Cell PRESS
Open ACCESS

Cell Reports
Article

# A Design Principle for a Posttranslational Biochemical Oscillator

Craig C. Jolley,[1,7] Koji L. Ode,[3,7] and Hiroki R. Ueda[1,2,3,4,5,6,*]
[1]Laboratory for Systems Biology
[2]Functional Genomics Unit
RIKEN Center for Developmental Biology, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[3]Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[4]Department of Biological Sciences, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan
[5]Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871 Japan
[6]Department of Mathematics, Graduate School of Science, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
[7]These authors contributed equally to this work.
*Correspondence: uedah-tky@umin.ac.jp
http://dx.doi.org/10.1016/j.celrep.2012.09.006

## SUMMARY

**Multisite phosphorylation plays an important role in biological oscillators such as the circadian clock. Its general role, however, has been elusive. In this theoretical study, we show that a simple substrate with two modification sites acted upon by two opposing enzymes (e.g., a kinase and a phosphatase) can show oscillations in its modification state. An unbiased computational analysis of this oscillator reveals two common characteristics: a unidirectional modification cycle and sequestering of an enzyme by a specific modification state. These two motifs cause a substrate to act as a coupled system in which a unidirectional cycle generates single-molecule oscillators, whereas sequestration synchronizes the population by limiting the available enzyme under conditions in which substrate is in excess. We also demonstrate the conditions under which the oscillation period is temperature compensated, an important feature of the circadian clock. This theoretical model will provide a framework for analyzing and synthesizing posttranslational oscillators.**

## INTRODUCTION

For a wide variety of organisms, biology runs on a schedule. The control of cellular and organismal processes by biochemical clocks has a number of potential advantages. The circadian clock system allows organisms to anticipate ecological changes that correlate with the day/night cycle, regulating both behavioral patterns and metabolic fluxes. The period of circadian rhythmicity is robust and fairly unaffected by the ambient temperature (Hastings and Sweeney, 1957; Pittendrigh, 1954).

Circadian clocks are thought to be driven by cell-autonomous transcriptional-translational oscillators (TTOs) (Dibner et al., 2010; Dunlap, 1999; Reppert and Weaver, 2002; Takahashi et al., 2008; Young and Kay, 2001). In mammalian clocks, the circadian transcriptional program is mediated by at least three clock-controlled DNA elements, known as morning-time (E-box/E′-box or E/E′-box), day-time (D-box), and night-time (Rev-Erb/ROR binding element or RRE) elements (Ueda et al., 2005). The E/E′-box-mediated transcriptional program has a critical role in the core autoregulatory loop of the mammalian circadian clock (Gekakis et al., 1998; Sato et al., 2006). In this core loop, bHLH-PAS transcriptional activators such as BMAL1 and CLOCK form heterodimers that directly and indirectly activate the transcription of their target genes, including the transcriptional repressors *Per* and *Cry*. PERs and CRYs in turn form repressor complexes that physically associate with the BMAL1/CLOCK complex to inhibit E/E′-box-mediated transcription, closing the negative feedback loop (Griffin et al., 1999; Kume et al., 1999; Okamura et al., 1999). This negative feedback, together with delayed expression of *Cry*, is critical for self-sustained circadian clock function (Hogenesch and Ueda, 2011; Ueda et al., 2005; Ukai-Tadenuma et al., 2011).

In the mammalian clock, however, transcription and translation are not the whole story; multisite phosphorylation also plays a central role (Gallego and Virshup, 2007). PER proteins undergo robust circadian changes in phosphorylation (Lee et al., 2001). Mutations in the potential phosphorylation sites of PERs or in their kinases, such as casein kinase I ε/δ (CKIε/δ), alter the circadian period (Lowrey et al., 2000; Meng et al., 2008; Toh et al., 2001; Xu et al., 2005, 2007). The oscillation period is also affected by pharmacological perturbation of CKI, CKII, or GSK3β (Chen et al., 2012; Hirota et al., 2008; Isojima et al., 2009; Meng et al., 2010; Tsuchiya et al., 2009; Walton et al., 2009). PER proteins can be phosphorylated at multiple sites by these kinases (Camacho et al., 2001; Gallego et al., 2006a; Maier et al., 2009; Schlosser et al., 2005; Takano et al., 2000; Tsuchiya et al., 2009; Vanselow et al., 2006). Furthermore, temperature-insensitive phosphorylation of a PER-derived peptide by CKIε/δ has been observed in vitro (Isojima et al., 2009). In addition to phosphorylation, the reverse reaction, dephosphorylation,

seems to play an important role in the mammalian circadian clock because the oscillation period is also affected by the pharmacological or RNA-interference-mediated perturbation of protein phosphatase 1 (PP1), which has been implicated in the dephosphorylation of PER (Gallego et al., 2006b; Lee et al., 2011; Schmutz et al., 2011) and PP5 (Partch et al., 2006).

The central role played by circadian posttranslational oscillators (PTOs) mediated by multisite phosphorylation is best understood in cyanobacteria. Purified clock proteins from the cyanobacterium *Synechococcus elongatus* can be used to reconstitute an in vitro posttranslational circadian oscillator (Johnson et al., 2008; Nakajima et al., 2005). This system contains three essential proteins: KaiA, KaiB, and KaiC. KaiC exhibits both autophosphorylation and autodephosphorylation activities; autophosphorylation is enhanced by KaiA, whereas KaiB inhibits KaiA. When the three proteins are incubated together with an excess of ATP, the phosphorylation level of KaiC oscillates with an ~24 hr period. The period of KaiC phosphorylation oscillations is fairly unaffected by the incubation temperature, suggesting that the KaiC PTO is a bona fide circadian oscillator. Studies have further shown that the phosphorylation cycle involves the sequential phosphorylation and dephosphorylation of two sites in KaiC (Nishiwaki et al., 2007; Rust et al., 2007).

It is even possible that posttranslational mechanisms might permit the mammalian circadian clock to operate in the absence of regulation by gene expression. When cultured fibroblast cells are treated with reagents that interfere with transcription, one would expect oscillations based on a TTO to be severely affected. Instead, circadian oscillations are robust to changes in transcription rate (Dibner et al., 2009). In addition, cyclic expression of *Bmal1* and *Clock* seems not to be necessary for circadian rhythmicity (von Gall et al., 2003). As for CRY, rhythmic, phase-specific expression of *Cry1* is critical for robust circadian oscillation (Ukai-Tadenuma et al., 2011), yet $Cry1^{-/-}$:$Cry2^{-/-}$ cells can exhibit weak circadian oscillations under constant expression of *Cry1* (Ukai-Tadenuma et al., 2011) or a constant supply of CRYs (Fan et al., 2007). Further evidence comes from red blood cells, which show a circadian variation in the cellular redox state despite the absence of transcription (O'Neill and Reddy, 2011). While these results do not constitute direct evidence of a posttranslational oscillator that can operate independently from transcription, they suggest that processes other than transcription and translation may be responsible for oscillations under some conditions.

One simple and attractive possibility is that, under some conditions, multisite phosphorylation systems can oscillate without additional regulation. Theoretical studies have indicated that multisite phosphorylation systems can exhibit a large number of stable states under very generic conditions (Markevich et al., 2004; Thomson and Gunawardena, 2009). Previous studies of oscillatory phosphorylation systems, such as the cyanobacterial clock (Clodong et al., 2007; Rust et al., 2007; van Zon et al., 2007), mitogen-activated protein kinase (MAPK) systems (Chickarmane et al., 2007; Liu et al., 2011; Qiao et al., 2007; Shankaran et al., 2009), and other biochemical oscillators (Ferrell et al., 2011; Kholodenko, 2006; Novák and Tyson, 2008) typically employed complex regulatory schemes. Cyanobacterial clock models (Clodong et al., 2007; Rust et al., 2007; van Zon et al., 2007) are constructed based on the autokinase and autophosphatase properties of KaiC. This unique feature makes the detailed conclusions drawn from these models difficult to apply to generic enzyme-substrate reactions. MAPK models are constructed based on a more general substrate-enzyme reaction scheme, but they often involve multiple substrates sharing an enzyme, protein synthesis, and degradation, or active feedback regulation of enzyme activity by a substrate. It is not clear whether such intricate regulation is required for oscillations, or whether a more generic system would suffice.

Here we show that autonomous, robust oscillations are possible in a system consisting of a single substrate with two phosphorylation sites. The substrate is modified by one kinase and one phosphatase, and no additional regulation is necessary. In a sense, this is the simplest possible biochemical oscillator. An extensive search of the model's parameter space reveals two design motifs for a simple biochemical oscillator. These features can be readily associated with known properties of PERs, CKIε/δ and other enzyme-substrate pairs described above. Under some conditions, the oscillation period is almost unaffected by systematic changes in the enzyme reaction rates induced by temperature differences, suggesting that temperature compensation is possible in this system. The simplicity and generality of the model make it a potentially useful design tool for de novo biochemical oscillators, and suggest that biological posttranslational oscillation may be more common than previously suspected.

## RESULTS

### Multisite Substrate Modifications Can Oscillate with Two Opposing Enzymes

To elucidate the minimum requirements for a PTO, we began by defining a set of components. Cyclic changes in substrate phosphorylation state require both a kinase and a phosphatase. Single-site systems will converge to a unique steady state under relevant conditions (see Extended Experimental Procedures; Figure S1), and a two-site system is the simplest one in which oscillations might be possible (Figure 1A). The model is described as a phosphorylation system, but the same results could be applied to any type of posttranslational modification. There are four distinct substrate states (Figure 1B): $S_{00}$, which is completely unphosphorylated; $S_{01}$ and $S_{10}$, which are singly phosphorylated; and $S_{11}$, which is doubly phosphorylated. The same kinase (E) and phosphatase (F) act on both sites. All reactions involve a 1:1 association of enzymes and substrates, with an intermediate enzyme-substrate complex (Figure 1C). The reactions are described with the use of mass-action kinetics. Although the model equations do not explicitly assume enzyme saturation by substrate, we assumed an excess of substrate over enzyme, and enzyme-substrate binding was usually saturated.

We found that sustained oscillations could be observed in certain parameter regions (Figure 1D). After an initial transient, integrations that began under different initial conditions rapidly converged to a common trajectory with the same amplitude and period, in other words, a limit-cycle oscillation (Figure 1E).

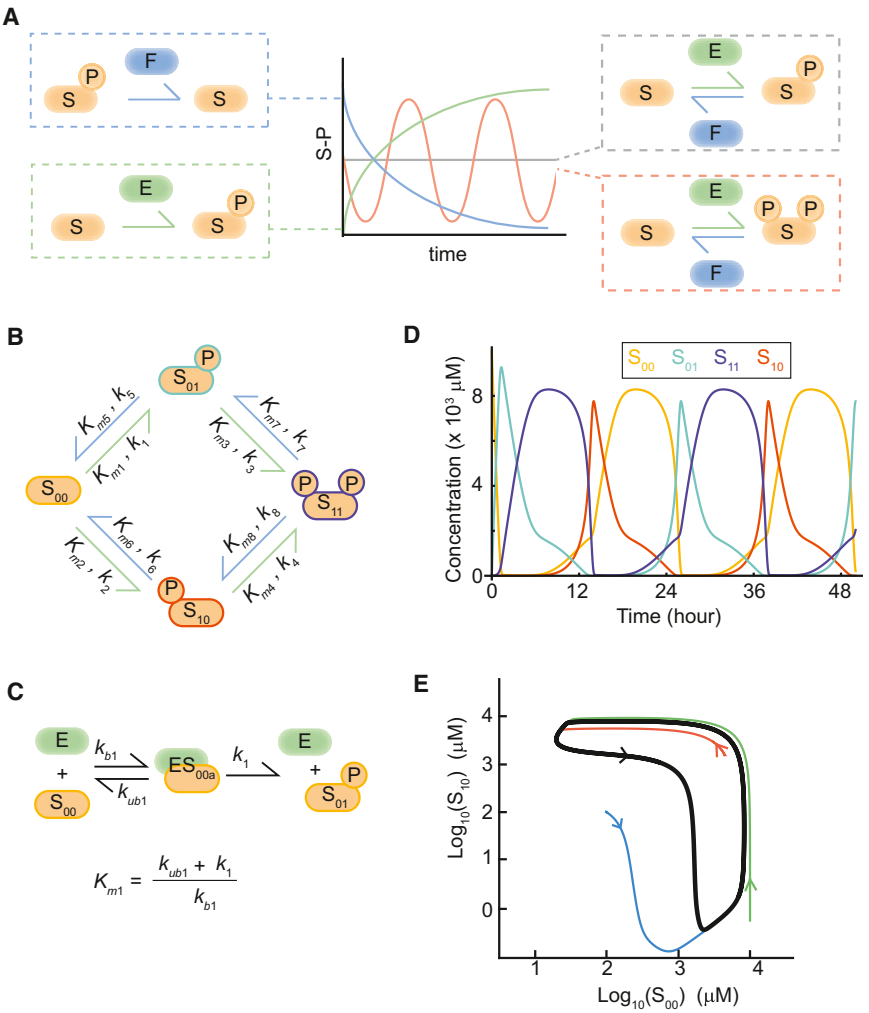



**Figure 1. Multisite Substrate Modifications Can Oscillate with Two Opposing Enzymes**

(A) A phosphatase (blue) or a kinase (green) acting on the substrate alone will yield monotonic convergence to a steady state, as indicated by blue and green curves. A kinase and phosphatase acting in concert are required for sustained oscillations (red).

(B) PTO reaction network. The substrate can be phosphorylated on two distinct sites, and the kinase and phosphatase act on all four phosphorylation states.

(C) Enzymatic reactions involve an intermediate enzyme-substrate complex.

(D) Time course of oscillations for an example parameter set.

(E) Oscillations projected into the $S_{00}$-$S_{10}$ plane. Different initial transients (colored arrows) converge rapidly to a single limit-cycle attractor (black).

See also Figure S1.

In our system, a limit-cycle oscillation is present only when the equilibrium state is unstable.

## Oscillatory Systems Possess Two Design Motifs

Once we had represented the system as a system of dynamical equations, our next aim was to determine which values of the system parameters would lead to sustained oscillations. The rate constants $k_1 \ldots k_8$ represent the number of substrate molecules converted to product molecules in a given reaction per enzyme per minute. The binding constants $K_{m1} \ldots K_{m8}$ are the substrate concentrations at which the rates of reaction 1–8 reach half of their maximum. If the binding is strong, then $K_m$ will be low because the enzyme-substrate binding is easily saturated. This system is too large and complex for straightforward analytical study, so we chose to identify solutions numerically through an extensive search of parameter space.

We generated solutions by selecting parameters from an exponential distribution bounded to the interval 1–1,000 $min^{-1}$ for $k_1$–$k_8$ and 0.01–1,000 μM for $K_{m1}$–$K_{m8}$. This range corresponds to reasonable values typically used to model systems such as MAPK (Chickarmane et al., 2007; Liu et al., 2011; Markevich et al., 2004; Qiao et al., 2007; Shankaran et al., 2009).

The "hit rate" of oscillating parameter sets was rather low, ~0.1%. The search was continued until ~$10^6$ oscillating parameter sets had been found (Figure 2A). Because this population may contain several distinct types of oscillators, we categorized the parameter sets using a clustering algorithm that accounts for the symmetry of the reaction network (see Extended Experimental Procedures). Two major clusters were present over a wide range of cluster diameters, and we chose a diameter for which these two clusters accounted for ~70% of solutions (Figures 2B and S2A). The parameter distributions for the major clusters are shown in Figure 2C. Both clusters exhibit approximate symmetry: fast and slow reactions or strong and weak binding constants tend to be located at symmetry-related positions. For example, in cluster 1, $k_1$, $k_8$, $K_{m2}$, and $K_{m7}$ all tend to be high, whereas $k_2$, $k_7$, $K_{m4}$, and $K_{m5}$ all tend to be low. The schematic diagrams shown in Figure 2D will be unchanged if the diagram is rotated 180° and the kinase and phosphatase are interchanged. Symmetry will rarely be exact for individual parameter sets; the symmetric pattern emerges only when a large number of parameter sets are compared.

Based on these histograms, consensus features of each major cluster can be identified (Figure 2D). Two design motifs are shared by both major clusters. The first is a unidirectional bias: in general, the rates of the "forward" (clockwise) reactions, $k_1$, $k_3$, $k_6$, and $k_8$, are higher than the corresponding "reverse" reactions, $k_5$, $k_7$, $k_2$, and $k_4$, respectively. The second is the presence of enzyme-sequestering steps. In cluster 1, the value of $K_{m4}$ is usually low, meaning that $S_{10}$ binds the kinase very strongly. The kinase-catalyzed $S_{10} \rightarrow S_{11}$ conversion is much slower than its reverse reaction; in this sense, the binding of the kinase by $S_{10}$ is unproductive. Although the $S_{10} \rightarrow S_{00}$ conversion is

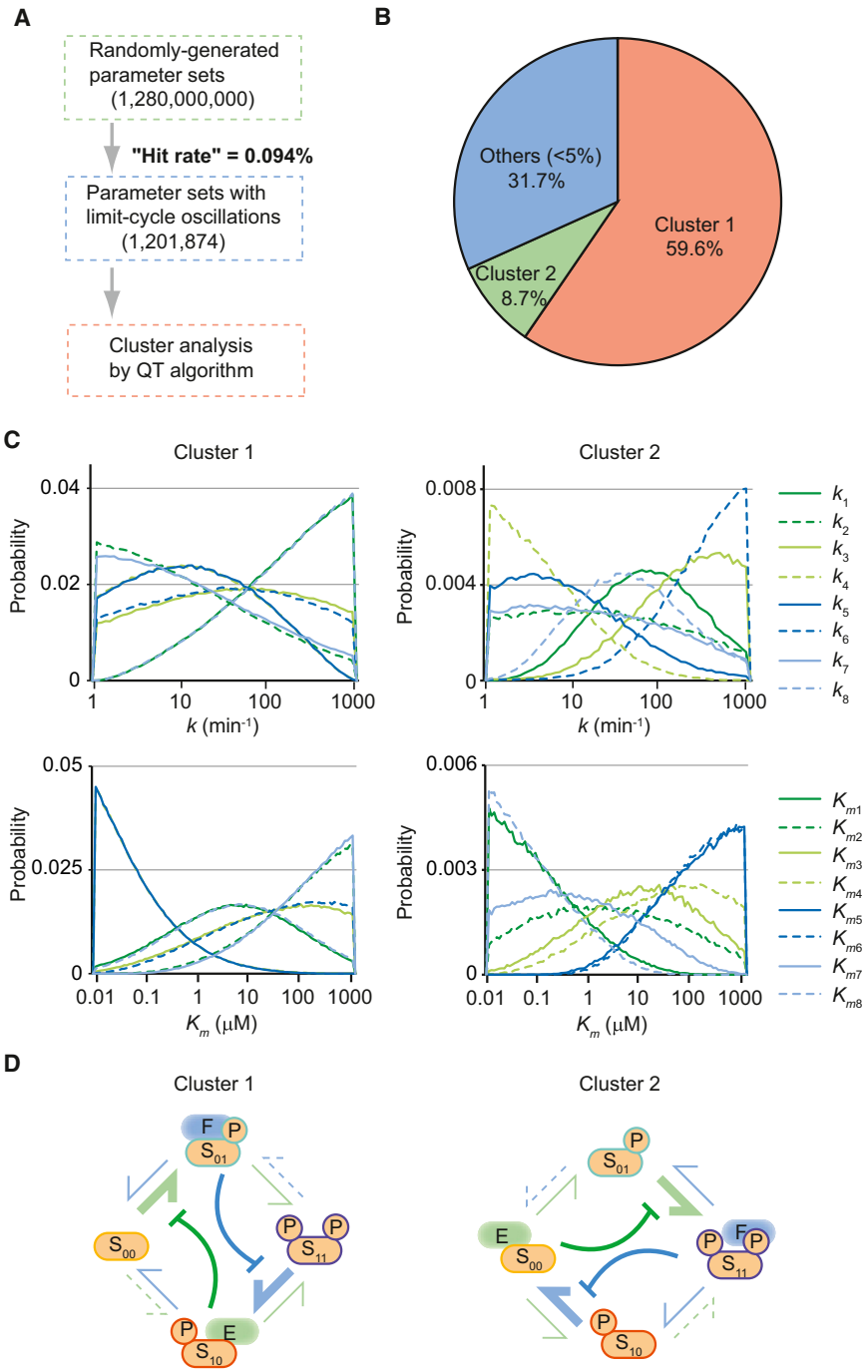



**Figure 2. Oscillatory Systems Possess Two Design Motifs**

(A) Workflow for motif identification.

(B) Results of QT clustering with a (unitless) diameter of 10.5.

(C) Parameter histograms for the two largest clusters. Although considerable variability exists, clear trends in rate and binding constants can be distinguished.

(D) Schematic representation of cluster motifs. Fast reactions are indicated by thick arrows; slow reactions are indicated by dotted arrows. Indirect negative regulation by enzyme sequestration is indicated by curved arrows with flat ends.

See also Figure S2.

fairly slow, it will proceed to completion because its reverse reaction ($S_{00} \rightarrow S_{10}$) is inhibited by the kinase sequestration. The next forward reaction, $S_{00} \rightarrow S_{01}$, requires the kinase and is unable to proceed until $S_{10}$ has been depleted by conversion to $S_{00}$. Once the $S_{10} \rightarrow S_{00}$ reaction has gone to completion, nearly the entire substrate population will be in the $S_{00}$ state and the kinase will no longer be sequestered by $S_{10}$. At this point, the kinase can bind $S_{00}$ and the inhibited $S_{00} \rightarrow S_{01}$ reaction can proceed. In the same way, the phosphatase is sequestered by $S_{01}$ (i.e., $K_{m5}$ is low), inhibiting the subsequent $S_{11} \rightarrow S_{10}$ step. These sequestration steps create checkpoints that ensure that the oscillatory cycle cannot proceed until one substrate state has been depleted and most of the substrate population has moved into the subsequent state.

The situation in cluster 2 is only slightly different. A low value of $K_{m1}$ means that $S_{00}$ binds the kinase strongly. Unlike the unproductive $S_{10}$-kinase complex, this complex is involved in a forward reaction and leads to the production of $S_{01}$. The subsequent kinase-catalyzed $S_{01} \rightarrow S_{11}$ reaction, however, cannot proceed until $S_{00}$ has been depleted, resulting in a checkpoint that requires nearly the entire population to accumulate as $S_{01}$. The low value of $K_{m8}$ means that the phosphatase is sequestered by $S_{11}$, and therefore most of the population has to accumulate as $S_{10}$ before the $S_{10} \rightarrow S_{00}$ step can proceed.

The remaining small clusters (~30% of the total) are generally similar to either cluster 1 or cluster 2. If all parameter sets assigned to neither cluster 1 nor cluster 2 are grouped together, their parameter distribution is similar to that of cluster 1 or cluster 2. For example, Figure S2B shows both cluster-2-type characteristics (high $k_3$ and $k_6$, low $k_4$ and $k_5$, and low $K_{m1}$ and $K_{m8}$) and cluster-1-type characteristics (high $k_1$ and $k_8$, low $k_2$ and $k_7$, and low $K_{m4}$ and $K_{m5}$). To a first approximation, therefore, solutions can be categorized as being similar to either cluster 1 or cluster 2. The presence of two distinct oscillatory clusters in parameter space can also be confirmed by principal components analysis (PCA) (Figures S2C and S2D).

To summarize, an extensive search of the 16-dimensional parameter space revealed two types of oscillatory parameter

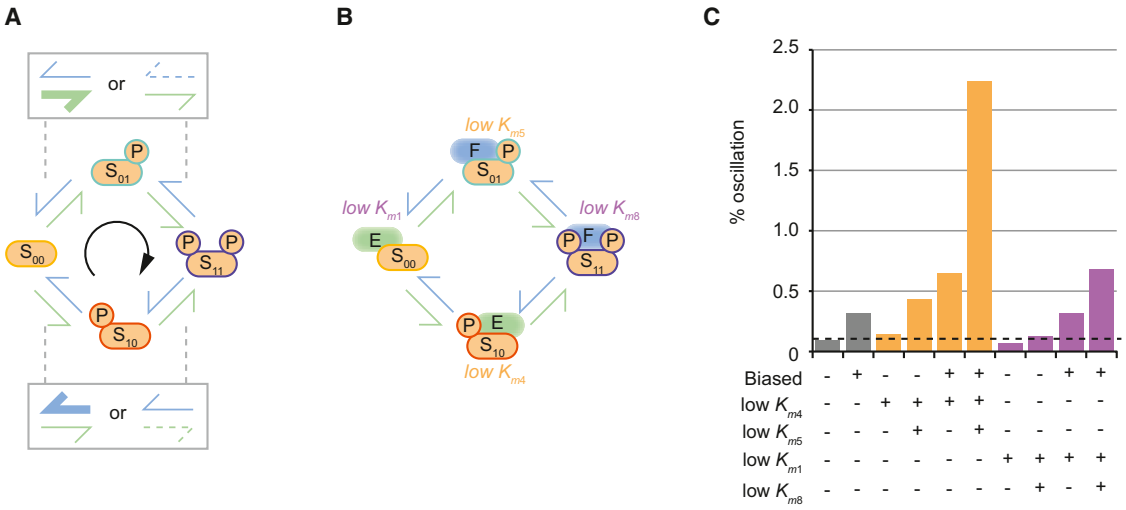



**Figure 3. Functional Significance of Two Design Motifs**

(A) Unidirectional modification cycle. Clockwise reactions are faster than counterclockwise ones, leading the system to prefer the ordering $S_{00}$,$S_{01}$,$S_{11}$,$S_{10}$,$S_{00}$….

(B) Indirect negative regulation by enzyme sequestration.

(C) Imposing the strong binding associated with enzyme sequestration during the random parameter search results in modest gains in the oscillator discovery rate. If this constraint is combined a clockwise rate constant bias, the discovery rate rises substantially.

See also Figure S3.

sets, illustrated in Figure 2D. Based on their parameter distributions, we have proposed two design motifs: a unidirectional bias of catalytic rate constants and indirect negative regulation through enzyme sequestration.

## Functional Significance of Two Design Motifs

The random-search results surely contain a large number of statistical correlations. How then can we determine whether the design motifs described above actually promote oscillation? One approach is to begin with a design motif and use it to constrain the search process (Figures 3A and 3B). If an alleged motif helps to improve the rate of oscillator discovery, then it is functionally significant. Enforcing a clockwise bias in the rate constants, for example, increases the hit rate 3.4-fold (Figure 3C). Constraining the search such that $K_{m4}$ is low improves the success rate 1.5-fold. If both constraints are combined, a 6.8-fold increase is achieved. Requiring both $K_{m4}$ and $K_{m5}$ to be low and maintaining the rate constant bias enforces the entire cluster 1 pattern; this gives an oscillator discovery rate that is 23.6 times the unbiased value. Qualitatively similar results can be obtained for cluster 2. Requiring a low value for $K_{m1}$ fails to improve the success rate, but combining constraints on $K_{m1}$ and/or $K_{m8}$ with the rate constant bias yields significant improvement.

Parameter histograms similar to those in Figure 2C were also constructed from the constrained search results (Figure S3). Although only part of the cluster 1 pattern was enforced (clockwise bias and low $K_{m4}$), the complete pattern was often present. For example, constraining $K_{m4}$ to be low resulted in a low value for $K_{m5}$, and the distributions of the catalytic rate constants are similar to those seen in Figure 2C. Although the distribution for $K_{m5}$ is broader than the distribution imposed for $K_{m4}$, and the presence of two strong sequestration motifs is not absolutely essential for oscillations, it is clear that oscillations are far more probable with both motifs present than with only one present.

Although the identification of design motifs from unbiased random search results can be somewhat subjective, the "inverse analysis" presented in this section shows that the motifs identified are functionally significant. Not only are unidirectional bias and enzyme sequestration typically present in oscillating parameter sets, demanding their presence also increases the number of oscillators found.

## Period- and Amplitude-Determining Processes

Once we had identified oscillator-enriched regions of the parameter space, we aimed to determine which specific parameters have a strong impact on system properties, in particular the oscillation period and amplitude. In Figure 4A, an example of a symmetric cluster 1 (see Table S1) is used as the reference parameter set for bifurcation calculations (Strogatz, 1994). The bifurcation diagrams in Figures 4A and S4A show the equilibrium points (either stable or unstable) of the system, as well as the period and amplitude of oscillations when they exist. Similar bifurcation methods can also be used to characterize the robustness of the oscillations (Figure S4B; Extended Experimental Procedures).

The slow steps by which the enzyme-sequestering substrate states are depleted can be expected to have a large effect on the period. The bifurcation diagram for $k_3$ confirms this intuition. Although the impact of $k_3$ on the oscillation amplitude is minimal, decreasing $k_3$ causes the period to increase dramatically, until oscillations terminate in a Hopf bifurcation at $k_3 \cong 7 \times 10^{-4}$ (Figure 4A). Figure 4B illustrates the impact of $k_3$ on period determination. When $k_3$ is reduced by half, the transition phase from $S_{01}$ to $S_{11}$ (yellow-shaded) nearly doubles.

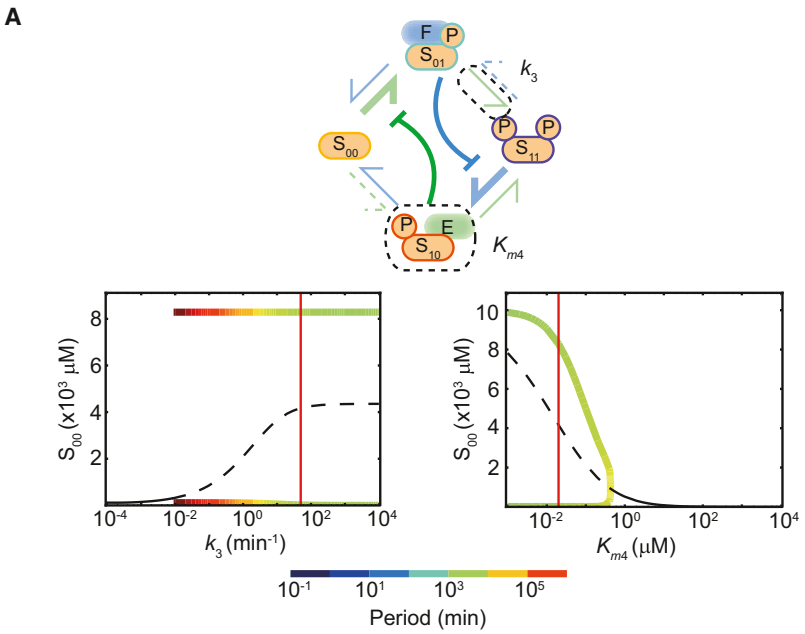
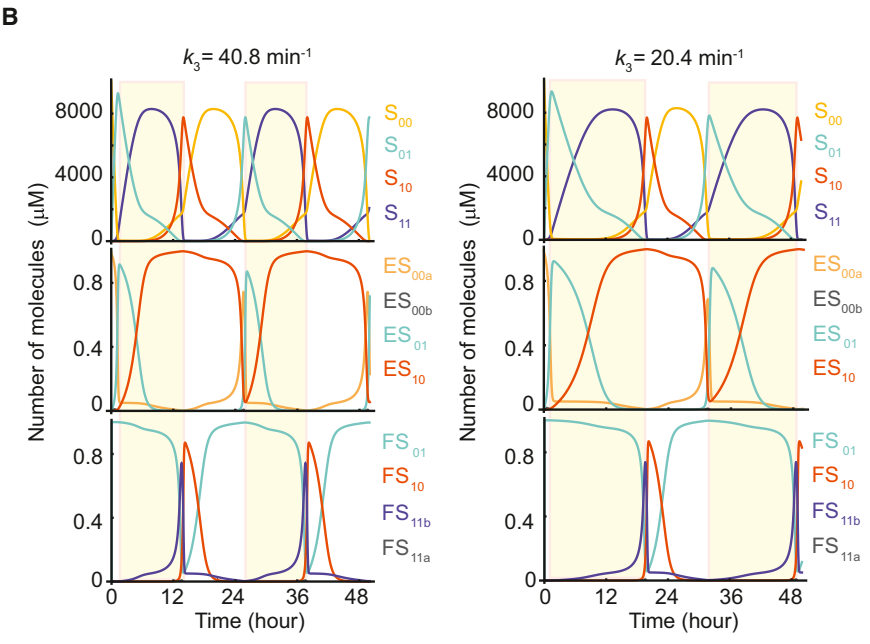



**Figure 4. Period- and Amplitude-Determining Processes**

(A) Two examples of single-parameter bifurcation diagrams. Solid black lines denote stable fixed points, and dashed black lines indicate an unstable fixed point surrounded by a limit cycle. Red vertical lines denote the parameter value in the symmetric example set; see Table S1 for precise values. Colored lines show oscillation extrema, with the period indicated by the line color; $k_3$ has a dramatic effect on the period, whereas $K_{m4}$ primarily affects the amplitude. A more extensive set of bifurcation diagrams is shown in Figure S4.

(B) Effects of decreasing a single rate constant. The oscillation cycle is divided roughly into two halves: the net conversion of $S_{10}$ to $S_{01}$ (via $S_{00}$) is shown on a white background, and the net conversion of $S_{01}$ to $S_{10}$ (via $S_{11}$) is shown on a shaded background. In the symmetric case (left), these phases are of equal length. If $k_3$ (the rate constant for the slow $S_{01} \rightarrow S_{11}$ conversion) is halved, the length of the $S_{01} \rightarrow S_{11} \rightarrow S_{10}$ phase (shaded background) nearly doubles, indicating that the reaction with rate $k_3$ is rate limiting for this phase. See also Figure S4 and Table S1.




$K_{m4}$, the binding constant of kinase to $S_{10}$, can be decreased apparently without limit, whereas the maximum oscillatory value of $K_{m4}$ is limited by a supercritical Hopf bifurcation (Figure 4A). Tight binding of the kinase by $S_{10}$ (and of the phosphatase by $S_{01}$) is an important part of the design motif described above, and the loss of sequestration rapidly abolishes oscillations.

## Robustness of the Oscillation Period against Stochastic Fluctuations

The mass-action approach assumes that chemical concentrations are continuous variables that can take on any positive value. At typical cellular volumes and concentrations, however, the discrete nature of molecules may be important and a stochastic approach becomes necessary. Stochastic simulations also allow us to quantify the robustness of oscillations against internal noise. Noise tolerance is an important requirement for realistic biological models.

The results of stochastic simulations for cluster 1 are shown in Figure 5 (see Table S2; similar plots for cluster 2 are shown in Figures S5A and S5B). For 600 molecules of each type of enzyme (maintaining the concentrations used above for a cell-like volume of $10^{-15}$ L), the oscillations are extremely robust. As the system shrinks (i.e., the total number of molecules decreases), the oscillator runs slowly relative to the deterministic limit (Figures 5A, S5C, and S5D; see also Extended Experimental Procedures). Even when only one of each type of enzyme was present, the period length was surprisingly robust (the standard deviation of the oscillation period was ~2% of the mean period). These results suggest that our model system is fairly robust to the internal noise that will prevail at the cellular scale.

## Coupling of Single-Molecule Oscillators Is a Design Principle

Further insight can be gained by tracking the state of individual molecules. Figure 5B shows the results of stochastic simulations. The top panels show the bulk population behavior, and the middle panels show smoothed trajectories of the four

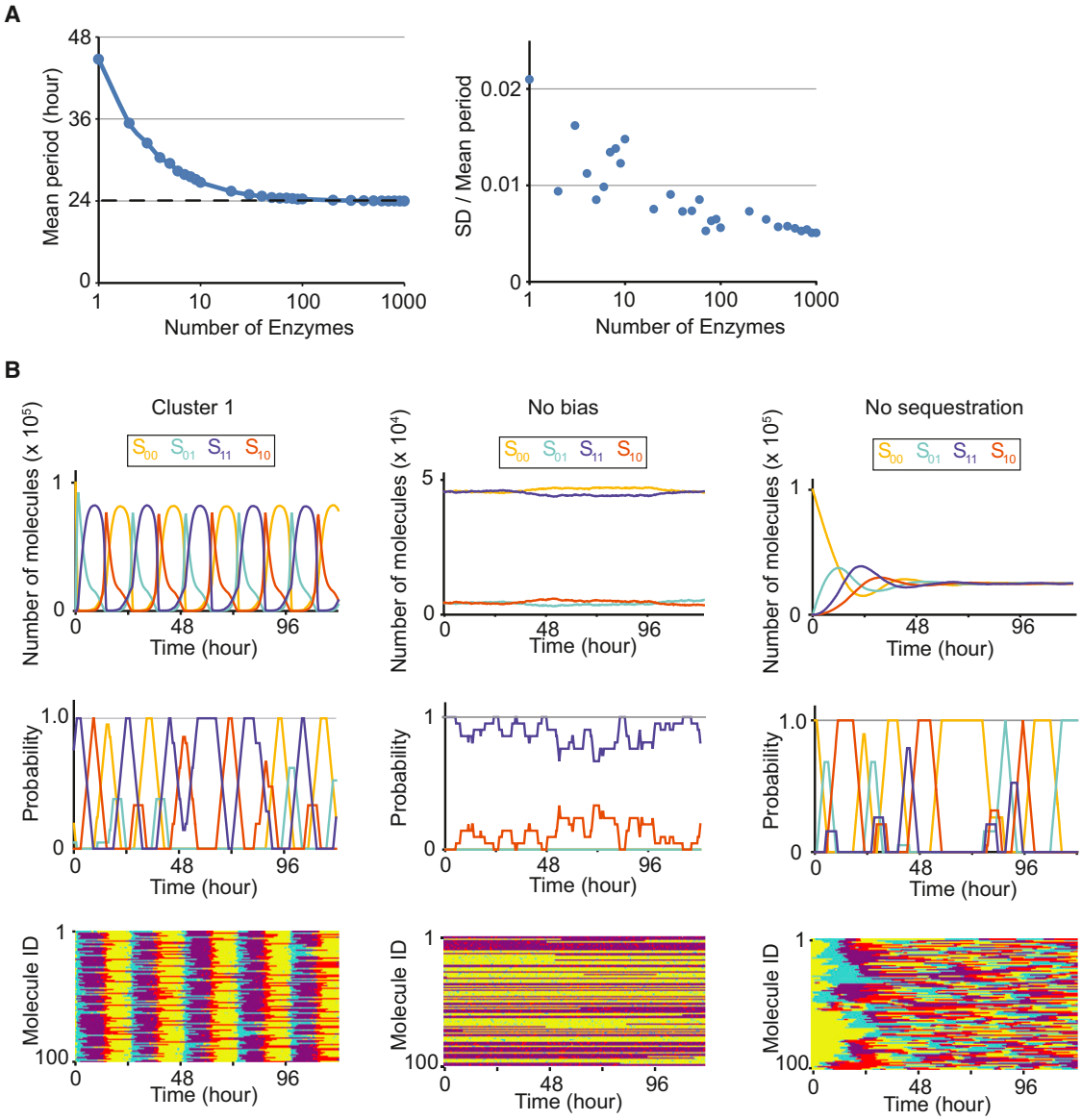



**Figure 5. Coupling of Single-Molecule Oscillators Is a Design Principle**

(A) Robustness of the oscillation period against stochastic fluctuations. Low-copy-number oscillations are slow (left) and variable (right) relative to larger systems. In the continuous limit, the period is precisely 24 hr with zero variability.

(B) Top row: Plots of the bulk population for the cluster 1 oscillator (left), similar binding constants lacking a rate constant bias (middle), and similar rate constants with a clockwise bias but no sequestration (right). Middle row: Smoothed traces showing the state of a single (arbitrarily chosen) substrate molecule. Bottom row: Results for 100 individually tracked molecules, each shown as a horizontal raster line, color-coded as in the plots above. If synchronization by enzyme sequestration is removed, the substrates still act as single-molecule oscillators but lack population-level coherence. If the unidirectional rate constant bias is eliminated, the substrates become trapped and fail to progress around the cycle. See also Figure S5 and Table S2.

phosphorylation states of an arbitrarily chosen substrate molecule. In the bottom panels of Figure 5B, each row of pixels represents a single substrate molecule, and different colors represent the four phosphorylation states.

The middle panel of the "Cluster 1" column (Figure 5B) shows cyclic transitions among the four phosphorylation states. Each individual substrate molecule acts as a noisy oscillator, synchronized with the rest of the population. In addition, the bottom panel of the Cluster 1 column indicates clear bands of $S_{11}$ (blue) and $S_{00}$ (yellow) dominance; individual substrates experience $S_{01}$ (cyan) transiently during the transition into the $S_{11}$ band, and $S_{10}$ (red) during the transition into the $S_{00}$ band. Also note that the $S_{00} \rightarrow S_{01}$ and $S_{11} \rightarrow S_{10}$ transitions seem fairly sharp, whereas the $S_{10} \rightarrow S_{00}$ and $S_{01} \rightarrow S_{11}$ transitions are more diffuse. This suggests that the population is coupled at the $S_{00} \rightarrow S_{01}$ and $S_{11} \rightarrow S_{10}$ transitions, where they are

synchronized by enzyme sequestration. The same is true for sequestration of phosphatase by $S_{01}$.

We next observed single-molecule behaviors when either one of the two design motifs was removed. When the unidirectional motif was removed (by setting $k_1$–$k_8$ equal but maintaining binding affinities), the individual substrate molecules no longer cycled through phospho-states in a defined order (the "No bias" column). When the sequestration motif was removed (by setting all binding affinities equal), phospho-state transitions failed to synchronize (the "No sequestration" column). Although the bulk population converged to a stable fixed point, individual substrate molecules still transitioned between phospho-states in a defined order (middle and bottom panels). These results confirm that (1) the design motif of unidirectional modification cycle constitutes a single-molecule oscillator, and (2) the design motif of enzyme sequestration serves to synchronize noisy single-molecule oscillators, suggesting that coupling of single-molecule oscillators is a design principle.

### The Period of a Simple PTO Can Be Temperature Compensated

To qualify as a bona fide circadian clock, an ~24 hr cellular rhythm must also exhibit temperature compensation, meaning that the oscillation period is fairly robust to changes in temperature. The temperature coefficient $Q_{10}$ measures the factor by which a process accelerates when the temperature is raised by 10°C. Typical $Q_{10}$ values for biochemical reactions range from 2 to 3; measured values for circadian clocks range from 0.8–1.4 (Dunlap et al., 2003).

In our model system, one can calculate $Q_{10}$ values for the oscillation period by assuming temperature coefficients $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ for the kinase and phosphatase, and rescaling the rate constants in accordance with a given temperature change (Extended Experimental Procedures; Figures 6A and S6A–S6C). $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ do not uniquely determine the oscillation $Q_{10}$ ($Q_{10}^{(cycle)}$); by analyzing a large number of parameter sets, one can calculate a histogram of $Q_{10}$ values (Figure 6B). In general, larger values of $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ cause the peak of $Q_{10}^{(cycle)}$ to be shifted to higher values, and the width of the $Q_{10}^{(cycle)}$ distribution increases with the difference between $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$. We considered the case with $Q_{10}^{(E)} = 1$ and $Q_{10}^{(F)} = 3$, corresponding to a temperature-insensitive kinase (such as CKIε/δ; Isojima et al., 2009) and a typical phosphatase. The resulting histogram includes populations that are strongly undercompensated ($Q_{10} >> 1$) as well as some that are overcompensated ($Q_{10} < 1$; Figure 6C).

Different degrees of temperature sensitivity correlate with changes in the distribution of individual parameters (Figures 6D, S6D, and S6E). In general, undercompensated oscillators will have higher rates for the temperature-insensitive kinase reactions and lower rates for the temperature-sensitive phosphatase reactions, such that the rate-determining steps are strongly affected by temperature. For well-compensated oscillators, the opposite is true (Figure 6E), and temperature compensation can be observed even in stochastic simulations (Figure 6F; Table S3). See the Extended Experimental Procedures for a more extensive discussion of this issue.

## DISCUSSION

### General Applicability of a Design Principle Composed of Two Design Motifs

Although the model developed in this study is very simple, we can extract a design principle of wider applicability. Two design motifs appear to be required for robust oscillations (Figure 7A). The first is a well-defined ordering of phosphorylation states, in which the rate constants in the forward direction around the loop are (in general) faster than those in the reverse direction. The other design motif is the presence of synchronizing checkpoints at which single-molecule oscillators that have progressed more quickly must stop and wait for the others. This is accomplished by enzyme sequestration: the cycle cannot proceed past a checkpoint until a strong-binding substrate population has been sufficiently depleted to permit competing reactions.

The general design principles uncovered in this study can also be applied to situations in which more than two phosphorylation sites or more than one kinase (and phosphatase) are present (Figure 7B). The principle may be also conserved when protein degradation and synthesis are incorporated in an appropriate position: if a substrate acts as a repressor for its own transcription and is degraded depending on phosphorylation, then such a transcription-degradation pathway will be compatible with the role of phosphatase reaction (i.e., the disappearance of phosphorylated substrate will lead to the appearance of unphosphorylated substrate; Figure 7C; Extended Experimental Procedures).

In this study we considered the behavior of an isolated PTO; oscillators in real biological systems are responsive to external signals that can modify their behavior. In this context, the relationship of our model to "integrators" (which accumulate a response to incoming signals) and "resonators" (which respond preferentially to signals with a given frequency) may be an interesting avenue for future study (Conrad et al., 2008; Guantes and Poyatos, 2006).

### Two Design Motifs Are Present in the Natural Circadian Clocks

The design principle mentioned above can easily be related to known properties of substrates and enzymes involved in circadian clock systems. In the cyanobacterial clock, the phosphorylation states of the clock protein KaiC have a well-defined cyclic order (Nishiwaki et al., 2007; Rust et al., 2007), similar to the first design motif (unidirectional modification cycle) in our model. In higher eukaryotes, CKIε/δ predominantly phosphorylates the consensus sequence pS-x-x-S*, where pS is a phosphorylated serine, x can be any amino acid, and S* is the target serine (Flotow et al., 1990). This motif promotes ordered phosphorylation: if the substrate contains a pS-x-x-S-x-x-S sequence, CKIε/δ will tend to phosphorylate the second serine before the third. In mammalian PER proteins, contiguous S-x-x-S motifs have been shown to be important for period determination (Toh et al., 2001; Vanselow et al., 2006; Xu et al., 2007).

For the second motif (enzyme sequestration), tight binding between the enzyme and substrate is required. Tight enzyme-substrate binding is found in several enzymes involved in circadian clock systems. In mammals, CKIε/δ forms a stable complex

Cell PRESS
Open ACCESS

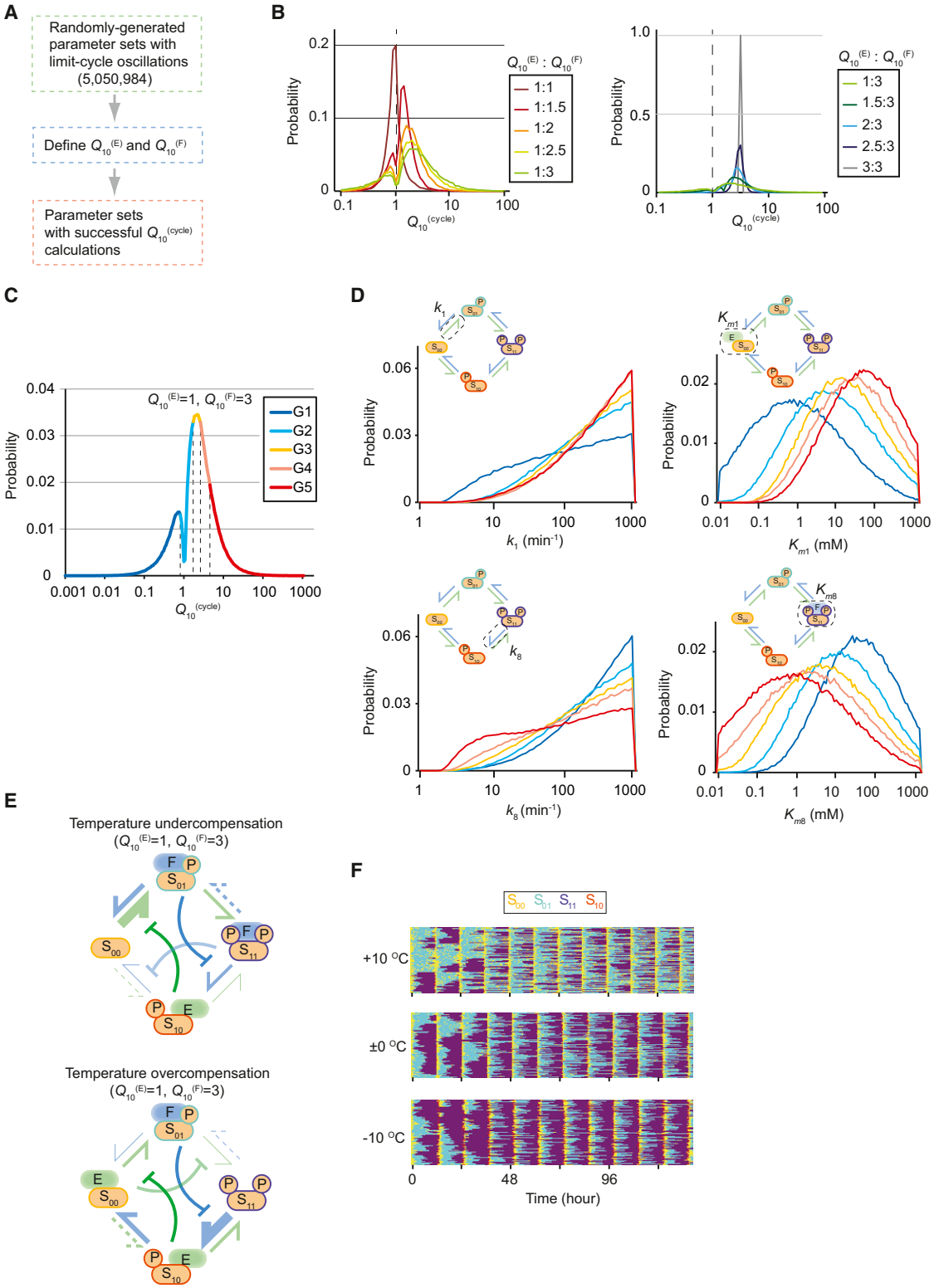



**Figure 6. The Period of a Simple PTO Can Be Temperature Compensated**

(A) Workflow diagram.

(B) Histogram of $Q_{10}^{(cycle)}$ values with different combinations of $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$; $Q_{10}$ for binding/unbinding reactions is assumed to be 3.0 for all calculations.

(C) Histogram of $Q_{10}^{(cycle)}$ values when $Q_{10}^{(E)}$ = 1, $Q_{10}^{(F)}$ = 3. Parameter sets were separated into one of five groups depending on their degree of temperature overcompensation (blue) or undercompensation (red).

with PERs that is important for clock function (Akashi et al., 2002; Camacho et al., 2001; Lee et al., 2004; Vielhaber et al., 2000). PP1 also interacts strongly with PERs (Gallego et al., 2006b). It is not known whether these tight bindings serve to synchronize substrates in different phases of an oscillation cycle, but both experimental and theoretical studies have highlighted enzyme sequestration as a source of nonlinearity in biochemical reactions (Blüthgen et al., 2006; Legewie et al., 2006, 2007; Markevich et al., 2004).

### Comparison with Other PTO Models

The design motifs uncovered in this study can also be seen in other mathematical models of PTOs. MAPK is typically modeled as a two-site system with strictly sequential phosphorylation, such that only three states exist (Huang and Ferrell, 1996; Kholodenko, 2000; Markevich et al., 2004). Feed-forward inhibition interactions similar to those described in this study can generate bistability, but oscillations require larger cascades or feedback loops. The cyanobacterial circadian clock has also been modeled extensively (Clodong et al., 2007; Hatakeyama and Kaneko, 2012; Rust et al., 2007; van Zon et al., 2007). These models also tend to converge on design features similar to ours, in that they typically employ some mechanism (often involving conformational changes or complex formation) to ensure that individual KaiC hexamers visit phosphorylation states in a well-defined order, and the oscillations of the individual hexamers are synchronized through competition for a limited pool of KaiA. A more detailed examination of individual models can be found in the Extended Experimental Procedures.

Although previous studies generally focused on correctly reproducing the behavior of a specific experimentally characterized system, here we focused on mapping out possible behaviors for a very general system. It is therefore quite striking that our approach, which began with no constraining assumptions about what kind of design features should promote oscillation, should lead to fairly similar conclusions.

### Strategies for Experimental Validation

Our modeling results can inform the design or discovery of PTOs in several ways. No unusual substrate properties are necessary, so a simple substrate, such as a short peptide, might be adequate.

The low probability of finding oscillations in our nonbiased search may raise questions about the feasibility of discovering PTOs experimentally. Results from the constrained-parameter search, however, show that one can increase the probability >20-fold by requiring sequential (de)phosphorylation and tight enzyme-substrate binding. These features are likely to be implemented (at least in part) in substrates and enzymes found in natural oscillatory systems. Accordingly, as a starting point for experimental investigation, CKIε/δ, a PER-derived peptide, and an appropriate phosphatase might be a suitable ex vivo PTO.

## CONCLUSIONS

This study illustrates the potential for robust oscillations to exist in a very simple kinase-phosphatase reaction system. This requires two design motifs: a well-ordered sequence of phosphorylation states and synchronization by enzyme sequestration. Although the vast majority of possible parameter sets examined here did not result in sustained oscillations, we were able to construct a typical example of an oscillatory parameter set with reaction rates and binding constants that were within a reasonable range for polypeptide kinases. The ultimate test of the utility of this proposed model, "proof by synthesis," lies ahead.

## EXPERIMENTAL PROCEDURES

### Modeling the PTO with Two Phosphorylation Sites

The PTO can be described using a system of 14 coupled ordinary differential equations (ODEs) describing the concentrations of four substrate phosphoforms, eight substrate-enzyme complexes, and two unbound enzymes (Figure 1B; the full system of equations is presented in Extended Experimental Procedures). In the case of branching reactions (e.g., phosphorylation of $S_{00}$ to form either $S_{01}$ or $S_{10}$), it is assumed that two different types of enzyme-substrate complexes are formed ($ES_{00a}$ or $ES_{00b}$, respectively), consistent with the presence of two distinct phosphorylation sites on the substrate molecule. Calculations were also performed with the opposite assumption (a single complex $ES_{00}$ that can lead to either $S_{01}$ or $S_{10}$), and the results were qualitatively similar (Figure S1).

### Biased Parameter Search

When a forward bias in the rate constants was desired (Figure 3), numbers were selected from the exponential distribution two at a time, with the larger number being assigned to a forward (clockwise) reaction ($k_1$, $k_3$, $k_6$, or $k_8$) and the smaller number assigned to the corresponding reverse (counterclockwise) reaction ($k_5$, $k_7$, $k_2$, or $k_4$, respectively; Figure 3A). Enzyme sequestration was encouraged by selecting constrained binding constants from an exponential distribution over 0.01–0.05 μM rather than 0.01–1,000 μM (Figure 3B).

### Defining "Typical" Parameter Sets

A stereotypical example was generated for each of the dominant clusters; the parameter values chosen are listed in Table S1, and oscillations for the cluster 1 example are depicted in Figure 4A. These example clusters were designed to exhibit symmetry, so that $k_i = k_j$ and $K_{mi} = K_{mj}$ for $(i,j) \in \{(1,8),(2,7),(3,6),(4,5)\}$. Rate constants were scaled to ensure an oscillation period of 24 hr.

### Bifurcation Analysis

Beginning from these stereotypical parameter sets, each parameter was systematically varied beyond the range of values over which oscillations are possible, and concentrations at the fixed points and the oscillation extrema were calculated. Although the starting parameter sets were symmetric, this symmetry was not maintained during bifurcation calculations. Note that, because of symmetry, bifurcation calculations were needed for only half of the parameters. For example, the diagram for $k_8$ can be obtained by

(D) Examples of parameter histograms for different oscillator groups, color-coded as in (C). The parameters $k_1$ and $k_8$ are located at symmetric positions in the network (Figure 1B) and would be equivalent if the roles of kinase and phosphatase were interchanged.
(E) Schematic diagrams of the extremes of temperature sensitivity. In the undercompensated case (top) the rate-limiting steps are catalyzed by the temperature-sensitive phosphatase, whereas in the overcompensated case (bottom) the temperature-insensitive kinase steps are rate limiting.
(F) Stochastic simulation results for a particular temperature-compensated parameter set.
See also Figure S6 and Table S3.

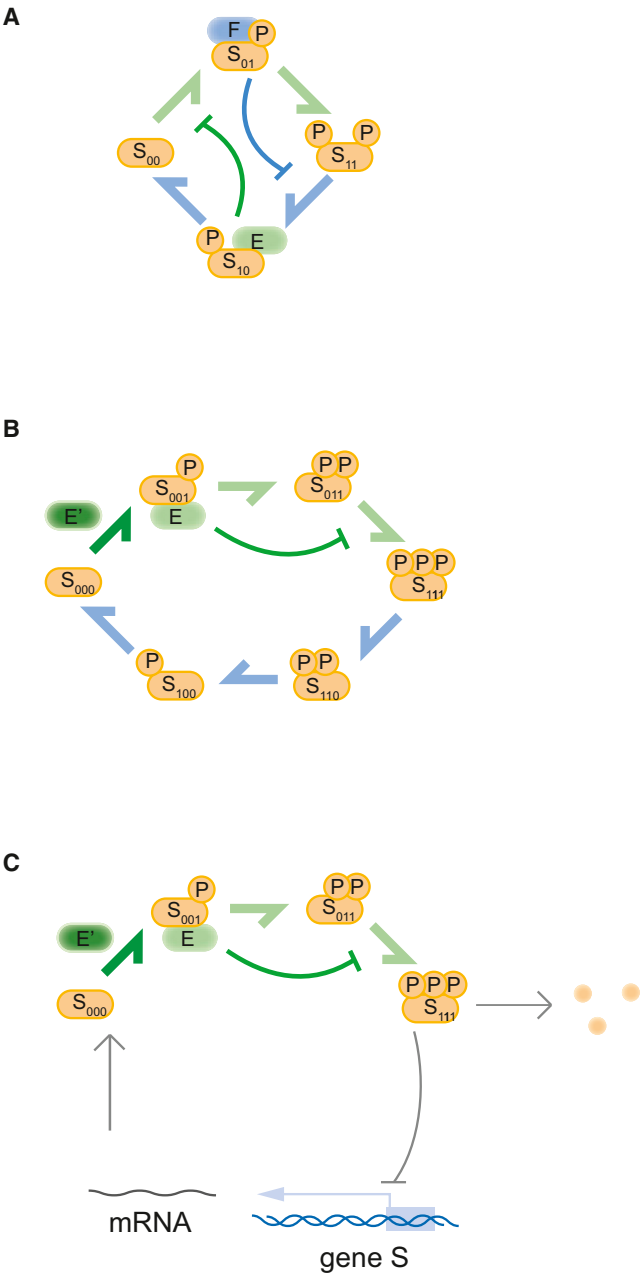



**Figure 7. Generalizations of the PTO Model**

(A) Simplified view of the PTO model examined in this work.

(B) The mechanisms outlined in this work could be realized in a more complex system with a larger number of phosphorylation states and more than two enzymes.

(C) Similar considerations could apply in a case in which the dephosphorylation of highly phosphorylated substrate is replaced by coupled processes of degradation and gene expression.

exchanging $S_{00}$ with $S_{11}$ and $S_{01}$ with $S_{10}$ in the $k_1$ calculation; the bifurcation points and period dependence will be identical.

### Stochastic Simulation

We studied the reaction networks using the stochastic simulation algorithm (Gillespie, 1977) as implemented in SPPARKS (Plimpton et al., 2009). To monitor the behavior of individual substrate molecules, we modified the stochastic simulations so that in addition to a large pool of "ordinary" substrate molecules, the system contains a smaller number of "labeled" substrate molecules whose progression through the four phosphorylation states can be tracked. The interactions of the four phospho-forms of these labeled substrate molecules with the kinase and phosphatase are identical to their unlabeled counterparts, but each can be followed in the stochastic simulation as a separate molecular population, allowing us to observe how a single substrate molecule experiences the oscillatory cycle.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2012.09.006.

## LICENSING INFORMATION



## ACKNOWLEDGMENTS

We thank Drs. Tatsuo Shibata and Ryokichi Tanaka for valuable comments. C.C.J. was supported by the RIKEN Foreign Postdoctoral Researcher Program. Large-scale calculations used the RIKEN Integrated Cluster of Clusters. This research was supported by an intramural Grant-in-Aid from the RIKEN Center for Developmental Biology and the Quantitative Biology Center (H.R.U.), the Uehara Memorial Foundation (H.R.U.), the Mitsubishi Foundation (H.R.U.), the President's Fund from RIKEN (H.R.U.), and a Grant-in-Aid for Scientific Research on Innovative Areas (No. 3306) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan (H.R.U.).

Received: June 11, 2012
Revised: August 22, 2012
Accepted: September 8, 2012
Published online: October 18, 2012

## REFERENCES

Akashi, M., Tsuchiya, Y., Yoshino, T., and Nishida, E. (2002). Control of intracellular dynamics of mammalian period proteins by casein kinase I epsilon (CKIepsilon) and CKIdelta in cultured cells. Mol. Cell. Biol. *22*, 1693–1703.

Blüthgen, N., Bruggeman, F.J., Legewie, S., Herzel, H., Westerhoff, H.V., and Kholodenko, B.N. (2006). Effects of sequestration on signal transduction cascades. FEBS J. *273*, 895–906.

Camacho, F., Cilio, M., Guo, Y., Virshup, D.M., Patel, K., Khorkova, O., Styren, S., Morse, B., Yao, Z., and Keesler, G.A. (2001). Human casein kinase Idelta phosphorylation of human circadian clock proteins period 1 and 2. FEBS Lett. *489*, 159–165.

Chen, Z., Yoo, S.H., Park, Y.S., Kim, K.H., Wei, S., Buhr, E., Ye, Z.Y., Pan, H.L., and Takahashi, J.S. (2012). Identification of diverse modulators of central and peripheral circadian clocks by high-throughput chemical screening. Proc. Natl. Acad. Sci. USA *109*, 101–106.

Chickarmane, V., Kholodenko, B.N., and Sauro, H.M. (2007). Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. J. Theor. Biol. *244*, 68–76.

Clodong, S., Dühring, U., Kronk, L., Wilde, A., Axmann, I., Herzel, H., and Kollmann, M. (2007). Functioning and robustness of a bacterial circadian clock. Mol. Syst. Biol. *3*, 90.

Conrad, E., Mayo, A.E., Ninfa, A.J., and Forger, D.B. (2008). Rate constants rather than biochemical mechanism determine behaviour of genetic clocks. J. R. Soc. Interface *5*(*Suppl 1*), S9–S15.

Dibner, C., Sage, D., Unser, M., Bauer, C., d'Eysmond, T., Naef, F., and Schibler, U. (2009). Circadian gene expression is resilient to large fluctuations in overall transcription rates. EMBO J. *28*, 123–134.

Dibner, C., Schibler, U., and Albrecht, U. (2010). The mammalian circadian timing system: organization and coordination of central and peripheral clocks. Annu. Rev. Physiol. *72*, 517–549.

Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271–290.

Dunlap, J.C., Loros, J.J., and Decoursey, P.J. (2003). Chronobiology: Biological Timekeeping, First Edition (Sunderland, MA: Sinauer Associates).

Fan, Y., Hida, A., Anderson, D.A., Izumo, M., and Johnson, C.H. (2007). Cycling of CRYPTOCHROME proteins is not necessary for circadian-clock function in mammalian fibroblasts. Curr. Biol. *17*, 1091–1100.

Ferrell, J.E., Jr., Tsai, T.Y., and Yang, Q. (2011). Modeling the cell cycle: why do certain circuits oscillate? Cell *144*, 874–885.

Flotow, H., Graves, P.R., Wang, A.Q., Fiol, C.J., Roeske, R.W., and Roach, P.J. (1990). Phosphate groups as substrate determinants for casein kinase I action. J. Biol. Chem. *265*, 14264–14269.

Gallego, M., and Virshup, D.M. (2007). Post-translational modifications regulate the ticking of the circadian clock. Nat. Rev. Mol. Cell Biol. *8*, 139–148.

Gallego, M., Eide, E.J., Woolf, M.F., Virshup, D.M., and Forger, D.B. (2006a). An opposite role for tau in circadian rhythms revealed by mathematical modeling. Proc. Natl. Acad. Sci. USA *103*, 10618–10623.

Gallego, M., Kang, H., and Virshup, D.M. (2006b). Protein phosphatase 1 regulates the stability of the circadian protein PER2. Biochem. J. *399*, 169–175.

Gekakis, N., Staknis, D., Nguyen, H.B., Davis, F.C., Wilsbacher, L.D., King, D.P., Takahashi, J.S., and Weitz, C.J. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. Science *280*, 1564–1569.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. *81*, 2340–2361.

Griffin, E.A., Jr., Staknis, D., and Weitz, C.J. (1999). Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. Science *286*, 768–771.

Guantes, R., and Poyatos, J.F. (2006). Dynamical principles of two-component genetic oscillators. PLoS Comput. Biol. *2*, e30.

Hastings, J.W., and Sweeney, B.M. (1957). On the mechanism of temperature independence in a biological clock. Proc. Natl. Acad. Sci. USA *43*, 804–811.

Hatakeyama, T.S., and Kaneko, K. (2012). Generic temperature compensation of biological clocks by autonomous regulation of catalyst concentration. Proc. Natl. Acad. Sci. USA *109*, 8109–8114.

Hirota, T., Lewis, W.G., Liu, A.C., Lee, J.W., Schultz, P.G., and Kay, S.A. (2008). A chemical biology approach reveals period shortening of the mammalian circadian clock by specific inhibition of GSK-3beta. Proc. Natl. Acad. Sci. USA *105*, 20746–20751.

Hogenesch, J.B., and Ueda, H.R. (2011). Understanding systems-level properties: timely stories from the study of clocks. Nat. Rev. Genet. *12*, 407–416.

Huang, C.Y., and Ferrell, J.E., Jr. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Natl. Acad. Sci. USA *93*, 10078–10083.

Isojima, Y., Nakajima, M., Ukai, H., Fujishima, H., Yamada, R.G., Masumoto, K.H., Kiuchi, R., Ishida, M., Ukai-Tadenuma, M., Minami, Y., et al. (2009). CKIepsilon/δ-dependent phosphorylation is a temperature-insensitive, period-determining process in the mammalian circadian clock. Proc. Natl. Acad. Sci. USA *106*, 15744–15749.

Johnson, C.H., Mori, T., and Xu, Y. (2008). A cyanobacterial circadian clockwork. Curr. Biol. *18*, R816–R825.

Kholodenko, B.N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur. J. Biochem. *267*, 1583–1588.

Kholodenko, B.N. (2006). Cell-signalling dynamics in time and space. Nat. Rev. Mol. Cell Biol. *7*, 165–176.

Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., and Reppert, S.M. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell *98*, 193–205.

Lee, C., Etchegaray, J.P., Cagampang, F.R., Loudon, A.S., and Reppert, S.M. (2001). Posttranslational mechanisms regulate the mammalian circadian clock. Cell *107*, 855–867.

Lee, C., Weaver, D.R., and Reppert, S.M. (2004). Direct association between mouse PERIOD and CKIepsilon is critical for a functioning circadian clock. Mol. Cell. Biol. *24*, 584–594.

Lee, H.M., Chen, R., Kim, H., Etchegaray, J.P., Weaver, D.R., and Lee, C. (2011). The period of the circadian oscillator is primarily determined by the balance between casein kinase 1 and protein phosphatase 1. Proc. Natl. Acad. Sci. USA *108*, 16451–16456.

Legewie, S., Blüthgen, N., and Herzel, H. (2006). Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability. PLoS Comput. Biol. *2*, e120.

Legewie, S., Schoeberl, B., Blüthgen, N., and Herzel, H. (2007). Competing docking interactions can bring about bistability in the MAPK cascade. Biophys. J. *93*, 2279–2288.

Liu, P., Kevrekidis, I.G., and Shvartsman, S.Y. (2011). Substrate-dependent control of ERK phosphorylation can lead to oscillations. Biophys. J. *101*, 2572–2581.

Lowrey, P.L., Shimomura, K., Antoch, M.P., Yamazaki, S., Zemenides, P.D., Ralph, M.R., Menaker, M., and Takahashi, J.S. (2000). Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. Science *288*, 483–492.

Maier, B., Wendt, S., Vanselow, J.T., Wallach, T., Reischl, S., Oehmke, S., Schlosser, A., and Kramer, A. (2009). A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock. Genes Dev. *23*, 708–718.

Markevich, N.I., Hoek, J.B., and Kholodenko, B.N. (2004). Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. J. Cell Biol. *164*, 353–359.

Meng, Q.J., Logunova, L., Maywood, E.S., Gallego, M., Lebiecki, J., Brown, T.M., Sládek, M., Semikhodskii, A.S., Glossop, N.R., Piggins, H.D., et al. (2008). Setting clock speed in mammals: the CK1 epsilon tau mutation in mice accelerates circadian pacemakers by selectively destabilizing PERIOD proteins. Neuron *58*, 78–88.

Meng, Q.J., Maywood, E.S., Bechtold, D.A., Lu, W.Q., Li, J., Gibbs, J.E., Dupré, S.M., Chesham, J.E., Rajamohan, F., Knafels, J., et al. (2010). Entrainment of disrupted circadian behavior through inhibition of casein kinase 1 (CK1) enzymes. Proc. Natl. Acad. Sci. USA *107*, 15240–15245.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science *308*, 414–415.

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

Novák, B., and Tyson, J.J. (2008). Design principles of biochemical oscillators. Nat. Rev. Mol. Cell Biol. *9*, 981–991.

O'Neill, J.S., and Reddy, A.B. (2011). Circadian clocks in human red blood cells. Nature *469*, 498–503.

Okamura, H., Miyake, S., Sumi, Y., Yamaguchi, S., Yasui, A., Muijtjens, M., Hoeijmakers, J.H., and van der Horst, G.T. (1999). Photic induction of mPer1 and mPer2 in cry-deficient mice lacking a biological clock. Science *286*, 2531–2534.

Partch, C.L., Shields, K.F., Thompson, C.L., Selby, C.P., and Sancar, A. (2006). Posttranslational regulation of the mammalian circadian clock by cryptochrome and protein phosphatase 5. Proc. Natl. Acad. Sci. USA *103*, 10467–10472.

Pittendrigh, C.S. (1954). On temperature independence in the clock system controlling emergence time in *Drosophila*. Proc. Natl. Acad. Sci. USA *40*, 1018–1029.

Plimpton, S., Battaile, C., Chandross, M., Holm, L., Thompson, A., Tikare, V., Wagner, G., Webb, E., Zhou, X., Garcia-Cardona, C., et al. (2009). Crossing the

mesoscale no-man's land via parallel kinetic Monte Carlo. Sandia Report SAND2009-6226.

Qiao, L., Nachbar, R.B., Kevrekidis, I.G., and Shvartsman, S.Y. (2007). Bistability and oscillations in the Huang-Ferrell model of MAPK signaling. PLoS Comput. Biol. *3*, 1819–1826.

Reppert, S.M., and Weaver, D.R. (2002). Coordination of circadian timing in mammals. Nature *418*, 935–941.

Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O'Shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science *318*, 809–812.

Sato, T.K., Yamada, R.G., Ukai, H., Baggs, J.E., Miraglia, L.J., Kobayashi, T.J., Welsh, D.K., Kay, S.A., Ueda, H.R., and Hogenesch, J.B. (2006). Feedback repression is required for mammalian circadian clock function. Nat. Genet. *38*, 312–319.

Schlosser, A., Vanselow, J.T., and Kramer, A. (2005). Mapping of phosphorylation sites by a multi-protease approach with specific phosphopeptide enrichment and NanoLC-MS/MS analysis. Anal. Chem. *77*, 5243–5250.

Schmutz, I., Wendt, S., Schnell, A., Kramer, A., Mansuy, I.M., and Albrecht, U. (2011). Protein phosphatase 1 (PP1) is a post-translational regulator of the mammalian circadian clock. PLoS ONE *6*, e21325.

Shankaran, H., Ippolito, D.L., Chrisler, W.B., Resat, H., Bollinger, N., Opresko, L.K., and Wiley, H.S. (2009). Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. Mol. Syst. Biol. *5*, 332.

Strogatz, S.H. (1994). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (Reading, MA: Addison-Wesley).

Takahashi, J.S., Hong, H.K., Ko, C.H., and McDearmon, E.L. (2008). The genetics of mammalian circadian order and disorder: implications for physiology and disease. Nat. Rev. Genet. *9*, 764–775.

Takano, A., Shimizu, K., Kani, S., Buijs, R.M., Okada, M., and Nagai, K. (2000). Cloning and characterization of rat casein kinase 1epsilon. FEBS Lett. *477*, 106–112.

Thomson, M., and Gunawardena, J. (2009). Unlimited multistability in multisite phosphorylation systems. Nature *460*, 274–277.

Toh, K.L., Jones, C.R., He, Y., Eide, E.J., Hinz, W.A., Virshup, D.M., Ptácek, L.J., and Fu, Y.H. (2001). An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome. Science *291*, 1040–1043.

Tsuchiya, Y., Akashi, M., Matsuda, M., Goto, K., Miyata, Y., Node, K., and Nishida, E. (2009). Involvement of the protein kinase CK2 in the regulation of mammalian circadian rhythms. Sci. Signal. *2*, ra26.

Ueda, H.R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., Iino, M., and Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. Nat. Genet. *37*, 187–192.

Ukai-Tadenuma, M., Yamada, R.G., Xu, H., Ripperger, J.A., Liu, A.C., and Ueda, H.R. (2011). Delay in feedback repression by cryptochrome 1 is required for circadian clock function. Cell *144*, 268–281.

van Zon, J.S., Lubensky, D.K., Altena, P.R., and ten Wolde, P.R. (2007). An allosteric model of circadian KaiC phosphorylation. Proc. Natl. Acad. Sci. USA *104*, 7420–7425.

Vanselow, K., Vanselow, J.T., Westermark, P.O., Reischl, S., Maier, B., Korte, T., Herrmann, A., Herzel, H., Schlosser, A., and Kramer, A. (2006). Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). Genes Dev. *20*, 2660–2672.

Vielhaber, E., Eide, E., Rivers, A., Gao, Z.H., and Virshup, D.M. (2000). Nuclear entry of the circadian regulator mPER1 is controlled by mammalian casein kinase I epsilon. Mol. Cell. Biol. *20*, 4888–4899.

von Gall, C., Noton, E., Lee, C., and Weaver, D.R. (2003). Light does not degrade the constitutively expressed BMAL1 protein in the mouse suprachiasmatic nucleus. Eur. J. Neurosci. *18*, 125–133.

Walton, K.M., Fisher, K., Rubitski, D., Marconi, M., Meng, Q.J., Sládek, M., Adams, J., Bass, M., Chandrasekaran, R., Butler, T., et al. (2009). Selective inhibition of casein kinase 1 epsilon minimally alters circadian clock period. J. Pharmacol. Exp. Ther. *330*, 430–439.

Xu, Y., Padiath, Q.S., Shapiro, R.E., Jones, C.R., Wu, S.C., Saigoh, N., Saigoh, K., Ptácek, L.J., and Fu, Y.H. (2005). Functional consequences of a CKIdelta mutation causing familial advanced sleep phase syndrome. Nature *434*, 640–644.

Xu, Y., Toh, K.L., Jones, C.R., Shin, J.Y., Fu, Y.H., and Ptácek, L.J. (2007). Modeling of a human circadian mutation yields insights into clock regulation by PER2. Cell *128*, 59–70.

Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. Nat. Rev. Genet. *2*, 702–715.

Cell PRESS
Open ACCESS

## EXTENDED EXPERIMENTAL PROCEDURES

### Full System of Equations (Mass-Action Kinetics)

(See "Modeling the PTO with Two Phosphorylation Sites" in Experimental Procedures section.)

$$\frac{d[S_{00}]}{dt} = -(k_{b1}+k_{b2})[S_{00}][E] + k_{ub1}[ES_{00a}] + k_{ub2}[ES_{00b}] + k_5[FS_{01}] + k_6[FS_{10}]$$

$$\frac{d[S_{01}]}{dt} = -k_{b3}[S_{01}][E] + k_{ub3}[ES_{01}] - k_{b5}[S_{01}][F] + k_{ub4}[FS_{01}] + k_1[ES_{00a}] + k_7[FS_{11a}]$$

$$\frac{d[S_{10}]}{dt} = -k_{b4}[S_{10}][E] + k_{ub4}[ES_{10}] - k_{b6}[S_{10}][F] + k_{ub6}[FS_{10}] + k_2[ES_{00b}] + k_8[FS_{11b}]$$

$$\frac{d[S_{11}]}{dt} = -(k_{b7}+k_{b8})[S_{11}][F] + k_{ub7}[FS_{11a}] + k_{ub8}[FS_{11b}] + k_3[ES_{01}] + k_4[ES_{10}]$$

$$\frac{d[ES_{00a}]}{dt} = k_{b1}[E][S_{00}] - (k_{ub1}+k_1)[ES_{00a}]$$

$$\frac{d[ES_{00b}]}{dt} = k_{b2}[E][S_{00}] - (k_{ub2}+k_2)[ES_{00b}]$$

$$\frac{d[ES_{01}]}{dt} = k_{b3}[E][S_{01}] - (k_{ub3}+k_3)[ES_{01}]$$

$$\frac{d[ES_{10}]}{dt} = k_{b4}[E][S_{10}] - (k_{ub4}+k_4)[ES_{10}]$$

$$\frac{d[FS_{01}]}{dt} = k_{b5}[F][S_{01}] - (k_{ub5}+k_5)[FS_{01}]$$

$$\frac{d[FS_{10}]}{dt} = k_{b6}[F][S_{10}] - (k_{ub6}+k_6)[FS_{10}]$$

$$\frac{d[FS_{11a}]}{dt} = k_{b7}[F][S_{11}] - (k_{ub7}+k_7)[FS_{11a}]$$

$$\frac{d[FS_{11b}]}{dt} = k_{b8}[F][S_{11}] - (k_{ub8}+k_8)[FS_{11b}]$$

$$\frac{d[E]}{dt} = -((k_{b1}+k_{b2})[S_{00}] + k_{b3}[S_{01}] + k_{b4}[S_{10}])[E] + (k_{ub1}+k_1)[ES_{00a}] + (k_{ub2}+k_2)[ES_{00b}] + (k_{ub3}+k_3)[ES_{01}] + (k_{ub4}+k_4)[ES_{10}]$$

$$\frac{d[F]}{dt} = -(k_{b5}[S_{01}] + k_{b6}[S_{10}] + (k_{b7}+k_{b8})[S_{11}])[F] + (k_{ub5}+k_5)[FS_{01}] + (k_{ub6}+k_6)[FS_{10}] + (k_{ub7}+k_7)[FS_{11a}] + (k_{ub8}+k_8)[FS_{11b}]$$

$[S_{00}]$, $[S_{01}]$, $[S_{10}]$ and $[S_{11}]$ represent concentrations of the four substrate phosphorylation states, $[E]$ and $[F]$ are the concentrations of free (unbound) kinase and phosphatase, and $[ES_{00a}]$, $[ES_{00b}]$, $[ES_{01}]$, $[ES_{10}]$, $[FS_{01}]$, $[FS_{10}]$, $[FS_{11a}]$ and $[FS_{11b}]$ are the concentrations of the eight distinct enzyme-substrate complexes. The catalytic rate constants (i.e., number of enzyme-substrate complexes converted to enzyme/product pairs per unit time) are $k_1$ through $k_8$, binding constants (i.e., number of substrate molecules bound into enzyme-substrate complexes per enzyme per unit time) are $k_{b1}$ through $k_{b8}$ and unbinding constants (i.e., number of enzyme-substrate complexes dissociating into enzyme/substrate pairs per unit time) are $k_{ub1}$ through $k_{ub8}$.

Closed, isolated systems in which all processes are reversible obey the detailed balance condition, which requires that the system's time evolution converge to a unique, stable equilibrium point (Kampen, 2007). In our system, the unbinding of product from enzyme is irreversible, so detailed balance is violated and oscillations are possible. Using chemical reaction network theory (Feinberg, 1987a) it can easily be shown that reversible product unbinding will eliminate the possibility of oscillations. For similar reasons, replacing the kinase and phosphatase by a single bifunctional enzyme capable of catalyzing both reactions reintroduces detailed balance and precludes oscillations.

### Simplified System with Michaelis-Menten Approximation

It can easily be shown that this system of equations is not independent, and can be reformulated as a system of 11 ODEs and three algebraic equations describing the conserved total levels of substrate, kinase, and phosphatase. The system can be simplified further if one makes the Michaelis-Menten approximation (Michaelis and Menten, 1913), in which the binding and unbinding of substrate-enzyme complexes is assumed to be in instantaneous equilibrium, so that the concentrations of the enzyme-substrate complexes $[ES_{00a}]$, $[ES_{00b}]$, $[ES_{01}]$, $[ES_{10}]$, $[FS_{01}]$, $[FS_{10}]$,$[FS_{11a}]$ and $[FS_{11b}]$ can be replaced with their steady-state expressions (Figure 1C).

$$K_{mx} = \frac{(k_{ubx}+k_x)}{k_{bx}},\ x \in 1,2,3,4,5,6,7,8$$

$$\frac{d[S_{00}]}{dt} = -\left(\frac{k_1}{K_{m1}} + \frac{k_2}{K_{m2}}\right)[S_{00}][E] + \frac{k_5[F][S_{01}]}{K_{m5}} + \frac{k_6[F][S_{10}]}{K_{m6}}$$

$$\frac{d[S_{01}]}{dt} = \frac{-k_3[S_{01}][E]}{K_{m3}} - \frac{k_5[F][S_{01}]}{K_{m5}} + \frac{k_1[E][S_{00}]}{K_{m1}} + \frac{k_7[F][S_{11}]}{K_{m7}}$$

$$\frac{d[S_{11}]}{dt} = -\left(\frac{k_7}{K_{m7}} + \frac{k_8}{K_{m8}}\right)[S_{11}][F] + \frac{k_3[E][S_{01}]}{K_{m3}} + \frac{k_4[E][S_{10}]}{K_{m4}}$$

$$[S_{10}] = S_{\text{tot}} - ([S_{00}] + [S_{01}] + [S_{11}] + (E_{\text{tot}} - [E]) + (F_{\text{tot}} - [F]))$$

$$[E] = \frac{E_{\text{tot}}}{(1 + [S_{00}]/K_{m1} + [S_{00}]/K_{m2} + [S_{01}]/K_{m3} + [S_{10}]/K_{m4})}$$

$$[F] = \frac{F_{\text{tot}}}{(1 + [S_{01}]/K_{m5} + [S_{10}]/K_{m6} + [S_{11}]/K_{m7} + [S_{11}]/K_{m8})}$$

The resulting equations represent the system as a system of three ODEs and three algebraic equations; it is possible to eliminate the algebraic equations by substitution, but the resulting nonlinear expressions are unwieldy and do little to aid in understanding the system, so we do not present them here. The binding and unbinding rate constants have been replaced by the binding constants $K_{m1}$ through $K_{m8}$. $K_{mi}$ is the substrate concentration at which the rate of reaction $i$ reaches half of its maximum; if binding is strong then $K_{mi}$ will be low because the enzyme-substrate binding is easily saturated. Numerical integration indicates that this reduced system gives identical results to the full dynamical system for physically-reasonable values of the rate constants.

### Stability of the Single-Site Phosphorylation System

(See "Multisite Substrate Modifications Can Oscillate with Two Opposing Enzymes" in Results section).

If only a single phosphorylation site is present, then the system of mass-action equations can be written as:

$$\frac{d[S_0]}{dt} = -k_{b0}[E][S_0] + k_{ub0}[ES_0] + k_F[FS_1]$$

$$\frac{d[S_1]}{dt} = -k_{b1}[F][S_1] + k_{ub1}[FS_1] + k_E[ES_0]$$

$$\frac{d[ES_0]}{dt} = k_{b0}[E][S_0] - (k_{ub0} + k_E)[ES_0]$$

$$\frac{d[FS_1]}{dt} = k_{b1}[F][S_1] - (k_{ub1} + k_F)[FS_1]$$

$$\frac{d[E]}{dt} = -k_{b0}[E][S_0] + (k_{ub0} + k_E)[ES_0]$$

$$\frac{d[F]}{dt} = -k_{b1}[F][S_1] + (k_{ub1} + k_F)[FS_1]$$

$[S_0]$ and $[S_1]$ represent the concentrations of the unphosphorylated and phosphorylated forms of the substrate; $[ES_0]$ and $[FS_1]$ are the concentrations of their complexes with the kinase (E) and the phosphatase (F), respectively. If we use the conservation relations $[E] = E_{\text{tot}} - [ES_0]$ and $[F] = F_{\text{tot}} - [FS_0]$, then the system can be simplified:

$$\frac{d[S_0]}{dt} = -k_{b0}(E_{tot} - [ES_0])[S_0] + k_{ub0}[ES_0] + k_F[FS_1]$$

$$\frac{d[S_1]}{dt} = -k_{b1}(F_{tot} - [FS_1])[S_1] + k_{ub1}[FS_1] + k_E[ES_0]$$

$$\frac{d[ES_0]}{dt} = k_{b0}(E_{tot} - [ES_0])[S_0] - (k_{ub0} + k_E)[ES_0]$$

$$\frac{d[FS_1]}{dt} = k_{b1}(F_{tot} - [FS_1])[S_1] - (k_{ub1} + k_F)[FS_1]$$

For typical conditions, we can make the Michaelis-Menten approximation and set:

$$[ES_0] = \frac{k_{b0}E_{tot}[S_0]}{k_{ub0} + k_E + k_{b0}[S_0]}$$

$$[FS_1] = \frac{k_{b1}F_{tot}[S_1]}{k_{ub1} + k_F + k_{b1}[S_1]}$$

The system can now be simplified to only two equations:

$$\frac{d[S_0]}{dt} = -\frac{k_{b0}k_E E_{tot}[S_0]}{k_{ub0}+k_E+k_{b0}[S_0]} + \frac{k_{b1}k_F F_{tot}[S_1]}{k_{ub1}+k_F+k_{b1}[S_1]}$$

$$\frac{d[S_1]}{dt} = -\frac{d[S_0]}{dt}$$

Clearly, $[S_0]+[S_1]$ is a constant of the motion, and we can write $[S_0] + [S_1] = S_{\text{eff}}$. This allows us to further simplify to:

$$\frac{d[S_0]}{dt} = -\frac{k_{b0}k_E E_{tot}[S_0]}{k_{ub0}+k_E+k_{b0}[S_0]} + \frac{k_{b1}k_F F_{tot}(S_{eff}-[S_0])}{k_{ub1}+k_F+k_{b1}(S_{eff}-[S_0])}$$

This is a one-dimensional system, and therefore cannot contain any attractors other than fixed points (Strogatz, 1994). Since all trajectories are bounded (due to the conservation relations) the only possible long-term behavior is convergence to a stable fixed point. In other words, limit-cycle oscillations are impossible as long as the Michaelis-Menten approximation is valid.

If the Michaelis-Menten approximation is not made, it appears that much less can be said. Using the tools of chemical reaction network theory (CRNT) (Craciun and Feinberg, 2005, 2006; Feinberg, 1987b, 1988), the system

$$S_0 + E \rightleftarrows ES_0 \rightarrow E + S_1$$

$$S_1 + F \rightleftarrows FS_1 \rightarrow F + S_0$$

can be shown to have a deficiency of 1 and to be injective, meaning that it exhibits a single unique physically-relevant steady state (i.e., one in which all concentrations are real and positive). A topological analysis based on CRNT is not, however, able to determine the conditions under which this steady state might be unstable and a limit cycle might be present. By comparison, the reaction network described in Figure 1 has a deficiency of 3 and is not injective; multistationarity cannot be ruled out on topological grounds alone. CRNT calculations were carried out using the CRN Toolbox (http://www.che.eng.ohio-state.edu/~feinberg/crnt/).

## Parameter Search

The system with dual phosphorylation sites is too large and complex for straightforward analytical study, so we chose to identify solutions numerically through an extensive search of parameter space. Our underlying assumption is that the biological plausibility of an emergent property (in this case, sustained oscillations) is proportional to the probability of observing this property in an ensemble of systems whose parameters are randomly selected. This is an approximation to the real situation, in which the existence of a biological feature depends on its emergence via a series of random mutations, but a random search over parameter values is far more theoretically tractable than a random search in genotype space.

Solutions were generated by selecting parameters from an exponential distribution bounded to the interval 1-1000 $\text{min}^{-1}$ for $k_1$ through $k_8$ and 0.01-1000 μM for $K_{m1}$ through $K_{m8}$. While the presence of oscillations can be determined by linearizing the system of equations about the fixed point and determining its stability (see Stability Analysis below), in practice it was more computationally efficient to simply integrate the mass-action equations and determine whether they converge to a constant value or a limit cycle. Integration used a fourth-order stiffly-stable Rosenbrock scheme (Press, 2007). Evidence of multistability was never observed, but multistability could not be ruled out conclusively. Integrations always began with $[S_{00}] = 10^4$, [E] = 1, [F] = 1, and all other concentrations set to zero; any stable attractors whose basin of attraction does not include this point would not have been observed.

## Stability Analysis

A necessary condition for the presence of sustained oscillations is that the dynamics of the system converge to a limit cycle rather than a fixed point. The conservation relations for S, E, and F make it clear that all trajectories are bounded – if the fixed points of the system can be shown to be unstable, motion on a higher-dimensional attractor is the only other possibility. (While chaotic motion on a strange attractor is possible in systems with more than two dimensions, sustained aperiodic trajectories were not observed in any of the parameter sets examined in this study. Once initial transients have disappeared, the system is effectively two-dimensional.) This can be done by finding a (real, positive) set of concentration values such that all time derivatives are equal to zero and linearizing the system about this point by calculating the Jacobian (Strogatz, 1994). If the real part of at least one eigenvalue of the Jacobian is positive, then the fixed point is unstable. For some parameter sets, an unstable fixed point coexists with a fixed point to which all trajectories converge, resulting in a false positive for instability (and hence oscillation) if the unstable solution is used to generate the Jacobian. Rather than attempt to exhaustively determine the roots of the nonlinear system of equations (which can in general be quite challenging) we observed that bona fide oscillatory regions in parameter space are often bounded by a Hopf bifurcation, in which two complex-conjugate eigenvalues of the Jacobian cross the imaginary axis at the same time. It is not uncommon, however, for the (positive) real parts of these eigenvalues to diverge significantly as the system moves further from the Hopf

bifurcation. By requiring that two of the eigenvalues be greater than zero, oscillatory and nonoscillatory parameter sets could be reliably distinguished.

## Clustering and PCA Analysis

Parameter sets were categorized using quality threshold (QT) clustering (Heyer et al., 1999). Given a metric for determining the distance between two points in parameter space, the QT algorithm finds the largest set of points such that all are within a distance $d$ of each other. These are then removed from consideration, and the largest cluster within the remaining collection of points is identified. This process continues until all points have been assigned to a cluster. This process becomes rather memory-intensive as the size of the data set to be clustered grows. A two-step process was employed for very large data sets in which the full QT algorithm was run on an initial seed set (~10,000 points) and each of the remaining points was assigned to the closest cluster. The results obtained by the two-step method for large data sets were qualitatively similar to those obtained by the full QT method for smaller data sets.

The distance metric was designed to take several factors into account. First, because the parameters were exponentially distributed in the random search, the distance metric was calculated on the logarithms of parameters. This ensures that (for example) a point with $k_1 = 900$ and a point with $k_1 = 1000$ are more similar than a point with $k_1 = 1$ and a point with $k_1 = 101$. In Figure 1B, the essential features of the reaction network are unchanged if $S_{10}$ and $S_{01}$ are exchanged. This means that, for any two points in parameter space, two possible distances between them exist, of which the shortest was chosen. Lastly, the rate constants $\{k_i\}$ have units of inverse time, and rescaling $\{k_i\}$ by a constant factor has no effect on the resulting dynamics except for the rescaling of the time axis. Therefore, when each pairwise distance was calculated, rate constants were rescaled to minimize the distance metric. These measures ensure that large distances between parameter sets will correspond to major differences in the oscillatory dynamics.

The choice of cluster diameter is somewhat arbitrary; two major clusters emerged for a range of cluster diameters and a diameter was chosen at which these two clusters accounted for ~70% of solutions (Figure 2B; Figure S2A). Each cluster was "aligned" by identifying the cluster member with the shortest average distance to others in the cluster and choosing the symmetry orientation of each cluster that minimizes the distance to this reference member. The parameter distributions in aligned clusters can then be calculated; representative histograms for the major clusters are shown in Figure 2C.

The remaining small clusters (~30% of the total) are generally similar either to Cluster 1 or Cluster 2; if all parameter sets assigned to neither Cluster 1 nor Cluster 2 are grouped together, their parameter distribution is similar to Cluster 1 or Cluster 2. For example, Figure S2B shows both Cluster-2-type characteristics (high $k_3$, $k_6$, low $k_4$, $k_5$, low $K_{m1}$, $K_{m8}$) and Cluster-1-type characteristics (high $k_1$, $k_8$, low $k_2$, $k_7$, low $K_{m4}$, $K_{m5}$). To a first approximation, therefore, solutions can be categorized as being similar to either Cluster 1 or Cluster 2.

The presence of two distinct oscillatory clusters in parameter space can also be confirmed using principal components analysis (PCA). If a covariance matrix is constructed from the logarithms of the oscillatory parameter values and normalized so that each column has zero mean and unit variance, the eigenvalues of this matrix can be used to reduce the dimensionality of parameter space by projecting it down onto the (orthogonal) directions that contain the largest share of the sample variance. When a random sample of oscillatory parameter sets is projected down onto the first two principal components (which together account for about 30% of the variance), two distinct clusters are visible, separated by an empty space (Figures S2C and S2D). Identifying the points on the PCA projection according to the cluster to which they were assigned by the QT algorithm makes it clear that Cluster 1 and Cluster 2 are qualitatively distinct and that no other major groupings are present.

## Detailed Analysis of Oscillation Mechanism

(See "Period- and Amplitude-Determining Processes" in Results section.)

Figure 4B divides the oscillation cycle of a typical Cluster 1 oscillator into two phases: one involves the net conversion of $S_{10}$ to $S_{01}$ (via $S_{00}$) and is shown against a white background; the other involves the conversion of $S_{01}$ back to $S_{10}$ via $S_{11}$ and is shown against a shaded background. The system is clearly symmetric (in the left side of Figure 4B) – the two phases look identical if $S_{10}$ and $S_{00}$ are exchanged with $S_{01}$ and $S_{11}$, respectively. In addition, the sequestration patterns of the kinase and phosphatase (lower two panels) have identical shapes, except for a 12 hr phase shift – both enzymes spend most of their time bound to a single substrate and associate with other binding partners only during brief antiphasic episodes.

At the beginning of the $S_{01} \rightarrow S_{11} \rightarrow S_{10}$ phase (colored background), $S_{01}$ is highly abundant and saturates the phosphatase almost completely, rendering it nearly inactive. Because $S_{01}$ is the only highly abundant species at this point, it also binds most of the kinase, despite the fairly low affinity of the kinase for $S_{01}$. Unlike the phosphatase, the kinase reacts efficiently when bound to $S_{01}$, causing $S_{01}$ to be converted to $S_{11}$ at a respectable rate. As $S_{11}$ accumulates, some of it is converted to very small amounts of $S_{10}$ –initially $S_{10}$ is hardly visible in the top panel (bulk protein concentrations), but the affinity of the kinase for $S_{10}$ is so high that it still rapidly sequesters most of the kinase away from $S_{01}$. Once the kinase sequestration by $S_{10}$ saturates (at around the middle of this phase), the depletion of $S_{01}$ slows substantially. Once the levels of $S_{01}$ have nearly reached zero, something interesting finally happens with the phosphatase. After spending nearly all of this phase bound unproductively to $S_{01}$, it is finally free to seek other partners; the most viable candidate is $S_{11}$, now the most abundant species in the system. The association of phosphatase and $S_{11}$ is, however, destined to be short-lived, as $S_{11}$ is precipitously converted to $S_{10}$, completing the net conversion of $S_{01}$ to $S_{10}$.

The $S_{10} \rightarrow S_{00} \rightarrow S_{01}$ phase follows precisely the same pattern. Initially, both the kinase and the phosphatase are bound primarily by the highly-abundant $S_{10}$. The $S_{10} \rightarrow S_{11}$ reaction is inefficient, so the kinase remains unproductively bound to $S_{10}$ through most of this phase. The phosphatase, however, rapidly converts $S_{10}$ to $S_{00}$. With the abundant but unattractive $S_{10}$ out of the picture, the phosphatase can then bind almost exclusively to the trace amounts of $S_{01}$ that have begun to appear, rendering it nearly-inactive for the rest of this phase. This causes the rate of the $S_{10} \rightarrow S_{00}$ conversion to slow substantially, with the slow decline of $S_{10}$ effectively determining the length of this phase. Once it has disappeared completely, the kinase has nowhere to turn except for the now-abundant $S_{00}$, which is rapidly consumed to produce $S_{01}$. Once $S_{01}$ again becomes the dominant species in the system, the cycle is ready to begin anew.

In the right panel, the symmetry of the system is broken by halving the rate of the $S_{01} \rightarrow S_{11}$ conversion. The concentration traces in the $S_{10} \rightarrow S_{00} \rightarrow S_{01}$ phase (white background) are nearly identical to the symmetric case, indicating that this particular reaction has a minimal impact on this phase. The $S_{01} \rightarrow S_{11} \rightarrow S_{10}$ phase (colored background), on the other hand, nearly doubles in length. Close examination shows that the shape of the transition in phosphatase binding near the end of this phase is unchanged – most of the lengthening comes from an extension of the $S_{01} \rightarrow S_{11}$ conversion. This explains the large impact of this parameter on the oscillation period – the phosphatase-catalyzed steps all wait for the sequestration by $S_{01}$ to be lifted, so changes that slow down the depletion of $S_{01}$ will impact the period much more than changes that affect the more transient reactions.

## Detailed Discussion of Bifurcations

(See "Period- and Amplitude-Determining Processes" in Results section).

Determining the precise nature of bifurcations in a complex dynamical system can sometimes be challenging, and our identification of the bifurcations illustrated in Figure 4A as Hopf bifurcations is based solely on numerical evidence. In each case, the nature of the bifurcations was determined by observing the behavior of the eigenvalues of the Jacobian at the fixed point as parameter values were varied. The symmetric reference parameter set for Cluster 1 (see Table S1) shows two nondegenerate, positive, real eigenvalues. As the bifurcation point is approached, these two eigenvalues approach each other and real the imaginary axis, so that they become complex conjugates with positive real parts. At the bifurcation point, these two eigenvalues simultaneously cross the imaginary axis, the typical signature of a Hopf bifurcation. Additional evidence comes from the scaling of the oscillation period with $r$, where $r$ is the distance from the bifurcation point. While the period can vary dramatically with parameter changes, it tends to remain constant as $r$ becomes very small, another typical feature of Hopf bifurcations. Finally, the amplitudes of oscillations typically decrease as $r^{1/2}$ as the bifurcation point is approached. This can be demonstrated for the $K_{m4}$ bifurcation shown on the right-hand side of Figure 4A, and for many of the other bifurcations shown in Figure S4. Unfortunately, this characteristic was observed in neither the $k_3$ bifurcation shown in Figure 4A nor the similar $k_1/k_8$ bifurcations for Cluster 2.

## Determining the Typical Robustness of Oscillation to Parameter Changes

(See "Period- and Amplitude-Determining Processes" in Results section).

In addition to detailed analysis of a few example cases, one can also use bifurcation analysis to examine the robustness of a larger collection of parameter sets. We chose ~1000 parameter sets each from Cluster 1 and Cluster 2 and determined the range over which each parameter could be varied before oscillations were lost. Figure S4B shows the average oscillatory ranges for the two major clusters. Because we only examined bifurcations lying between $10^{-3}$ and $10^{4}$, these are conservative estimates; Figures 4A and S4 suggest that parameters can sometimes be varied far outside this range without affecting oscillations. On average, the parameters of Cluster 1 oscillators can tolerate about 1000-fold variation before oscillations are lost; Cluster 2 oscillators are slightly less robust and can tolerate roughly 800-fold variation.

## Sensitivity of the PTO Period to Total Substrate Concentration

(See "Period- and Amplitude-Determining Processes" in Results section.)

Close examination of the bifurcation diagram for $S_{tot}$ (Figure S4A) raises an important concern that should be addressed. While oscillations are lost for $S_{tot} \lesssim 3$; there is no upper bound to the oscillatory regime. For $S_{tot} \gg 3$, the period is almost exactly linear in the $S_{tot}/E_{tot}$ ratio (or, equivalently, $S_{tot}/F_{tot}$). Adjusting this ratio amounts to rescaling the time axis of the simulations; all steps of the oscillation cycle are equally affected. When the substrate/enzyme ratio is large, enzymes will be operating near saturation, meaning that all catalytic rates are roughly equal to the product of substrate and enzyme concentrations, and a change will affect all reactions equally. Near the bifurcation point, enzymes are no longer saturated and this relationship breaks down.

The period of the circadian clock has been shown to be largely insensitive to the global rate of mRNA transcription (Dibner et al., 2009). Is this robustness captured in our simple model? If transcriptional repression decreases the concentrations of substrate, kinase, and phosphatase by the same factor, then the dynamics of the oscillator will be completely unchanged. Some studies (Coté and Brody, 1986), however, indicate that the circadian period may also be insensitive to a selective change in the expression level of PER, a major phosphorylation target of CKIε/δ and the most likely circadian analog to the substrate in our model system.

This difficulty can be avoided if the effective substrate/enzyme ratio is kept at a constant level regardless of the total expression level of each substrate and enzyme. For example, subcellular compartmentalization could be dependent on the formation of a tightly bound complex between enzymes and substrates – if a "date" is required for admission, then stoichiometry will be well-controlled. In

the circadian system, localization of PER to the nucleus requires participation in a multisubunit complex, often including CKIε/δ and phosphatase (Lee et al., 2001).

### Stochastic Effects on the Stability and Period of PTO

(See "Robustness of the Oscillation Period against Stochastic Fluctuations" in Results section).

To get a sense of whether the results depicted in Figure 5 are typical of Clusters 1 and 2 more generally, we used the stochastic simulation algorithm (SSA) to estimate what fraction of oscillators remain viable as the number of enzymes is decreased; roughly 10% remain oscillatory even with only one of each type of enzyme present (Figure S5C). In this sense, Cluster 1 is slightly more robust than Cluster 2, although the differences are small. Interestingly, the two clusters show opposing behaviors as molecular noise becomes more pronounced – the period of Cluster 1 oscillators tends to increase, while that of Cluster 2 oscillators tends to decrease (Figure S5D).

### Biological Significance of the Tunable Period Length

The bifurcation diagram for $k_3$ in Figure 4A makes it clear that a wide range of periods are available in certain parameter regimes. This implies the presence of broadly-applicable design features that might apply both in the circadian field and other oscillatory systems with shorter or longer time scales. Comparison with the bifurcation diagrams shown in Figure S4 shows that, for most parameters, the period variation in the vicinity of the reference parameter set is rather modest – an oscillator based on this scheme is likely to be reasonably mutationally robust. Although tuning of the oscillation period is possible (and likely to be desirable) on evolutionary timescales, Figure 5A shows that the periods of individual oscillators are likely to be fairly consistent, even in the presence of stochastic noise. These features are consistent with the current understanding of circadian rhythmicity: the (mammalian) circadian clock is robust against temperature (Hastings and Sweeney, 1957; Pittendrigh, 1954) and global transcription rate (Dibner et al., 2009), but changes in the activity of CKIε/δ are a fragile point (Gallego and Virshup, 2007; Isojima et al., 2009; Lee et al., 2009). CKI-catalyzed reactions likely play a period-determining role similar to that of the reactions with rates $k_3$ and $k_6$ in our much simpler system.

### Calculating Temperature Sensitivity

For biochemical reactions, temperature sensitivity is often described using the temperature coefficient $Q_{10}$:

$$Q_{10} = \left(\frac{R_2}{R_1}\right)^{10/T_2 - T_1},$$

where $R_1$ and $R_2$ are the rates of a process at temperatures $T_1$ and $T_2$. If $Q_{10} = 2$, the reaction rate doubles over an interval of 10 degrees. Note that, if a reaction has a constant $Q_{10}$, its rate will depend exponentially on the temperature. If the reaction rate is scaled such that $R_2 = \alpha R_1$, the rescaling constant can be expressed as $\alpha = Q_{10}^{\Delta T/10}$.

Rather than explicitly estimating activation energies, the effects of temperature were introduced by assuming $Q_{10}$ values for the kinase and the phosphatase ($Q_{10}^{(E)}$ and $Q_{10}^{(F)}$) and using these to calculate rescaling factors $\alpha_E$ and $\alpha_F$ at a given temperature difference $\Delta T$. The resulting changes in the oscillation period $\tau(\Delta T)$ could usually be fit very well with an exponential curve of the form $\tau = \tau_0 e^{\beta \Delta T}$. The temperature coefficient of the entire cycle can then be expressed as:

$$Q_{10}^{(\mathrm{cycle})} = \left(\frac{R_2}{R_1}\right)^{10/\Delta T} = \left(\frac{\tau_1}{\tau_2}\right)^{10/\Delta T} = e^{-\beta \Delta T\, 10/\Delta T} = e^{-10\beta}$$

For a given parameter set, the rate constants were rescaled for nine points evenly spaced along a temperature range of ± 5°, the period was calculated for each temperature point, and an exponential fit was used to obtain $Q_{10}^{(\mathrm{cycle})}$. Binding-unbinding reactions had a minor impact on temperature sensitivity; we assumed a fairly high $Q_{10}$ value of 3.0. Most fits were of excellent quality; those with $R^2 < 0.9$ were discarded (Figures S6A–S6C).

### Detailed Explanation of Temperature Compensation

(See "The Period of a Simple PTO Can Be Temperature Compensated" in Results section.)

Bona fide circadian clocks involve more than merely autonomous oscillations. A cellular rhythm is generally not considered a circadian clock unless it also exhibits temperature compensation, in which the oscillation period is roughly unchanged by the ambient temperature. The temperature sensitivity of a system is often described using the temperature coefficient $Q_{10}$, which describes the factor by which the temperature changes as the temperature is raised by 10°C. Most biochemical reactions have a $Q_{10}$ value of 2-3, meaning that their rates increase between two and three times in response to a 10°C temperature change. For a temperature-compensated process, $Q_{10} \cong 1$; measured $Q_{10}$ values for circadian clocks range from about 0.8 to 1.4 (Dunlap et al., 2004).

To further deepen the analogy between our simple model system and the circadian clock, we examined its potential for temperature compensation. This was done by assuming temperature coefficients $Q_{10}^{(E)}$ for the kinase and $Q_{10}^{(F)}$ for the phosphatase, and using these to modify rate constants in accordance with an assumed external temperature (see Experimental Procedures; Figures S6A and S6B).

It is obvious that if $Q_{10}^{(E)} = Q_{10}^{(F)} = 2$, every reaction rate in the simple PTO is doubled when the temperature changes by 10°C. In that case, the period is cut in half and temperature compensation is impossible; equal $Q_{10}$ values only permit temperature compensation if both are equal to one. Because we are most interested in the potential behavior of systems with temperature-insensitive kinases such as CKIε/δ, we have generally assumed $Q_{10}^{(E)} = 1$. Of course it is entirely possible that $Q_{10}^{(F)} = 1$, and this would provide a very elegant mechanism for temperature compensation. For the sake of generality, however, we consider the case in which $Q_{10}^{(F)} = 3$, and ask whether temperature compensation is possible even in the presence of a fairly temperature-sensitive phosphatase.

The distribution of $Q_{10}^{(cycle)}$, the temperature coefficient of the oscillation period was calculated as shown in Figure 6A: $Q_{10}^{(cycle)}$ calculations were attempted for $\sim 5 \times 10^6$ oscillating parameter sets. Many parameter sets show oscillations that are robust to temperature change, while most were unsuccessful. The systematic changes in enzyme reaction rates (induced by temperature differences) can involve fairly significant motion in parameter space, often bringing the system outside the oscillatory regime. For $Q_{10}^{(E)} = 1$ and $Q_{10}^{(F)} = 3$, $Q_{10}^{(cycle)}$ could be calculated for $\sim 10^6$ parameter sets. In general, larger values of $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ cause the peak of $Q_{10}^{(cycle)}$ to be shifted to higher values, and the width of the $Q_{10}^{(cycle)}$ distribution increases as the difference between $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ increases (Figures 6B and S6C). It is possible for the oscillation period to be temperature compensated or even overcompensated as long as $Q_{10}^{(E)} \neq Q_{10}^{(F)}$; the probability of temperature compensation increases when values of $Q_{10}^{(E)}$ or $Q_{10}^{(F)}$ is close to 1.

We next sought to determine whether temperature-compensated oscillators share characteristic parameter distributions. For the case with $Q_{10}^{(E)} = 1$ and $Q_{10}^{(F)} = 3$ (i.e., a temperature-insensitive kinase and a temperature-sensitive phosphatase), the collection of parameter sets was sorted according to their calculated $Q_{10}^{(cycle)}$ values and separated into five groups numbered from G1 (most overcompensated; period increases with increasing temperature) to G5 (most undercompensated; period decreases with increasing temperature) (Figure 6C). Parameter distributions similar to those in Figure 2C could then be calculated for each group. A few representative traces are shown in Figure 6D; see also Figures S6D and S6E. The parameter distributions vary smoothly from G1 to G5 and show a striking symmetry property: the G1 distributions are nearly identical to the G5 distributions upon an exchange of $S_{00}$ and $S_{11}$ (reversing the roles of the kinase and phosphatase).

Figure 6E distills these parameter distributions into a schematic form. In the undercompensated case (G5), the fastest reactions in the system are the kinase-driven forward reactions ($S_{00} \rightarrow S_{01} \rightarrow S_{11}$), while the phosphatase-driven reactions ($S_{11} \rightarrow S_{10} \rightarrow S_{00}$) are slower and nearly irreversible. The temperature-insensitive kinase reactions have a negligible impact on the period, while the temperature-sensitive phosphatase steps are rate-limiting, leading to overall strong temperature dependence. In addition, $K_{m8}$ is typically lower than $K_{m6}$ in G5; this means that the phosphatase often will not bind to $S_{10}$ until $S_{11}$ has been largely depleted, introducing an additional sequestration that favors temporal ordering of the slow phosphatase steps. $K_{m1}$, however, is not typically lower than $K_{m3}$, so the fast kinase steps tend to happen concurrently. In the overcompensated G1 case, the opposite is true. The temperature-insensitive kinase reactions are slow, irreversible, and temporally-ordered, giving them a disproportionate impact on the period, while the temperature-sensitive phosphatase steps are fast and concurrent, and have little effect on the period. It should be stressed that overcompensation is probably not due to an excessive period dependence on the temperature-insensitive kinase reactions; if the period depended *only* on these reactions then the cycle would be perfectly compensated.

Both overcompensation and extreme undercompensation (i.e., $Q_{10}^{(cycle)} > Q_{10}^{(F)}$) can best be understood by examining the bifurcation diagrams in Figures 4B and S4. If the system is near a bifurcation point, then small changes in parameters can result in large period changes. The fact that ~80% of $Q_{10}^{(cycle)}$ calculations failed suggests that a large fraction of the randomly-discovered oscillators are near the edge of the oscillatory regime. This is a consequence of the geometry of high-dimensional space: only ~19% of the area of a two-dimensional circle is located within 10% of its edge, while ~81% of the volume of a 16-dimensional ball is located within 10% of its boundary. The oscillatory parameter regime is likely to be nonspherical, further increasing the fraction of its volume located close to a boundary. This intuition is supported by the fact that most of the oscillators with extreme $Q_{10}^{(cycle)}$ values also showed very small oscillation amplitudes, as is often true near bifurcations. Most of the tails of the distribution can be attributed to low-amplitude, marginally-stable oscillators. We would therefore expect a viable temperature-insensitive oscillator to be one in which the temperature-insensitive steps are rate-limiting (as described above) and the system is located far from the boundaries of the oscillatory regime, so that amplitudes remain large and undercompensation is avoided.

Temperature compensation can also be observed in stochastic simulations. The $Q_{10}^{(cycle)}$ values obtained from deterministic calculations are not always valid in the stochastic case, because system size can affect the oscillation period (Figure 5A). A systematic search revealed several parameter sets for which stochastic simulations are nearly temperature-insensitive; one example is shown in Figure 6F (see Table S3). In this calculation, $S_{00}$ and $S_{10}$ are present only in very brief spikes, consistent with fast reactions. Most of the time in the cycle is taken up by the slow, nearly-irreversible $S_{01} \rightarrow S_{11}$ conversion, which is catalyzed by the temperature-insensitive kinase.

In summary, precise temperature compensation may not be particularly common, but it is certainly possible. Roughly 3.4% of oscillators show a $Q_{10}$ value between 0.9 and 1.1. Most importantly, temperature compensation does not require highly unusual parameter values or modified network topologies. Instead, a range of temperature sensitivities is possible, and temperature insensitivity occupies an intermediate position in this range. To ensure temperature compensation, the temperature-insensitive steps should be rate-limiting – this is realized through slow, irreversible reactions and a tendency toward temporal ordering through differential binding constants. Interestingly, the phosphorylation of PERs by CKIε/δ is apparently a rate-determining reaction in the

mammalian circadian clock and is temperature insensitive when PER-derived peptide is used as a substrate in vitro (Isojima et al., 2009).

## Detailed Discussion of Other Posttranslational Oscillator Models

(See ''Comparison with Other PTO Models'' in the Discussion section.)

Several models have been described for bistability and oscillations in the MAPK system, and our system exhibits important similarities and differences with these. For example, Markevich et al. (2004) describe a single level of a MAPK cascade, in which a single kinase (MAPKK) and a single phosphatase (MKP) catalyze the interconversions of MAPK enzymes with zero, one, or two phosphate modifications. The most important difference between this scheme and the one described here is that MAPK models (following the original formulation by Huang and Ferrell (Huang and Ferrell, 1996)), typically do not distinguish between the two possible singly-phosphorylated forms of MAPK, assuming instead that phosphorylation follows a strict ordering. Interestingly, Markevich et al. describe a pair of feed-forward inhibition motifs that look strikingly similar to those presented in Figure 2. Unless a distinction is made between the two singly-phosphorylated states, however, a directional rate bias (our first design motif) is not possible, and oscillations are not observed. Several studies have observed that, while bistability is possible within a single level of the cascade, feedback loops in a multilevel system are typically required for oscillations (Chickarmane et al., 2007; Kholodenko, 2000; Shankaran et al., 2009). Liu and coworkers (Liu et al., 2011) have shown that oscillations can be generated in a similar strictly-ordered two-site phosphorylation system with implicit feedback similar to the enzyme sequestration observed in our study, with the major difference that their system requires a fourth component (in addition to oscillating substrate, kinase, and phosphatase) that is actively produced and degraded (Liu et al., 2011).

Another essential difference between our study and the various MAPK models in the literature is that these models are all intended as descriptions of an experimentally-characterized system and therefore have access to parameters that have been (at least approximately) experimentally-determined. For example, Qiao and co-workers (Qiao et al., 2007) used a Monte Carlo parameter search similar to ours to classify the dynamical behaviors available in the complete MAPK cascade system, with no explicit feedback between the MAPK level and the MAPKKK level. By sampling $2 \times 10^5$ exponentially-distributed parameter sets, they were able to determine the relative frequencies of single-valued, hysteretic, and oscillatory input-output relationships. In contrast to our approach, however, in which parameter values were selected from a range covering several orders of magnitude, their parameter values were selected to be within 25-fold of the experimentally estimated values, meaning that they characterized the behaviors present within a reasonable neighborhood in parameter space. Our approach, in contrast, has been to sample the parameter space as completely as possible and analyze the resulting parameter combinations to determine which regions harbor interesting dynamical behaviors.

Our model can also be compared fruitfully to various models that exist for the in vitro reconstitution of the cyanobacterial circadian clock. Van Zon et al. (van Zon et al., 2007) describe a model in which each monomer in a KaiC hexamer can be either phosphorylated or unphosphorylated, and the ring as a whole is in either an active state (favoring phosphorylation) or an inactive state (favoring dephosphorylation). Highly-phosphorylated hexamers easily become inactivated and weakly-phosphorylated hexamers tend to become activated, leading to a preferred ordering of the system's states (our first design motif). The population of single-hexamer oscillators is synchronized through the action of KaiA, which is assumed to bind most strongly to weakly-phosphorylated hexamers, promoting phosphorylation (they refer to this mechanism as ''differential affinity''). This leads to a situation similar to our sequestration-based synchronization – the hexamers that are ahead of the pack are less successful in competing for KaiA and will tend to slow down, while the laggards will tend to speed up. The supplementary information of their study formulates differential affinity in very general terms that are easily applicable to our system.

The circadian clock model of Van Zon and co-workers also discusses a hypothesis for temperature compensation that has important similarities to ours. They begin with the assumption that the various (de)phosphorylation rates are all temperature-insensitive, while other factors (such as binding/unbinding and enzyme/substrate affinities) may be temperature sensitive. In order for the system as a whole to remain temperature-compensated, the temperature-insensitive steps must be rate limiting; from this observation they infer requirements that the (un)binding and conformational change rates should be much faster than the catalytic rates and that affinities between the Kai proteins should be high enough that binding is typically saturated. A more detailed study of temperature compensation in the KaiABC system was presented by Hatakeyama and Kaneko, who argued that temperature compensation could be achieved (even if individual reactions are temperature-sensitive) as long as reaction rates are limited by an enzyme that is shared across several reactions (Hatakeyama and Kaneko, 2012).

In our study, we assumed only that the kinase was temperature-insensitive, and examined conditions under which temperature compensation can still be achieved with a temperature-sensitive phosphatase. This complicates the picture substantially, but the requirement of nearly-saturated binding and fast noncatalytic processes is consistent with our results.

Another approach to the modeling of the cyanobacterial circadian clock is illustrated by Clodong and co-workers (Clodong et al., 2007), who began with a loop topology similar to the one used by van Zon et al., except that the distinction between active and inactive conformations is replaced by a distinction between KaiC hexamers without KaiB (which will tend to be phosphorylated) and KaiBC complexes which will tend to be dephosphorylated. They then used a ''rewiring'' algorithm to add additional positive or negative regulatory connections to this loop topology. The topology that generated the highest-amplitude, most robust oscillations did so using a feed-forward inhibition synchronization scheme similar to ours, in which KaiBC complexes with low phosphorylation number

inhibit the phosphorylation of KaiC complexes. This can be accomplished if KaiA binds strongly to low-phosphorylated KaiBC complexes, making it unavailable to promote the phosphorylation of the KaiC hexamers.

The model of Rust et al. (Rust et al., 2007) describes a two-site system similar to ours, in which the two phosphorylation states of the substrate (KaiC) have distinct roles. Several differences exist in the detailed implementation, however – their model lacks a separate phosphatase (relying instead on the autophosphatase activity of KaiC) and the autokinase activity of KaiC is modulated by the activities of the regulatory proteins KaiA and KaiB. KaiA promotes the preferential phosphorylation of KaiC on T432 (KaiC-T), followed by phosphorylation on S431 (KaiC-ST); the phosphorylation on S431 to form KaiC-S is slower, and its dephosphorylation from KaiC-ST is inhibited by KaiA. Once low levels of KaiC-S do form, however, they recruit KaiB, which inhibits the activity of KaiA. In this way, KaiC-S indirectly catalyzes its own formation, as well as inhibiting the formation of KaiC-T and KaiC-ST. Similarities to our model exist in the general principles of kinetic bias (i.e., T432 is phosphorylated more quickly than S431) and synchronization by feedforward inhibition, although the molecular details are quite different.

The various cyanobacterial clock models show differences among themselves, but the basic design principles identified in this study can be discerned in each. In all cases, a kinetic bias (sometimes enforced by the presence or absence of KaiB or by an active or inactive conformational state of KaiC) is used to give states a temporal ordering, consistent with our first design motif. In addition, all of the models described involve some variation on the theme of synchronization by feedforward inhibition, with high concentrations of a particular phosphorylation state inhibiting the progress of reactions that are a few steps ahead in the cycle. The most striking difference lies in the physical system – cyanobacterial clock models focus on a specific biochemical system and their parameter choices are guided by experimental evidence, while our model focuses on a very generic system and design elements were discovered via a broad, unconstrained parameter search. Most importantly, our results demonstrate that these design principles are not a result of specific features of the KaiABC system (such as auto(de)phosphorylation activity, multimeric arrangement of substrates, or competitive regulation) but that they also appear in a much more general context.

## More General Formulation of Design Principles

(See "General Applicability of a Design Principle Composed of Two Design Motifs" in Discussion section.)

The general design principles uncovered in this study can also be applied to situations in which more than two phosphorylation sites are present (Figure 7B). Suppose that a protein has $n$ phosphorylation sites. Out of the $2^n$ possible phosphorylation states, a subsequence of $k \leq 2^n$ states exists such that the rate constants in the system predispose individual substrates to visit these states in the order $S_1,S_2,..,S_k,S_1,S_2,\ldots$. This is equivalent to our first design motif. Enzyme sequestration requires that some phosphorylation state $S_j$ binds strongly and sequesters the enzyme that catalyzes the downstream reaction $S_k \rightarrow S_{k+1}$, where $k > j$. This might also be the same enzyme that catalyzes $S_j \rightarrow S_{j+1}$, as in the case of our Cluster 2, where ($S_j$,$S_k$,$S_{k+1}$) corresponds to ($S_{00}$,$S_{01}$,$S_{11}$) or ($S_{11}$,$S_{10}$,$S_{00}$). Alternatively, it may be a different enzyme, as in the case of Cluster 1, where ($S_j$,$S_k$,$S_{k+1}$) corresponds to ($S_{01}$,$S_{11}$,$S_{10}$) or ($S_{10}$,$S_{00}$,$S_{01}$). It is not necessary that the system have only one type of kinase and one type of phosphatase, as long as these conditions for enzyme competition are present. As can be seen from the examples developed in this study, oscillations will be most robust when several overlapping modules exist. Geometric considerations (Manrai and Gunawardena, 2008; Thomson and Gunawardena, 2009) have shown that competition for enzyme between different substrate phosphorylation states can lead to multistability in multisite phosphorylation systems; our results lead one to imagine the presence of higher-dimensional attractors as well.

We have described the model as a phosphorylation system because our initial interest in this system was motivated by the prevalence of multisite phosphorylation in circadian systems. However, we note that the same mathematical framework could be applied to any type of multisite posttranslational modification such as methylation or acetylation.

## Compatibility of PTO with TTO

While the system presented in this paper would be an adequate model of an in vitro experiment, realization *in cellulo* would require the inclusion of transcription, translation, and degradation. Accounting for these processes would complicate the model substantially and is beyond the scope of this paper, but a few comments are in order. Figure 1B suggests that the system can be abstracted into two processes – one is a production of $S_{11}$ from $S_{00}$ (with $S_{01}$ as an intermediate), and the other is the disappearance of $S_{11}$ and reappearance of $S_{00}$. In our model, the second process is accomplished by phosphatase-catalyzed steps. This could also easily be accomplished by a transcriptional-translational loop in which $S_{11}$ is degraded and acts as a transcriptional repressor of the gene encoding the substrate. The newly-translated substrate will be unphosphorylated, so this scheme includes the necessary ingredients of $S_{00}$ appearance, $S_{11}$ disappearance, and coupling between these two processes (Figure 7C). In this case, the presence of the phosphatase-catalyzed pathway from $S_{11}$ to $S_{00}$ is superfluous, and its function can be replaced by degradation of $S_{11}$ and synthesis of $S_{00}$. Thus, our PTO model can easily be generalized to a TTO model, and many of the same considerations will apply.

This scheme is somewhat analogous to the current canonical model of the circadian clock in various organisms. Typically, the substrates that are phosphorylated at multiple sites and degraded upon progressive phosphorylation are also involved in the core negative feedback loop – PER in fly and mammals, and FRQ in fungi (Dibner et al., 2010; Dunlap, 1999; Reppert and Weaver, 2002; Takahashi et al., 2008; Young and Kay, 2001). These proteins directly or indirectly repress their own transcription (Aronson et al., 1994; Darlington et al., 1998; Kume et al., 1999), though it should be noted that the repressor activity of mammalian PER seems

to be relatively weak (Sangoram et al., 1998) and both mammalian PER and fungi FRQ can act as positive regulators under some circumstances (Lee et al., 2000; Ogawa et al., 2011; Shearman et al., 2000). Multi-site phosphorylation, followed by degradation, may also be experienced by these proteins (Baker et al., 2009; Chiu et al., 2011; Shanware et al., 2011; Tang et al., 2009; Vanselow et al., 2006). The PTO oscillator presented in this paper is thus likely to be compatible with well-characterized TTO systems, and some of the same design principles may apply.

## SUPPLEMENTAL REFERENCES

Aronson, B.D., Johnson, K.A., Loros, J.J., and Dunlap, J.C. (1994). Negative feedback defining a circadian clock: autoregulation of the clock gene frequency. Science *263*, 1578–1584.

Baker, C.L., Kettenbach, A.N., Loros, J.J., Gerber, S.A., and Dunlap, J.C. (2009). Quantitative proteomics reveals a dynamic interactome and phase-specific phosphorylation in the Neurospora circadian clock. Mol. Cell *34*, 354–363.

Chickarmane, V., Kholodenko, B.N., and Sauro, H.M. (2007). Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. J. Theor. Biol. *244*, 68–76.

Chiu, J.C., Ko, H.W., and Edery, I. (2011). NEMO/NLK phosphorylates PERIOD to initiate a time-delay phosphorylation circuit that sets circadian clock speed. Cell *145*, 357–370.

Clodong, S., Dühring, U., Kronk, L., Wilde, A., Axmann, I., Herzel, H., and Kollmann, M. (2007). Functioning and robustness of a bacterial circadian clock. Mol. Syst. Biol. *3*, 90.

Coté, G.G., and Brody, S. (1986). Circadian rhythms in *Drosophila melanogaster*: analysis of period as a function of gene dosage at the *per* (period) locus. J. Theor. Biol. *121*, 487–503.

Craciun, G., and Feinberg, M. (2005). Multiple equilibria in complex chemical reaction networks: I. The injectivity property. SIAM J. Appl. Math. *65*, 1526–1546.

Craciun, G., and Feinberg, M. (2006). Multiple equilibria in complex chemical reaction networks: II. The species-reaction graph. SIAM J. Appl. Math. *66*, 1321–1338.

Darlington, T.K., Wager-Smith, K., Ceriani, M.F., Staknis, D., Gekakis, N., Steeves, T.D., Weitz, C.J., Takahashi, J.S., and Kay, S.A. (1998). Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. Science *280*, 1599–1603.

Dibner, C., Sage, D., Unser, M., Bauer, C., d'Eysmond, T., Naef, F., and Schibler, U. (2009). Circadian gene expression is resilient to large fluctuations in overall transcription rates. EMBO J. *28*, 123–134.

Dibner, C., Schibler, U., and Albrecht, U. (2010). The mammalian circadian timing system: organization and coordination of central and peripheral clocks. Annu. Rev. Physiol. *72*, 517–549.

Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271–290.

Dunlap, J.C., Loros, J.J., and DeCoursey, P.J. (2004). Chronobiology: Biological Timekeeping (Sunderland, MA: Sinauer Associates).

Feinberg, M. (1987a). Chemical reaction network structure and the stability of complex isothermal reactors. 1. The deficiency-zero and deficiency-one theorems. Chem. Eng. Sci. *42*, 2229–2268.

Feinberg, M. (1987b). Chemical reaction network structure and the stability of complex isothermal reactors. 1. The deficiency-zero and deficiency-one theorems. Chem. Eng. Sci. *42*, 2229–2268.

Feinberg, M. (1988). Chemical reaction network structure and the stability of complex isothermal reactors. 2. Multiple steady states for networks of deficiency one. Chem. Eng. Sci. *43*, 1–25.

Gallego, M., and Virshup, D.M. (2007). Post-translational modifications regulate the ticking of the circadian clock. Nat. Rev. Mol. Cell Biol. *8*, 139–148.

Hastings, J.W., and Sweeney, B.M. (1957). On the mechanism of temperature independence in a biological clock. Proc. Natl. Acad. Sci. USA *43*, 804–811.

Hatakeyama, T.S., and Kaneko, K. (2012). Generic temperature compensation of biological clocks by autonomous regulation of catalyst concentration. Proc. Natl. Acad. Sci. USA *109*, 8109–8114.

Heyer, L.J., Kruglyak, S., and Yooseph, S. (1999). Exploring expression data: identification and analysis of coexpressed genes. Genome Res. *9*, 1106–1115.

Huang, C.Y., and Ferrell, J.E., Jr. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Natl. Acad. Sci. USA *93*, 10078–10083.

Isojima, Y., Nakajima, M., Ukai, H., Fujishima, H., Yamada, R.G., Masumoto, K.H., Kiuchi, R., Ishida, M., Ukai-Tadenuma, M., Minami, Y., et al. (2009). CKIepsilon/δ-dependent phosphorylation is a temperature-insensitive, period-determining process in the mammalian circadian clock. Proc. Natl. Acad. Sci. USA *106*, 15744–15749.

Kampen, N.G.v. (2007). Stochastic Processes in Physics and Chemistry, Third Edition (Amsterdam/Boston: Elsevier).

Kholodenko, B.N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur. J. Biochem. *267*, 1583–1588.

Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., and Reppert, S.M. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell *98*, 193–205.

Lee, C., Etchegaray, J.P., Cagampang, F.R., Loudon, A.S., and Reppert, S.M. (2001). Posttranslational mechanisms regulate the mammalian circadian clock. Cell *107*, 855–867.

Lee, H., Chen, R., Lee, Y., Yoo, S., and Lee, C. (2009). Essential roles of CKIdelta and CKIepsilon in the mammalian circadian clock. Proc. Natl. Acad. Sci. USA *106*, 21359–21364.

Lee, K., Loros, J.J., and Dunlap, J.C. (2000). Interconnected feedback loops in the Neurospora circadian system. Science *289*, 107–110.

Liu, P., Kevrekidis, I.G., and Shvartsman, S.Y. (2011). Substrate-dependent control of ERK phosphorylation can lead to oscillations. Biophys. J. *101*, 2572–2581.

Manrai, A.K., and Gunawardena, J. (2008). The geometry of multisite phosphorylation. Biophys. J. *95*, 5533–5543.

Markevich, N.I., Hoek, J.B., and Kholodenko, B.N. (2004). Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. J. Cell Biol. *164*, 353–359.

Michaelis, L., and Menten, M.L. (1913). Die Kinetik der Invertinwirkung. Biochem. Z. *49*, 333–369.

Ogawa, Y., Koike, N., Kurosawa, G., Soga, T., Tomita, M., and Tei, H. (2011). Positive autoregulation delays the expression phase of mammalian clock gene Per2. PLoS ONE *6*, e18663.

Pittendrigh, C.S. (1954). On temperature independence in the clock system controlling emergence time in *Drosophila*. Proc. Natl. Acad. Sci. USA *40*, 1018–1029.

Press, W.H. (2007). Numerical Recipes: The Art of Scientific Computing, Third Edition (Cambridge, UK/New York: Cambridge University Press).

Qiao, L., Nachbar, R.B., Kevrekidis, I.G., and Shvartsman, S.Y. (2007). Bistability and oscillations in the Huang-Ferrell model of MAPK signaling. PLoS Comput. Biol. *3*, 1819–1826.

Reppert, S.M., and Weaver, D.R. (2002). Coordination of circadian timing in mammals. Nature *418*, 935–941.

Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O'Shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science *318*, 809–812.

Sangoram, A.M., Saez, L., Antoch, M.P., Gekakis, N., Staknis, D., Whiteley, A., Fruechte, E.M., Vitaterna, M.H., Shimomura, K., King, D.P., et al. (1998). Mammalian circadian autoregulatory loop: a timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription. Neuron *21*, 1101–1113.

Shankaran, H., Ippolito, D.L., Chrisler, W.B., Resat, H., Bollinger, N., Opresko, L.K., and Wiley, H.S. (2009). Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. Mol. Syst. Biol. *5*, 332.

Shanware, N.P., Hutchinson, J.A., Kim, S.H., Zhan, L., Bowler, M.J., and Tibbetts, R.S. (2011). Casein kinase 1-dependent phosphorylation of familial advanced sleep phase syndrome-associated residues controls PERIOD 2 stability. J. Biol. Chem. *286*, 12766–12774.

Shearman, L.P., Sriram, S., Weaver, D.R., Maywood, E.S., Chaves, I., Zheng, B., Kume, K., Lee, C.C., van der Horst, G.T., Hastings, M.H., and Reppert, S.M. (2000). Interacting molecular loops in the mammalian circadian clock. Science *288*, 1013–1019.

Strogatz, S.H. (1994). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (Reading, MA: Addison-Wesley).

Takahashi, J.S., Hong, H.K., Ko, C.H., and McDearmon, E.L. (2008). The genetics of mammalian circadian order and disorder: implications for physiology and disease. Nat. Rev. Genet. *9*, 764–775.

Tang, C.T., Li, S., Long, C., Cha, J., Huang, G., Li, L., Chen, S., and Liu, Y. (2009). Setting the pace of the Neurospora circadian clock by multiple independent FRQ phosphorylation events. Proc. Natl. Acad. Sci. USA *106*, 10722–10727.

Thomson, M., and Gunawardena, J. (2009). Unlimited multistability in multisite phosphorylation systems. Nature *460*, 274–277.

van Zon, J.S., Lubensky, D.K., Altena, P.R., and ten Wolde, P.R. (2007). An allosteric model of circadian KaiC phosphorylation. Proc. Natl. Acad. Sci. USA *104*, 7420–7425.

Vanselow, K., Vanselow, J.T., Westermark, P.O., Reischl, S., Maier, B., Korte, T., Herrmann, A., Herzel, H., Schlosser, A., and Kramer, A. (2006). Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). Genes Dev. *20*, 2660–2672.

Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. Nat. Rev. Genet. *2*, 702–715.

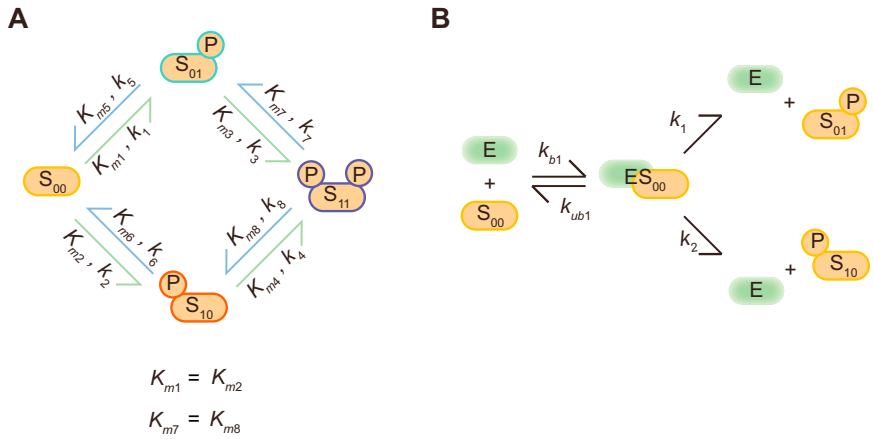



**Figure S1. PTO Model with Shared Intermediate Complexes at Branching Reactions, Related to Figure 1**

(A) PTO reaction network.

(B) In the model described in the main text, it was assumed that branching reactions (e.g., the conversion of $S_{00}$ to either $S_{01}$ or $S_{10}$) involved two distinct intermediates (e.g., $ES_{00a}$ for conversion to $S_{01}$ and $ES_{00b}$ for conversion to $S_{10}$). In the alternative approach shown here, branching reactions can proceed from a single complex $ES_{00}$, leading to either $S_{01}$ or $S_{10}$. Results from calculations with this assumption are qualitatively similar to the model shown in Figures 1B and 1C.

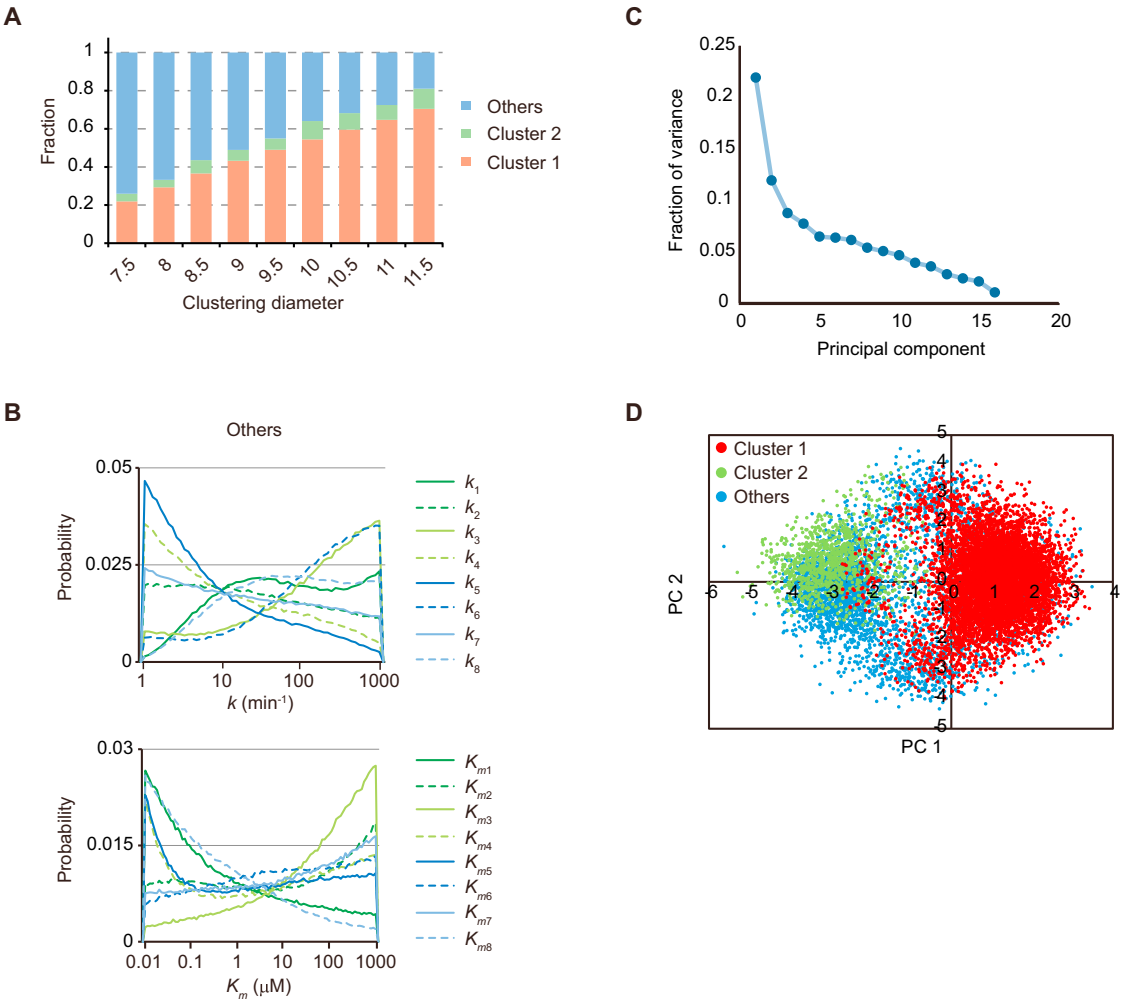



**Figure S2. Clustering of Oscillatory Parameter Sets, Related to Figure 2**

(A) Results of QT clustering with various clustering diameters.

(B) Parameter histograms obtained when all solutions not assigned to clusters 1 and 2 were pooled together.

(C) PCA of the random parameter search results. The first few principal components capture a reasonably large fraction of the variance.

(D) Projection of 10,000 parameter sets onto their first two principal components. Points are color-coded according to the clusters assigned using the QT algorithm. PCA is able to identify cluster 1 fairly clearly, and a clear separation exists between cluster 1 and the remaining parameter sets. No other major groupings are apparent.

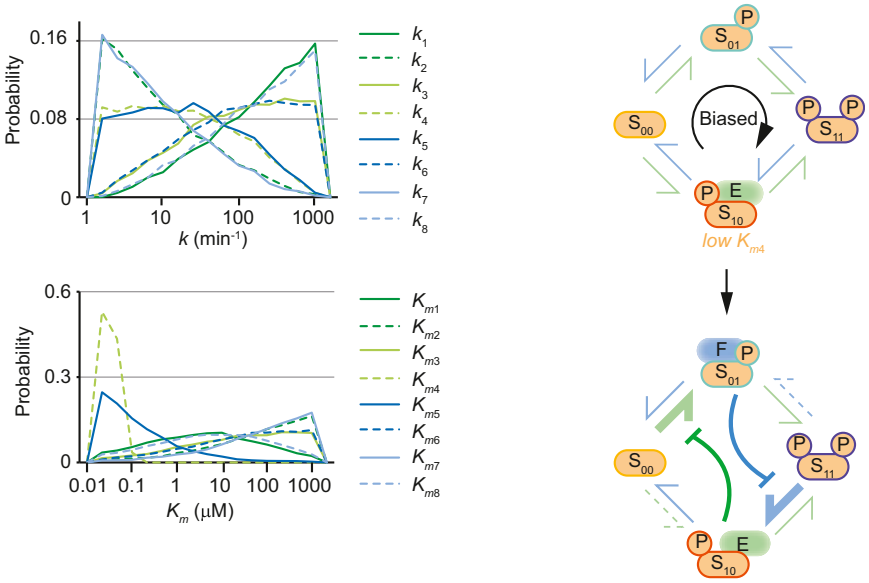



**Figure S3. Parameter Histogram for Oscillatory Parameter Sets Found when Clockwise Bias and Low $K_{m4}$ Were Enforced, Related to Figure 3**

The parameter histograms use the same coloring scheme as Figure 2C in the main text, and show clear similarities to the parameter histograms for cluster 1.

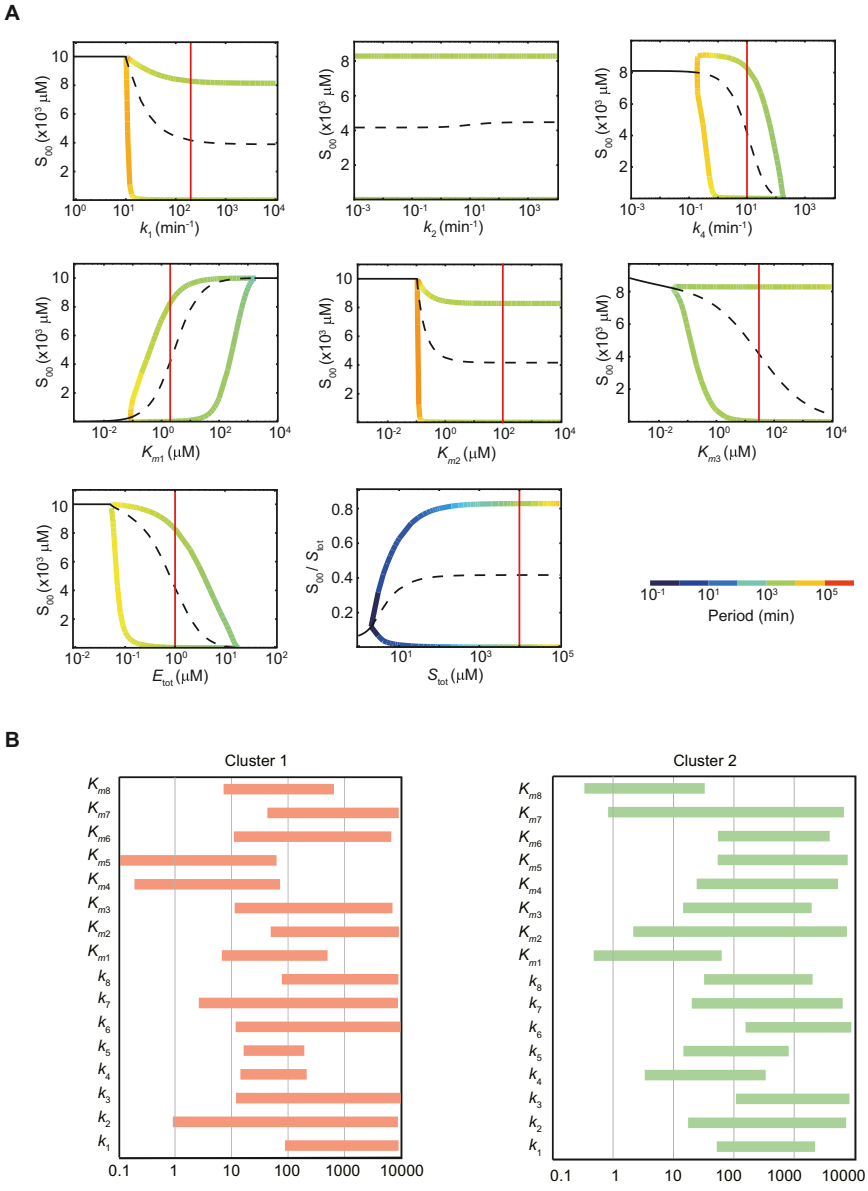



**Figure S4. Effects of Parameter Variation on Oscillatory Behavior, Related to Figure 4**

(A) Single-parameter bifurcation diagrams. Solid black lines denote stable fixed points, and dashed black lines indicate an unstable fixed point surrounded by a limit cycle. Red vertical lines denote the parameter value in the symmetric example set; see Table S1 for precise values.

(B) Oscillation ranges for parameter sets in clusters 1 and 2. Bifurcations were sought between $10^{-3}$ and $10^{4}$ for ~1,000 parameter sets in each cluster; the colored bars represent the average ranges over which each parameter can be varied without abolishing oscillations.

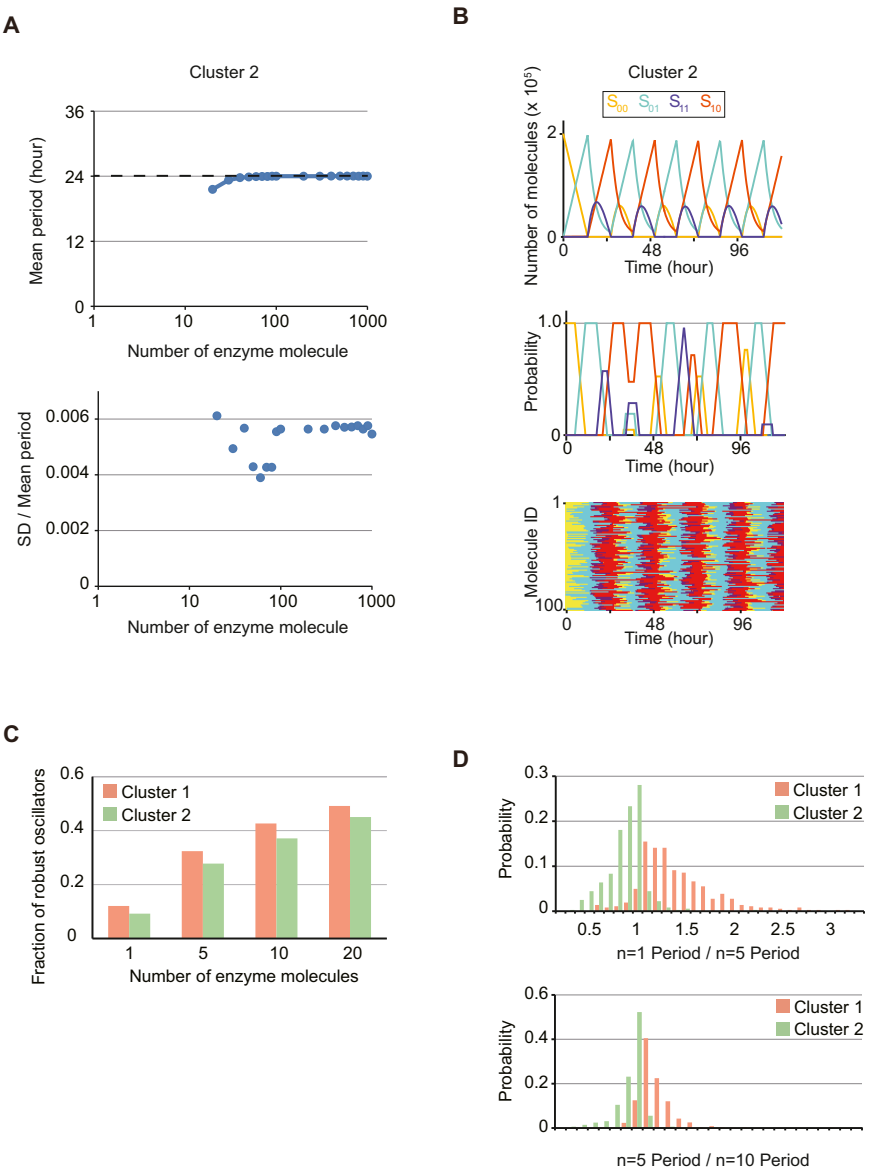



**Figure S5. Stochastic Simulation of a Stereotypical Cluster 2 Oscillator, Related to Figure 5**

(A) Effects of system size on the oscillation period. Oscillations were not observed for <20 enzyme molecules. Once they appear, the oscillations are quite robust and the period rapidly converges on the bulk average value.

(B) Top row: Plots of the bulk population obtained in stochastic stimulations for the cluster 2 oscillator. Middle row: Smoothed traces showing the probability that a single (arbitrarily chosen) substrate molecule will occupy different phosphorylation states as a function of time. Bottom row: Results for 100 individually tracked molecules, each shown as a horizontal raster line, color-coded as in the plots above.

(C) Fraction of parameter sets in clusters 1 and 2 that remain oscillatory as the system size is decreased. Even if only one of each type of enzyme is present, roughly 10% of parameter sets in each cluster remain robust against stochastic noise.

(D) Systematic changes in oscillator period with stochastic noise. As the system size is decreased (i.e., stochastic noise increases), cluster 1 oscillators tend to slow down (see also Figure 5A), whereas cluster 2 oscillators tend to speed up.

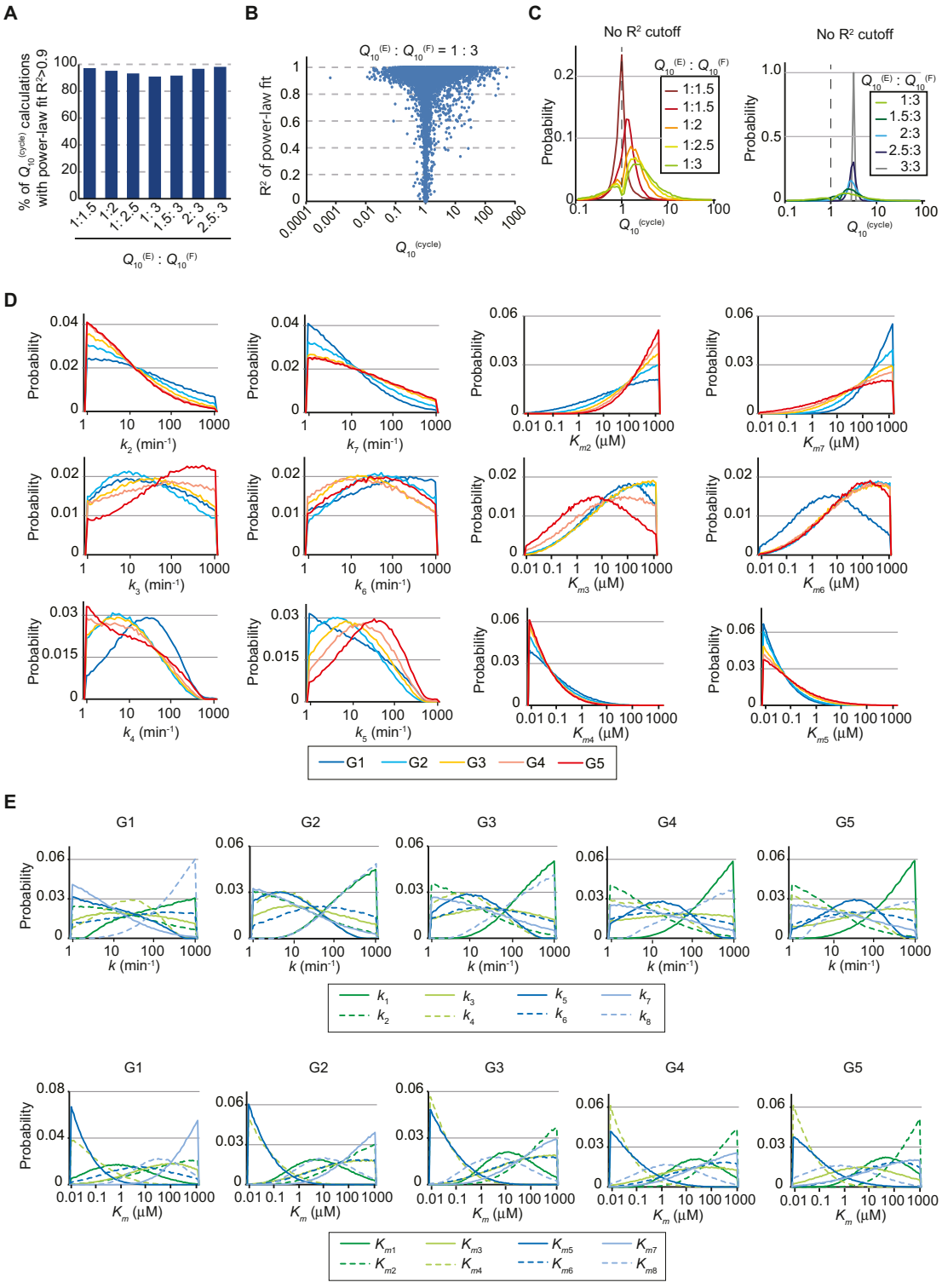



**Figure S6. Dependence of Temperature Sensitivity on Parameter Distributions, Related to Figure 6**

(A and B) Quality of the exponential fit for the calculated period at varying temperatures. Most fits are of excellent quality, regardless of the values of $Q_{10}^{(E)}$ and $Q_{10}^{(F)}$ (A), whereas poorly fit parameter sets tend to show a cycle $Q_{10}$ close to 1.0 (B).

(C) Histogram of cycle $Q_{10}$ values when the poorly fit parameter sets ($R^2 < 0.9$) are included.

(D) Changes in parameter distributions as a function of temperature overcompensation or undercompensation. Histogram traces are color-coded as in Figure 6C.

(E) The same histograms as in (D), organized by group rather than by parameter.