Bayesian Modeling Reveals Ultrasensitivity Underlying Metabolic Compensation in the Cyanobacterial Circadian Clock

Lu Hong¹, Danylo O Lavrentovich^{2,†}, Archana Chavan³, Eugene Leypunskiy¹, Eileen Li⁴, Charles Matthews^{4,‡}, Andy LiWang^{3,5,6,7,8,9}, Michael J Rust^{10,11,12,*} & Aaron R Dinner^{2,11,13,*}

1. Graduate Program in Biophysical Sciences, University of Chicago, Chicago, IL 60637, USA
2. Department of Chemistry, University of Chicago, Chicago, IL 60637, USA
3. School of Natural Sciences, University of California, Merced, CA 95343, USA
4. Department of Statistics, University of Chicago, Chicago, IL 60637, USA
5. Quantitative and Systems Biology, University of California, Merced, CA 95343, USA
6. Center for Circadian Biology, University of California, San Diego, La Jolla, CA 92093, USA
7. Chemistry and Chemical Biology, University of California, Merced, CA 95343, USA
8. Health Sciences Research Institute, University of California, Merced, CA 95343, USA
9. Center for Cellular and Biomolecular Machines, University of California, Merced, CA 95343, USA
10. Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, USA
11. Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637, USA
12. Institute for Genomics and Systems Biology, University of Chicago, Chicago, Illinois 60637, USA
13. James Franck Institute, University of Chicago, Chicago, IL 60637, USA

*Corresponding Authors: mrust@uchicago.edu (MJR), dinner@uchicago.edu (ARD)

†Current address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

‡Current address: School of Mathematics, University of Edinburgh, Edinburgh, EH9 3FD, UK

Keywords: Bayes factor/kinetic modeling/Markov chain Monte Carlo/robustness/substrate competition

Running Title: Bayesian modeling reveals ultrasensitivity

Abstract

Mathematical models can enable a predictive understanding of mechanism in cell biology by quantitatively describing complex networks of interactions, but such models are often poorly constrained by available data. Owing to its relative biochemical simplicity, the core circadian oscillator in *Synechococcus elongatus* has become a prototypical system for studying how collective dynamics emerge from molecular interactions. The oscillator consists of only three proteins, KaiA, KaiB, and KaiC, and near-24-h cycles of KaiC phosphorylation can be reconstituted in vitro. Here, we formulate a molecularly-detailed but mechanistically agnostic model of the KaiA-KaiC subsystem and fit it directly to experimental data within a Bayesian parameter estimation framework. Analysis of the fits consistently reveals an ultrasensitive response for KaiC phosphorylation as a function of KaiA concentration, which we confirm experimentally. This ultrasensitivity primarily results from the differential affinity of KaiA for competing nucleotide-bound states of KaiC. We argue that the ultrasensitive stimulus-response relation is critical to metabolic compensation by suppressing premature phosphorylation at nighttime.

Synopsis

This study takes a data-driven kinetic modeling approach to characterizing the interaction between KaiA and KaiC in the cyanobacterial circadian oscillator and understanding how the oscillator responds to changes in cellular metabolic conditions.

- An extensive dataset of KaiC autophosphorylation measurements was gathered and fit to a detailed yet mechanistically agnostic kinetic model within a Bayesian parameter estimation framework.
- KaiA concentration tunes the sensitivity of KaiC autophosphorylation and the period of the full oscillator to %ATP.
- The model reveals an ultrasensitive dependence of KaiC phosphorylation on KaiA concentration as a result of differential KaiA binding affinity to ADP- vs. ATP-bound KaiC.
- Ultrasensitivity in KaiC phosphorylation contributes to metabolic compensation by suppressing premature phosphorylation at nighttime.

Introduction

Achieving a predictive understanding of biological systems and chemical reaction networks is challenging because complex behavior can emerge from even a small number of interacting components. Classic examples include the propagation of action potentials in neurobiology and chemical oscillators such as the Belousov-Zhabotinsky reaction. The collective dynamics in such systems cannot be easily intuited through qualitative reasoning alone, and thus mathematical modeling has long played an important role in summarizing and interpreting existing observations and formulating testable, quantitative hypotheses.

In general, mathematical modeling can be classified as either “forward” or “reverse.” In forward modeling, known interactions are expressed mathematically, which allows a researcher to draw out the logical implications of the model and its underlying assumptions (Gunawardena, 2014). In reverse modeling, experimental data are used to infer unknown interactions through a statistical approach (Villaverde and Banga, 2014). Many forward modeling studies are highly phenomenological; such studies excel in showing how effects like feedback (Novák and Tyson, 2008) and ultrasensitivity (Ferrell and Ha, 2014a,b) can give rise to collective dynamics, including bistable switching, oscillation, and adaptation (Ma et al., 2009). The simplicity of this class of models, however, makes quantitative prediction and experimental verification difficult. Reverse modeling, on the other hand, has found success in untangling complex interactions in -omic data (Machado et al., 2011; Wu et al., 2017) and signaling pathways such as the eukaryotic circadian clock (Forger and Peskin, 2003) and the JAK2/STAT5 signaling pathway (Hug et al., 2013). However, the complexity of such models raises issues of identifiability, i.e., whether a model topology and/or parameter values can be uniquely determined given the input data (Bellman and Åström, 1970; Cobelli and DiStefano, 1980). Furthermore, the nonlinear dynamics typical of such models give rise to non-convex optimization problems that pose significant technical and computational challenges.

The circadian clock from the cyanobacterial species *Synechococcus elongatus* PCC 7942 (Johnson et al., 2011) represents a unique opportunity to combine elements of both forward and reverse modeling. The core oscillator is post-translational (Tomita et al., 2005) and consists of just three proteins: KaiA, KaiB, and KaiC. A stable rhythm in KaiC phosphorylation with a period of nearly 24-h emerges spontaneously from these components, driven by KaiA-dependent autokinase reactions followed by a KaiB-mediated delayed negative feedback loop that favors dephosphorylation. The phosphorylation cycle can be reconstituted in vitro while still retaining the hallmarks of circadian rhythms in living organisms (Nakajima et al., 2005; Yoshida et al., 2009; Rust et al., 2011; Leypunskiy et al., 2017). Previous work has clearly articulated the basic biochemical events in the phosphorylation cycle (Johnson et al., 2011; Swan et al., 2018), allowing specification of a model topology with few ambiguities.

Despite the apparent simplicity of the system, the dynamics of the Kai oscillator are sufficiently complex that reverse modeling can provide useful insights. KaiC molecules can exist in multiple phosphorylation states and nucleotide-bound states, and how these states affect KaiC’s interaction with KaiA (Mori et al., 2018) and KaiB (Phong et al., 2013; Lin et al., 2014) is not fully understood. A related unresolved issue is the effect of the solution nucleotide pool (ATP and ADP) on the oscillator. In *S. elongatus*, the day/night cycle is reflected in the cellular metabolic state, including changes in the adenylate nucleotide pool %ATP (defined as $100\%[\text{ATP}]/([\text{ATP}] + [\text{ADP}])$), which acts as a timing cue and plays an important role in controlling the amplitude and phase of the phosphorylation cycle (Rust et al., 2011; Phong et al., 2013; Leypunskiy et al., 2017). KaiC is an ATPase (Terauchi et al., 2007) and phosphotransferase (Nishiwaki and Kondo, 2012), and its activities are regulated by which nucleotides are bound. The nucleotide-bound state is in turn regulated by KaiA, which acts as a nucleotide-exchange factor (Nishiwaki-Ohkawa et al., 2014). The kinetics of nucleotide exchange, the affinities of KaiC for nucleotides, and the heterogeneity of nucleotide-bound states in the KaiC hexamer have been measured (Nishiwaki-Ohkawa et al., 2014; Abe et al., 2015), but it is experimentally challenging to monitor all of the relevant quantities simultaneously over the course of the cycle.

Here we take a data-driven Bayesian modeling approach (Figure 1A) to elucidate the regulatory relations between KaiA, nucleotides in solution, KaiC phosphorylation, and KaiC nucleotide-bound state, with the goal of deducing dynamical rules that can predict the behavior of the system. The resulting model

does not include KaiB; it focuses on describing the dynamics of phosphorylation during the daytime part of the clock cycle. To provide a training set for this model, we collected kinetic time series characterizing the metabolic sensitivity of the KaiC phosphorylation kinetics (in the absence of KaiB) over a wide range of KaiA concentrations ($[KaiA]$) and %ATP. Although such data do not give us direct access to all relevant states of the KaiA-KaiC subsystem, they place constraints on the underlying molecular interactions. Bayesian parameter estimation (Mackay and Kay, 2003) has been used to systematically quantify parameter uncertainties and compare models in many fields (Geweke, 1989; Wasserman, 2000; Hou et al., 2012), including systems biology (Flaherty et al., 2008; Klinke, 2009; Toni Tina et al., 2009; Xu et al., 2010; Schmidl et al., 2012; Eydgahi et al., 2013; Pullen and Morris, 2014; Mello et al., 2018). Here it allows us to estimate parameter values, quantify the importance of specific model elements, and make mechanistic predictions from the model.

The Markov chain Monte Carlo (MCMC) sampling method that we use to fit the model to the data yields an ensemble of parameter sets, rather than a single best fit. We find that, even with extensive training data, many microscopic parameters in the model are not tightly constrained and their values vary widely across the ensemble of fits. Despite this, we show that this ensemble of fits robustly makes predictions that are borne out in experimental tests (Brown and Sethna, 2003; Gutenkunst et al., 2007). In particular, the model reveals an ultrasensitive dependence of phosphorylation on the concentration of KaiA, with strong nonlinearity at low $[KaiA]$, conditions that likely apply near the nighttime to daytime transition point, when a large fraction of KaiA molecules are inhibited. The ultrasensitive response primarily arises from a differential affinity of KaiA for different nucleotide-bound states of KaiC. This mechanism is analogous to substrate competition (Ferrell and Ha, 2014b), where kinetic competition of multiple enzyme substrates leads to ultrasensitivity.

Lastly, we consider the implications of these results for the full oscillator, in which KaiC rhythmically switches between phosphorylation and dephosphorylation. Incorporation of the ultrasensitive response to KaiA into a mathematical model of the full oscillator suggests that this effect both stabilizes the period against changes in the nucleotide pool and allows oscillations to persist even when KaiB binds KaiA relatively weakly. Consistent with this prediction, we find that a substantial amount of KaiA is not bound by KaiB even when the clock is dephosphorylating. These results shed new light on metabolic compensation, a property that allows robust 24-h oscillation in spite of changes in %ATP conditions (Johnson and Egly, 2014). Taken together, our results show how the Bayesian framework combined with extensive training data can be used to discover unanticipated mechanisms and direct experimental investigations.

Results

A molecularly motivated model of KaiA-KaiC dynamics

To probe the response of KaiC phosphorylation to a wide range of metabolic conditions, we made kinetic measurements of KaiC phosphorylation at three %ATP conditions and six [KaiA] conditions while holding the KaiC concentration constant (**Figure 1B**). KaiC is a homohexamer and each subunit has two domains, termed CI and CII. Both CI and CII domains have ATPase activity (Hayashi et al., 2003; Pattanayek et al., 2004; Terauchi et al., 2007), while the CII domain is in addition a bidirectional phosphotransferase (Nishiwaki and Kondo, 2012; Egli et al., 2012) with two phosphorylation sites (Xu et al., 2004; Rust et al., 2007; Nishiwaki et al., 2007). Each KaiC subunit thus has four phosphoforms: the unphosphorylated (U), phosphoserine-431 (S), phosphothreonine-432 (T), and doubly phosphorylated (D) states. The measurement resolves the kinetics of all four KaiC phosphoforms.

Our strategy is to fit these data with a model of the KaiC catalytic cycle with a minimum of simplifying assumptions. To this end, we formulate a model based on mass action kinetics. We explicitly keep track of three properties of the CII domain of each KaiC subunit: its phosphorylation status (right superscripts in **Figure 1C**), nucleotide-bound state (right subscript), and whether or not KaiA is bound (left superscript). We do not consider CI or the hexameric nature of KaiC explicitly (see **SI** for further discussion). There are thus 16 possible KaiC states, 8 of which are shown in **Figure 1C**, along with the phosphotransfer, nucleotide exchange, KaiA (un)binding, and hydrolysis reactions that connect the states (see **Figure S1A** for the full model structure). We also hypothesized that nucleotides might interact directly with KaiA, which could allow KaiA's activity to directly depend on nucleotides in solution. However, we did not detect any direct interaction between KaiA and ATP or ADP using NMR spectroscopy, so we do not allow for this scenario in the model (**Figure S2**). Below, we step through the four classes of reactions that we include; further details can be found in Materials and Methods.

Phosphotransfer KaiC is a bidirectional phosphotransferase (Egli et al., 2012; Nishiwaki and Kondo, 2012), which means that it can transfer a γ -phosphate group from a bound ATP to a phosphorylation site, but unlike a typical phosphatase, it regenerates ATP from ADP during dephosphorylation, i.e.,

where $(X, Y) \in \{(U, T), (U, S), (T, D), (S, D)\}$. This mechanism implies that the nucleotide-bound state of KaiC has a significant impact on the direction of its phosphotransferase activity: an ATP-bound KaiC presumably cannot dephosphorylate, and an ADP-bound KaiC cannot phosphorylate.

Nucleotide exchange KaiA binding to the CII domain (Kim et al., 2008; Pattanayek and Egli, 2015) stimulates KaiC autophosphorylation (Iwasaki et al., 2002; Williams et al., 2002; Kageyama et al., 2006). Recent work has shown that KaiA can bind to KaiC and act as a nucleotide-exchange factor (Nishiwaki-Ohkawa et al., 2014) by facilitating conformational changes at the subunit interface that promote solvent exposure of the nucleotide-binding pocket (Hong et al., 2018). It is currently unclear whether this nucleotide exchange activity is responsible for all of KaiA's effect on KaiC or whether it alters the KaiC catalytic cycle in other ways (see **SI** for further analysis of this issue). The reversible binding of KaiA

contributes two classes of rate constants, k_a and k_b .

Figure 1: Phosphorylation data are fit by a mechanistically agnostic kinetic model. A) An outline of the data-driven Bayesian model fitting approach employed in this work. B) To constrain the model, measurements of KaiC phosphorylation kinetics were collected at six [KaiA] and three %ATP conditions. The curves represent the best fit model prediction. C) A schematic of the mass action kinetics model. The model elaborates on the autophosphorylation reactions of KaiC by explicitly keeping track of the time evolution of the KaiC phosphoforms, nucleotide-bound states, and KaiA binding mediated by phosphotransfer, nucleotide exchange, and ATP hydrolysis. Note that the KaiA binding reactions are second-order, but KaiA concentration ($[A]$) is written as part of the effective first-order rate constant. See the main text for a discussion of the state and rate constant nomenclature and **Figure S1A** for a schematic of the full model. D) The posterior distributions for $\log K_d^{\text{TP}}/\mu\text{M}$ (base 10). The horizontal axis represents the affinity for ADP-bound KaiC, and the vertical axis represents the affinity for ATP-bound KaiC; the four colors correspond to the KaiC phosphoforms, as in panel C. The asterisks represent the best fit, and the contour lines represent the 95% and 68% highest posterior density regions (HDR). The dashed line represents the $K_d^{\text{TP}} = K_d^{\text{DP}}$ line, so that densities above the line indicate higher affinity for the ADP-bound species and densities below the line indicate higher affinity for the ATP-bound species.

Because the CII domain of KaiC releases its bound nucleotide very slowly in the absence of KaiA (Nishiwaki-Ohkawa et al., 2014), we ignore the possibility of KaiA-independent nucleotide exchange in the model. Under the assumptions that i) the apo state is in a quasi-steady state, ii) the ADP and ATP on-rates are identical, and iii) ATP release is slow, nucleotide exchange can be modeled as a one-step reaction

where

$$k_{\text{TP}}^A = k_r^{\text{DP}} \frac{[\text{ATP}]}{[\text{ATP}] + [\text{ADP}]} \quad (4)$$

and k_r^{DP} is the ADP dissociation rate constant. Nucleotide exchange thus contributes one class of rate constant, k_r^{DP} . See Materials and Methods for the derivation of (4).

ATP hydrolysis Finally, we allow for irreversible ATP hydrolysis in the CII domain

which contributes one class of rate constants, k_h . Because each KaiC molecule consumes relatively little ATP on the timescale of a day (Terauchi et al., 2007), we assume the solution ATP and ADP concentrations are constant.

Species-dependent rates Given the six classes of rate constants, k_p , k_d , k_a , k_b , k_r^{DP} , and k_h , we make the model maximally general, or mechanistically agnostic, by allowing each rate constant to potentially depend on the specific molecular state involved in the reaction. For example, the KaiA dissociation rate constant is allowed to vary depending on the nucleotide-bound state and phosphoform background of KaiC, and thus the dissociation rate constants for the ADP-bound U ($k_b^{\text{U,DP}}$) and ATP-bound T phosphoforms ($k_b^{\text{T,TP}}$) are two independent model parameters. In this way, the parameter fitting and model comparison procedures automatically test specific biochemical hypotheses about the functions of KaiA and KaiC. For example, allowing the KaiA off-rates to depend on the nucleotide-bound states is equivalent to the hypothesis that KaiA has different dwell times for ATP- versus ADP-bound states of KaiC. In fact, because each reaction has an independent rate constant, except for the thermodynamic constraints of detailed balance, the fitting procedure effectively allows for simultaneous testing of all possible two-way interactions of the three categories of KaiC properties, without a priori preference for any particular mechanism.

The data constrain the parameters to widely varying degrees

We estimate the model parameters through a Bayesian framework. In this framework, we maximize the posterior probability, which is proportional to the product of the prior distribution and the likelihood function. Here, we interpret the prior as representing subjective beliefs on the model parameters before experimental inputs, while the likelihood function quantifies the goodness of fit. Bayesian parameter estimation reduces to least-squares fitting under the assumption of normally distributed residuals and uniform priors. In practice, we find that direct numerical optimization of the posterior usually results in fits that are trapped in low probability local maxima (**Figure S3B**). Thus we instead draw parameters from the prior distribution and then use a heuristic combination of Markov chain Monte Carlo (MCMC) sampling and optimization (Powell's algorithm) to explore the parameter space. The MCMC method that we use (Goodman and Weare, 2010; Foreman-Mackey et al., 2013) efficiently searches the parameter space by simulating an ensemble of parameter sets in parallel; the spread of the ensemble reflects the geometry of the posterior distribution and is used to guide the directions of Monte Carlo moves. See Materials and Methods for a more mathematical treatment of the fitting procedure and comparison of different numerical optimization and sampling methods.

We use this approach to fit the phosphorylation data (**Figure 1B**) together with previously published data on dephosphorylation (Rust et al., 2011), ATP hydrolysis rate (Terauchi et al., 2007), and the KaiA dwell time for each KaiC phosphoform (Kageyama et al., 2006; Mori et al., 2018) (see Materials and Methods). Overall, the model achieves excellent agreement with the training data (**Figure 1B** and **Figure S4A–C**). In the following analyses, we refer to model predictions using the best fit parameter values, and quantify the uncertainties using the posterior distribution (see **SI** for further discussion on the convergence of the simulation).

We find that certain parameters, such as the hydrolysis rates in the U and T phosphoforms and the KaiA off-rates from the U phosphoform, are tightly constrained, while many others, mainly involving S and D phosphoforms, are less constrained, in the sense that their posterior distributions span multiple orders of magnitude, exhibit multimodality, or cannot be reproduced over multiple independent runs (**Figure S1B**). Some parameters are highly correlated and certain combinations of the parameters are much better constrained than the individual parameters. For example, the posterior distributions for the KaiA binding affinities (**Figure 1D**) appear better constrained than the on/off rates (**Figure S5B**).

Taken together, these results are consistent with the notion that collective fits of multiparameter models are generally “sloppy,” meaning that the sensitivities of different combinations of parameters can range over orders of magnitude with no obvious gaps in the spectrum (Brown and Sethna, 2003; Gutenkunst et al., 2007). As we will see, we can nonetheless make useful predictions using the ensemble of model parameters, because the model behavior is constrained along the stiffest directions of the posterior distribution. By contrast, direct parameter measurements need to be both complete and precise to achieve similar predictive validity (Gutenkunst et al., 2007). We further characterize the structure of the parameter space in **SI** and **Figure S5**.

KaiC (de)phosphorylation goes through transient kinetic intermediates

Given the model, we can interpret the underlying molecular events in KaiC phosphorylation. Here we consider the phosphorylation kinetics at the standard reaction condition (3.5 μ M KaiC, 1.5 μ M KaiA, 100% ATP; **Figure 2A** and **B**, solid curves); we examine the effect of varying [KaiA] and %ATP in the following sections. At the beginning of the phosphorylation reaction, KaiC molecules are predominantly in the ADP-bound U state (C_{DP}^U), the end product of the dephosphorylation pathway in the absence of KaiA (**Figure 2A**). With the addition of KaiA, the C_{DP}^U state becomes rapidly depleted within the first 10 minutes of the reaction and enters the C_{TP}^U state. Consistent with the kinetic ordering observed in the full oscillator, the C_{TP}^U population is primarily converted into the T phosphoform over the S phosphoform. The exact pathway underlying the preference for the T phosphoform is not well constrained by the data, but it appears to be the result of more than just a difference in the relative U $\rightarrow$ T and U $\rightarrow$ S phosphorylation rates; there is an interplay between KaiC phosphorylation and KaiA (un)binding kinetics (see **SI** and **Figure S6**). The ADP- and KaiA-bound T phosphoform species are unstable kinetic intermediates, and the population accumulates at the C_{TP}^T bottleneck for the first 4 hours. As phosphorylation reaches completion, the T phosphoform is converted first into $A_{C_{TP}}^D$ through the unstable ADP-bound intermediates, and then to the C_{TP}^D state; the populations of the $A_{C_{TP}}^D$ and C_{TP}^D states are comparable at steady state. We note here, however, that previous measurements indicate that approximately 30% of CII nucleotide-binding pockets should be ADP-bound in the presence of KaiA at steady state (Nishiwaki-Ohkawa et al., 2014), which suggests that the stability of the ADP-bound species is systematically underestimated by the model fit.

During the phosphorylation reaction, the amount of free KaiA is initially transiently depleted due to association with the ADP-bound U phosphoform (**Figure 2B**). Afterwards, KaiA primarily associates with the ATP-bound S and D phosphoforms as they appear, but does not bind to the T phosphoform strongly. Therefore, despite the lack of KaiB in the model, not all KaiA is free during the phosphorylation phase, and the amount of free KaiA depends on both the affinities of the nucleotide-bound states and the mixture of KaiC phosphorylation states (**Figure 1D**).

Figure 2: The model captures the kinetics of KaiC phosphorylation. A) The time evolution of all 16 KaiC species in a phosphorylation reaction with 100% ATP and either 1.5 μM (solid curves) or 0.2 μM (dashed curves) KaiA. The corresponding KaiA kinetics, broken down according to the phosphoform of the bound KaiC, is shown in B). The gray regions represent the 95% posterior interval. Refer to **Figure S1A** for the KaiC state nomenclature. C) KaiA concentration can tune the sensitivity of the KaiC phosphorylation level to %ATP. The points represent the measured total percentage phosphorylation levels at $t = 12.25$ h (see **Figure 1B** for the full kinetics), and the curves represent the model prediction at the same time point. D) KaiA concentration can tune the sensitivity of the clock period to %ATP. The period of the full KaiABC oscillator is calculated from fluorescence polarization measurement (see **Figure S7A–C** for further analysis).

We were surprised to see that the model fit predicts that the KaiA binding affinity for the ATP-bound T phosphoform is lower than those for the S and D phosphoforms. This is apparently in contradiction with experimental results that show that S-phosphomimetic mutants reduce A-loop exposure and weaken KaiA binding, while T-phosphomimetic mutants have opposite effects (Tseng et al., 2014; Chang et al., 2011). However, there is some ambiguity in the experimental literature, with various results employing different experimental methods and using proteins from different organisms, that has yet to be resolved. Our model stresses the importance of the nucleotide-bound state, especially that of the U phosphoform. KaiC nucleotide-bound state has generally not been measured in KaiA interaction studies and may be different in phosphomimetic mutants (see also Kageyama et al., 2006; Qin et al., 2010a; Murakami et al., 2016).

The dephosphorylation pathway is simpler because KaiA is not involved. In the model, KaiC by itself has no nucleotide exchange activity, and thus phosphorylated KaiC molecules in the absence of KaiA enter a cycle of dephosphorylation by the transfer of phosphoryl groups from the phosphorylation sites back to bound ADP molecules, followed by ATP hydrolysis and removal of inorganic phosphate, until the protein reaches the C_{DP}^U state (**Figure S4D**). The ADP-bound forms of the T, S, and D phosphoforms are only transiently populated, suggesting that the dephosphorylation bottlenecks are the ATP hydrolysis reactions, which make bound ADP available as a cofactor for dephosphorylation, rather than the phosphotransfer reaction itself. The kinetic preference for the $D \rightarrow S$ dephosphorylation pathway is the direct result of faster dephosphorylation via the $D \rightarrow S$ compared to the $D \rightarrow T$ reaction (**Figure S1B**; compare the posterior distribution of k_d^{DS} with that of k_d^{DT}). During this process, KaiC can occasionally autophosphorylate, but it is driven irreversibly towards the dephosphorylated state by ATP hydrolysis. We note here that the independence of the dephosphorylation reaction from solution ADP (Rust et al., 2011) is a built-in feature of the model, since

solution %ATP only affects the nucleotide exchange rate, which is assumed in the model to be zero in the absence of KaiA.

KaiA concentration tunes clock sensitivity to %ATP

The model further allows us to summarize and interpret the effect of %ATP and KaiA on KaiC phosphorylation observed in the training dataset (**Figure 2C**). Consistent with previous measurements (Rust et al., 2011; Phong et al., 2013), these results indicate that the near-steady-state ($t = 12$ h) total phosphorylation level of KaiC ($\%C^P = \%C^T + \%C^S + \%C^D$) is lower in the presence of ADP. Since we simultaneously vary %ATP and [KaiA], the data reveal that this inhibitory effect can be tuned by [KaiA]. In particular, the system is most insensitive to %ATP at either very low or very high [KaiA], while the %ATP sensitivity is the highest around [KaiA] = 0.75 μ M.

Interestingly, the %ATP sensitivity of the system cannot be fully abolished even at saturating KaiA concentrations (**Figure 2C**). This effect can be interpreted qualitatively in terms of the structure of the model. When there is more KaiA in solution, more KaiC goes through intermediate states that are in complex with KaiA, by Le Châtelier's principle. This shifts a larger fraction of the KaiC population through states that allow for exchange of bound ADP for ATP, which promotes phosphorylation. On the other hand, as the %ATP decreases, the ATP to ADP exchange rate decreases according to (4). When nucleotide exchange becomes less efficient, more KaiC stays in ADP-bound states, which are prone to dephosphorylation. In summary, [KaiA] and %ATP both act on the phosphorylation kinetics via the nucleotide exchange step, where %ATP directly regulates the exchange rate constant and sets its upper bound, while [KaiA] controls the population of exchange-competent KaiC and thus the effective exchange rate. Therefore, the effects of KaiA and increasing solution %ATP are not equivalent; because the effective exchange rate cannot exceed the limit set by %ATP, even a saturating amount of KaiA cannot fully compensate for low %ATP.

Given that the metabolic sensitivity of the KaiA-KaiC subsystem can be tuned by KaiA concentration, we asked whether metabolic sensitivity of the full oscillator period may also be tuned by KaiA. To address this question, we characterized the dependence of the period of the in vitro KaiABC oscillator on [KaiA] and %ATP using an optical assay (Leypunskiy et al., 2017; Heisler et al., 2019) that allows automated, parallelized monitoring of the fluorescence polarization of labeled KaiB (**Figure 2D** and **Figure S7A–C**). Consistent with the hypothesis, we found that low KaiA concentration enhances the period sensitivity to %ATP compared to the standard condition (1.5 μ M KaiA). These results suggest that the KaiA activity, and how it is controlled, plays a critical role in determining the clock period stability.

KaiC phosphorylation exhibits ultrasensitive dependence on KaiA levels

In addition to inferring kinetics of states not easily accessible to experiments, the model allows us to interpolate between the training data points and study the relation between KaiC phosphorylation, [KaiA], and %ATP at a much finer resolution. This analysis shows an ultrasensitive dependence of the steady-state $\%C^P$ on KaiA concentration (**Figure 3A** and **D left**). Specifically, we see a threshold-hyperbolic stimulus-response relation (Gomez-Urbe et al., 2007; Ferrell and Ha, 2014a), where KaiC phosphorylation is highly suppressed near the sub-micromolar [KaiA] regime, but then follows a right-shifted hyperbolic stimulus-response function once [KaiA] exceeds a threshold. Importantly, the threshold depends on %ATP. The model makes similar predictions for the steady-state T, S, and D phosphoforms as well (**Figure S8A**). However, because of the $T \rightarrow D$ and $S \rightarrow D$ phosphotransfer reactions, the stimulus-response relations of T and S are not monotonic functions of [KaiA] because high [KaiA] and high %ATP conditions stabilize the D phosphoform at the expense of the T and S phosphoforms.

Previous studies of KaiA-KaiC interactions examined the response of KaiC at relatively high KaiA concentrations (≥ 1.2 μ M), comparable to the total amount of KaiA used in an oscillating reaction. Ma and

Figure 3: Substrate competition explains KaiC phosphorylation ultrasensitivity. A) The predicted stimulus-response relation of the total steady-state KaiC phosphorylation level as a function of %ATP and [KaiA]. B) Experimentally-determined stimulus-response function of KaiC at three %ATP conditions; the curves are based on refitting the best fit of the full model to the steady-state measurements. C) Similar to B) but for KaiC S431A, which has only one phosphorylation site; the curves are based on independent fits to a simple phenomenological substrate competition model. D) Cross sections of the stimulus-response relation at three %ATP, computed using the full model (left) and model -n (right). E) Posterior distributions for the shapes of the stimulus-response functions at 25% ATP predicted by the full model and model -n. The contours represent the 68% and 95% HDRs, and the gray stars represent the model best fits. The shape of the stimulus-response functions are quantified using two metrics: EC10, which quantifies threshold-like behavior, and EC90 - EC10, which quantifies switch-like behavior. The shape of the experimentally-determined stimulus-response function at 25% ATP is shown as the yellow star. The dashed line represents $(EC10, EC90 - EC10) = (K/9, 80K/9)$, which characterizes the shape of a hyperbolic stimulus-response function $[A]/(K + [A])$ that has no switching or thresholding. F) The stimulus-response functions of KaiC S431A/T432A (AA; left) and S431E/T432E (EE; right) phosphomimetic mutants to probe the effect of kinetic competition on KaiC phosphorylation. G) The relations between EC50 (the midpoint of a stimulus-response function) and KaiC AA/EE concentrations, quantified using the curves shown in F).

315 Ranganathan (2012) investigated the steady-state stimulus-response relation, but did not consider the ef-
 316 fect of %ATP or fully characterize the low [KaiA] regime. Previous reports of initial phosphorylation rates

Table 1: Effects of differential KaiA (un)binding kinetics

Model	Log likelihood			Bayes factor*
	Phosphorylation	Dephosphorylation	Hydrolysis	
Full model	422.9	346.8	-0.8	1
-n [†]	249.2	275.4	-0.3	10.4
-p [‡]	392.4	303.2	-2.4	2.2
-n,-p	204.4	266.4	-2.6	19.7

* We define the Bayes factor as the ratio of the marginal likelihood function of the full model over that of the simplified models. We adopt the convention that a Bayes factor larger than 3.2 is substantial evidence against the model (Vyshemirsky and Girolami, 2008).

[†] -n: on/off rates decoupled from nucleotide-bound state.

[‡] -p: on/off rates decoupled from phosphorylation state.

suggest that they exhibit a hyperbolic dependence on [KaiA] (Rust et al., 2007; Lin et al., 2014), similar to simple Michaelis-Menten enzyme-substrate systems. However, this does not imply that the steady-state stimulus-response relation is hyperbolic as well.

To assess the robustness of the model prediction across the ensemble, we use two metrics proposed by Gunawardena (2005) to quantify the shape of the predicted stimulus-response curves for %C^P at any fixed %ATP: we use EC10 to measure the extent to which the curve acts as a threshold and EC90 - EC10 to measure the extent to which the curve acts as a switch. Here, EC_x is the KaiA concentration required to reach x% of the steady-state phosphorylation level at saturation. **Figure 3E** shows the distribution of these quantities in the ensemble at 25% ATP. Overall these statistics are tightly constrained by the training data set, and are clearly distinct from those from hyperbolic stimulus-response relations (**Figure 3E**, dashed gray line).

Given the robustness of the prediction, we sought to experimentally verify the shape of the stimulus-response function. We measured KaiC phosphorylation at $t = 24$ h at various concentrations of [KaiA] at three %ATP conditions (**Figure 3B** and **Figure S8B**). Consistent with model prediction, the experimentally-derived stimulus-response relations are ultrasensitive with an %ATP-dependent phosphorylation threshold, and the stimulus-response relation of the S phosphoform at 100% ATP is non-monotonic. We then quantified the shape of the stimulus-response curve for %C^P at 25% ATP using the same two metrics defined above (**Figure 3E**, yellow star). At 25% ATP, the shape of the experimentally-derived stimulus-response curve is close to that of the model prediction, but the model fit is systematically less threshold-like (i.e., smaller EC10) and less switch-like (i.e., larger EC90 - EC10). This inconsistency is likely due to a combination of training data under-determining the shape of the curve at the sub-micromolar range (compare **Figure 2C** with **Figure 3B**) and the fitting method under-estimating uncertainties (see **SI**).

Lastly, the saturation phosphorylation levels in the steady-state measurements appear systematically lower than those implied by the training data set (compare **Figure 3B** with D left). This may be a result of batch-to-batch variations in protein and nucleotide quantification. This difference can be corrected by refitting the full model to the steady-state measurement (**Figure 3B** and **Figure S8C**). The refit results suggest that errors in protein and nucleotide concentrations primarily affect the kinetic properties of the S phosphoform in the model (**Figure S8D**), but the refit does not change the qualitative conclusions.

A substrate competition mechanism underlies ultrasensitivity in KaiC phosphorylation

What is the mechanism of ultrasensitivity in KaiC phosphorylation? Given that each KaiC subunit has two phosphorylation sites, a plausible explanation is multisite phosphorylation, whereby the concentration of the maximally phosphorylated species exhibits an ultrasensitive dependence on the kinase concentration (Gunawardena, 2005) (or in this case, the nucleotide-exchange factor concentration), even if each consecutive phosphorylation step follows mass action kinetics. To examine this possibility, we measured the stimulus-response relation of the KaiC S431A mutant, which has only one phosphorylation site, and the results show ultrasensitivity comparable to that of the WT protein (**Figure 3C**). Furthermore, because KaiC is its own phosphatase, it violates the assumption of distributivity (i.e., at most one modification takes place before the dissociation of the enzyme and substrate) (Gunawardena, 2005). Multisite phosphorylation thus cannot explain the observed ultrasensitivity.

In the ensemble of parameter sets, the KaiA dissociation constant of the ADP-bound (but not ATP-bound) U phosphoform (C_{DP}^U) is in or below the nanomolar range, much smaller than that of any other species of KaiC (**Figure 1D**). This is consistent with recent single molecule observations suggesting that the unphosphorylated form of KaiC can bind very tightly to KaiA (Mori et al., 2018) and native mass spectrometry measurements suggesting that KaiA binding to KaiC is enhanced by ATP hydrolysis, which would be needed to produce ADP-bound KaiC (Yunoki et al., 2019). Here, we argue that the key to understanding the origin of ultrasensitivity in the model lies in the differential binding affinity of KaiA to the ADP- and ATP-bound states of KaiC. We note here that since the model does not consider the hexameric structure of KaiC, we cannot rule out possible hexameric cooperative effects that may contribute to ultrasensitivity.

In the model, the differential KaiA binding affinity leads to the following dynamics: KaiA promotes phosphorylation by catalyzing the exchange of the bound ADP for ATP, but this process is in a kinetic competition with ATP hydrolysis, which returns KaiC to the ADP-bound state. At the beginning of the phosphorylation reaction, almost all the KaiA is bound to C_{DP}^U (**Figure 2A** and **B**) due to its high abundance and high affinity for KaiA (**Figure 1D**). When [KaiA] is low, the competition between nucleotide exchange and hydrolysis in the U phosphoform reaches a steady-state where $[C_{DP}^U]$ stays above [KaiA] (**Figure 2A**, dashed curves). Therefore, KaiA stays trapped by C_{DP}^U and the phosphorylation products (mostly T) cannot undergo nucleotide exchange. In the absence of KaiA, the autophosphatase activity of KaiC dominates, and the phosphorylation products revert back to the U phosphoform.

When [KaiA] is high, however, the competition between nucleotide exchange and hydrolysis in the U phosphoform pushes C_{DP}^U below [KaiA] (**Figure 2A**, solid curves), which frees KaiA to catalyze the nucleotide exchange reactions of the phosphorylation products. Once the flux of phosphorylation, KaiA binding, and nucleotide exchange outweighs that of hydrolysis, dephosphorylation, and KaiA unbinding, the phosphorylation products stay phosphorylated at steady state. Furthermore, the formation of phosphorylation product positively feeds back to deplete C_{DP}^U , further removing a KaiC state that traps KaiA and leading to rapid saturation of phosphorylation past the [KaiA] threshold. The [KaiA] threshold for phosphorylation depends on %ATP (**Figure 3A**), because when %ATP is low, more KaiA is needed to counteract the reduced ADP-to-ATP exchange rate.

This mechanism is a form of substrate competition (Ferrell and Ha, 2014b; Buchler and Louis, 2008), where the kinetic competition of multiple substrates for enzyme binding leads to ultrasensitivity. Here, KaiA acts as the enzyme, while the ADP-bound U phosphoform and the T (as well as S and D to a lesser extent due to phosphorylation ordering) phosphoform are the substrates that compete for KaiA binding. However, the fact that the phosphorylated and unphosphorylated forms of KaiC can interconvert through phosphotransfer reactions distinguishes the Kai system from a typical substrate competition scheme, where the substrates cannot interconvert.

The model suggests that the U phosphoform plays a special role in generating ultrasensitivity due to the significant difference in the affinity of KaiA for the ATP- vs. ADP-bound states of KaiC (**Figure 1D**). This

observation leads to two testable predictions. First, the amount of KaiA required to activate phosphorylation should be higher when more U phosphoform is present. We tested this prediction experimentally by measuring the stimulus-response relation of KaiC S431A in the presence of KaiC S431A/T432A (AA), which mimics the U phosphoform, or KaiC S431E/T432E (EE), which mimics the D phosphoform. The KaiC AA and EE mutants act as competitors for the KaiA-KaiC interaction (**Figure 3F**). Consistent with the hypothesis, the EC₅₀ (i.e., the midpoint of the ultrasensitive switch) is positively correlated with the concentration of KaiC AA, while varying KaiC EE has little effect (**Figure 3G**).

Second, the substrate competition mechanism suggests that the model should exhibit weaker nonlinearity if KaiA has the same affinity to ATP- vs. ADP-bound states of a given KaiC phosphoform. To test this prediction, we constructed simplified models where KaiA on/off rates are set to be independent of the nucleotide-bound state (model -n) or phosphorylation state (model -p) and fit the new models to the experimental data *ab initio*. Consistent with the prediction, decoupling KaiA on/off rates from the nucleotide-bound states results in a significant loss of ultrasensitivity (**Figure 3D** right and **Figure S9A**). Model -p by contrast behaves similarly to the full model (**Figure S9C**); consistent with the substrate competition mechanism, the ADP-bound states of KaiC in model -p have higher affinity to KaiA than the ATP-bound states, regardless of the phosphorylation state (**Figure S9C**). We quantify the effects of such model reductions by computing the Bayes factor, which is a metric for systematic model comparison that favors goodness of fit but penalizes model complexity and parameter fine tuning (MacKay and Kay, 2003); it is similar to the Bayesian information criterion (Schwarz, 1978), but makes no asymptotic assumptions. The analysis shows that the loss of ultrasensitivity in model -n degrades the fit quality significantly, while model -p is only marginally worse than the full model (**Table 1**). Interestingly, a model where the KaiA on/off rates are completely independent of the state of KaiC (model -n,-p; **Figure S9D** and **E**) is much worse than either model -n or model -p (**Table 1**). We conclude that the nucleotide-bound state of KaiC plays a key role in regulating its interaction with KaiA and thus in determining phosphorylation kinetics.

Substrate competition underlies metabolic compensation

Finally, we consider the implications of the ultrasensitivity for the full oscillator. For the sake of clarity, we make a distinction in this section among three subpopulations of KaiA: the sequestered KaiA, which refers to inactive KaiA in a KaiABC complex; the active KaiA, which refers to (free or bound) KaiA not sequestered by KaiB; and the free KaiA, which is not associated with either KaiB or KaiC.

We first consider the role ultrasensitivity plays in regulating KaiA activity. It is well-established that KaiB plays an essential role in regulating KaiA during nighttime. At dusk, the buildup of KaiC D and S phosphoforms triggers the binding of KaiB to CI (Rust et al., 2007; Chang et al., 2012; Mutoh et al., 2013; Phong et al., 2013; Lin et al., 2014; Tseng et al., 2017; Snijder et al., 2017; Mukaiyama et al., 2018) and subsequently the sequestration of KaiA by CI-bound KaiB (Kageyama et al., 2006; Qin et al., 2010a). In the absence of active KaiA, the CII domain autodephosphorylates, and the KaiABC ternary complex disassembles (Snijder et al., 2017) at dawn as KaiC reaches its dephosphorylated state (Tomita et al., 2005), freeing KaiA and readying the clock for the next cycle.

This understanding of the negative feedback loop implies that the sequestration of KaiA by KaiB is a source of nonlinearity in the system that is critical for generating oscillation. Indeed, in many models of the Kai oscillator, the complete sequestration of KaiA during dephosphorylation is either a built-in or required feature for stable oscillation (e.g., Yoda et al., 2007; van Zon et al., 2007; Phong et al., 2013; Pajmans et al., 2017b). Our observation that phosphorylation is suppressed nonlinearly at low [KaiA] suggests that complete sequestration of KaiA by KaiB is not necessary to prevent phosphorylation. Indeed, there is mounting evidence that KaiB sequestration by itself is insufficient to completely inactivate KaiA during dephosphorylation. Specifically, measurements using native mass spectrometry, co-immunoprecipitation (co-IP), and native PAGE suggest that there is a significant amount of KaiA₂C₆ complex (Kageyama et al., 2006; Brettschneider et al., 2010) and free KaiA (Qin et al., 2010a) throughout the entire phosphorylation cycle.

To confirm that KaiA is not fully sequestered by KaiBC complexes, we used co-IP of FLAG-tagged KaiB to monitor the amount of uncomplexed KaiA in supernatant, which we interpret to be a measure of active KaiA concentration (**Figure 4A** and **Figure S7D**). The experiment shows that there is indeed a sizable amount of active KaiA in solution in the first half of the dephosphorylation stage, although the experiment does not allow us to assign absolute concentrations. Taken together, these results suggest that either the binding of KaiA to KaiBC is more labile or has lower affinity than previously assumed, or that the sequestration kinetics are slow compared to the length of the dephosphorylation stage. In either case, substantial amounts of KaiA appear to remain free of KaiABC complexes during oscillation.

The fact that the phosphorylation threshold scales with %ATP suggests that ultrasensitivity may also lead to insensitivity of the period of the Kai oscillator to %ATP (Phong et al., 2013), a phenomenon termed “metabolic compensation” (Johnson and Egly, 2014). We examine this issue using a simple model of the Kai oscillator proposed by Phong et al. (2013), which we hereafter refer to as the Phong model. The Phong model explicitly keeps track of the monomer phosphorylation cycle and uses KaiB binding to the S phosphoform to generate negative feedback (**Figure 4B**). In the Phong model, the KaiA sequestration affinity is effectively infinite. In light of the co-IP experiment, we modify the model by assuming that the KaiA sequestration reaction is in a quasi-equilibrium with a dissociation constant for KaiA binding to the KaiBC complex, K_D (**Figure 4C**; see **SI** for mathematical details). When K_D is small (i.e., $< 10^{-3}$ μM), the modified model exhibits the same robust oscillations as the original model over a large range of %ATP, but the range of %ATP that allows for stable oscillation shrinks as K_D increases (**Figure 4E** top), and the model is unstable when K_D is in the micromolar range regardless of %ATP.

In the original Phong model, the dependence of KaiC phosphorylation on KaiA is described by a Michaelis-Menten-like function with no ultrasensitivity. In this scenario, a small increase in active KaiA leads to a proportional increase in phosphorylation, making the dephosphorylation phase of the clock strongly dependent on the strength of KaiB-mediated KaiA sequestration. To test if ultrasensitivity can increase the robustness of oscillations in the model, we introduce a phenomenological patch to the model in the form of an ultrasensitive KaiA threshold to the phosphorylation rate function, which varies as a function of %ATP and U phosphoform concentration (**Figure 4D**; see **SI** for mathematical details). Given that the ultrasensitivity is a result of substrate competition, this modification effectively introduces an inhibitory interaction between the U phosphoform and KaiA (**Figure 4B**, dashed arrow). This modification amounts to the assumption that the EC50 measured at steady state (**Figure 3A**) in the absence of KaiB corresponds to the active KaiA concentration required to re-enter the phosphorylation phase at the trough of the circadian oscillation. Remarkably, the resulting model can generate stable oscillations over a larger range of both %ATP and K_D conditions (**Figure 4E** bottom). This observation suggests that ultrasensitivity in KaiC phosphorylation plays a role in clock stability that complements the function of KaiB-dependent KaiA sequestration.

Why does ultrasensitivity in KaiC phosphorylation allow for metabolic compensation? The binding of KaiB to KaiC, and thus the sequestration and inactivation of KaiA, depends on S431 phosphorylation of KaiC (i.e., the S and D phosphoforms). At low %ATP, the maximal S and D concentrations, $[C_{\text{max}}^{S+D}]$, are lower (**Figure 4F**). Thus the maximal amount of KaiA sequestered by the KaiBC complex is smaller. This is problematic for the stability of the clock at low %ATP, since the active KaiA can promote premature KaiC U $\rightarrow$ T phosphorylation of some molecules, which can lead to phase decoherence, manifest as decaying oscillation (**Figure S7E**). The ultrasensitive stimulus-response that we report here implies that a finite amount of KaiA must be liberated from KaiB before there is a noticeable impact on KaiC phosphorylation. In other words, the inhibitory effect of ultrasensitivity is a synchronization mechanism. Importantly, the EC50 of the stimulus-response function scales with %ATP, such that the capacity of phosphorylation suppression by C^U is enhanced at low %ATP, which compensates for weaker KaiA sequestration (**Figure 4F**). This relation likely also contributes to the scaling of the phosphorylation limit cycle size with %ATP (**Figure 4G**). At higher %ATP, the EC50 is smaller and thus more KaiA needs to be sequestered to trigger dephosphorylation, which implies that higher concentrations of the S and D phosphoforms need to accumulate to enable KaiB binding. Since KaiC phosphorylation is ordered, this means that the T phosphoform concentration scales with %ATP as well.

Figure 4: Ultrasensitivity enables metabolic compensation. A) The time series of the total input KaiC phosphorylation level (purple, left scale) and residual KaiA concentration not precipitated with KaiB-FLAG (gray, right scale). B) A schematic of the oscillator model by Phong et al. (2013). Here, $B^S C^S$ and $B^D C^D$ represent the KaiB-bound S and D phosphoforms, respectively, which can sequester KaiA. The dashed line represents the effect of introducing ultrasensitivity to the model. C) A cartoon representation of introducing a KaiA sequestration affinity, K_D , into the Phong model. The original model has an effectively infinite sequestration affinity (dashed curve). D) A cartoon representation of introducing a KaiA threshold to the Michaelis-Menten-type phosphorylation rate constant in the Phong model. E) The period of the oscillator model as a function of %ATP and K_D , a measure of KaiA sequestration affinity, without (top) or with (bottom) a phosphorylation threshold. All model simulations were done with 3.5 μM of KaiC and 1.5 μM of KaiA. White regions indicate unstable or no oscillation. F) The extent to which KaiA can be sequestered by KaiB depends on the maximal S and D phosphoform concentration, $[C_{\text{max}}^{S+D}]$, achieved over the phosphorylation cycle. The scaling of the EC50 of the phosphorylation stimulus-response function, which is a measure of the capacity of the U phosphoform to suppress KaiA activity, compensates for the scaling of $[C_{\text{max}}^{S+D}]$ with %ATP. G) The scaling of EC50 and $[C_{\text{max}}^{S+D}]$ contribute to the scaling of the phosphorylation cycle size with %ATP. As %ATP increases, EC50 decreases and thus higher concentrations of the S and D phosphoforms are required to sequester active KaiA and trigger dephosphorylation. These dynamics enable increased accumulation of the T phosphoform at higher %ATP.

Discussion

In this work we undertook a data-driven kinetic modeling approach to understand the metabolic sensitivity of the KaiA-KaiC subsystem, part of the *S. elongatus* circadian oscillator. We constructed a detailed yet mechanistically agnostic kinetic model, which was fit to extensive experimental measurements of KaiC phosphorylation kinetics within a Bayesian parameter estimation framework. Approaches that are similar in spirit have been pursued in eukaryotic systems (e.g., Forger and Peskin, 2003; Locke et al., 2005; Mirsky et al., 2009; Relógio et al., 2011; Kim and Forger, 2012). However, owing to the greater complexity of eukaryotic clocks, these studies combined direct experimental measurements, cost function optimization, and hand tuning of selected parameters to account for unknown or unconstrained biochemical processes. Because the Kai system can be studied as a well-defined mixture of purified components, the participating molecular species are known, and all the parameters in the model can be treated in a consistent manner to enable objective comparison of mechanisms underlying collective oscillations.

This data-driven approach is to be contrasted with the more common hypothesis-driven, forward modeling approach, whereby a model is built to examine how features of the oscillator arise from proposed mechanisms. This hypothesis-driven approach has been employed extensively in the study of the cyanobacterial clock. These studies have revealed insights into specific aspects of the oscillator function, such as entrainment (Brettschneider et al., 2010; Leypunskiy et al., 2017), synchronization (Yoda et al., 2007; van Zon et al., 2007; Sasai, 2019), irreversibility (Cao et al., 2015), and robustness against variations in temperature (Hatakeyama and Kaneko, 2012; François et al., 2012; Kidd et al., 2015; Murayama et al., 2017), ATP/ADP concentration (Phong et al., 2013; Paijmans et al., 2017a; del Junco and Vaikuntanathan, 2019), protein copy numbers (Brettschneider et al., 2010; Lin et al., 2014; Chew et al., 2018), and environmental noise in general (Pittayakanchit et al., 2018; Monti et al., 2018). This hypothesis-driven approach is pedagogically powerful but gives little indication of the range of the parameter space consistent with a proposed mechanism, which makes it difficult to quantify the uncertainties of model predictions and validate them experimentally.

Not all parameters in our model were fully constrained by the data, as expected given the complexity of the model (Gutenkunst et al., 2007). Nevertheless, the ensemble of parameter sets still led to consistent predictions. In particular, the model revealed unexpected ultrasensitivity in KaiC phosphorylation as a function of KaiA, which we confirmed experimentally. The source of ultrasensitivity in the model is a substrate competition mechanism that arises from the differential affinity of ADP- and ATP-bound KaiC for KaiA. Previous studies have considered the importance of the differential affinity of KaiA for KaiC states but have focused on phosphorylation (Paijmans et al., 2017b; Mori et al., 2018). We note here that the ultrasensitivity in KaiC phosphorylation that we discovered acts to regulate KaiA activity at a different phase of the cycle than the ultrasensitivity in KaiB-dependent KaiA sequestration that arises from opposing S and T phosphorylations within hexamers (Lin et al., 2014). Presumably, the presence of nonlinearities and delayed feedback at multiple steps in a molecular oscillator allows the system to achieve greater robustness (Kim and Forger, 2012; Jolley et al., 2012; Dovzhenok et al., 2015; Pett et al., 2016).

We hypothesized that ultrasensitivity in KaiC phosphorylation plays a role in stabilizing the oscillator at low %ATP conditions by suppressing premature phosphorylation during the dephosphorylation stage and thus promoting phase coherence. Currently, the Kai oscillator model most robust against yet tunable by metabolic conditions appears to be that of Paijmans et al. (2017a,b). In the Paijmans model, metabolic compensation is achieved both at the hexamer and ensemble level. At the hexamer level, the onset of dephosphorylation is primarily controlled by the antagonistic effects of the T and S phosphoforms. Since fewer subunits in the T phosphoform accumulate at low %ATP, fewer subunits in the S phosphoform are needed to trigger dephosphorylation; therefore, the reduced amplitude of oscillation counteracts the slower phosphorylation rate at low %ATP. At the ensemble level, low %ATP limits the fraction of hexamers that are able to trigger dephosphorylation before the onset of KaiB-mediated delayed inhibition; this makes the dephosphorylation phase shorter, which compensates for the longer phosphorylation phase. It is worth noting that the Paijmans model is not oscillatory when %ATP reaches below 50% partly due to phase decoherence during dephosphorylation, an issue that can potentially be addressed with ultrasensitivity in KaiC phosphorylation. In our model the coupling between KaiA binding affinity and KaiC nucleotide-bound states is

critical in generating ultrasensitivity, a feature that is missing in the Paijmans model. It remains an open question whether a hexameric model with no such coupling can nevertheless produce ultrasensitivity in KaiC phosphorylation (see **SI** for further comparison between this work and the Paijmans model).

In *S. elongatus*, the Kai oscillator is embedded in a transcription-translation feedback loop (Kitayama et al., 2008; Zwicker et al., 2010; Qin et al., 2010b). However, with the exception of peroxiredoxin oxidation cycles (O'Neill et al., 2011; Edgar et al., 2012), cell-autonomous circadian rhythms in eukaryotes are thought to be generated by interlocked transcription-translation feedback loops (Novák and Tyson, 2008); the cooperative autoregulation of transcription is a key source of nonlinearity and robustness in the circuit (e.g., Leloup et al., 1999; Gonze et al., 2002; Leloup and Goldbeter, 2003; Locke et al., 2005; Brown et al., 2012). Our results raise the possibility that post-translational protein modifications and protein-protein interactions may also contribute to robustness by introducing ultrasensitivity, even if these processes do not generate self-sustaining rhythms that can be decoupled from transcription. In general, it is clear that post-translational steps such as (de)phosphorylation (Gallego and Virshup, 2007; Reischl and Kramer, 2011; Zhou et al., 2015; Fustin et al., 2018), protein degradation (Gallego and Virshup, 2007; Reischl et al., 2007), and complex formation (Kim and Forger, 2012) play an important role in eukaryotic circadian oscillators, but to our knowledge there is currently no complete experimental characterization of the stimulus-response relations of these processes.

Materials and Methods

Computational methods

Treatment of nucleotide exchange Here we derive (4) in Results. The nucleotide exchange process is in principle a two-step reaction that includes an *apo* intermediate state of KaiC, i.e.,

where we have omitted the free ATP and ADP from the chemical equation. Here, k_r^{TP} and k_r^{DP} are the dissociation rate constants and k_{on}^{TP} and k_{on}^{DP} are the binding rate constants for ATP and ADP, respectively. Since KaiC requires nucleotides for hexamerization (Hayashi et al., 2003, 2006; Mutoh et al., 2013), the *apo* state of KaiC is presumably both thermodynamically and kinetically unstable in the presence of saturating amount of nucleotide (5 mM in our experiments). Therefore, under the assumption that the KaiC *apo* state is in a quasi-steady state throughout the reactions, we can eliminate the *apo* state and model nucleotide exchange as a one-step reaction

where

$$k_{TP}^A = k_r^{DP} \frac{[ATP]}{[ATP] + K_{on}[ADP]} \quad (8)$$

$$k_{DP}^A = k_r^{TP} \left(1 - \frac{[ATP]}{[ATP] + K_{on}[ADP]} \right) \quad (9)$$

and $K_{on} = k_{on}^{DP} / k_{on}^{TP}$ is a ratio of the two nucleotide binding rate constants.

We make two further simplifying assumptions. First, we assume the on rates are completely diffusion controlled and are thus the same for ATP and ADP, which allows us to set $K_{on} = 1$. Second, based on fit results (Figure S4F) showing that the posterior for k_r^{TP} has a long tail to negative infinity in log space, we follow the approach proposed by Transtrum and Qiu (2014) and set $k_r^{TP} = 0$; i.e., the dwell time of ATP-bound states are sufficiently long such that a bound ATP cannot be released without first giving up its γ -phosphate group. This assumption implies that the only ways for KaiC to enter an ADP-bound state are through hydrolysis and phosphorylation, and solution ADP has no effect on the system except to slow down the ADP to ATP exchange process. With these two assumptions, we eliminate (9) and (8) reduces to (4).

Model parameterization In Results, we introduced a model parameterization scheme in which rate constants for phosphotransfer, nucleotide exchange, KaiA (un)binding, and ATP hydrolysis reactions depend on the participating molecular species. Although we use this independent-rate scheme to interpret the model, including computing the sensitivity ODEs, during the fitting itself we represent species-dependent effects by modifying each of the six basic rate constants (k_p , k_d , k_a , k_b , k_r^{DP} , and k_h) by multiplicative Δk factors. For example, the KaiA dissociation rate $k_b^{T,TP} = k_b \Delta k_b^{T,TP}$ is represented by the product between a base rate k_b and a modifier $\Delta k_b^{T,TP}$ (compare Figure S1 and Figure S10). The multiplicative-factor scheme introduces 38 Δk parameters. Because of the requirement for detailed balance (see below), only 34 of these parameters are free; these free parameters are listed on Table 2. The advantage of the multiplicative parameterization scheme is that it facilitates ℓ^1 regularization, discussed below.

Detailed balance All elementary reactions, except ATP hydrolysis and nucleotide exchange, are assumed to occur in equilibrium, and thus the free energy change over each reversible cycle must be equal to zero.

Table 2: Full model parameters and their priors

Category	Species-dependent effect	Parameters	Prior	Unit
Basic	N/A	$k_h, k_p, k_d, k_a, k_b, k_r^{DP}$	$10^{\mathcal{N}(\mu, 3)}$ †	s^{-1} §
Nucleotide exchange	KaiA & phos.*	$\Delta k_{TP}^{A,T}, \Delta k_{TP}^{A,S}, \Delta k_{TP}^{A,D}$		
Hydrolysis	phos.	$\Delta k_h^T, \Delta k_h^S, \Delta k_h^D$		
	KaiA & phos.	$\Delta k_h^{A,U}, \Delta k_h^{A,T}, \Delta k_h^{A,S}, \Delta k_h^{A,D}$		
KaiA on	nuc.* & phos.	$\Delta k_a^{U,DP}, \Delta k_a^{D,DP}, \Delta k_a^{D,TP}$		
KaiA off	nuc. & phos.	$\Delta k_b^{U,DP}, \Delta k_b^{T,DP}, \Delta k_b^{S,DP}, \Delta k_b^{D,DP}, \Delta k_b^{T,TP}, \Delta k_b^{S,TP}, \Delta k_b^{D,TP}$	$10^{\text{Laplace}(\mu, 1)}$ †	N/A
	phos.	$\Delta k_p^{US}, \Delta k_d^{SU}, \Delta k_p^{TD}, \Delta k_d^{DT}, \Delta k_p^{SD}, \Delta k_d^{DS}$		
(De)phosphorylation	KaiA & phos.	$\Delta k_p^{A,UT}, \Delta k_d^{A,TU}, \Delta k_p^{A,TD}, \Delta k_d^{A,DT}, \Delta k_p^{A,SD}, \Delta k_d^{A,DS}, \Delta k_p^{A,US}, \Delta k_d^{A,SU}$		
Global error	N/A	σ^2	Inv-Gamma (1, 0.01)	μM^2
Initial conditions	N/A	$[C_{TP}^U]_0, [C_{DP}^U]_0, [C_{TP}^T]_0, [C_{DP}^T]_0, [C_{TP}^S]_0, [C_{DP}^S]_0, [C_{TP}^D]_0, [C_{DP}^D]_0$	Dirichlet(a) ‡	μM

* phos., phosphoform; nuc., nucleotide-bound state

† The mean of the priors, μ , is zero unless specified by **Table 5**.

§ or $s^{-1} \cdot \mu M^{-1}$ for the second-order rate constant k_a .

‡ $a = (20, 100, 1, 1, 1, 1, 1, 1)$; points drawn from the distribution are scaled by the total KaiC concentration. The support of the Dirichlet distribution implies that only seven of the eight initial conditions are free fitting parameters.

In practice, this means that the product of all rate constants in the forward direction of each cycle listed on **Table 3** must be equal to that in the reverse direction (see also **Figure S10A**). This introduces an additional algebraic constraint for each such cycle, which is used to eliminate one free Δk parameter. In total, one can eliminate four such arbitrarily chosen parameters.

Fitting data set To constrain the model parameters, we collected experimental measurements that characterized different aspects of the KaiA-KaiC subsystem, which are summarized in **Table 4**.

The dephosphorylation reaction taken from Rust et al. (2011) constrains the dephosphorylation rates and the ATP hydrolysis rates of KaiC in the absence of KaiA, because the model structure dictates that

Table 3: Detailed balance conditions

Cycle	Detailed balance condition
$\{C_{TP}^S, C_{DP}^D, {}^AC_{DP}^D, {}^AC_{TP}^S\}$	$\delta k_a^{S,TP} = \Delta k_b^{S,TP} \frac{\Delta k_d^{A,DS}}{\Delta k_p^{A,SD}} \frac{\Delta k_a^{D,DP}}{\Delta k_b^{D,DP}}$
$\{C_{TP}^T, C_{DP}^D, {}^AC_{DP}^D, {}^AC_{TP}^T\}$	$\delta k_a^{T,TP} = \Delta k_b^{T,TP} \frac{\Delta k_d^{A,DT}}{\Delta k_p^{A,TD}} \frac{\Delta k_a^{D,DP}}{\Delta k_b^{D,DP}}$
$\{C_{TP}^U, C_{DP}^T, {}^AC_{DP}^T, {}^AC_{TP}^U\}$	$\delta k_a^{T,DP} = \Delta k_b^{T,DP} \frac{\Delta k_p^{A,UT}}{\Delta k_d^{A,TU}}$
$\{C_{TP}^U, C_{DP}^S, {}^AC_{DP}^S, {}^AC_{TP}^U\}$	$\delta k_a^{S,DP} = \Delta k_b^{S,DP} \frac{\Delta k_p^{A,US}}{\Delta k_d^{A,SU}}$

dephosphorylation requires alternating phosphotransfer and hydrolysis reactions. There is currently no direct measurement of CII hydrolysis rate in the presence of KaiA. However, the maximum ADP production rate of KaiC in the presence of 1.2 μM of KaiA was reported to be $29.8 \pm 5.1 \text{ KaiC}^{-1} \cdot \text{day}^{-1}$ (Terauchi et al., 2007), which we take as an upper bound on the average CII hydrolysis rate in phosphorylation reactions for all $[\text{KaiA}] = 0.375, 0.75, \text{ and } 1.50 \mu\text{M}$ conditions.

Because the phosphorylation reactions were measured in the presence of varying %ATP and higher [ADP] inhibits phosphorylation by slowing down nucleotide exchange (see equation 4), they provide indirect constraints on the nucleotide exchange rate. Similarly, because the reactions were measured in the presence of varying $[\text{KaiA}]$ conditions, they provide direct constraint on the phosphorylation rates of KaiC with and without KaiA, as well as the KaiA binding affinity, i.e., the ratio of KaiA on/off rates. Although there are direct experimental measures of KaiA binding and dissociation (Kageyama et al., 2006; Mori et al., 2018), these results cannot be directly mapped onto model rate constants. This is primarily because the KaiC nucleotide-bound fractions are not reported in these experiments, or, in the case of phosphomimetic mutants, it is unclear if the mutations affect nucleotide binding affinities. As a consequence, the experimental constraints on KaiA on/off rates enter through the priors rather than the likelihood function, in contrast to the other data (see below).

Initial conditions For each phosphorylation reaction in **Table 4**, we solve the ODE model with the corresponding $[\text{KaiA}]$ and %ATP condition. The predicted phosphoform composition, as well as the ATP hydrolysis rate when appropriate, is compared to the experimental measurements in the Bayesian parameter estimation formalism described below. However, since the experimental data do not resolve the initial conditions for all 16 KaiC species in the model, we have chosen to directly estimate the initial concentrations as free model parameters. We take $t = 0$ to be the time point at which KaiA is mixed with KaiC, and thus all eight KaiA-bound KaiC states are assumed to have zero concentration at the onset of the experiment. Because total KaiC concentration is conserved, this introduces seven additional free parameters (see **Table 2**).

For the dephosphorylation reaction, we do not estimate the initial conditions. To mimic the way the experiment was done, the dephosphorylation reaction is simulated in two stages. In the first stage, we assume that 3.4 μM of dephosphorylated KaiC is phosphorylated in the presence of 1.3 μM KaiA and 100% ATP for 20 h. Since the protein is initially dephosphorylated, we assume that $[C_{TP}^U]_0 = 3.4 \mu\text{M}$ while the concentrations of all other KaiC species are set to zero. In the second stage, we simulate the autodephosphorylation reaction after KaiA pull-down, which corresponds to eliminating all free KaiA as well as KaiA-bound KaiC species from the simulation. The amount of KaiC lost in the pull-down experiment was not reported in the original experiment (Rust et al., 2011). We therefore make the assumption that the amount of KaiC lost in the pull-down experiment in the simulation is exactly equal to that in the experiment for the purpose of

Table 4: Fitting data set

Measurement (source)	Temperature (°C)	[KaiA]* (μM)	%ATP	Time points	Phosphoform
Phosphorylation (this work)	30	0.375, 0.75, 1.50, 3.00, 4.50, 6.00	10, 25, 100	8	U, T, D [†]
Dephosphorylation (Rust et al., 2011)	30	1.4	100	21	
ADP production (Terauchi et al., 2007)	30	1.2	100	1	N/A
KaiA on/off rates (Kageyama et al., 2006)	25	Variable	100	N/A	Likely U
KaiA dwell time (Mori et al., 2018)	25–28	1.0	N/A	N/A	T, S, D [‡]

* We report here on the KaiA monomer concentration. However, since KaiA functions as a dimer, all KaiA concentration is divided by two in the models.

[†] The conservation of mass constraint implies that one of the four phosphoforms is not a free state variable. We have chosen the S phosphoform to be the constrained state variable.

[‡] Phosphomimetic mutants

computing the likelihood function.

Bayesian parameter estimation We directly fit numerically integrated ODEs to experimental data in the Bayesian parameter estimation framework (Wasserman, 2000). The best fit model parameters, $\hat{\theta}$, are obtained from the maximum a posteriori estimator:

$$\hat{\theta} = \arg \max_{\theta} p(\theta|\mathcal{D}), \quad (10)$$

where $p(\theta|\mathcal{D})$ is the posterior distribution of the parameters θ , conditioned on the training data set $\mathcal{D}$. Using Bayes' theorem, the posterior distribution can be written as

$$p(\theta|\mathcal{D}) = \frac{\mathcal{L}(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})}. \quad (11)$$

Here, $p(\theta)$ is the prior distribution, which represents subjective belief in the model parameters θ prior to experimental input; $\mathcal{L}(\mathcal{D}|\theta)$ is the likelihood function, which represents a probabilistic model of the experimental data set $\mathcal{D}$ given a particular model $\mathcal{M}$ (implicit in the formulas) that depends on the parameters θ ; $p(\mathcal{D})$ is the evidence, which is analogous to the partition function in statistical mechanics. Note that the evidence $p(\mathcal{D})$ does not depend on the parameter choice, and is thus an irrelevant constant for the purpose of parameter estimation. The specific choices for the functional forms of the likelihood function and priors are discussed further below.

Model priors The priors for all model parameters used in Bayesian parameter estimation are given in **Table 2**. Here, the choice of the prior distributions is primarily motivated by the need for regularization (see below). In addition, as discussed above, the experimental measurements on KaiA binding kinetics are

Table 5: Priors incorporating KaiA on/off constraints

Parameter	Prior mean (μ)	Experimental measurements	Source
k_a	$\log k_{a,\text{exp}}$	$k_{a,\text{exp}} = 0.0279 \text{ s}^{-1} \cdot \mu\text{M}^{-1}$	Kageyama et al. (2006)
k_b	$\log k_{b,\text{exp}}$	$k_{b,\text{exp}} = 0.0663 \text{ s}^{-1}$	
$\Delta k_b^{\text{T,DP}}, \Delta k_b^{\text{T,TP}}$	$-\log \tau_{b,\text{exp}}^{\text{T}} k_{b,\text{exp}}$	$\tau_{b,\text{exp}}^{\text{T}} = 1.0 \text{ s}$	Mori et al. (2018)
$\Delta k_b^{\text{S,DP}}, \Delta k_b^{\text{S,TP}}$	$-\log \tau_{b,\text{exp}}^{\text{S}} k_{b,\text{exp}}$	$\tau_{b,\text{exp}}^{\text{S}} = 0.43 \text{ s}$	
$\Delta k_b^{\text{D,DP}}, \Delta k_b^{\text{D,TP}}$	$-\log \tau_{b,\text{exp}}^{\text{D}} k_{b,\text{exp}}$	$\tau_{b,\text{exp}}^{\text{D}} = 0.26 \text{ s}$	

incorporated into the priors rather than the likelihood function (Table 5). Note that all the rate constants and their multiplicative factors are estimated in the log space (base 10). This ensures that all rate constants are positive.

ℓ^1 regularization As model complexity grows, the constraint of experimental data on the underlying mechanism weakens. To address this problem, we impose sparsity on the species-dependent effects (i.e., the Δk factors) using ℓ^1 regularization (Tibshirani, 1996). This is accomplished in the Bayesian parameter estimation framework by using a Laplace prior centered at zero in the log parameter space (or one in the real space). Intuitively, the Laplace prior imposes sparsity by forcing the (marginalized) posterior distribution for each $\log \Delta k$ to peak at zero unless there is experimental evidence in the fitting data set to suggest otherwise. Since the Δk s are multiplicative factors modifying the six basic rate constants, $\log \Delta k = 0$ implies that the species-dependent rate is identical to that of the base rate. This method is directly analogous to the use of the lasso estimator in the context of linear least squares models. To see this, consider the Laplace distribution

$$p(\theta; b) = \frac{1}{2b} e^{-\|\theta\|_1/b} \quad (12)$$

where $\|\theta\|_1$ is the ℓ^1 -norm of θ . Then from (11) the negative log-posterior distribution becomes

$$-\ln p(\theta|\mathcal{D}) = -\ln \mathcal{L}(\mathcal{D}|\theta) + \lambda \|\theta\|_1 + \text{const.} \quad (13)$$

where $\lambda = 1/b$. In a linear model $Y = X\beta + \epsilon$ where Y is the response vector, X is the design matrix, β is the parameter vector, and ϵ is the error vector, the negative log-likelihood function reduces to the sum of squares $\|Y - X\beta\|_2^2/N$, where N is the number of dependent variables. Thus, maximizing the posterior is equivalent to minimizing the sum of squares with an ℓ^1 penalty, which is the lasso estimator.

Likelihood function To determine the functional form of the likelihood function, consider a kinetic experiment where measurements on some observables y are made at a set of time points $\{t_i\}$ with uncertainties $\{\sigma_i\}$. If we assume that the experimental errors σ_i are independent and normally distributed, then the likelihood function is given by

$$\mathcal{L}(\mathcal{D}|\theta) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-[y_{\text{exp}}(t_i) - y_{\text{model}}(t_i; \theta)]^2 / 2\sigma_i^2} \quad (14)$$

In other words, the likelihood function gives the probability for observing a given data set provided that the model prediction is true. In practice, all posterior evaluations are done in the log space (base e) for numerical stability. Thus, taken together, (11) can be rewritten as,

$$\ln p(\theta|\mathcal{D}) = -\sum_i \frac{[y_{\text{exp}}(t_i) - y_{\text{model}}(t_i; \theta)]^2}{2\sigma_i^2} + \ln p(\theta) + \text{const.} \quad (15)$$

For the sake of simplicity, we assume that there is a single global error, σ , for all (de)phosphorylation measurements, which is then estimated during fitting as a hyperparameter (see **Table 2**).

The choice of the Gaussian likelihood function applies to all (de)phosphorylation data sets, but not the hydrolysis constraint, which only provides an upper bound on the average hydrolysis rate per day (Terauchi et al., 2007). Therefore, for the hydrolysis data a “half harmonic” is used as the log-likelihood function:

$$\ln \mathcal{L}(\mathcal{D}|\theta) = \begin{cases} -\frac{([\text{ADP}]_{\text{model}}(\theta) - [\text{ADP}]_{\text{exp}})^2}{2\sigma_h^2}, & [\text{ADP}]_{\text{exp}} \leq [\text{ADP}]_{\text{model}} \\ 0, & 0 \leq [\text{ADP}]_{\text{model}} < [\text{ADP}]_{\text{exp}} \end{cases} \quad (16)$$

The total amount of ADP produced by the model during a phosphorylation reaction over 12 h, $[\text{ADP}]_{\text{model}}$, is given by the sum of all P_i production over time plus all ADP-bound KaiC species at $t = 12$ h.

The log-likelihood values from the appropriate phosphorylation, dephosphorylation, and hydrolysis reactions are added together to determine the log-likelihood of the data set for each given model parameter choice. Since the phosphorylation data set is much larger than the dephosphorylation data set, the fitting procedure tends to favor fitting the phosphorylation data set at the expense of fitting the dephosphorylation data set. To overcome this problem, the log-likelihood function for the dephosphorylation reaction is multiplied by a factor of 4 to increase the weight of the dephosphorylation data points.

Model fitting procedure To determine the posterior mode and the uncertainties associated with the estimate, we employ a heuristic combination of ensemble MCMC sampling and numerical optimization methods (**Figure S3A**). This fitting procedure can be divided into four steps that are analogous to those in a genetic algorithm:

1. Initialization. An ensemble MCMC method evolves a set of random walkers (i.e., parameter sets) simultaneously; we thus begin by drawing 224 walkers from the prior distribution $p(\theta)$ and use these walkers for simulated annealing (Kirkpatrick et al., 1983; Kirkpatrick, 1984). In annealing, instead of sampling $p(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta)p(\theta)$ (in the log space), a flattened distribution $\mathcal{L}(\mathcal{D}|\theta)^\beta p(\theta)$ is sampled with an annealing schedule of $\beta = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$. Note that instead of letting $\beta \rightarrow \infty$, the simulation ends with the target distribution at $\beta = 1.0$. Each temperature is sampled over 20,000 steps.
2. Selection. The fitnesses of the walkers are determined by their log-posterior values (equation 15). 10 walkers from the best 300 parameter sets sampled in the previous step are chosen and subjected to numerical optimization to find the nearby local maximums, which are then used to seed a sampling run in the next step. In the spirit of elitist selection, the best walker is always included for the next generation.
3. Recombination and mutation. The initial walkers for the sampling run are generated by adding a Gaussian noise $\mathcal{N}(0, 0.001)$ to the 10 optimized walkers, and the number of initial walkers centered around each optimized walker θ_j is given by

$$224 \frac{p(\theta_j|\mathcal{D})^\beta}{\sum_k p(\theta_k|\mathcal{D})^\beta}. \quad (17)$$

That is, the proportion of the initial walkers generated from each optimized walker is weighted by its posterior value with a temperature factor of $\beta = 0.6$; the temperature factor is chosen to allocate more walkers to optimized walkers with lower posterior values. The sampling run consists of 50,000 steps. Note that the purpose of the Gaussian noise is to ensure that the proposal distribution is valid for any pair of walkers for the ensemble MCMC method (see below), rather than to control the mutation strength, as is done in evolution strategy (Beyer and Schwefel, 2002).

4. Termination. The best walkers from the sampling run are compared to the optimized walkers. If the best walkers do not escape to new local maximum(s) with higher posterior value(s), then the procedure is terminated after an additional 50,000 sampling steps. If, however, new local maximum(s) are discovered during sampling, the algorithm loops back to the selection step. This process is repeated until no better local maximum is discovered at the end of sampling. Unless otherwise specified, only the last 30,000 sampling steps (downsampled every 100 steps) are used for post-analysis.

In general, the number of walkers in ensemble MCMC needs to be larger than the number of free parameters; here the number 224 is chosen to optimize parallel performance on a local computer cluster (8 nodes $\times$ 28 CPU cores/node).

We found that this procedure outperformed either ensemble MCMC or numerical optimization by itself (compare **Figure S3A** and **B**). For the full model we ran this procedure three times to assess the reproducibility of the fit (see **SI** for further discussion).

Markov chain Monte Carlo One major challenge in efficient MCMC sampling in systems biology is that the target distributions are often poorly scaled. In the context of ODE kinetic modeling, this means that different reaction rates and their associated uncertainties can be separated by several orders of magnitude. This is almost certainly true for the KaiA-KaiC subsystem because, among other things, the experimentally measured rates of KaiA binding and dissociation are much faster than the ATP hydrolysis rate of KaiC. Without *a priori* knowledge of the natural time scales, conventional MCMC schemes are inefficient in such sampling problems, because only very small displacements are accepted at appreciable rates. In this work we employ an ensemble MCMC method developed by Goodman and Weare (2010). The advantage of the Goodman-Weare algorithm is that it is affine invariant, which means that it performs equally well for isotropic and poorly scaled measures, providing that the two can be related by a linear transformation of the coordinate system. This appears to be the case for the present problem since the Goodman-Weare algorithm vastly outperforms a standard Metropolis-Hastings scheme with a (preconditioned) Gaussian proposal distribution (**Figure S3B**).

In brief, the Goodman-Weare algorithm evolves an ensemble of walkers, rather than a single one. At each step, individual walker positions are updated sequentially. For a given walker θ_k at step τ , a walker θ_j is drawn randomly from the rest of the ensemble and a new position, η , on the line connecting θ_k and θ_j is proposed by a “stretch move”

$$\eta = \theta_j + z [\theta_k(\tau) - \theta_j] \quad (18)$$

where z is a random number drawn from the distribution

$$Z \sim g(z; \alpha) = \begin{cases} 1/\sqrt{z}, & z \in [1/\alpha, \alpha] \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

The “stretch factor” α is a tunable parameter that controls the step size. In an N -dimensional parameter space, the new walker η is accepted with the probability

$$q = \min \left(1, z^{N-1} \frac{p(\eta|\mathcal{D})}{p(\theta_k(\tau)|\mathcal{D})} \right) \quad (20)$$

which guarantees that the scheme obeys detailed balance and thus converges to the target distribution $p(\theta|\mathcal{D})$ as $\tau \rightarrow \infty$. Note that no derivative of the posterior distribution is required to draw from the proposal distribution. In this work we use $\alpha = 1.1$, which gives an average acceptance rate of 47% in steps 3 and 4 of the fitting procedure.

Numerical optimization The numerical optimization method used in this work is a modified version of Powell’s method (Powell, 1964; Press et al., 2007). Briefly, given an initial guess and direction set, which

is usually the Cartesian coordinate set, Powell's method performs a line search to minimize the objective function, here $-\ln p(\theta|\mathcal{D})$, sequentially along each vector in the direction set. The direction set is then updated by replacing the direction of largest decrease in the objective function in the current iteration with the displacement vector from the estimated minimum at the beginning to that at the end of the line minimizations, provided that certain technical conditions are met to avoid the build-up of linear dependence in the direction set. This process is repeated until a convergence threshold is met. Note that unlike the original method, the modified Powell's method does not guarantee that the vectors in the direction set are mutually conjugate.

Similar to the Goodman-Weare algorithm, Powell's method is derivative-free. For the current system Powell's method converges faster than the Nelder-Mead method (Nelder and Mead, 1965), another commonly used derivative-free method, although the Nelder-Mead method appears less prone to becoming trapped in local metastable states (**Figure S3C**).

Software implementation The fitting procedure is implemented in an in-house Python (Oliphant, 2015) script that interfaces with several existing Python modules: numerical integration of the model ODEs is done using the `odespy` package (Langtangen and Wang, 2015) with the BDF method; the Goodman-Weare algorithm is implemented in `emcee` (version 2.2.1) (Foreman-Mackey et al., 2013); Powell's method and the Nelder-Mead method are implemented in `scipy` (version 1.2.1) (Eric Jones et al., 2001). The derivative evaluation step in ODE integration is accelerated using `numba` (Lam et al., 2015), and the script is parallelized using `mpi4py` 2.0.0 (Dalcín et al., 2005, 2008; Dalcín et al., 2011).

The most computationally expensive step in the fitting procedure is the MCMC sampling, because each move requires multiple ODE evaluations to compute the posterior function. With 224 walkers and 8 nodes (each with 28 Intel E5-2680v4 2.4GHz cores), the speed of MCMC sampling is 46,000 steps/hour.

Model comparison In the preceding sections all definitions of probability distributions implicitly assume that there is a model $\mathcal{M}$ with a well-defined functional form, whose parameters θ need to be determined. For the sake of model comparison, we make this assumption explicit and rewrite (11) as

$$p(\theta|\mathcal{D}, \mathcal{M}) = \frac{\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})p(\theta|\mathcal{M})}{p(\mathcal{D}|\mathcal{M})}. \quad (21)$$

To compare two models $\mathcal{M}_i$ and $\mathcal{M}_j$, we need to compare the posterior probabilities for each model, usually in the form of their ratios

$$\frac{p(\mathcal{M}_i|\mathcal{D})}{p(\mathcal{M}_j|\mathcal{D})} = \frac{p(\mathcal{D}|\mathcal{M}_i) p(\mathcal{M}_i)}{p(\mathcal{D}|\mathcal{M}_j) p(\mathcal{M}_j)}. \quad (22)$$

Assuming that we have no prior preference for any model, the ratio becomes the Bayes factor

$$B_{ij} = \frac{p(\mathcal{D}|\mathcal{M}_i)}{p(\mathcal{D}|\mathcal{M}_j)}, \quad (23)$$

which we adopt as the metric for model comparison.

The primary difficulty in computing the Bayes factor is thus estimating the evidence, or the marginal likelihood function, for each $\mathcal{M}_i$. There are several methods for computing the evidence (Vyshemirsky and Girolami, 2008). Here we derive a formula compatible with the ensemble MCMC scheme that is directly analogous to free energy perturbation (Zwanzig, 1954). For the sake of simplicity, we drop the model index i from this point on. First, note that for a given model $\mathcal{M}$,

$$p(\mathcal{D}|\mathcal{M}) = \int \mathcal{L}(\mathcal{D}|\theta, \mathcal{M})p(\theta|\mathcal{M}) d\theta = \langle \mathcal{L}(\mathcal{D}|\theta, \mathcal{M}) \rangle_{p(\theta|\mathcal{M})} \quad (24)$$

where the first equality follows from the law of total probability and the second equality assumes that the prior $p(\theta|\mathcal{M})$ is normalized (as a probability density function of θ). Equation (24) suggests that the marginal

likelihood function can be computed by estimating the average of the likelihood function $\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})$ against the prior. Using MCMC to estimate this integral is inefficient since there is very little overlap between the likelihood function and the prior for the models of interest. Instead, we define

$$q_\lambda(\theta) = \mathcal{L}(\mathcal{D}|\theta, \mathcal{M})^\lambda p(\theta|\mathcal{M}) \quad \text{and} \quad Z_\lambda = \int q_\lambda(\theta) d\theta$$

for $0 \leq \lambda \leq 1$ and then note that (24) can be recast as

$$p(\mathcal{D}|\mathcal{M}) = \frac{Z_1}{Z_0} = \left(\frac{Z_{\lambda_0}}{Z_{\lambda_1}} \frac{Z_{\lambda_1}}{Z_{\lambda_2}} \dots \frac{Z_{\lambda_{N-1}}}{Z_{\lambda_N}} \right)^{-1} \quad (25)$$

for $0 = \lambda_0 < \lambda_1 < \dots < \lambda_N = 1$, and the λ s are chosen to allow for sufficient overlap between successive $q_\lambda(\theta)$ s. Each fraction in (25) is given by

$$\frac{Z_{\lambda_{n-1}}}{Z_{\lambda_n}} = \frac{\int \mathcal{L}(\mathcal{D}|\theta, \mathcal{M})^{\lambda_{n-1}-\lambda_n} q_{\lambda_n}(\theta) d\theta}{\int q_{\lambda_n}(\theta) d\theta} = \langle \mathcal{L}(\mathcal{D}|\theta, \mathcal{M})^{\lambda_{n-1}-\lambda_n} \rangle_{q_{\lambda_n}} \quad (26)$$

Therefore,

$$p(\mathcal{D}|\mathcal{M}) = \prod_{n=1}^N \langle \mathcal{L}(\mathcal{D}|\theta, \mathcal{M})^{\lambda_{n-1}-\lambda_n} \rangle_{q_{\lambda_n}}^{-1} = \prod_{n=1}^N \left\langle e^{\ln q_{\lambda_{n-1}}(\theta) - \ln q_{\lambda_n}(\theta)} \right\rangle_{q_{\lambda_n}}^{-1} \quad (27)$$

where the averages $\langle \cdot \rangle_{q_{\lambda_n}}$ can be approximated with MCMC. Equation (27) is a version of the free energy perturbation formula. Note that (27) requires that the likelihood function $\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})$ be properly normalized (as a probability density function of $\mathcal{D}$), but does not require the prior $p(\theta|\mathcal{M})$ to be normalized, as any missing normalization constant cancels out in each term of the product.

For each simplified model in **Table 1** and **Table S1**, the ensemble of walkers from the last time step of the model fitting procedure is used to initialize an MCMC sampling run with $\lambda = 1.00$. The lambda value is reduced by 0.01 at each subsequent stage until λ reaches 0.01. Each stage is sampled for 2,000 time steps using the Goodman-Weare ensemble sampler Goodman and Weare (2010). Only the last 1,000 time steps from each stage is used to compute the ensemble average in (27). The Bayes factors are then computed as the ratios of the evidence for the full model to each simplified model.

Refitting The steady-state KaiC phosphorylation measurements (**Figure 3B** and **Figure S8B**) are fit to the full model using Powell's method, starting from the best fit based on the training data set. The priors on the kinetic parameters (**Table 2**) are centered on the best fit values, so that the refit model can be interpreted as the "minimal" perturbation to the best fit that enables agreement with the steady-state measurement.

Curve fitting The experimentally determined stimulus-response relations for KaiC S431A (**Figure 3C** and **F**) are fit to the simple inhibitor ultrasensitivity scheme described in Box 5 of Ferrell and Ha (2014b). Using their notation, the amount of phosphorylated protein substrate (%XP) as a function of kinase concentration ($[K]$) is given by

$$\%XP([K]) = P_{\max} \frac{K_1[I] + K_1K_2 - K_1[K] + 2K_2[K] - K_1\sqrt{[I]^2 + 2(K_2 - [K])[I] + (K_2 + [K])^2}}{2K_1[I] - 2(K_1 - K_2)(K_1 + [K])} + b \quad (28)$$

Here, $P_{\max}$, $[I]$, K_1 , K_2 , and b are free model parameters. Unlike the Hill function, EC50 is not an explicit parameter of (28) and thus needs to be determined numerically. Note that (28) can be reduced to a right-shifted hyperbolic function as $K_2 \rightarrow 0$:

$$\%XP([K]) = P_{\max} \frac{[I] - [K] - |[I] - [K]|}{2[I] - 2(K_1 + [K])} + b, \quad (29)$$

which is equivalent to a threshold-hyperbolic stimulus-response function,

$$\%XP([K]) = \frac{P_{\max}([K] - [I])}{K_1 + ([K] - [I])} H([K] - [I]) + b, \quad (30)$$

where H is the unit step function.

Stimulus-response relations are fit using the NonlinearModelFit function in Mathematica 12.0. We stress here that the curve fits are purely phenomenological and are thus not intended to be interpreted in terms of the biochemical assumptions underlying the model.

Experimental methods

Protein expression and purification KaiA, KaiB, KaiB-FLAG, and KaiC were expressed and purified as previously described (Phong et al., 2013) with two modifications to the protocol: anion exchange chromatography was performed using HiTrap Q columns (GE Healthcare), and KaiC was purified using Ni-NTA affinity chromatography followed by size-exclusion chromatography, omitting the anion exchange step. The expression, purification, and 6-iodoacetofluorescein (6-IAF) labeling of KaiB K25C mutant as a fluorescence reporter in the plate reader assay is described in Leypunskiy et al. (2017). All mutants of KaiC were constructed using QuikChange II XL Site-Directed Mutagenesis Kit (Aligent). For the KaiC AA and EE mutants, the His-tags were not cleaved during purification; this ensures that these mutant proteins have shifted mobility in SDS-PAGE, allowing their bands to separate from those of KaiC S431A.

The U-[^{15}N] labeled N-terminal (residues 1–135) and C-terminal (residues 181–284) domains of KaiA were expressed in BL21(DE3) *E. coli* (Novagen) in minimal (M9) media supplemented with ^{15}N -enriched NH_4Cl . For expression of C-terminal domain, M9 media enriched with ^{15}N - NH_4Cl was prepared using 98% deuterated water (D_2O). The proteins were purified by Ni-NTA affinity chromatography followed by size-exclusion chromatography using Superdex 75 1660 prep grade column, as described previously (Chang et al., 2011, 2012; Tseng et al., 2014). N-terminal KaiA eluted as ~ 15 kDa monomer (Vakoukakis et al., 2001), while C-terminal KaiA eluted as ~ 23 kDa homodimer (Vakoukakis and LiWang, 2004).

GFP was expressed as N-terminal 6xHis-tag fusion from the pET28a plasmid in the BL21 (DE3) strain of *E. coli*. Harvested cells were sonicated for lysis and clarified lysate was loaded onto a HisTrap FF column (GE Healthcare). The His tag was cleaved by overnight incubation at 4°C with SUMO protease (Invitrogen), after which the sample was loaded again onto a HisTrap FF column to recover the cleaved products. The cleaved proteins were further purified on a 5 mL HiTrap Q HP column (GE Healthcare) and then a Superdex 200 10/300 GL size-exclusion column. The eluted fractions were concentrated in a sample buffer (20 mM HEPES [pH 7.4], 150 mM KCl, 2.5 mM MgCl_2 , 2 mM DTT), aliquoted, and snap frozen in liquid nitrogen for storage in -80°C .

In vitro clock reactions All in vitro clock reactions were done in the standard reaction buffer (20 mM Tris-HCl [pH 8], 150 mM NaCl, 5 mM MgCl_2 , 0.5 mM EDTA, 10% glycerol, 50 $\mu\text{g/mL}$ Kanamycin). KaiC concentration was 3.5 μM in all experiments unless otherwise specified; KaiB concentration was 3.5 μM for all oscillatory reactions, and 6-IAF-labeled KaiB K25C concentration was 0.2 μM for plate reader assays. KaiA concentration and %ATP were determined by each individual experiment, while the total nucleotide concentration (i.e., [ATP] + [ADP]) was held constant at 5 μM . Phosphorylation kinetics was resolved using SDS-PAGE on 10% acrylamide gels (37.5:1 acrylamide:bis-acrylamide) run for 4.5 h at 30 mA constant current at 12°C ; the gels were stained in SimplyBlue SafeStain (Invitrogen) and then imaged using Bio-Rad ChemiDoc Imager. The oscillatory reactions (Figure 2D) were also monitored using the plate reader assay described in Leypunskiy et al. (2017).

NMR spectroscopy A Bruker 600 MHz AVANCE III spectrometer equipped with a TCI cryoprobe was used for all of the NMR experiments of the N- and C-terminal domains of KaiA (**Figure S2**). Chemical shifts were referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonate (10 μ M). Data were processed using NMRPipe and visualized using NMRDraw (Delaglio et al., 1995). NMR samples were prepared with 100 μ M monomer concentration of protein in 20 mM Tris-HCl [pH 8], 150 mM NaCl, 5 mM MgCl₂, and 5% D₂O buffer. All experiments were performed at 30 °C. Samples were incubated with 1mM ATP or ADP, when needed, for 30 minutes before spectral measurement.

Immunoprecipitation Immunoprecipitation of KaiB-FLAG and associated protein complexes in a clock reaction (**Figure 4A**) was done as previously described (Phong et al., 2013). The supernatant was analyzed by SDS-PAGE on 4–20% Criterion TGX Stain-Free Precast Gels (BioRad) and stained with SYPRO Ruby (BioRad). 1.5 μ M GFP was added to the reaction mixture at the beginning of the time course and serve as an internal standard to correct for changes in protein concentration due to handling. The relative supernatant KaiA concentration was determined as a ratio of KaiA band intensity in each lane to the GFP band intensity and is normalized as percentage of the largest ratio in the time course (**Figure S7D**).

Acknowledgments

We thank Connie Phong and Haneul Yoo for their protein samples, and Jonathan Weare for helpful discussions. This work was supported by National Institutes of Health awards GM109455, GM107369, GM107521, and EY025957, Department of Energy Office of Advanced Scientific Computing Research contract DE-AC02-06CH11347 and award DE-SC0014205, and a Howard Hughes Medical Institute Faculty Scholarship (to MJR). AL was also supported by the Center for Cellular and Biomolecular Machines at University of California, Merced (NSF Grant HRD-1547848). Computations were performed on resources provided by the University of Chicago Research Computing Center, and the Extreme Science and Engineering Discovery Environment (Towns et al., 2014) (NSF Grant ACI-1548562) Bridges (PSC) computing nodes through allocation TG-MCB180007.

Supplementary Information

Additional biochemistry of the KaiC (de)phosphorylation reactions

In this work we construct a general model of the KaiA-KaiC subsystem based on a set of assumptions of basic clock biochemistry; that is, KaiC is an ATPase and a reversible phosphotransferase with two phosphorylation sites at S431 and T432, while KaiA is a nucleotide-exchange factor that promotes the exchange of bound ADP for ATP in CII nucleotide-binding pockets. Through model fitting, we demonstrate that this set of assumptions is sufficient to explain the (de)phosphorylation kinetics of KaiC and its dependence on %ATP and [KaiA]. Our results, however, do not imply that the model is biochemically exhaustive; in this section and the next we briefly discuss aspects of KaiC enzymology that we do not consider in the model.

First, the current model does not account for the CI domain. The CI domain of KaiC is required for the hexamerization of *S. elongatus* KaiC (Hayashi et al., 2004b, 2006) and its ATPase activity is necessary for the allosteric transition into the dephosphorylation phase of the circadian cycle (Phong et al., 2013; Tseng et al., 2017). However, the hydrolysis state of the CI domain has no significant effect on the CII (de)phosphorylation reactions (Phong et al., 2013), and in the current study we are not concerned with KaiB-dependent processes. Therefore the current model does not keep track of the CI hydrolysis state or the allosteric coupling between the CI and CII domains.

Second, the model does not explicitly consider the function of Mg^{2+} . The presence of Mg^{2+} is required for the assembly of the KaiC hexamer (Hayashi et al., 2006; Mutoh et al., 2013), and computational analyses indicate that release of Mg^{2+} independent of the bound nucleotide is highly energetically unfavorable (Hong et al., 2018). Given these results, we assume that Mg^{2+} and nucleotide cannot act independently of each other, and the model implicitly assumes that each bound nucleotide is always in complex with a Mg^{2+} ion. A recent study, however, shows that the absence of Mg^{2+} , especially in buffers with no EDTA, promotes KaiC autophosphorylation (Jeong et al., 2019). Moreover, some structures of KaiC have modeled two Mg^{2+} ions in the nucleotide-binding pocket, which has been interpreted to mean that KaiC kinase activity relies on a two-metal-ion phosphotransfer mechanism (Pattanayek et al., 2009). Currently, the functions of Mg^{2+} have not been characterized kinetically or mechanistically at a level necessary to constrain a molecularly detailed model such as ours.

The hexameric structure of KaiC

The current model does not consider any hexameric effect. There is evidence to suggest that intersubunit interaction regulates KaiC autokinase activity (Kitayama et al., 2013) as well as the ultrasensitive dependence of KaiBC complex formation on the KaiC hexamer phosphoform composition (Lin et al., 2014). However, explicitly accounting for the hexameric nature of KaiC, as in Li et al. (2009) or Lin et al. (2014), would lead to a significant increase in the number of model parameters, which likely cannot be constrained by available data and makes interpretation of the model difficult. Therefore, we only keep track of monomeric KaiC species, and the rate constants should be considered averages over hexameric background configurations, weighted by their nonequilibrium state populations.

This leads to two simplifications concerning the KaiA-KaiC interactions. The first issue relates to the stoichiometry of KaiAC complexes. During phosphorylation, KaiA dimers bind to KaiC hexamers with either a 1:1 or 2:1 stoichiometry (Hayashi et al., 2004a; Yunoki et al., 2019). Because the model does not consider the hexameric structure of KaiC explicitly, this stoichiometry is not enforced, and each KaiC monomer can bind independently to KaiA.

The second issue relates to the effect of the hexameric phosphorylation state on KaiA (un)binding kinetics. KaiA and KaiB compete with each other for KaiC binding (Lin et al., 2014), even though their binding

sites are on opposing sides of KaiC. This has been interpreted as a result of cooperative allosteric transition between a kinase mode of KaiC, stabilized by KaiA binding and the T phosphoform, and a phosphatase mode of KaiC, stabilized by KaiB binding and the S phosphoform (Lin et al., 2014). An implication of this proposed mechanism is that KaiC hexamers in the phosphatase mode have uniformly diminished affinity for KaiA at the CII interface, regardless of the phosphorylation state of the subunits. Given that the current model is trained using primarily the phosphorylation data set, the predicted KaiA dwell time (**Figure S4B**) and dissociation constant (**Figure 1D**) likely reflect the property of KaiC subunits in the kinase mode, whereas experiments with phosphomimetic mutants mimicking the S and D phosphorylation states presumably probe the system kinetics in the phosphatase mode. Indeed, in the KaiC EE titration experiment (**Figure 3F** right), the presence of KaiC EE has virtually no effect on the KaiC S431A stimulus-response curve, which may be partly due to the fact that the KaiC EE hexamers are in the phosphatase mode, a condition not considered in our model.

Correlation structure in the posterior distribution

As discussed in the literature Gutenkunst et al. (2007), often ratios of parameters are better constrained than the parameters themselves. The parameter pairs that have a correlation coefficient larger than 0.9 in log space are shown in **Figure S5A–C**. Such correlations typically reflect that thermodynamic, rather than kinetic, properties of the system are constrained. These include the free energy of phosphotransfer between the S and D phosphoforms (**Figure S5A**) and the free energy for KaiA binding to the ATP-bound states of KaiC (**Figure S5B**; compare with **Figure 1D**). Interestingly, there is a linear relation among $k_{TP}^{A,T}$, $k_d^{A,TU}$, and $k_b^{T,DP}$ (**Figure S5C**); this implies that the data constrain the flux out of the $^A C_{DP}^T$ state, but the exact pathway is underdetermined.

More generally, we characterize the extent to which the model parameters, or linear combinations thereof, are constrained by the data using the principal components of the posterior distribution; that is, the eigenvectors of the covariance matrix from MCMC sampling. The eigenvalues of the covariance matrix span multiple orders of magnitude with no obvious gap (**Figure S5D** left), except for the stiffest direction (**Figure S5G**), which is almost entirely aligned with σ^2 , the global error hyperparameter. In addition, the directions of the principal components are in general not aligned with the directions of the bare coordinates (**Figure S5D** right), and there is no obvious interpretation for the directions of most of the principal components (**Figure S5G**). These features of the ensemble indicate that the model is “sloppy” (Gutenkunst et al., 2007), and many model parameters are poorly constrained by the data. Nevertheless, as we demonstrate in Results, the model can still be used to make consistent predictions because the variabilities in the ensemble of parameter sets obtained from MCMC sampling align with the softest degrees of freedom of the posterior distribution.

Convergence of the model fit

We assessed the quality of the fit in three ways. First, we repeated the full procedure three times to assess reproducibility of the fitting procedure (**Figure S1B**). The three independent runs yielded marginalized posterior distributions that are remarkably consistent and tightly constrained for some parameters but diverge over several orders of magnitude for others. The best fits from the three runs have log posterior values of 720, 714, and 705, respectively. Unless otherwise specified, in this work we base our analyses on the run that produced the best fit with the highest posterior value. The ruggedness of the posterior distribution demonstrates that given the model and training data set, the parameter estimation problem is far from the asymptotic (i.e., large sample) regime. Moreover, in our fitting procedure the vast majority of the walkers from the annealing step is discarded for the sake of improving the MCMC sampling efficiency (the acceptance rate is $\leq 7\%$ without pruning). Given the presence of multiple local maxima, this choice likely resulted in an underestimation of the uncertainties in parameter values.

Table S1: Effects of KaiA on KaiC function

Model*	Log likelihood			Bayes factor
	Phosphorylation	Dephosphorylation	Hydrolysis	
Full model	422.9	346.8	-0.8	1
-P [†]	390.0	338.6	-0.8	1.4
-H [‡]	355.5	308.8	-0.9	2.5
-P,-H	278.2	283.6	-0.2	7.7

* We do not consider a model where KaiA binding does not promote nucleotide exchange, because such models are incapable of autophosphorylation by construction.

[†] -P: KaiA binding decoupled from (de)phosphorylation rates.

[‡] -H: KaiA binding decoupled from KaiC hydrolysis rates.

Second, we compared the model predictions with a test data set that probed the phosphorylation reaction at two non-standard KaiC concentrations (**Figure S4E**). The fit quality on the test data set is somewhat worse compared to the training set (compare with **Figure 1B**). In particular, the model overestimates the D phosphoform concentration at 1.75 μ M KaiC and underestimates the T and D phosphoform concentrations at 7 μ M KaiC. This result suggests some degree of overfitting. This, however, is not a significant issue because we base our conclusions on the ensemble of walkers rather than the behavior of the best fit.

Lastly, we assessed the convergence of the MCMC simulations using the integrated autocorrelation times for the 48 principal components of the posterior distribution (**Figure S5E**). The autocorrelation times for the largest and smallest principal components are 8,500 and 3,300 steps, respectively, which gives rough estimates of the times between independent samples for the slowest and fastest degrees of freedom. These estimates are far shorter than the length of the final MCMC runs in the fitting procedure (100,000 steps). We also checked the autocorrelation time for the KaiA binding affinities (**Figure S5F**), which is within the bound given by the principal components.

KaiA function cannot be solely explained in terms of nucleotide exchange

Because of the generality of the model, the function of KaiA is not restricted to that of a nucleotide-exchange factor. In particular, the phosphotransfer and ATP hydrolysis rates are allowed to depend on KaiA binding. There is some experimental evidence to support such effects—KaiA binding inhibits dephosphorylation (Xu et al., 2003) and the addition of KaiA increases the ATPase activity of KaiC (Terauchi et al., 2007; Murakami et al., 2008). However, the biochemical mechanisms underlying these effects are not clear from the experiments; for example, does KaiA increase KaiC ATPase activity by reconfiguring the KaiC active site, or does KaiA binding indirectly promote ATP hydrolysis by shifting the KaiC population towards phosphoforms that have high ATPase activity?

To test whether KaiA binding has a direct effect on KaiC catalytic activities, we construct simplified models where hydrolysis and/or phosphotransfer is independent of KaiA binding and compare the resulting models to the full model using the Bayes factor (**Table S1**). We find that decoupling either phosphotransfer (model -P) or hydrolysis (model -H) from KaiA binding decreases the evidence for the simplified models, but the effects are weak, especially in comparison to a model where both classes of reactions are decoupled from KaiA binding (model -P,-H). These results indicate that the function of KaiA cannot be solely explained by nucleotide exchange, but we cannot conclusively distinguish between models -P and -H.

The experimental data admit two S phosphorylation pathways

We analyzed a random selection of 500 walkers to understand the implications of their variations for the mechanisms of ordered phosphorylation. To simplify the analysis, we converted each walker to two single-site models in which either T431 or S432 was available for phosphorylation but not both. We asked how important each reaction rate constant is to the overall T and S phosphoform concentrations in the single-site models using the relative first-order sensitivities computed at the standard reaction condition (i.e., 100% ATP with 1.5 μ M KaiA). We focus on the initial phosphorylation rates because the steady-state rates are determined by balances of many contributing processes, making them harder to interpret. The parameter sensitivities at $t = 1$ h are used as proxies for the sensitivities of the initial phosphorylation rates.

Because each single-site model has 18 parameters, there are 36 sensitivities for the two phosphoforms. To characterize this high-dimensional space, we used spectral clustering (**Figure S6A**). Overall, the parameter sensitivities are much more constrained by the data for the T phosphoform than the S phosphoform, which is unsurprising given the relatively low concentrations of the S phosphoform under all experimental conditions in the training dataset. The clustering furthermore indicates that there are two plausible kinetic ordering mechanisms, which differ primarily in terms of the phosphorylation pathways taken by the S phosphoform (**Figure S6A** and **B**). In both clusters, the $U \rightarrow T$ transition is most sensitive to k_p^{UT} (i.e., the phosphorylation rate in the absence of KaiA).

In the first cluster (319 parameter sets), the $U \rightarrow S$ transition is most sensitive to $k_p^{A,US}$ in the presence of KaiA. This is primarily because k_p^{US} is very small in this cluster relative to $k_p^{A,US}$ (**Figure S6C**); however, since, in this cluster, the $U \rightarrow S$ transition is dominated by the KaiA-bound states, the S phosphoform has negative sensitivity to the KaiA dissociation rate constant $k_b^{U,TP}$. This suggests that in the first cluster, KaiA (un)binding to the U phosphoform is important in determining the initial phosphorylation rate of S. The best fit belongs to this first cluster.

The phosphorylation pathway suggested by the second cluster (181 parameter sets) is more complex. In this cluster, the $U \rightarrow S$ transition is mostly independent of KaiA, similar to the $U \rightarrow T$ reaction. However, the S phosphoform is limited by the dephosphorylation reaction k_d^{SU} , which is much faster than the corresponding phosphorylation rate (**Figure S6C**). In addition, the S phosphoform is sensitive to the rate constant for KaiA binding, $k_a^{S,DP}$, which is important for facilitating nucleotide exchange for the ADP-bound S phosphoform, but tends to be slower in the second cluster (**Figure S6C**). Therefore, faster dephosphorylation and slower KaiA binding is important for determining the initial S phosphorylation rate in the second cluster.

A comparison of **Figure S6C** with **Figure S1B** (blue distributions) shows that the two clusters correspond to the bimodal posterior distributions for the rate constants k_p^{US} , $k_p^{A,US}$, k_d^{SU} , and $k_a^{S,DP}$. The two clusters, however, do not cleanly separate along the two modes of $k_a^{S,TP}$ and $k_b^{S,TP}$; the kinetic significance of this bimodal distribution is unclear.

As discussed above, the posterior distribution is fairly rugged and thus the fitting procedure is not fully reproducible over independent runs. As a result, there are likely multiple potential kinetic ordering mechanisms that remain unexplored through this analysis. Regardless, the analysis suggests that kinetic ordering is likely a result of an interplay between (de)phosphorylation and KaiA (un)binding kinetics, rather than purely the product of equilibrium free energies of phosphotransfer.

Comparison to the Paijmans model

Among all the computational work on the Kai oscillator, the model most similar to the current work in terms of the treatment of the KaiA-KaiC subsystem is that by Paijmans et al. (2017b), although the latter is a full oscillator model including KaiB, the CI domain, and the allosteric transition between the active (i.e., phosphorylation phase) and inactive (dephosphorylation phase) KaiC conformations. The Paijmans model

and the full model in this work are both molecularly detailed, and describe how phosphotransfer, ATP hydrolysis, KaiA (un)binding, and nucleotide exchange reactions control the phosphoform and nucleotide-bound states of KaiC. However, there are some significant differences between these two models, which we examine below.

The Paijmans model is more general than this work in two ways. First, the Paijmans model explicitly considers the hexameric nature of KaiC. There is no intersubunit coordination of phosphotransfer in the Paijmans model, but it explicitly considers the binding of one KaiA dimer to a KaiC hexamer, which is assumed to uniformly accelerate nucleotide exchange in all six subunits. In this work, however, we do not consider the hexameric states of KaiC, and each KaiC monomer is allowed to bind to a KaiA dimer independently. In this way the affinities and kinetics of KaiA binding in this work may not be directly comparable to those in the Paijmans model. Second, the Paijmans model allows for the exchange of bound ATP for ADP, such that KaiA accelerates the exchange rates of both ATP and ADP while leaving the binding affinity unchanged. In our model, however, we assume that there is no exchange of bound ATP for ADP, effectively assuming that the affinity of ATP is infinite (i.e., $K_{\text{ATP/ADP}}^{\text{CII}} = 0$ in the Paijmans model terminology). The treatment of nucleotide exchange in both models are otherwise similar, in that both assume that the ATP/ADP on rates are identical, that the apo state is in a quasi-steady state, and that there is no KaiA-independent nucleotide exchange.

The current work goes beyond the Paijmans model in the following ways. First, we determine the rate constants under the framework of Bayesian parameter estimation, which enables more rigorous uncertainty quantification, while the parameters in the Paijmans model were hand-tuned to reproduce selected experimental observations. Second, for simplicity the Paijmans model does not consider any possible coupling between ATP hydrolysis and KaiC phosphorylation states, between nucleotide exchange and KaiC phosphorylation states, between ATP hydrolysis and KaiA binding, or between KaiC nucleotide-bound states and KaiA binding. Although many of the species-dependent effects are not fully constrained by data in this work, as we describe in Results, the ultrasensitivity in KaiC phosphorylation depends critically on the coupling between KaiC nucleotide-bound states and KaiA binding. It is an open question whether a model that lacks such effects but explicitly accounts for the hexameric nature of KaiC can generate ultrasensitivity.

The difference in the treatment of KaiA binding affinity implies that some detailed balance conditions are incompatible between the two models. In the Paijmans model, the binding affinity of KaiA to KaiC hexamers during the phosphorylation phase depends on the phosphoform composition of the subunits, and each subunit i in the phosphoform X_i other than U contributes an additive factor of $\delta g_{\text{bind}}^{\text{CII-KaiA}}(X_i)$ to the changes in KaiA binding free energy, $\Delta G_{\text{bind}}^{\text{CII-KaiA}}$. Due to detailed balance, the fact that KaiA binds to different KaiC phosphoforms with differential affinities implies that KaiA binding changes the free energy of phosphotransfer [see eq. 8 in Paijmans et al. (2017b)]. This condition is also present in our model, but is complicated by the nucleotide-bound states of KaiC. Using the multiplicative-factor parametrization scheme (see Materials and Methods), the detailed balance conditions in this work can be related to those in the Paijmans model by

$$\delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{T}) - \delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{U}) = -kT \ln \frac{\Delta k_{\text{b}}^{\text{T,DP}}}{\delta k_{\text{a}}^{\text{T,DP}}} \quad (31)$$

$$\delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{S}) - \delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{U}) = -kT \ln \frac{\Delta k_{\text{b}}^{\text{S,DP}}}{\delta k_{\text{a}}^{\text{S,DP}}} \quad (32)$$

$$\delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{D}) - \delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{T}) = -kT \ln \frac{\Delta k_{\text{b}}^{\text{D,DP}}}{\Delta k_{\text{a}}^{\text{D,DP}}} \frac{\delta k_{\text{a}}^{\text{T,TP}}}{\Delta k_{\text{b}}^{\text{T,TP}}} \quad (33)$$

$$\delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{D}) - \delta g_{\text{bind}}^{\text{CII-KaiA}}(\text{S}) = -kT \ln \frac{\Delta k_{\text{b}}^{\text{D,DP}}}{\Delta k_{\text{a}}^{\text{D,DP}}} \frac{\delta k_{\text{a}}^{\text{S,TP}}}{\Delta k_{\text{b}}^{\text{S,TP}}} \quad (34)$$

In general, this set of equations are inconsistent. That is, one cannot express the $\delta g_{\text{bind}}^{\text{CII-KaiA}}(X)$ s in the Paijmans model in terms of the Δk s in our model. The only condition under which these equations can be

made consistent is when

$$\frac{\Delta k_b^{T,DP}}{\delta k_a^{T,DP}} = \frac{\Delta k_b^{T,TP}}{\delta k_a^{T,TP}} \text{ and } \frac{\Delta k_b^{S,DP}}{\delta k_a^{S,DP}} = \frac{\Delta k_b^{S,TP}}{\delta k_a^{S,TP}}$$

that is, when the nucleotide-bound states have no effect on KaiA binding affinities to the T and S phosphoforms.

Phenomenological modifications to the Phong model

In the model by Phong et al. (2013), KaiA sequestration is determined by the equation

$$[A]_{\text{active}} = \max(0, [A]_{\text{total}} - m[B^S] - n[B^D]). \quad (35)$$

Here, $[A]_{\text{total}}$ is the total KaiA concentration; $[B^S]$ and $[B^D]$ are the concentrations of the S and D phosphoforms in complex with KaiB, respectively; m and n are model parameters describing the binding stoichiometries between KaiA and the KaiBC complex. In this way, the KaiA binding affinity to the inhibitory complex is effectively infinite, with no KaiA dissociation until $[B^S]$ and $[B^D]$ drop below a threshold. To make the representation of KaiA sequestration more realistic, we introduce a KaiA dissociation constant K_D (Figure 4C). Assuming that the KaiA sequestration reaction is in a quasi-equilibrium, we replace (35) with

$$[A]_{\text{active}} = \frac{1}{2} \left([A]_{\text{total}} - SD - K_D + \sqrt{4[A]_{\text{total}}K_D + ([A]_{\text{total}} - SD - K_D)^2} \right) \quad (36)$$

where we have defined $SD = m[B^S] + n[B^D]$.

To introduce ultrasensitivity to the Phong model, we first note that the four phosphorylation rate constants for the $U \rightarrow T$, $U \rightarrow S$, $T \rightarrow D$, and $S \rightarrow D$ transitions are given by Michaelis-Menten kinetics with ADP serving as a competitive inhibitor,

$$k_{\text{phos}} = \frac{k_{\text{phos}}^A [A]_{\text{active}}}{K_{1/2} + [A]_{\text{active}}} \frac{1}{1 + K_I [\text{ADP}]/[\text{ATP}]} \quad (37)$$

where k_{phos}^A varies with the specific phosphorylation reaction. To introduce ultrasensitivity, we add a threshold term T (Figure 4D),

$$k_{\text{phos}} = \frac{k_{\text{phos}}^A ([A]_{\text{active}} - T)}{K_{1/2} + ([A]_{\text{active}} - T)} \frac{1}{1 + K_I [\text{ADP}]/[\text{ATP}]} H([A]_{\text{active}} - T) \quad (38)$$

where

$$T = (0.3 + 1.5e^{-5\% \text{ATP}/100\%})(1 + 0.08[C^U]/\mu\text{M}).$$

and H is the Heaviside function. The first part of the expression for the phosphorylation threshold, $0.3 + 1.5e^{-5\% \text{ATP}/100\%}$, describes how the threshold changes as a function of %ATP; the constants are determined by approximating the $[KaiA]$ threshold in Figure 3A as an exponential function. The second part of the expression, $1 + 0.08[C^U]/\mu\text{M}$, describes how the threshold changes as a function of C^U concentration. This formula is defined by

$$\frac{a/\mu\text{M} + b([C^U]/\mu\text{M} - 3.5)}{a/\mu\text{M} + b(-3.5)}$$

where parameters a and b are determined by taking a linear fit, $a + b[AA]$, of the data from Figure 3G (yellow line); that is, $1 + 0.08[C^U]/\mu\text{M}$ gives the fractional changes to the phosphorylation threshold as a result of any U phosphoform (in the form of AA phosphomimetic mutant) additional to the 3.5 μM KaiC S431A in the experiment.

We use the original Phong model parameters in all analyses of the model with one exception. In the final model with both a K_D and a phosphorylation threshold (Figure 4E bottom), the period is systematically longer than 24 h due to a slow down in phosphorylation. To fix this problem, we change k_{ds}^A and k_{ds}^0 , the two rate constants controlling the $D \rightarrow S$ transition, to $0.94 k_{ds}^A$ and $1.1 k_{ds}^0$.

References

- 1097
- 1098 Abe, J., Hiyama, T. B., Mukaiyama, A., Son, S., Mori, T., Saito, S., Osako, M., Wolanin, J., Yamashita, E., Kondo, T. and Akiyama, S. (2015)
- 1099 Atomic-scale origins of slowness in the cyanobacterial circadian clock. *Science*, **349**, 312–316.
- 1100 Bellman, R. and Åström, K. J. (1970) On structural identifiability. *Mathematical Biosciences*, **7**, 329–339.
- 1101 Beyer, H.-G. and Schwefel, H.-P. (2002) Evolution strategies – A comprehensive introduction. *Natural Computing*, **1**, 3–52.
- 1102 Brettschneider, C., Rose, R. J., Hertel, S., Axmann, I. M., Heck, A. J. R. and Kollmann, M. (2010) A sequestration feedback determines
- 1103 dynamics and temperature entrainment of the KaiABC circadian clock. *Molecular Systems Biology*, **6**, 389.
- 1104 Brown, K. S. and Sethna, J. P. (2003) Statistical mechanical approaches to models with many poorly known parameters. *Physical Review*
- 1105 *E*, **68**, 021904.
- 1106 Brown, S. A., Kowalska, E. and Dallmann, R. (2012) (Re)inventing the Circadian Feedback Loop. *Developmental Cell*, **22**, 477–487.
- 1107 Buchler, N. E. and Louis, M. (2008) Molecular Titration and Ultrasensitivity in Regulatory Networks. *Journal of Molecular Biology*, **384**,
- 1108 1106–1119.
- 1109 Cao, Y., Wang, H., Ouyang, Q. and Tu, Y. (2015) The free-energy cost of accurate biochemical oscillations. *Nature Physics*, **11**, 772–778.
- 1110 Chang, Y.-G., Kuo, N.-W., Tseng, R. and LiWang, A. (2011) Flexibility of the C-terminal, or CII, ring of KaiC governs the rhythm of the
- 1111 circadian clock of cyanobacteria. *Proceedings of the National Academy of Sciences*, **108**, 14431–14436.
- 1112 Chang, Y.-G., Tseng, R., Kuo, N.-W. and LiWang, A. (2012) Rhythmic ring–ring stacking drives the circadian oscillator clockwise. *Proceed-*
- 1113 *ings of the National Academy of Sciences*, **109**, 16847–16851.
- 1114 Chew, J., Leypunskiy, E., Lin, J., Murugan, A. and Rust, M. J. (2018) High protein copy number is required to suppress stochasticity in the
- 1115 cyanobacterial circadian clock. *Nature Communications*, **9**, 3004.
- 1116 Cobelli, C. and DiStefano, J. J. (1980) Parameter and structural identifiability concepts and ambiguities: A critical review and analysis.
- 1117 *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, **239**, R7–R24.
- 1118 Dalcín, L., Paz, R. and Storti, M. (2005) MPI for Python. *Journal of Parallel and Distributed Computing*, **65**, 1108–1115.
- 1119 Dalcín, L., Paz, R., Storti, M. and D’Elía, J. (2008) MPI for Python: Performance improvements and MPI-2 extensions. *Journal of Parallel*
- 1120 *and Distributed Computing*, **68**, 655–662.
- 1121 Dalcin, L. D., Paz, R. R., Kler, P. A. and Cosimo, A. (2011) Parallel distributed computing using Python. *Advances in Water Resources*, **34**,
- 1122 1124–1139.
- 1123 del Junco, C. and Vaikuntanathan, S. (2019) Robust oscillations in multi-cyclic models of biochemical clocks. *arXiv:1909.02534 [cond-mat,*
- 1124 *physics:physics]*.
- 1125 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system
- 1126 based on UNIX pipes. *Journal of Biomolecular NMR*, **6**, 277–293.
- 1127 Dovzhenok, A. A., Baek, M., Lim, S. and Hong, C. I. (2015) Mathematical Modeling and Validation of Glucose Compensation of the
- 1128 *Neurospora* Circadian Clock. *Biophysical Journal*, **108**, 1830–1839.
- 1129 Edgar, R., Green, E., Zhao, Y., Van, O., Olmedo, M., Qin, X., Xu, Y., Pan, M., Valekunja, U., Feeney, K., Maywood, E., Hastings, M., Baliga, N.,
- 1130 Merrow, M., Millar, A., Johnson, C., Kyriacou, C., O’Neill, J. and Reddy, A. (2012) Peroxiredoxins are conserved markers of circadian
- 1131 rhythms. *Nature*, **485**, 459–464.
- 1132 Egli, M., Mori, T., Pattanayek, R., Xu, Y., Qin, X. and Johnson, C. H. (2012) Dephosphorylation of the Core Clock Protein KaiC in the
- 1133 Cyanobacterial KaiABC Circadian Oscillator Proceeds via an ATP Synthase Mechanism. *Biochemistry*, **51**, 1547–1558.
- 1134 Eric Jones, Oliphant, T. E. and Pearu Peterson (2001) SciPy: Open Source Scientific Tools for Python.
- 1135 Eydgahi, H., Chen, W. W., Muhlich, J. L., Vitkup, D., Tsitsiklis, J. N. and Sorger, P. K. (2013) Properties of cell death models calibrated and
- 1136 compared using Bayesian approaches. *Molecular Systems Biology*, **9**.
- 1137 Ferrell, J. E. and Ha, S. H. (2014a) Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. *Trends in Biochemical*
- 1138 *Sciences*, **39**, 496–503.
- 1139 — (2014b) Ultrasensitivity part II: Multisite phosphorylation, stoichiometric inhibitors, and positive feedback. *Trends in Biochemical*
- 1140 *Sciences*, **39**, 556–569.
- 1141 Flaherty, P., Radhakrishnan, M. L., Dinh, T., Rebres, R. A., Roach, T. I., Jordan, M. I. and Arkin, A. P. (2008) A Dual Receptor Crosstalk
- 1142 Model of G-Protein-Coupled Signal Transduction. *PLOS Computational Biology*, **4**, e1000185.

- 1143 Foreman-Mackey, D., Hogg, D. W., Lang, D. and Goodman, J. (2013) Emcee: The MCMC Hammer. *Publications of the Astronomical Society*
1144 *of the Pacific*, **125**, 306.
- 1145 Forger, D. B. and Peskin, C. S. (2003) A detailed predictive model of the mammalian circadian clock. *Proceedings of the National Academy*
1146 *of Sciences*, **100**, 14806–14811.
- 1147 François, P., Despierre, N. and Siggia, E. D. (2012) Adaptive Temperature Compensation in Circadian Oscillations. *PLOS Computational*
1148 *Biology*, **8**, e1002585.
- 1149 Fustin, J.-M., Kojima, R., Itoh, K., Chang, H.-Y., Ye, S., Zhuang, B., Oji, A., Gibo, S., Narasimamurthy, R., Virshup, D., Kurosawa, G., Doi,
1150 M., Manabe, I., Ishihama, Y., Ikawa, M. and Okamura, H. (2018) Two CK1 δ transcripts regulated by m6A methylation code for two
1151 antagonistic kinases in the control of the circadian clock. *Proceedings of the National Academy of Sciences*, **115**, 5980–5985.
- 1152 Gallego, M. and Virshup, D. M. (2007) Post-translational modifications regulate the ticking of the circadian clock. *Nature Reviews Molec-*
1153 *ular Cell Biology*, **8**, 139–148.
- 1154 Geweke, J. (1989) Bayesian Inference in Econometric Models Using Monte Carlo Integration. *Econometrica*, **57**, 1317–1339.
- 1155 Gomez-Urbe, C., Verghese, G. C. and Mirny, L. A. (2007) Operating Regimes of Signaling Cycles: Statics, Dynamics, and Noise Filtering.
1156 *PLOS Computational Biology*, **3**, e246.
- 1157 Gonze, D., Halloy, J. and Goldbeter, A. (2002) Robustness of circadian rhythms with respect to molecular noise. *Proceedings of the*
1158 *National Academy of Sciences*, **99**, 673–678.
- 1159 Goodman, J. and Weare, J. (2010) Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational*
1160 *Science*, **5**, 65–80.
- 1161 Gunawardena, J. (2005) Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proceedings of the National*
1162 *Academy of Sciences*, **102**, 14617–14622.
- 1163 — (2014) Models in biology: ‘accurate descriptions of our pathetic thinking’. *BMC Biology*, **12**, 29.
- 1164 Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R. and Sethna, J. P. (2007) Universally Sloppy Parameter Sensitivities
1165 in Systems Biology Models. *PLOS Computational Biology*, **3**, e189.
- 1166 Hatakeyama, T. S. and Kaneko, K. (2012) Generic temperature compensation of biological clocks by autonomous regulation of catalyst
1167 concentration. *Proceedings of the National Academy of Sciences*, **109**, 8109–8114.
- 1168 Hayashi, F., Ito, H., Fujita, M., Iwase, R., Uzumaki, T. and Ishiura, M. (2004a) Stoichiometric interactions between cyanobacterial clock
1169 proteins KaiA and KaiC. *Biochemical and Biophysical Research Communications*, **316**, 195–202.
- 1170 Hayashi, F., Itoh, N., Uzumaki, T., Iwase, R., Tsuchiya, Y., Yamakawa, H., Morishita, M., Onai, K., Itoh, S. and Ishiura, M. (2004b) Roles of
1171 Two ATPase-Motif-containing Domains in Cyanobacterial Circadian Clock Protein KaiC. *Journal of Biological Chemistry*, **279**, 52331–
1172 52337.
- 1173 Hayashi, F., Iwase, R., Uzumaki, T. and Ishiura, M. (2006) Hexamerization by the N-terminal domain and intersubunit phosphorylation
1174 by the C-terminal domain of cyanobacterial circadian clock protein KaiC. *Biochemical and Biophysical Research Communications*, **348**,
1175 864–872.
- 1176 Hayashi, F., Suzuki, H., Iwase, R., Uzumaki, T., Miyake, A., Shen, J.-R., Imada, K., Furukawa, Y., Yonekura, K., Namba, K. and Ishiura, M.
1177 (2003) ATP-induced hexameric ring structure of the cyanobacterial circadian clock protein KaiC. *Genes to Cells*, **8**, 287–296.
- 1178 Heisler, J., Chavan, A., Chang, Y.-G. and LiWang, A. (2019) Real-Time In Vitro Fluorescence Anisotropy of the Cyanobacterial Circadian
1179 Clock. *Methods and Protocols*, **2**, 42.
- 1180 Hong, L., Vani, B. P., Thiede, E. H., Rust, M. J. and Dinner, A. R. (2018) Molecular dynamics simulations of nucleotide release from the
1181 circadian clock protein KaiC reveal atomic-resolution functional insights. *Proceedings of the National Academy of Sciences*, 201812555.
- 1182 Hou, F., Goodman, J., Hogg, D. W., Weare, J. and Schwab, C. (2012) An Affine-invariant Sampler for Exoplanet Fitting and Discovery in
1183 Radial Velocity Data. *The Astrophysical Journal*, **745**, 198.
- 1184 Hug, S., Raue, A., Hasenauer, J., Bachmann, J., Klingmüller, U., Timmer, J. and Theis, F. J. (2013) High-dimensional Bayesian parameter
1185 estimation: Case study for a model of JAK2/STAT5 signaling. *Mathematical Biosciences*, **246**, 293–304.
- 1186 Iwasaki, H., Nishiwaki, T., Kitayama, Y., Nakajima, M. and Kondo, T. (2002) KaiA-stimulated KaiC phosphorylation in circadian timing
1187 loops in cyanobacteria. *Proceedings of the National Academy of Sciences*, **99**, 15788–15793.
- 1188 Jeong, Y. M., Dias, C., Diekman, C., Brochon, H., Kim, P., Kaur, M., Kim, Y.-S., Jang, H.-I. and Kim, Y.-I. (2019) Magnesium Regulates the
1189 Circadian Oscillator in Cyanobacteria. *Journal of Biological Rhythms*, 0748730419851655.
- 1190 Johnson, C. H. and Egli, M. (2014) Metabolic Compensation and Circadian Resilience in Prokaryotic Cyanobacteria. *Annual Review of*
1191 *Biochemistry*, **83**, 221–247.

Johnson, C. H., Stewart, P. L. and Egli, M. (2011) The Cyanobacterial Circadian System: From Biophysics to Bioevolution. *Annual Review of Biophysics*, **40**, 143–167.

Jolley, C. C., Ode, K. L. and Ueda, H. R. (2012) A Design Principle for a Posttranslational Biochemical Oscillator. *Cell Reports*, **2**, 938–950.

Kageyama, H., Nishiwaki, T., Nakajima, M., Iwasaki, H., Oyama, T. and Kondo, T. (2006) Cyanobacterial Circadian Pacemaker: Kai Protein Complex Dynamics in the KaiC Phosphorylation Cycle In Vitro. *Molecular Cell*, **23**, 161–171.

Kidd, P. B., Young, M. W. and Siggia, E. D. (2015) Temperature compensation and temperature sensation in the circadian clock. *Proceedings of the National Academy of Sciences*, **112**, E6284–E6292.

Kim, J. K. and Forger, D. B. (2012) A mechanism for robust circadian timekeeping via stoichiometric balance. *Molecular Systems Biology*, **8**, 630.

Kim, Y.-I., Dong, G., Carruthers, C. W., Golden, S. S. and LiWang, A. (2008) The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. *Proceedings of the National Academy of Sciences*, **105**, 12825–12830.

Kirkpatrick, S. (1984) Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics*, **34**, 975–986.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983) Optimization by Simulated Annealing. *Science*, **220**, 671–680.

Kitayama, Y., Nishiwaki, T., Terauchi, K. and Kondo, T. (2008) Dual KaiC-based oscillations constitute the circadian system of cyanobacteria. *Genes & Development*, **22**, 1513–1521.

Kitayama, Y., Nishiwaki-Ohkawa, T., Sugisawa, Y. and Kondo, T. (2013) KaiC intersubunit communication facilitates robustness of circadian rhythms in cyanobacteria. *Nature Communications*, **4**, 2897.

Klinke, D. J. (2009) An empirical Bayesian approach for model-based inference of cellular signaling networks. *BMC Bioinformatics*, **10**, 371.

Lam, S. K., Pitrou, A. and Seibert, S. (2015) Numba: A LLVM-based Python JIT Compiler. In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, LLVM '15, 7:1–7:6. New York, NY, USA: ACM.

Langtangen, H. P. and Wang, L. (2015) Odespy software package.

Leloup, J.-C. and Goldbeter, A. (2003) Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences*, **100**, 7051–7056.

Leloup, J.-C., Gonze, D. and Goldbeter, A. (1999) Limit Cycle Models for Circadian Rhythms Based on Transcriptional Regulation in *Drosophila* and *Neurospora*. *Journal of Biological Rhythms*, **14**, 433–448.

Leypunskiy, E., Lin, J., Yoo, H., Lee, U., Dinner, A. R. and Rust, M. J. (2017) The cyanobacterial circadian clock follows midday in vivo and in vitro. *eLife*, **6**, e23539.

Li, C., Chen, X., Wang, P. and Wang, W. (2009) Circadian KaiC Phosphorylation: A Multi-Layer Network. *PLOS Comput Biol*, **5**, e1000568.

Lin, J., Chew, J., Chockanathan, U. and Rust, M. J. (2014) Mixtures of opposing phosphorylations within hexamers precisely time feedback in the cyanobacterial circadian clock. *Proceedings of the National Academy of Sciences*, **111**, E3937–E3945.

Locke, J. C. W., Millar, A. J. and Turner, M. S. (2005) Modelling genetic networks with noisy and varied experimental data: The circadian clock in *Arabidopsis thaliana*. *Journal of Theoretical Biology*, **234**, 383–393.

Ma, L. and Ranganathan, R. (2012) Quantifying the Rhythm of KaiB-C Interaction for In Vitro Cyanobacterial Circadian Clock. *PLOS ONE*, **7**, e42581.

Ma, W., Trusina, A., El-Samad, H., Lim, W. A. and Tang, C. (2009) Defining Network Topologies that Can Achieve Biochemical Adaptation. *Cell*, **138**, 760–773.

Machado, D., Costa, R. S., Rocha, M., Ferreira, E. C., Tidor, B. and Rocha, I. (2011) Modeling formalisms in Systems Biology. *AMB Express*, **1**, 45.

MacKay, D. J. C. and Kay, D. J. C. M. (2003) *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.

Madras, N. and Sokal, A. D. (1988) The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk. *Journal of Statistical Physics*, **50**, 109–186.

Mello, B. A., Pan, W., Hazelbauer, G. L. and Tu, Y. (2018) A dual regulation mechanism of histidine kinase CheA identified by combining network-dynamics modeling and system-level input-output data. *PLOS Computational Biology*, **14**, e1006305.

Mirsky, H. P., Liu, A. C., Welsh, D. K., Kay, S. A. and Doyle, F. J. (2009) A model of the cell-autonomous mammalian circadian clock. *Proceedings of the National Academy of Sciences*, **106**, 11107–11112.

Monti, M., Lubensky, D. K. and ten Wolde, P. R. (2018) Robustness of Clocks to Input Noise. *Physical Review Letters*, **121**, 078101.

1239 Mori, T., Sugiyama, S., Byrne, M., Johnson, C. H., Uchihashi, T. and Ando, T. (2018) Revealing circadian mechanisms of integration and
1240 resilience by visualizing clock proteins working in real time. *Nature Communications*, **9**, 3245.

1241 Mukaiyama, A., Furuie, Y., Abe, J., Yamashita, E., Kondo, T. and Akiyama, S. (2018) Conformational rearrangements of the C1 ring in
1242 KaiC measure the timing of assembly with KaiB. *Scientific Reports*, **8**, 8803.

1243 Murakami, R., Miyake, A., Iwase, R., Hayashi, F., Uzumaki, T. and Ishiura, M. (2008) ATPase activity and its temperature compensation of
1244 the cyanobacterial clock protein KaiC. *Genes to Cells*, **13**, 387–395.

1245 Murakami, R., Mutoh, R., Ishii, K. and Ishiura, M. (2016) Circadian oscillations of KaiA-KaiC and KaiB-KaiC complex formations in an in
1246 vitro reconstituted KaiABC clock oscillator. *Genes to Cells*, **21**, 890–900.

1247 Murayama, Y., Kori, H., Oshima, C., Kondo, T., Iwasaki, H. and Ito, H. (2017) Low temperature nullifies the circadian clock in cyanobacteria
1248 through Hopf bifurcation. *Proceedings of the National Academy of Sciences*, **114**, 5641–5646.

1249 Mutoh, R., Nishimura, A., Yasui, S., Onai, K. and Ishiura, M. (2013) The ATP-Mediated Regulation of KaiB-KaiC Interaction in the
1250 Cyanobacterial Circadian Clock. *PLOS ONE*, **8**, e80200.

1251 Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T. and Kondo, T. (2005) Reconstitution of Circadian
1252 Oscillation of Cyanobacterial KaiC Phosphorylation in Vitro. *Science*, **308**, 414–415.

1253 Nelder, J. A. and Mead, R. (1965) A Simplex Method for Function Minimization. *The Computer Journal*, **7**, 308–313.

1254 Nishiwaki, T. and Kondo, T. (2012) Circadian Autodephosphorylation of Cyanobacterial Clock Protein KaiC Occurs via Formation of ATP
1255 as Intermediate. *Journal of Biological Chemistry*, **287**, 18030–18035.

1256 Nishiwaki, T., Satomi, Y., Kitayama, Y., Terauchi, K., Kiyohara, R., Takao, T. and Kondo, T. (2007) A sequential program of dual phospho-
1257 rylation of KaiC as a basis for circadian rhythm in cyanobacteria. *The EMBO Journal*, **26**, 4029–4037.

1258 Nishiwaki-Ohkawa, T., Kitayama, Y., Ochiai, E. and Kondo, T. (2014) Exchange of ADP with ATP in the CII ATPase domain promotes
1259 autophosphorylation of cyanobacterial clock protein KaiC. *Proceedings of the National Academy of Sciences*, **111**, 4455–4460.

1260 Novák, B. and Tyson, J. J. (2008) Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, **9**, 981–991.

1261 Oliphant, T. E. (2015) *Guide to NumPy*. USA: CreateSpace Independent Publishing Platform, 2nd edn.

1262 O'Neill, J. S., van Ooijen, G., Dixon, L. E., Troein, C., Corellou, F., Bouget, F.-Y., Reddy, A. B. and Millar, A. J. (2011) Circadian rhythms persist
1263 without transcription in a eukaryote. *Nature*, **469**, 554–558.

1264 Pajmams, J., Lubensky, D. K. and ten Wolde, P. R. (2017a) Period Robustness and Entrainability of the Kai System to Changing Nucleotide
1265 Concentrations. *Biophysical Journal*, **113**, 157–173.

1266 Pajmams, J., Lubensky, D. K. and ten Wolde, P. R. (2017b) A thermodynamically consistent model of the post-translational Kai circadian
1267 clock. *PLOS Computational Biology*, **13**, e1005415.

1268 Pattanayek, R. and Egli, M. (2015) Protein–Protein Interactions in the Cyanobacterial Circadian Clock: Structure of KaiA Dimer in Com-
1269 plex with C-Terminal KaiC Peptides at 2.8 Å Resolution. *Biochemistry*, **54**, 4575–4578.

1270 Pattanayek, R., Mori, T., Xu, Y., Pattanayek, S., Johnson, C. H. and Egli, M. (2009) Structures of KaiC Circadian Clock Mutant Proteins: A
1271 New Phosphorylation Site at T426 and Mechanisms of Kinase, ATPase and Phosphatase. *PLOS ONE*, **4**, e7529.

1272 Pattanayek, R., Wang, J., Mori, T., Xu, Y., Johnson, C. H. and Egli, M. (2004) Visualizing a Circadian Clock Protein: Crystal Structure of KaiC
1273 and Functional Insights. *Molecular Cell*, **15**, 375–388.

1274 Pett, J. P., Korenčič, A., Wesener, F., Kramer, A. and Herzel, H. (2016) Feedback Loops of the Mammalian Circadian Clock Constitute
1275 Repressilator. *PLOS Computational Biology*, **12**, e1005266.

1276 Phong, C., Markson, J. S., Wilhoite, C. M. and Rust, M. J. (2013) Robust and tunable circadian rhythms from differentially sensitive
1277 catalytic domains. *Proceedings of the National Academy of Sciences*, **110**, 1124–1129.

1278 Pittayakanchit, W., Lu, Z., Chew, J., Rust, M. J. and Murugan, A. (2018) Biophysical clocks face a trade-off between internal and external
1279 noise resistance. *eLife*, **7**, e37624.

1280 Powell, M. J. D. (1964) An efficient method for finding the minimum of a function of several variables without calculating derivatives.
1281 *The Computer Journal*, **7**, 155–162.

1282 Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2007) *Numerical Recipes 3rd Edition: The Art of Scientific Computing*.
1283 Cambridge, UK ; New York: Cambridge University Press, 3 edition edn.

1284 Pullen, N. and Morris, R. J. (2014) Bayesian Model Comparison and Parameter Inference in Systems Biology Using Nested Sampling.
1285 *PLOS ONE*, **9**, e88419.

1286 Qin, X., Byrne, M., Mori, T., Zou, P., Williams, D. R., Mchaourab, H. and Johnson, C. H. (2010a) Intermolecular associations determine the
1287 dynamics of the circadian KaiABC oscillator. *Proceedings of the National Academy of Sciences*, **107**, 14805–14810.

1288 Qin, X., Byrne, M., Xu, Y., Mori, T. and Johnson, C. H. (2010b) Coupling of a Core Post-Translational Pacemaker to a Slave Transcription/Translation Feedback Loop in a Circadian System. *PLOS Biol*, **8**, e1000394.

1290 Reischl, S. and Kramer, A. (2011) Kinases and phosphatases in the mammalian circadian clock. *FEBS Letters*, **585**, 1393–1399.

1291 Reischl, S., Vanselow, K., Westermarck, P. O., Thierfelder, N., Maier, B., Herzel, H. and Kramer, A. (2007) β -TrCP1-Mediated Degradation
1292 of PERIOD2 Is Essential for Circadian Dynamics. *Journal of Biological Rhythms*, **22**, 375–386.

1293 Relógio, A., Westermarck, P. O., Wallach, T., Schellenberg, K., Kramer, A. and Herzel, H. (2011) Tuning the Mammalian Circadian Clock:
1294 Robust Synergy of Two Loops. *PLOS Computational Biology*, **7**, e1002309.

1295 Rust, M. J., Golden, S. S. and O'Shea, E. K. (2011) Light-Driven Changes in Energy Metabolism Directly Entrain the Cyanobacterial
1296 Circadian Oscillator. *Science*, **331**, 220–223.

1297 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-
1298 Protein Circadian Clock. *Science*, **318**, 809–812.

1299 Sasai, M. (2019) Effects of Stochastic Single-Molecule Reactions on Coherent Ensemble Oscillations in the KaiABC Circadian Clock. *The
1300 Journal of Physical Chemistry B*, **123**, 702–713.

1301 Schmidl, D., Hug, S., Li, W. B., Greiter, M. B. and Theis, F. J. (2012) Bayesian model selection validates a biokinetic model for zirconium
1302 processing in humans. *BMC Systems Biology*, **6**, 95.

1303 Schwarz, G. (1978) Estimating the Dimension of a Model. *The Annals of Statistics*, **6**, 461–464.

1304 Snijder, J., Schuller, J. M., Wiegard, A., Lössl, P., Schmelling, N., Axmann, I. M., Plitzko, J. M., Förster, F. and Heck, A. J. R. (2017) Structures
1305 of the cyanobacterial circadian oscillator frozen in a fully assembled state. *Science*, **355**, 1181–1184.

1306 Swan, J. A., Golden, S. S., LiWang, A. and Partch, C. L. (2018) Structure, function, and mechanism of the core circadian clock in cyanobac-
1307 teria. *Journal of Biological Chemistry*, **293**, 5026–5034.

1308 Terauchi, K., Kitayama, Y., Nishiwaki, T., Miwa, K., Murayama, Y., Oyama, T. and Kondo, T. (2007) ATPase activity of KaiC determines the
1309 basic timing for circadian clock of cyanobacteria. *Proceedings of the National Academy of Sciences*, **104**, 16377–16381.

1310 Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*,
1311 **58**, 267–288.

1312 Tomita, J., Nakajima, M., Kondo, T. and Iwasaki, H. (2005) No Transcription-Translation Feedback in Circadian Rhythm of KaiC Phospho-
1313 rylation. *Science*, **307**, 251–254.

1314 Toni Tina, Welch David, Strelkova Natalja, Ipsen Andreas and Stumpf Michael P.H (2009) Approximate Bayesian computation scheme
1315 for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, **6**, 187–202.

1316 Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G. D., Roskies, R.,
1317 Scott, J. R. and Wilkins-Diehr, N. (2014) XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering*, **16**, 62–74.

1318 Transtrum, M. K. and Qiu, P. (2014) Model Reduction by Manifold Boundaries. *Physical Review Letters*, **113**, 098701.

1319 Tseng, R., Chang, Y.-G., Bravo, I., Latham, R., Chaudhary, A., Kuo, N.-W. and LiWang, A. (2014) Cooperative KaiA–KaiB–KaiC Interactions
1320 Affect KaiB/SasA Competition in the Circadian Clock of Cyanobacteria. *Journal of Molecular Biology*, **426**, 389–402.

1321 Tseng, R., Goularte, N. F., Chavan, A., Luu, J., Cohen, S. E., Chang, Y.-G., Heisler, J., Li, S., Michael, A. K., Tripathi, S., Golden, S. S., LiWang,
1322 A. and Partch, C. L. (2017) Structural basis of the day-night transition in a bacterial circadian clock. *Science*, **355**, 1174–1180.

1323 Vakonakis, I. and LiWang, A. C. (2004) Structure of the C-terminal domain of the clock protein KaiA in complex with a KaiC-derived
1324 peptide: Implications for KaiC regulation. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10925–
1325 10930.

1326 Vakonakis, I., Risinger, A. T., Latham, M. P., Williams, S. B., Golden, S. S. and LiWang, A. C. (2001) Letter to the Editor: Sequence-
1327 specific ¹H, ¹³C and ¹⁵N resonance assignments of the N-terminal, 135-residue domain of KaiA, a clock protein from *Synechococcus*
1328 *elongatus*. *Journal of Biomolecular NMR*, **21**, 179–180.

1329 Villaverde, A. F. and Banga, J. R. (2014) Reverse engineering and identification in systems biology: Strategies, perspectives and chal-
1330 lenges. *Journal of The Royal Society Interface*, **11**, 20130505.

1331 Vyshemirsky, V. and Girolami, M. A. (2008) Bayesian ranking of biochemical system models. *Bioinformatics*, **24**, 833–839.

1332 Wasserman, L. (2000) Bayesian Model Selection and Model Averaging. *Journal of Mathematical Psychology*, **44**, 92–107.

Williams, S. B., Vakonakis, I., Golden, S. S. and LiWang, A. C. (2002) Structure and function from the circadian clock protein KaiA of *Synechococcus elongatus*: A potential clock input mechanism. *Proceedings of the National Academy of Sciences*, **99**, 15357–15362.

Wu, P., Cheng, C., Kaddi, C. D., Venugopalan, J., Hoffman, R. and Wang, M. D. (2017) –Omic and Electronic Health Record Big Data Analytics for Precision Medicine. *IEEE Transactions on Biomedical Engineering*, **64**, 263–273.

Xu, T.-R., Vyshemirsky, V., Gormand, A., von Kriegsheim, A., Girolami, M., Baillie, G. S., Ketley, D., Dunlop, A. J., Milligan, G., Houslay, M. D. and Kolch, W. (2010) Inferring Signaling Pathway Topologies from Multiple Perturbation Measurements of Specific Biochemical Species. *Science Signaling*, **3**, ra20–ra20.

Xu, Y., Mori, T. and Johnson, C. H. (2003) Cyanobacterial circadian clockwork: Roles of KaiA, KaiB and the kaiBC promoter in regulating KaiC. *The EMBO Journal*, **22**, 2117–2126.

Xu, Y., Mori, T., Pattanayek, R., Pattanayek, S., Egli, M. and Johnson, C. H. (2004) Identification of key phosphorylation sites in the circadian clock protein KaiC by crystallographic and mutagenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 13933–13938.

Yoda, M., Eguchi, K., Terada, T. P. and Sasai, M. (2007) Monomer-Shuffling and Allosteric Transition in KaiC Circadian Oscillation. *PLOS ONE*, **2**, e408.

Yoshida, T., Murayama, Y., Ito, H., Kageyama, H. and Kondo, T. (2009) Nonparametric entrainment of the in vitro circadian phosphorylation rhythm of cyanobacterial KaiC by temperature cycle. *Proceedings of the National Academy of Sciences*, **106**, 1648–1653.

Yunoki, Y., Ishii, K., Yagi-Utsumi, M., Murakami, R., Uchiyama, S., Yagi, H. and Kato, K. (2019) ATP hydrolysis by KaiC promotes its KaiA binding in the cyanobacterial circadian clock system. *Life Science Alliance*, **2**, e201900368.

Zhou, M., Kim, J. K., Eng, G. W. L., Forger, D. B. and Virshup, D. M. (2015) A Period2 Phosphoswitch Regulates and Temperature Compensates Circadian Period. *Molecular Cell*, **60**, 77–88.

van Zon, J. S., Lubensky, D. K., Altena, P. R. H. and ten Wolde, P. R. (2007) An allosteric model of circadian KaiC phosphorylation. *Proceedings of the National Academy of Sciences*, **104**, 7420–7425.

Zwanzig, R. W. (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics*, **22**, 1420–1426.

Zwicker, D., Lubensky, D. K. and ten Wolde, P. R. (2010) Robust circadian clocks from coupled protein-modification and transcription–translation cycles. *Proceedings of the National Academy of Sciences*, **107**, 22540–22545.

Figure S1: Overview of the model. A) A schematic of the full mass action kinetic model. Here, each arrow represents a reaction, and the associated rate constant is represented using the notation introduced in the main text. The thickness of the arrows is proportional to the best fit rate on a log scale (base 10) at 100% ATP and 1.5 μM KaiA. B) The posterior distributions for all rate constants, initial conditions, and the global error hyperparameter. The rate constants have the unit $\text{s}^{-1} \cdot (\mu\text{M})^{-1}$ and the horizontal axis has a log scale (base 10). The three distributions represent the results from three independent runs; the log posterior values for the best fits from the three runs are listed. The red lines represent the best fit from the best run (i.e., the blue distributions). See Materials and Methods and **Figure S10** for further details on the model parameterization method.

Figure S2: No evidence of direct nucleotide-KaiA interaction. ^1H - ^{15}N HSQC spectroscopy of the N-terminal fragment (residues 1–135) of KaiA in the presence and absence of ATP (A) or ADP (B) show no significant differences in chemical shifts, while spectra of the C-terminal fragment (residues 181–284) show subtle line broadening in the presence of ATP (C) and ADP (D), suggesting weak, if any, interaction between the nucleotide and the C-terminal fragment. Given these results, we do not include any direct KaiA-nucleotide interaction in the model.

Figure S3: Performance of the fitting procedure. A) The time evolution of the log posterior values over the four steps of the fitting procedure (see Materials and Methods). For step 1 and 3, the individual Markov chains that do not produce walkers used in the next step are shown in beige. B) A comparison of the performance of Powell's method, a derivative-free numerical optimization method, Goodman-Weare (GM), an ensemble MCMC method, and conventional Metropolis-Hastings (MH) algorithm with a Gaussian trial distribution. For the Metropolis-Hastings algorithm the covariance matrix of the trial distribution is given by the global covariance of the fit (i.e., the last step in panel A), scaled by a factor of 0.005 to give an average acceptance rate of 19.8%. A set of 224 walkers drawn from the prior distribution are used to initialize the simulations for all three methods; the 224 walkers are evolved independently for Powell's method and Metropolis-Hastings, and in an ensemble for Goodman-Weare. Chains that do not reach log posterior above 450 are shown in beige. C) A comparison of the performance of Powell's method with Nelder-Mead, a simplex-based numerical optimization method. The simulations are initialized using the same walkers as in step 2 of A).

Figure S4: Behavior of the model. A) Model fit to the dephosphorylation dataset. B) The best fit KaiA dwell time as a function of KaiC phosphoform and nucleotide-bound state. The error bars represent the 95% posterior interval, and the dashed lines represent the experimental measurements, which did not resolve the nucleotide-bound states. C) Inorganic phosphate production per KaiC monomer over the course of a phosphorylation reaction. The gray region represents the experimental bounds on the KaiC hydrolysis rate with 1.2 μ M KaiA and no KaiA. See Materials and Methods for the source of the experimental data in A–C. D) The kinetics of the dephosphorylation reaction in the absence of KaiA, broken down into the eight individual KaiC species. The gray region represents the 95% posterior interval. Refer to **Figure S1A** for the KaiC state names. E) The predicted phosphorylation kinetics at 7 and 1.75 μ M KaiC, both at 100% ATP and 1.5 μ M KaiA, compared to experimental measurements. Note that these two time series are not part of the training set. F) The posterior distributions for k_r^{TP} and k_r^{DP} , the dissociation rates for ATP and ADP, respectively, in an early iteration of the model. The rate constants have a unit of s^{-1} and the horizontal axis has a log scale (base 10). The long tail to the left of the posterior distribution for k_r^{TP} suggests that the model can be simplified by setting the rate to zero.

Figure S5: Correlation structure in the MCMC ensemble. A), B), and C) show parameters with a correlation coefficients larger than 0.9. In B), “X” represents the KaiC phosphoforms. In C), the projections of the 3D scatter plot onto pairwise correlations are shown in gray. D) The principal component/eigenvalue spectrum of the covariance matrix (left), and the alignment of the principal components with the coordinates (right). Here, I denotes the intersection of the principal component ellipsoid with the coordinates and P denotes the projection of the principal components onto the corresponding coordinates (Gutenkunst et al., 2007). E) The integrated autocorrelation time for the 48 principal components (PC); the principal components are indexed from the largest to the smallest. The integrated autocorrelation time is calculated using an automated windowing procedure (Madras and Sokal, 1988) from the autocorrelation function averaged over the ensemble. F) The integrated autocorrelation time for the KaiA dissociation constants as a function of KaiC phosphoform and nucleotide-bound states. G) The ten largest vector components, ordered by absolute value, for the first and last three principal components.

Figure S6: The mechanism of kinetic ordering is not well-constrained. A) Spectral clustering on the relative sensitivity of the T and S phosphoform concentrations at $t = 1$ h to rate constants in the T- and S-site models, respectively. Only the parameters with significant (> 0.2) relative sensitivities in either cluster are shown in the plot. "X" stands for either the T (horizontal axis) or S (vertical axis) phosphoform. The sensitivities are calculated using 500 sampled parameter sets chosen randomly from the ensemble. The clustering analysis was done using the FindClusters function in Mathematica 12.0. B) Model diagrams that highlight the reactions that have the highest relative sensitivities in the first (left) and second (right) clusters; the D phosphoform is not shown. C) Selected model parameter values in the two clusters. A comparison with blue distributions in **Figure S1B** indicates that the clustering based on sensitivity can be mapped onto the modes of the posterior distribution.

Figure S7: The metabolic compensation property of the Kai oscillator. The fluorescence polarization measurement of the oscillatory reactions are fit to a curve $FP(t) = A \cos(2\pi T^{-1}t + \phi) + bt + c$ to extract A) the period (T) and B) the normalized amplitude ($100A/c$; dimensionless) of the oscillator as a function of [KaiA] and %ATP. Reactions with an amplitude $A < 0.5$ are considered to be non-oscillatory. C) Representative traces demonstrating the effect of %ATP at 1.25 μM KaiA; The polarization data are shifted vertically to avoid overlaps and horizontally to align the first peaks. D) SDS-PAGE gel image of the supernatant from the KaiB-FLAG immunoprecipitation experiment. E) A Comparison of the metabolic compensation property of the Phong model without (left) or with (right) a phosphorylation threshold at $K_D = 10^{-3} \mu M$. The model exhibits phase decoherence at low %ATP without a phosphorylation threshold.

Figure S8: KaiC stimulus-response relations. A) The steady-state stimulus-response relations for T, S, and D phosphoforms predicted by the model. B) The experimentally determined stimulus-response functions of the T, S, and D phosphoforms at three %ATP conditions; the curves are based on refitting the best fit to the steady-state measurements. C) The model-predicted stimulus-response relation of the total steady-state KaiC phosphorylation level as a function of %ATP and [KaiA] after refitting to the steady-state measurements. D) The differences in the log parameter values (base 10) of the best fit before and after refit. The differences are ordered by magnitude and only the 10 parameters (in the multiplicative-factor scheme) with the largest changes are shown.

Figure S9: KaiA binding affinities of simplified models. A) The posterior distributions for the KaiA dissociation constants as a function of KaiC phosphoform in model -n, where the KaiA on/off rates are decoupled from the nucleotide-bound states of KaiC. The dashed lines represent the best fit. B) The posterior distributions for the KaiA dissociation constants as a function of KaiC phosphoform in model -p, where the KaiA on/off rates are decoupled from the KaiC phosphoform; the dashed lines represent the best fit. C) Cross sections of the stimulus-response relation at three %ATP, computed using model -p. The inset represents posterior distribution for the shapes of the stimulus-response function at 25% ATP. The contours represent the 68% and 95% HDRs, and the gray star represents the model best fit. The shape of the stimulus-response function is quantified using two metrics: EC10, which quantifies threshold-like behavior, and EC90 - EC10, which quantifies switch-like behavior. The shape of the experimentally-determined stimulus-response function at 25% ATP is shown as the yellow star. The dashed line represents $(\text{EC}_{10}, \text{EC}_{90} - \text{EC}_{10}) = (K/9, 80K/9)$, which characterizes the shape of a hyperbolic stimulus-response function $[A]/(K + [A])$ that has no switching or thresholding. D) Similar to B), but for model -n,-p, where there is a single KaiA on/off rate in the model. E) Similar to C), but for model -n,-p.

Figure S10: Overview of the model with the multiplicative-factor parameterization scheme. Panels A) and B) are analogous to those in **Figure S1**, but the rate constants are represented as products of the factors that are actually optimized in the MCMC simulations. In B), the δk parameters are fixed parameters determined by detailed balance conditions. The parentheses denote species-dependent effects; A: KaiA-bound state, P: phosphoform, N: nucleotide-bound state. See Materials and Methods for further description of the detailed balance conditions and the model parameterization method.