at 20 °C (Fig. 3a). Second, to validate that the augmented KaiB and/or KaiC levels were responsible for the improved rhythmicity at cool temperatures, we co-expressed the native (non-optimized) sequences of *kaiB* (strain $KaiB^{WT}$/OX) or *kaiBC* (strain $KaiBC^{WT}$/OX) from an isopropylthiogalactoside (IPTG)-inducible *trc* promoter to enhance endogenous production of KaiB and KaiC (Fig. 3c). Figure 3d shows that KaiB is overexpressed in the $KaiB^{WT}$/OX and

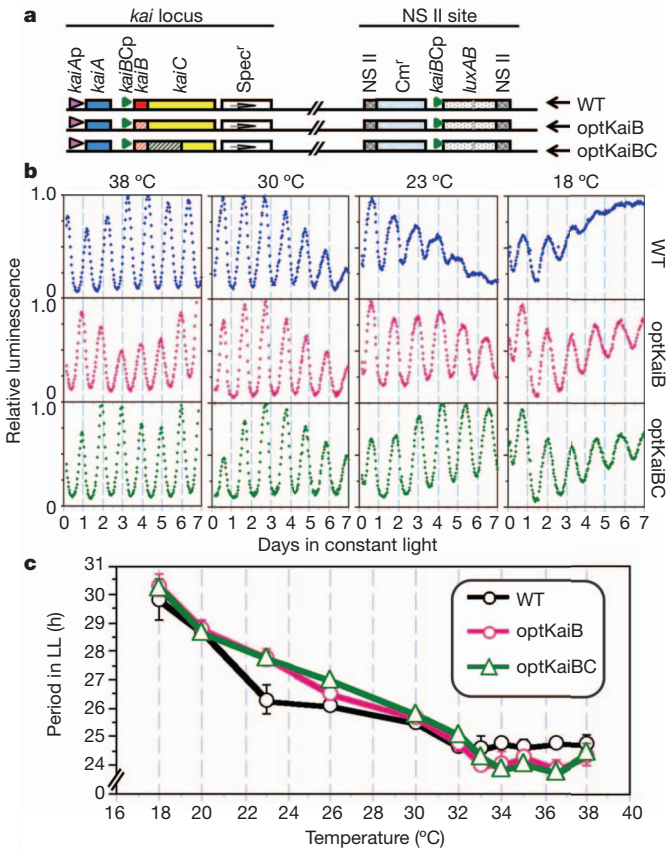

Figure 2 | Conditional circadian phenotypes of the *kai*-optimized strains. **a**, Diagrams of genes in the wild-type *kaiABC* (WT), *kaiB*-optimized (optKaiB) and *kaiBC*-optimized (optKaiBC) strains (see Methods for a detailed description). Cm^r, chloramphenicol resistance; *kaiBC*p, *kaiBC* promoter; *luxAB*, reporter cassette containing the *luxA* and *luxB* genes; NS II, neutral site II; $Spec^r$, spectinomycin resistance. **b**, Luminescence rhythms of wild-type and *kai*-optimized strains in constant light at the indicated temperatures. *In vivo* luminescence rhythms were monitored from a group of 12 colonies for each strain, and a representative example is shown for each group. **c**, Free-running periods of luminescence rhythms in constant light from different strains over a temperature range of 18–38 °C. Data are mean ± s.e.m.

KaiBC^{WT}/OX strains, even in the absence of the IPTG inducer (owing to the *trc* promoter being slightly leaky[28]), whereas KaiC levels are not altered. In the presence of a very low concentration of the inducer (5 μM IPTG), KaiB levels are enhanced in KaiB^{WT}/OX, and both KaiB and KaiC levels are increased in KaiBC^{WT}/OX (Fig. 3d). KaiA levels are not substantially affected under any of these conditions. In rhythm assays at 18 °C, optKaiB and optKaiBC exhibit improved rhythmicities compared with wild type, as noted above, but so does the KaiBC^{WT}/OX strain in the absence and presence of the inducer IPTG (Fig. 3e and Supplementary Table 5). (Much higher concentrations of IPTG lead to arhythmicity as noted previously[1,28]; see Supplementary Fig. 3.) Moreover, although the data depicted in Fig. 3e are normalized luminescence data, unnormalized data for an equivalent experiment at 20 °C are shown in Supplementary Fig. 4. Conversely, the KaiB^{WT}/OX strain rhythm damps rapidly. There is a clear correlation between the strains that exhibit sustained rhythmicity in constant conditions (an indicator of the endogenous circadian system) and a favourable KaiB abundance relative to KaiC abundance (optKaiB, optKaiBC and KaiBC^{WT}/OX) versus the strains in which rhythmicity damps rapidly and that express either a low (wild type) or a high (KaiB^{WT}/OX) KaiB level relative to the KaiC level (Fig. 3d, e). These experimental manipulations of KaiB and KaiC levels using non-optimized sequences strongly suggest that codon optimization of *kaiB* and *kaiC* affects the robustness of rhythmicity at cool temperatures primarily by affecting KaiB and KaiC expression, thereby altering the relative levels of KaiB, KaiC and KaiA protein abundances in the cells. The stoichiometry among the Kai proteins is known to be crucial for expression of the cyanobacterial *in vitro* oscillator[29], and it is therefore likely to be a key determining factor in the expression of rhythmicity *in vivo* that we report here.

If alternative (that is, 'optimal') *kaiBC* sequences promote rhythmicity at cooler temperatures, why have not they been naturally selected? Although 30 °C is the optimal growth temperature for *S. elongatus* (Supplementary Fig. 5), 18–23 °C is certainly a temperature range that this freshwater, temperate cyanobacterium could experience in its environment. We therefore tested the growth rates of *S. elongatus* in 12-h light/12-h dark cycles at 37, 34, 30, 25, 20 and 18 °C (Fig. 4 and Supplementary Fig. 5). Consistent with our previous results using competition assays between wild type and the arhythmic strain CLAb, as well as the highly damped strain CLAc[3] (both CLAb and CLAc result from point mutations in the *kaiC* gene[1]), the wild-type strain grew at a faster rate in light–dark at 30 °C than the CLAb or CLAc strains (Fig. 4b and Supplementary Table 7). The optKaiBC strain grew at about the same rate as wild type or perhaps slightly faster. At cooler temperatures, however, the results were markedly different. At 18 °C and 20 °C when the circadian rhythm of gene expression is damped in wild type (Figs 2b, 3e and Supplementary Fig. 4), the wild-type strain grew significantly faster than optKaiBC

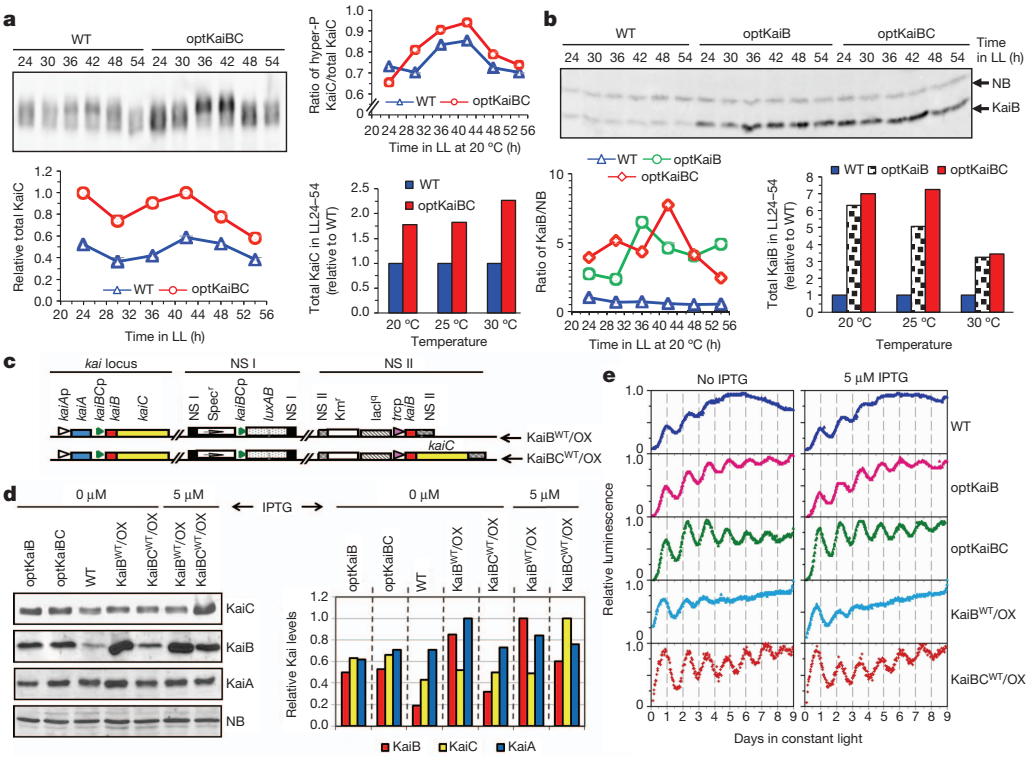

Figure 3 | Kai protein expression and circadian regulation of cells expressing wild-type or optimized versions of *kaiBC*. **a**, KaiC immunoblots in wild-type and optKaiBC strains in constant light (LL) at 20 °C. Densitometry reveals the KaiC phosphorylation profiles (right; hyper-P, hyperphosphorylated); the KaiC abundance patterns (bottom left, relative to the maximum density band, optKaiBC at 42 h constant light); and the total KaiC abundance at three temperatures (20, 25 and 30 °C; bottom right and Supplementary Fig. 1) in the optKaiBC strain from 24–54 h constant light relative to wild type. **b**, KaiB immunoblots in wild-type, optKaiB and optKaiBC strains in constant light at 20 °C. A nonspecific band (NB) was used as an internal reference for KaiB density. Densitometry reveals relative KaiB abundance (ratio of KaiB density to the nonspecific band, bottom left) and total KaiB abundance at three temperatures (20, 25 and 30 °C; bottom right and Supplementary Fig. 2) in the optKaiB and optKaiBC strains from 24– 54 h constant light relative to wild type. **c**, Diagrams of the *kaiB*^{WT}- or *kaiBC*^{WT}-co-expressing strains. Km^r, kanamycin resistance; *trcp*, *trc* promoter. **d**, Increase of KaiB and/or KaiC levels in *kaiB*^{WT}- or *kaiBC*^{WT}-co-expressing strains (KaiB^{WT}/OX and KaiBC^{WT}/OX) at 12 h constant light with/without 5 μM IPTG. Left panel shows immunoblot assays for KaiB, KaiC and KaiA; equal loading was confirmed by the nonspecific band. Right panel depicts the densitometry of relative KaiA, KaiB and KaiC protein abundances. **e**, Phenocopying of the cool-temperature rhythmicity of the *kaiB*- or *kaiBC*-optimized strains in the wild-type strain by increased expression of the *kaiB*^{WT} and *kaiC*^{WT} genes. Luminescence was recorded in constant light at 18 °C from cultures of wild-type, codon-optimized (optKaiB and optKaiBC), and *kai*-co-expressing (KaiB^{WT}/OX and KaiBC^{WT}/OX) strains in the presence or absence of IPTG (5 μM). Representative traces are shown for each case.

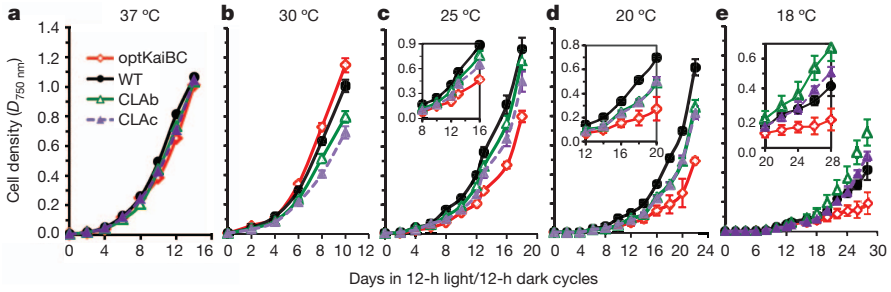

Figure 4 | Optimizing the *kaiBC* sequence causes slower growth rate at cool temperatures. **a**–**e**, Wild-type, optKaiBC, CLAb (arhythmic) and CLAc (damped oscillation) strains were grown in 12-h light/12-h dark cycles at 37 °C (**a**), 30 °C (**b**), 25 °C (**c**), 20 °C (**d**) or 18 °C (**e**), with constant air bubbling and shaking. Cell densities were monitored by measuring attenuance at 750 nm ($D_{750\,nm}$) every 2 days. Data are mean ± s.e.m. from 2–6 independent experiments for each strain and condition. For a better comparison at 18, 20 and 25 °C, the insets are a magnified portion of the specified times. (For doubling time calculations, see Supplementary Table 7 and Supplementary Fig. 5.)

(Fig. 4d, e and Supplementary Table 7). Even the arhythmic CLAb and damped CLAc strains outperformed optKaiBC at 18–25 °C (Fig. 4c–e and Supplementary Table 7). Therefore, at cool temperatures in light–dark cycles, strains having damped (wild type and CLAc) or arhythmic (CLAb) phenotypes under free-running conditions outgrew the strain that expressed robust rhythms (optKaiBC).

Biases in codon usage are generally thought to be under directional selection for an optimal balance between translational efficiency and accuracy, in which a higher CAI is always better[5,7]. Other examples of selection for non-optimal codon usage (mediated by mRNA secondary structure and/or tRNA availability) are rare and poorly characterized[8,30]. Our study presents a counter-example to the standard view, and suggests the action of either selection against optimal codon usage, or stabilizing selection in which both low and high extremes in codon-usage bias have higher fitness costs. Indeed, the non-optimal codon usage seems to be a molecular mechanism in which post-transcriptional events allow *S. elongatus* cells to switch between circadian and non-circadian regulation of gene expression depending on the environmental conditions, and our results are reminiscent of earlier observations of such 'conditionality' for circadian rhythms[2,31].

Why is conditionality an important characteristic of circadian rhythms? We tend to think of circadian clocks as rhythmic activators, but they are also rhythmic repressors. Under some conditions this repressor/activator balance may inhibit growth and circadian regulation may not be adaptive. At cooler temperatures, the free-running period of the circadian rhythm of *S. elongatus* can be as long as 30 h (Fig. 2c), suggesting that the coupling of gene expression and rhythmic regulation might be maladaptive for growth at lower temperatures. For example, we previously reported that *S. elongatus* mutant strains with 30 h period lengths can entrain to 24 h light–dark cycles, but they do so with a considerably later phase relationship that is maladaptive[16]; a result that fits with our current observations. As another example of conditionality, our previous competition experiments at 30 °C found that the arhythmic CLAb strain is more fit than wild type under constant illumination, but is rapidly outcompeted by wild type in light–dark cycles[3] (Fig. 4 shows that the growth of CLAb is poorer than wild type in light–dark cycles at 20–30 °C). Therefore, both illumination and temperature are environmental parameters that demonstrate the 'conditional' advantages of circadian regulation under some conditions but not others.

Our data show that optimizing the codon usage for the circadian *kaiBC* genes of *S. elongatus* does not disrupt endogenous circadian regulation (as in the case of the *Neurospora* FRQ protein[25]), but instead the sequence optimization enhances circadian regulation in a range of cool temperatures that are relevant for the ecology of this organism. Despite this enhancement of intrinsic rhythmicity, however, optimal codon usage at the *kai* locus impairs cell growth at cooler temperatures. Therefore, our data suggest selection against an optimal codon usage because it is incompatible with a post-transcriptionally modulated conditional suppression of circadian rhythmicity at cool temperatures. In *S. elongatus*, circadian conditionality allows expression of robust endogenous rhythmicity in the range of temperatures that permit vigorous growth and suppress rhythmicity at temperatures in which growth of this species is minimal (Supplementary Fig. 5). Our observations provide a new example of post-transcriptional regulation of circadian clock genes that confers an adaptive response to different environmental conditions.

METHODS SUMMARY

Frequencies of codon usage were analysed from all putative proteins in the whole genome of *Synechococcus elongatus* PCC 7942, all 59 ribosomal genes in the genome (Supplementary Table 2), and 16 highly expressed genes that show high microarray values from both data sets (Supplementary Table 3). The fractions of codon usage were based on usage frequencies per 1,000 codons (Supplementary Table 1). The overall relative synonymous codon usage (RSCU) values of the genome were calculated from 3,261 coding sequences from the genome (Supplementary Table 8), and infrequently used codons in *kaiB* or *kaiC* were changed to those coding for the specific amino acids with higher RSCU values (Supplementary Tables 9 and 10) or higher codon-usage fractions (Supplementary Table 1) in the genome. DNA fragments containing optimized *kaiB* or *kaiC* coding sequences (*optKaiB* or *optKaiC*) were synthesized and cloned into a plasmid containing the wild-type *kai* cluster DNA to replace the corresponding wild-type DNA sequences. The CAI values of all genes in the genome and the codon-optimized versions of the *kai* genes were calculated based on RSCU values.

Full Methods and any associated references are available in the online version of the paper.

Received 29 May 2012; accepted 30 January 2013.
Published online 17 February 2013.

1. Ishiura, M. *et al.* Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria. *Science* **281,** 1519–1523 (1998).
2. Njus, D., McMurry, L. & Hastings, J. W. Conditionality of circadian rhythmicity: synergistic action of light and temperature. *J. Comp. Physiol.* **117,** 335–344 (1977).
3. Woelfle, M. A., Ouyang, Y., Phanvijhitsiri, K. & Johnson, C. H. The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr. Biol.* **14,** 1481–1486 (2004).
4. Drummond, D. A. & Wilke, C. O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* **134,** 341–352 (2008).
5. Bulmer, M. The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129,** 897–907 (1991).
6. Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Rev. Genet.* **12,** 32–42 (2011).
7. Shah, P. & Gilchrist, M. A. Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proc. Natl Acad. Sci. USA* **108,** 10231–10236 (2011).
8. Eyre-Walker, A. & Bulmer, M. Reduced synonymous substitution rate at the start of enterobacterial genes. *Nucleic Acids Res.* **21,** 4599–4603 (1993).
9. Akashi, H. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* **136,** 927–935 (1994).
10. Shah, P. & Gilchrist, M. A. Effect of correlated tRNA abundances on translation errors and evolution of codon usage bias. *PLoS Genet.* **6,** e1001128 (2010).
11. Zhou, T., Weems, M. & Wilke, C. O. Translationally optimal codons associate with structurally sensitive sites in proteins. *Mol. Biol. Evol.* **26,** 1571–1580 (2009).

12. Ikemura, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* **151,** 389–409 (1981).
13. Sharp, P. M. & Li, W. H. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15,** 1281–1295 (1987).
14. Lynn, D. J., Singer, G. A. & Hickey, D. A. Synonymous codon usage is subject to selection in thermophilic bacteria. *Nucleic Acids Res.* **30,** 4272–4277 (2002).
15. Paul, S., Bag, S. K., Das, S., Harvill, E. T. & Dutta, C. Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol.* **9,** R70 (2008).
16. Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S. & Johnson, C. H. Resonating circadian clocks enhance fitness in cyanobacteria. *Proc. Natl Acad. Sci. USA* **95,** 8660–8664 (1998).
17. Liu, Y. *et al.* Circadian orchestration of gene expression in cyanobacteria. *Genes Dev.* **9,** 1469–1478 (1995).
18. Ito, H. *et al.* Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc. Natl Acad. Sci. USA* **106,** 14168–14173 (2009).
19. Vijayan, V., Zuzow, R. & O'Shea, E. K. Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc. Natl Acad. Sci. USA* **106,** 22564–22568 (2009).
20. Vijayan, V., Jain, I. H. & O'Shea, E. K. A high resolution map of a cyanobacterial transcriptome. *Genome Biol.* **12,** R47 (2011).
21. Woelfle, M. A., Xu, Y., Qin, X. & Johnson, C. H. Circadian rhythms of superhelical status of DNA in cyanobacteria. *Proc. Natl Acad. Sci. USA* **104,** 18819–18824 (2007).
22. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324,** 255–258 (2009).
23. Tuller, T., Waldman, Y. Y., Kupiec, M. & Ruppin, E. Translation efficiency is determined by both codon bias and folding energy. *Proc. Natl Acad. Sci. USA* **107,** 3645–3650 (2010).
24. Gu, W., Zhou, T. & Wilke, C. O. A universal trend of reduced mRNA stability near translation-initiation site in prokaryotes and eukaryotes. *PLOS Comput. Biol.* **6,** e1000664 (2010).
25. Zhou, M. *et al.* Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature* http://dx.doi.org/10.1038/nature11833 (this issue).
26. Pittendrigh, C. S. On temperature independence in the clock system controlling emergence time in *Drosophila*. *Proc. Natl Acad. Sci. USA* **40,** 1018–1029 (1954).
27. Kondo, T. *et al.* Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. *Proc. Natl Acad. Sci. USA* **90,** 5672–5676 (1993).
28. Xu, Y., Mori, T. & Johnson, C. H. Cyanobacterial circadian clockwork: roles of KaiA, KaiB and the *kaiBC* promoter in regulating KaiC. *EMBO J.* **22,** 2117–2126 (2003).
29. Nakajima, M., Ito, H. & Kondo, T. *In vitro* regulation of circadian phosphorylation rhythm of cyanobacterial clock protein KaiC by KaiA and KaiB. *FEBS Lett.* **584,** 898–902 (2010).
30. Konigsberg, W. & Godson, G. N. Evidence for use of rare codons in the *dna*G gene and other regulatory genes of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **80,** 687–691 (1983).
31. Liu, Y., Garceau, N. Y., Loros, J. J. & Dunlap, J. C. Thermally regulated translational control of FRQ mediates aspects of temperature responses in the *Neurospora* circadian clock. *Cell* **89,** 477–486 (1997).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful for the suggestions of M. Woelfle and the technical assistance of D. Zelli and C. Chintanaphol. This research was supported by grants from the National Institute of General Medical Science to C.H.J. (R01 GM067152 and R01 GM088595) and to Y.L. (GM068496 and GM062591), the Welch Foundation (I-1560) to Y.L., and the National Science Foundation (DEB-0844968) and the Searle Scholars Program to A.R. P.S. acknowledges support from a Burroughs Wellcome Fund Career Award and a David & Lucille Packard Foundation Fellowship awarded to Joshua B. Plotkin.

Author Contributions Y.X. and P.M. collected data; Y.X., P.M. and Y.L. analysed the experimental data; Y.X., P.S. and A.R. analysed the bioinformatic data; Y.L. and C.H.J. designed the original conceptual basis for the study; Y.X. and C.H.J. designed the experimental bases for the study; Y.X., P.S. and C.H.J. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.H.J. (carl.h.johnson@vanderbilt.edu).

METHODS

Evaluation of codon usage. To evaluate the frequency of codon usage of the central clock genes *kaiABC*, we analysed coding sequences from different groups. One group is from 2,400,255 residue sequences from all putative proteins in the whole genome of *Synechococcus elongatus* PCC 7942. The second group is from all 59 ribosomal genes in the genome (Supplementary Table 2). We also analysed microarray data sets from two independent laboratories[18,19]. First, we calculated total microarray signals from one circadian cycle in constant light for each of these genes. Then, all of these genes were rearranged from strongest to weakest mRNA abundances based on the total microarray signal values in constant light. Finally, we selected the top 16 genes that show high microarray values from both data sets (Supplementary Table 3), and combined all of these coding sequences to the third group as putative highly expressed genes. The codon usage frequency was analysed with a web-based program from the Sequence Manipulation Suite of Bioinformatics.org (http://www.bioinformatics.org). The fractions of codon usage were based on usage frequencies per 1,000 codons (Supplementary Table 1).

Codon optimization of *kai* genes. Relative synonymous codon usage (RSCU) is defined as the ratio of the observed frequency of codons to the expected frequency, given that all the synonymous codons for the same amino acids are used equally[13,32]. In the absence of any codon-usage bias, the RSCU value would be 1.00. A codon that is used less frequently than expected will have a value of less than 1.00, and vice versa for a codon used more frequently than expected. The overall RSCU values of *S. elongatus* PCC 7942 were calculated from 3,261 coding sequences (990,021 codons) from the genome (Supplementary Table 8). Infrequently used codons in *kaiB* or *kaiC* coding sequences were changed to those coding for the specific amino acids with higher RSCU values (Supplementary Tables 9 and 10) or higher codon-usage fractions (Supplementary Table 1) in the genome. As shown in Supplementary Fig. 6 and Supplementary Table 9, 67 out of 102 codons were optimized for the entire *kaiB* coding region, whereas for the *kaiC* gene, the infrequently used codons were mainly optimized in the amino-terminal KaiC-I domain that is encoded by the 5′ half of the *kaiC* gene (Supplementary Fig. 7 and Supplementary Table 10).

Synthesis and construction of optimized *kai* genes. DNA fragments containing optimized *kaiB* or *kaiC* coding sequences with wild-type flanking sequences were commercially synthesized and cloned into the SmaI site of pUC57 (EZBiolab) to produce pUCoptKaiB or pUCoptKaiC, respectively. On the basis of pUCoptKaiB or pUCoptKaiC templates, the *optKaiB* or *optKaiC* fragments were resynthesized using 12–18 thermal cycles with pfuUltra High-fidelity DNA polymerase (Stratagene) and primers containing the corresponding wild-type flanking sequences. After purification, the *optKaiB* or *optKaiC* fragments were EZcloned (Stratagene) into a plasmid containing the wild-type *kai* cluster DNA to replace the corresponding wild-type DNA sequences. After the unchanged parental plasmid was digested at 37 °C for 1 h with Dpn I (New England Biolabs), the circular, nicked optimized double-stranded DNA was transformed into *Escherichia coli* to generate pKai-optKaiB and pKai-optKaiC, respectively. To optimize the *kaiB* and *kaiC* genes, the resynthesized *optKaiB* fragment was EZcloned into the pKai-optKaiC plasmid to replace the corresponding wild-type *kaiB* coding sequences and produce pKai-optKaiBC. All of these *kai*-optimized constructs were confirmed by DNA sequencing analysis. In case of direct comparison, the wild-type versions of the *kaiB*, *kaiC* and *kaiBC* genes or coding sequences are indicated as $kaiB^{WT}$, $kaiC^{WT}$ and $kaiBC^{WT}$, whereas the codon-optimized versions are denoted as $kaiB^{opt}$, $kaiC^{opt}$ and $kaiBC^{opt}$, respectively.

Computation of CAI and 5′ mRNA folding energy. To calculate CAI values of genes, we began by calculating the RSCU of all ribosomal genes[13] (Supplementary Table 2). Using these RSCU values, we calculated the CAI values of all genes in the *S. elongatus* genome and the codon-optimized versions of the *kai* genes. Using a sliding window of 20 codons, Fig. 1a shows that optimized versions of both *kaiB* ($kaiB^{opt}$) and *kaiC* ($kaiC^{opt}$) genes have higher CAI values along the entire length of the genes than the wild-type versions, and higher than the average CAI of ribosomal genes. The CAI of the wild-type *kaiB* and *kaiC* genes is less than the average CAI of the genome, in spite of their high abundance in the transcriptome. As expected, the average CAI of ribosomal genes is much higher than the average. As an example, the CAI of $kaiB^{opt}$ was optimized to 0.95 as shown by the red line in Fig. 1b using the RSCU of ribosomal genes.

In addition to codon usage as it relates to the relative expression of various tRNA genes, the translational efficiency of the production of a protein depends on the 5′ folding energy of its mRNA. To calculate the minimum free energy (ΔG) of folding of the 5′ region of mRNAs, we used the first 40 nucleotides of the coding sequences (1–40 nucleotides). In addition, we used the RNA folding algorithm in RNA fold of the Vienna RNA package 2.0 (ref. 33). We calculated ΔG at 37 °C for each gene in the *S. elongatus* genome using default parameters. Then, to calculate the effect of temperature on folding energy for the *kai* genes (wild-type and optimized versions), we varied the temperature from 18 to 40 °C. The folding energy of the $kaiB^{WT}$ and $kaiC^{WT}$ genes is much higher than the rest of the genome, indicating selection for faster translation initiation (Fig. 1c). Because the folding energy of an mRNA depends on the temperature, calculation of the minimum free energy of folding (ΔG) at varying temperatures showed that ΔG for both $kaiB^{WT}$ and $kaiB^{opt}$ increases with temperature, and that optimizing codon usage also increases the ΔG of the $kaiB^{opt}$ gene at all temperatures (Fig. 1d).

Generation of *kai*-optimized and *kai*-co-expressing cyanobacterial strains. *S. elongatus* PCC 7942 was used as the cyanobacterial host strain. A kaiBCp::luxAB luminescence reporter of clock-controlled gene expression was integrated into either neutral site I (NS I) with a spectinomycin selection marker (or in other experiments, into NS II with a chloramphenicol selection marker), in which the expression of the *Vibrio harveyi* luciferase structure gene cassette *luxAB* is driven by the promoter of the *kaiBC* genes (*kaiBC*p) and serves as a real-time reporter of promoter activity[1,27,28]. The *kaiABC*-null strain was created by replacement of the *kaiABC* DNA region with a kanamycin-resistance gene[1]. The wild-type *kaiABC* cluster or codon-optimized *kaiABC* cluster containing a $kaiB^{opt}$, $kaiC^{opt}$ or $kaiBC^{opt}$ coding region and a spectinomycin-resistance cassette was reintroduced into the endogenous *kai* locus by replacing the kanamycin-resistance gene of the *kaiABC*-null strain to give rise to transgenic strains with either the wild-type or the optimized version of the *kaiABC* cluster. This replacement was done in such a way that the *kaiABC* cluster was recreated to be exactly the same as the wild-type cluster with no insertions or deletions except a selection marker downstream of the genomic *kaiC* sequence. The transgenic cyanobacterial strain containing the wild-type version of the *kai* cluster was named the wild-type strain, whereas the strains containing codon-optimized versions of the *kai* cluster with $kaiB^{opt}$, $kaiC^{opt}$ or $kaiBC^{opt}$ were called optKaiB, optKaiC or optKaiBC, respectively (Fig. 2a). For co-expression of wild-type *kaiB*, *kaiC* or *kaiBC* genes *in vivo*, the trcp::kaiB^{WT}, trcp::kaiC^{WT} or trcp::kaiBC^{WT} constructs with a kanamycin-resistance marker[28] were transformed into the NS II region of a wild-type luminescence reporter strain to generate *kai*-co-expressing strains KaiB^{WT}/OX, KaiC^{WT}/OX or KaiBC^{WT}/OX (Fig. 3c), in which the co-expression of wild-type *kaiB*, *kaiC* or *kaiBC* genes from NS II was under the control of an IPTG-derepressible heterologous *trc* promoter, and the original *kaiABC* cluster remains at the original wild-type site. These codon-optimized strains containing $kaiB^{opt}$, $kaiC^{opt}$ or $kaiBC^{opt}$ and *kai*-co-expressing strains were confirmed by PCR, DNA sequencing, immunobloting and luminescence analyses. As indicated in Fig. 2a, the wild-type coding region of the *kaiB* gene was replaced with $kaiB^{opt}$ coding sequences in the optKaiB strain, whereas the *kaiBC* coding region was replaced with both the $kaiB^{opt}$ and $kaiC^{opt}$ (that is, $kaiBC^{opt}$) coding sequences in the optKaiBC strain. At NS II, a *kaiBC* promoter-driving *luxAB* expression cassette with a chloramphenicol selection marker was integrated to function as a luminescence reporter of clock-controlled promoter activity. In the $kaiB^{WT}$- or $kaiBC^{WT}$-co-expressing strains in Fig. 3c, three sites in the chromosome are depicted: (1) the wild-type *kaiABC* cluster ($kaiABC^{WT}$); (2) the *luxAB* luminescence reporters in NS I; and (3) an IPTG-derepressible promoter driving expression of $kaiB^{WT}$ (KaiB^{WT}/OX strain) or $kaiBC^{WT}$ (KaiBC^{WT}/OX strain) in NS II.

Measurement of luminescence rhythms. Cyanobacterial strains were grown in modified BG11 (ref. 34) liquid media with air bubbling or BG11 agar plates supplemented with appropriate antibiotics (20 μg ml^{-1} spectinomycin; 10 μg ml^{-1} kanamycin; 7.5 μg ml^{-1} chloramphenicol) at 30 °C under continuous cool-white illumination (constant light; 40–50 μE m^{-2} s^{-1}). Before the cells were released into constant light for the luminescence assay, a 12-h dark exposure was given to synchronize the rhythms of the individual cells in the population. For induction of co-expressed *kaiB*, *kaiC* or *kaiBC* genes, the *trc*p inducer IPTG was added at the beginning of constant light to final concentrations of 0, 2, 5, 10, 100 or 1,000 μM. Luminescence was measured with the kaiBCp::luxAB reporter of clock-controlled gene expression that serves as a real-time reporter of promoter activity[1,27,28]. For measurement of *in vivo* luminescence rhythms at different temperatures (18–38 °C), at least 12 independent colonies for each condition were monitored as previously described[28,35]. Our standard light intensity for the luminescence rhythm assay was 40–50 μE m^{-2} s^{-1}, but we observed that the rate of damping in the wild type was influenced by both the temperature (as described in this paper) and light intensity (as tested under a range of light intensities).

Calculation of free-running period, Q_{10} value, and damping rate of luminescence rhythms. The period of luminescence rhythms was analysed with ChronoAnalysis II, version 10.1 (courtesy of T. Roenneberg), and the Q_{10} value for evaluation of temperature compensation over a wide range of temperatures was calculated with the following equation: $Q_{10} = [(1/\tau_2)/(1/\tau_1)]^{10/(T_2 - T_1)}$; in which τ_1 denotes period at the lower temperature (T_1), and τ_2 denotes period at the higher temperature (T_2)[36]. Damping rate is the number of days required for the amplitude of the rhythm to decrease to 1/*e* of the starting value. The damping rates were calculated with the LUMICYCLE data analysis program (Actimetrics;

courtesy of D. Ferster). The program fits the data to a sine wave multiplied by an exponential decay factor[37].

Immunoblot assays for Kai abundance. After two 12-h light–dark cycles, liquid cultures at $D_{750\,nm} = 0.3$ were released to constant light at indicated temperatures. About 30 ml of cells were collected at different time points and an appropriate amount of fresh medium was added to the culture flask to maintain an equal cell density during the time-course experiment. For *kaiB*- or *kaiBC*-co-expressing strains, the cultures were treated with/without 5 μM IPTG at constant light time zero, for 12 h before cell collection. Total proteins were extracted as previously described[35]. Total extracts were separated by SDS–PAGE (15% gel for KaiB and 10% gel for KaiA and KaiC) and transferred onto nitrocellulose membranes. Proteins were transferred to nitrocellulose for immunoblotting using polyclonal rabbit antisera raised against KaiA or KaiB[35] or using polyclonal mouse antisera raised against KaiC[28]. Equal loading of extracts was confirmed by Coomassie blue staining in the gel, Ponceau red staining on the membrane, and/or by the density of nonspecific bands on the immunoblots. The immunoblot signals for relative Kai protein density/abundance were analysed with National Institutes of Health Image J software.

Determination of growth rate and doubling time. Growth rates of cyanobacterial strains, including wild type, optKaiBC, as well as two arhythmic/damped *kaiC* mutants (CLAb and CLAc)[1,3], were measured in parallel at 18, 20, 25, 30, 34 and 37 °C. Initial cultures were grown in liquid BG-11 medium at 30 °C under constant illumination (50 $\mu E\,m^{-2}\,s^{-1}$) in a shaking water bath at 100 r.p.m. and with air bubbling into the cultures. Cell densities were monitored by measuring the attenuance at 750 nm ($D_{750\,nm}$). When cell densities reached $D_{750\,nm} = 0.8$, cultures were diluted to $D_{750\,nm} = 0.005$, and grown in 12-h light–dark cycles (12 h light (50 $\mu E\,m^{-2}\,s^{-1}$) followed by 12 h darkness) in water baths set to 18, 20, 25, 30, 34 or 37 °C with shaking (100 r.p.m.) and air bubbling. Cell densities were determined at $D_{750\,nm}$ over a time course as indicated. When $D_{750\,nm}$ values of cell cultures exceed 0.9, the attenuance measurement is not linear with cell density. Therefore, for samples with a $D_{750\,nm}$ that was larger than 0.9, the samples were diluted to a $D_{750\,nm}$ that was within the linear range before attenuance determination (and the plotted D value is then corrected for the dilution). Two-to-six independent experiments were performed for each strain, and the growth curves were plotted as average $D_{750\,nm}$ values over time in light–dark. Doubling time was calculated by fitting exponential curves to the growth curves. Growth rate was generated by fitting growth data to exponential curves (cell density at a specific time = initial cell density $\times\, e^{(\text{growth rate} \times \text{time})}$) (see http://mathworld.wolfram.com/LeastSquaresFittingExponential.html). Doubling time was calculated as: doubling time (h) = (ln(2)/growth rate) × 24.

Statistical analyses. A two tailed Student's *t*-test was used for statistical analyses. $*P < 0.05$; $**P < 0.01$.

32. Sharp, P. M. & Li, W. H. An evolutionary perspective on synonymous codon usage in unicellular organisms. *J. Mol. Evol.* **24,** 28–38 (1986).
33. Schuster, P., Fontana, W., Stadler, P. F. & Hofacker, I. L. From sequences to shapes and back: a case study in RNA secondary structures. *Proc. R. Soc. Lond. B* **255,** 279–284 (1994).
34. Bustos, S. A. & Golden, S. S. Expression of the *psbDII* gene in *Synechococcus* sp. strain PCC 7942 requires sequences downstream of the transcription start site. *J. Bacteriol.* **173,** 7525–7533 (1991).
35. Xu, Y., Mori, T. & Johnson, C. H. Circadian clock-protein expression in cyanobacteria: rhythms and phase setting. *EMBO J.* **19,** 3349–3357 (2000).
36. Xu, Y. *et al.* Intramolecular regulation of phosphorylation status of the circadian clock protein KaiC. *PLoS ONE* **4,** e7509 (2009).
37. Izumo, M., Johnson, C. H. & Yamazaki, S. Circadian gene expression in mammalian fibroblasts revealed by real-time luminescence reporting: temperature compensation and damping. *Proc. Natl Acad. Sci. USA* **100,** 16089–16094 (2003).