Jo, Hang-Hyun; Kim, Yeon Jeong; Kim, Jae Kyoung; Foo, Mathias; Somers, David E.; Kim, Pan-Jun

Waveforms of molecular oscillations reveal circadian timekeeping mechanisms

Published in: Communications Biology

DOI:
10.1038/s42003-018-0217-1

Published: 01/01/2018

Document Version
Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:
Jo, H-H., Kim, Y. J., Kim, J. K., Foo, M., Somers, D. E., & Kim, P-J. (2018). Waveforms of molecular oscillations reveal circadian timekeeping mechanisms. Communications Biology, 1, 1-11. [207]. https://doi.org/10.1038/s42003-018-0217-1

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
Waveforms of molecular oscillations reveal circadian timekeeping mechanisms

Hang-Hyun Jo1,2,3, Yeon Jeong Kim4, Jae Kyoung Kim5, Mathias Foo6, David E. Somers4 & Pan-Jun Kim7,8,9,10

Circadian clocks play a pivotal role in orchestrating numerous physiological and developmental events. Waveform shapes of the oscillations of protein abundances can be informative about the underlying biochemical processes of circadian clocks. We derive a mathematical framework where waveforms do reveal hidden biochemical mechanisms of circadian timekeeping. We find that the cost of synthesizing proteins with particular waveforms can be substantially reduced by rhythmic protein half-lives over time, as supported by previous plant and mammalian data, as well as our own seedling experiment. We also find that previously enigmatic, cyclic expression of positive arm components within the mammalian and insect clocks allows both a broad range of peak time differences between protein waveforms and the symmetries of the waveforms about the peak times. Such various peak-time differences may facilitate tissue-specific or developmental stage-specific multicellular processes. Our waveform-guided approach can be extended to various biological oscillators, including cell-cycle and synthetic genetic oscillators.
A variety of light-sensing organisms feature circadian clocks, which generate endogenous molecular oscillations with ~24 h periodicity and thereby control numerous physiological and behavioral events\(^1\)–\(^4\). Despite the identification of biochemical mechanisms of circadian timekeeping in various organisms\(^5\)–\(^9\), our understanding of a design principle of these clock mechanisms is far from complete. For example, the mammalian clock protein BMAL1 exhibits the abundance oscillations\(^10\), but these oscillations are not empirically required for the generation of circadian rhythms per se, leaving their biological roles still unclear\(^11\)–\(^13\). As another example, the plant circadian system involves post-translational regulations, such as the degradation of PSEUDO RESPONSE REGULATOR 5 (PRR5) protein by ZEITLUPE (ZTL) protein\(^14,15\); however, a previous mathematical modeling suggests that such post-translational interactions may not be strictly required for the formation of the rhythms of any core clock components\(^16\), raising a question about the fundamental role of these interactions.

The temporal trajectory of mRNA or protein concentration exhibiting a circadian rhythm can be characterized by its shape or waveform. A waveform of a protein expression profile, apart from its few characteristic quantities (period, amplitude, and peak phase)\(^17\)–\(^21\), has long been underappreciated, but recently recognized for its potential importance to clock function\(^16,22,23\). A cuspidate waveform, which shows a notable acuteness around its peak phase, was speculated to confer high-resolution timing of downstream biological events around the peak phase\(^16\). In addition, according to plant–clock experiments, precise changes in the waveform of GIGANTEA (GI) expression were sufficient to alter hypocotyl growth as a downstream phenotype\(^23\). Moreover, a specific circadian waveform seems crucial for the molecular arithmetic processes involved in daily starch degradation\(^24\). Although not in the circadian context, there are interesting reports that modifying the waveform shape of neuro-stimulating signals changes the efficiency of entraining the neural spiking activities\(^25\). Nevertheless, the reverse yet complementary view of the waveforms as a window to the inner biochemical mechanisms of circadian clocks has not yet been taken into consideration for systematic investigation.

Here, we report that the waveforms of clock protein profiles can serve as an information source of previously underexplored, biochemical mechanisms of circadian timekeeping. These mechanisms can be exemplified by the above PRR5–ZTL interaction and BMAL1 abundance oscillation. Interestingly, our waveform analysis predicts the considerable benefit of rhythmic regulation of protein degradation in reducing the biosynthetic cost of the waveform formation. Our mathematical framework is supported by previous, as well as our new, experimental data. This study can be extended to time-course data from various biological oscillators, such as cell cycle systems and synthetic genetic oscillators.

**Results**

**Relationship between waveforms and cost.** In a circadian system, the dynamics of protein production governs the protein concentration profile \(x(t)\) over time and thereby its waveform. This dynamics can often be described by the following equation:

\[ \frac{dx(t)}{dt} = g(t) - r(t)x(t), \]  

(1)

where \(g(t)\) and \(r(t)\) denote protein synthesis and degradation rates, respectively, as depicted in Fig. 1a. \(g(t)\) is proportional to an mRNA concentration and a translation rate. Protein degradation with a rate \(r(t)\) is driven by post-translational mechanisms. An oscillatory waveform of \(x(t)\) satisfies \(x(t) = x(t + T)\) with \(T = 24\) h in diurnal light and dark cycles or \(T \approx 24\) h in constant light or darkness. We stress that to maintain \(x(t)\)’s rhythmicity, \(g(t)\) or \(r(t)\) should not remain constant but change over time. We will consider the relationships between \(x(t), g(t), r(t)\), and later, the cost \(c\) of protein production defined as

\[ c \equiv \frac{\Delta x}{T} = \langle g(t) \rangle = \langle r(t)x(t) \rangle, \]  

(2)

where \(\Delta x\) denotes the amount of proteins synthesized over the period \(T\) and \(\langle \cdot \rangle\) represents a time average, e.g., \(\langle g(t) \rangle \equiv (1/T) \int_0^T g(t) dt\). The equalities \(\Delta x/T = \langle g(t) \rangle\) and \(\langle g(t) \rangle = \langle r(t)x(t) \rangle\) are derived from Eq. (1) and \(x(t) = x(t + T)\). In other words, the cost \(c\) is defined as an average protein amount synthesized per time, which is equal to an average protein amount degraded per time. Because the circadian protein levels are periodic over time, the proteins must be synthesized as much as they

---

**Fig. 1** Schematic diagrams of protein synthesis and turnover, and the resulting protein profiles in the circadian system. a) Proteins are synthesized through mRNA-to-protein translation, and destined for degradation. B) Cyclic protein abundances are represented by waveforms. For each waveform, the arrow indicates the point when \(R(t) = r_{\text{min}}\) in Eq. (4), and the shaded area corresponds to the interval between the steepest decline and the trough. The right waveform \((r_{\text{min}} \approx 0.69\) h\(^{-1}\)) has larger \(r_{\text{min}}\) than the left waveform \((r_{\text{min}} \approx 0.30\) h\(^{-1}\)). For the definition of each notation in a, b, refer to Eq. (1).
are degraded. We will show step by step that the biosynthetic cost \( c \) of a protein waveform helps us decipher circadian degradation mechanisms, mainly through the examples from the plant circadian system. Then, we will focus on other cases such as the mammalian system.

In the case of the plant Arabidopsis thaliana, more than 20 clock genes have been discovered, and many of their mRNAs undergo high-amplitude cycling in their abundance\(^{28,29}\). This mRNA-level oscillation is a result of transcriptional control by other clock gene products or by light signals. In the core plant clock, the protein synthesis rate \( g(t) \), which is largely proportional to the transcript concentration, would likely exhibit similar oscillatory patterns. On the other hand, the characteristics of the degradation rate \( r(t) \) remain rather elusive for plant clock proteins, with only a limited number of experimental reports \(^{28-32}\). Given the clearly time-dependent nature of the protein synthesis rate, the degradation rate may not have to be also time-dependent, as demonstrated by the previous mathematical modeling\(^{16}\). Existing experimental data, nonetheless, indicate that plant clock proteins often seem to have time-specific or phase-specific degradation rates\(^{28-31}\), raising a question about the beneficial effect of such rhythmic regulation of protein stability.

One study suggests that rhythmic degradation rates allow nontrivial phase differences between transcript and protein profiles\(^{18}\). However, given that the phases of transcript profiles have relatively little functional significance, this previous study is unlikely to be about biologically beneficial effects of the rhythmic degradation rates.

We begin with the following observation: because \( x(t) \geq 0 \), \( r(t) \geq 0 \), and \( g(t) \geq 0 \), Eq. (1) leads to

\[
\begin{align*}
  r(t) & \geq R(t) = \max_{t} \left\{ \frac{-x'(t)}{x(t)} , 0 \right\}, \\
  & \quad \text{The above inequality is always satisfied with arbitrary } g(t) \geq 0. \text{ In other words, regardless of any specific form of a transcript profile, the protein waveform } x(t) \text{ imposes a stringent constraint on the protein degradation rate } r(t), \text{ through a lower bound } R(t) \text{ in Eq. (3). Therefore, a protein waveform itself can be informative about the degradation rate.}
\end{align*}
\]

Can waveforms indicate the effect of time-specific or phase-specific degradation rates observed in empirical data? In order to address this issue, we start with a contradictory scenario that the degradation rate \( r(t) \) is constant over time, i.e., \( r(t) = r \), and examine its consequence. From Eq. (3),

\[
  r \geq r_{\text{min}} = \max_{t} R(t). \quad (4)
\]

Here, \( r_{\text{min}} \) the strict lower bound of the degradation rate \( r \), is essentially determined only at a single time point \( t = t_R \) with \( t_R \equiv \arg \max_{t} R(t) \), throughout this work, time \( t \) in a periodic function \( f(t) = f(t + T) \) is represented by a unique value within the range \( 0 \leq t < T \), unless specified. Because \( R(t) \equiv \max_{t} \left\{ -x'(t)/x(t) , 0 \right\} \), \( t_R \) in practice would be a point that approaches the trough of \( x(t) \) after the \( x(t) \)'s steepest decline \( t_R \) is placed between \( t_4 \) and \( t_6 \), where \( t_4 \equiv \arg \max_{t} \left\{ -x'(t) \right\} \) and \( t_6 \equiv \arg \min_{t} x(t) \), as shown in Fig. 1b). It is surprising that only one such a single time point, which will be henceforth referred to as a single hotspot, plays a critical role in determining a range of the constant degradation rate \( r \). Typically, the sharper a waveform \( x(t) \) is, the larger \( r_{\text{min}} \) at the hotspot (Fig. 1b).

For each plant clock protein, we can calculate the lower bound of its degradation rate, \( r_{\text{min}} \). Figures 2a and 3a exhibit the empirical PRR7 and PRR5 protein profiles in equal length light–dark (12L:12D) cycles\(^{26}\). Here, time points in light–dark cycles are counted from dawn (zeitgeber time). Using each protein profile \( x(t) \), we obtain \( R(t) \) in Eq. (3), and then by Eq. (4), \( r_{\text{min}} \approx 0.88 \text{ h}^{-1} \) for PRR7 \((t_R = 21 \text{ h}) \) and \( r_{\text{min}} \approx 1.69 \text{ h}^{-1} \) for PRR5 \((t_R = 22.3 \text{ h}) \), as in Figs. 2b and 3b. It means that if the degradation rates are constant over time, the PRR7 and PRR5 half-lives at any given time points cannot be longer than \(~47\) and \(~25\) min, respectively. Provided that there are some erroneous data points in the experimental profiles, the PRR7 and PRR5 half-lives might be up to \(~13\) and \(~51\) min longer than the above, respectively (Methods). In any cases, these half-lives appear to be very short, compared to other documented protein half-lives\(^{33,34}\). As previously mentioned, such a large degradation rate over the entire course of a day is attributed to only a single hotspot \( t = t_R \), under the assumption that the degradation rate is constant over time.

Next, we demonstrate that such a constant and large degradation rate can incur too large a cost of the protein production. In Eq. (2), the cost \( c \) of protein production is defined as an average protein amount synthesized per time, which is equal to an average protein amount degraded per time. For a constant degradation rate \( r(t) = r \), one obtains from Eqs. (2) and (4)

\[
  c = r\langle x(t) \rangle \geq c_{\text{g}} \equiv r_{\text{min}} \langle x(t) \rangle. \quad (5)
\]

Therefore, given the protein profile \( x(t) \), the lower bound of the cost \( c \) (i.e., \( c_{\text{g}} \)) is directly proportional to \( r_{\text{min}} \). PRR7 or PRR5, which exhibits large \( r_{\text{min}} \) would pay an accordingly high production cost if the degradation rate is constant. More specifically, from \( c_T = \Delta x \geq \frac{r_{\text{min}}}{x(t)} \), PRR7 and PRR5 must be synthesized per day at least \(~21\) and \(~41\) times more than actual protein level \( \langle x(t) \rangle \), respectively. In other words, these protein syntheses are far excessive compared to the actual protein abundance levels.

**Time-dependent degradation rates and cost reduction.** The above excessive cost of protein production can be effectively alleviated by time-varying degradation rates. If the degradation rate \( r(t) \) is no longer constant, \( r(t) \neq t = t_R \) is allowed to be smaller than \( r_{\text{min}} \), as far as Eq. (3) is satisfied. This fact leads to the possibility that the cost \( c \) can be lower than \( c_{\text{g}} = r_{\text{min}} \langle x(t) \rangle \). Hence, the cost can be reduced below the case of a constant degradation rate. A time-dependent degradation rate is enabled in nature by rhythmic post-translational regulation, such as PRR5 degradation by ZTL in the plant clock. Both PRR5 and ZTL levels oscillate over time, and this ZTL oscillation possibly contributes to the rhythmic degradation rate of PRR5. Including PRR5, plant clock proteins often seem to have phase-specific half-lives. These experimental data allow us to evaluate our hypothesis that rhythmic degradation rates help reduce protein production costs.

Before the calculation of the protein production costs to examine our hypothesis, we stress that all experimental degradation rates of the plant PRR7 and PRR5 proteins and of the mouse PERIOD2 (PER2) protein\(^{28-30,35}\) are found to satisfy the fundamental relation \( r(t) \geq R(t) \) in Eq. (3) (see Figs. 2b, 3b, and 4b). PER2 is an essential component of the mammalian clock\(^{10,35-38}\), and its synthesis and turnover dynamics approximately follows Eq. (1), thereby satisfying Eq. (3). To further test the validity of Eq. (3), we performed a cycloheximide (CHX) experiment and measured the PRR7 half-life at a time point that lacks preexisting half-life data (Fig. 2c, d and Methods). Again, the PRR7 half-life at this time point \((t = 18 \text{ h}) \) from our own experiment is in good agreement with Eq. (3) (Fig. 2b, c). Integration of these new and previous experimental data offers a rough estimate of protein production costs, as in the following paragraphs.

Calculation of protein production cost \( c \) requires information on both degradation rate \( r(t) \) and waveform \( x(t) \) over time, as \( c =

\[
  c = \int_{0}^{T} r(t) x(t) \text{d}t
\]

\[
  = \int_{0}^{T} r(t) x(t) \text{d}t
\]
Experimental degradation rates of PER2 cover a relatively wide range of time and are thus informative enough to envisage the overall trend of \( r(t) \). Therefore, only based on these experimental degradation rates and \( R(t) \), without mRNA profile data, we can make a rough estimate of the PER2 degradation rate over the entire circadian period, as demonstrated in Fig. 4c (Methods).

The estimated, phase-specific degradation rates of clock proteins in Figs. 2f, 3d and 4c show the characteristic curves that peak around the hotspots \( (t = t_b) \) and decline elsewhere. These patterns are the hallmarks of the rhythmic degradation rates that can reduce the protein production costs below the cases of constant degradation rates; except for the hotspots that must have large degradation rates \( (\approx r_{\text{min}}) \) by Eq. (3), if degradation rates remain small for most time, proteins do not have to be much synthesized to compensate for the degradation \( (\text{Eq. (2)}) \) and hence the production costs will become reduced.

Using the above degradation rate curves of several clock proteins, we now compute the actual protein production cost \( c \) by \( c = \langle r(t)x(t) \rangle \) in Eq. (2). Compared to the cases of constant degradation rates, the PER7, PRR5, and PER2 production costs
Fig. 3 PRR5 in the plant clock. For the definition of each notation, refer to Eqs. (1), (3), (4) and (18). a Existing experimental data of PRR5 protein levels (x(t)), filled circles; normalized by the peak level of their spline curve\(^3\). b r(t) (red solid line; calculated from x(t) in a), \(r_{\text{min}}\) (gray solid line), and empirical r(t) values (circles). The vertical axis unit is h\(^{-1}\). The r(t) values are from previous experimental data\(^2\). In agreement with Eq. (3), there exists no r(t) smaller than R(t). c Existing experimental data of PRR5 mRNA levels (\(g(t)\), filled circles; normalized by the peak level of their spline curve)\(^7\). d Estimated r(t) over time (green dashed line; green circles for direct calculation from empirical r(t), x(t), and \(g(t)\) using Eqs. (18)-(20) with constant k), along with \(r_{\text{min}}\) in b. The vertical axis unit is h\(^{-1}\). All experimental data here pertain to 12L:12D cycles, except for r(t) at t = 19 h in b, which was collected from a different light condition due to the scarcity of experimental data (Supplementary Methods). White and black segments in a, c correspond to light and dark intervals, respectively.

Fig. 4 PER2 in the mammalian clock. For the definition of each notation, refer to Eqs. (1), (3) and (4). a Existing experimental data of PER2 protein levels (x(t)), normalized by the peak level; moving window average of experimental data\(^3\). b r(t) (red solid line; calculated from x(t) in a), \(r_{\text{min}}\) (gray solid line), and empirical r(t) values (circles). The vertical axis unit is h\(^{-1}\). The r(t) values are from previous experimental data\(^2\). In agreement with Eq. (3), there exists no r(t) smaller than R(t). c Estimated r(t) over time (green dashed line; green circles for experimental r(t) data in b), along with \(r_{\text{min}}\) in b. The vertical axis unit is h\(^{-1}\).

Indeed, decrease by at least ~70%, ~83%, and ~52%, respectively, as summarized in Table 1. If we consider a possible deviation of the degradation rate in Fig. 2c, the PRR7 production cost decreases by ~68% to ~73% (Fig. 2). Interestingly, in the case of alga Ostreococcus taurii, rhythmic protein degradation is known to be very crucial for circadian timekeeping\(^9\). For its clock proteins CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and TIMING OF CAB EXPRESSION 1 (TOC1), the full time series of experimental r(t) is available\(^9\), and our analysis suggests that the rhythmic r(t) saves ~30% and ~41% of the CCA1 and TOC1 production costs, respectively (Methods). These results well support our hypothesis that rhythmic control of clock protein half-lives is beneficial to the cost reduction of protein production. This cost saving effect would be valid even if other benefits from the rhythmic half-lives are not clear. We thus predict the statistical tendency that the sharper a waveform is at the hotspot (i.e., the larger is \(r_{\text{min}}\), and therefore is \(c_{\text{g}}\)), the more likely a protein half-life is to be phase-specific.

| Protein | \(r_{\text{min}}\) (h\(^{-1}\)) | \(c_{\text{g}}\) (h\(^{-1}\)) | \(c\) (h\(^{-1}\)) | Cost reduction (%) |
|---------|---------------------|----------------------|-----------------|------------------|
| PRR7    | 0.88                | 0.40                 | 0.12            | -70              |
| PRR5    | 1.69                | 0.77                 | 0.13            | -83              |
| PER2    | 0.47                | 0.23                 | 0.11            | -52              |

For the definitions of \(c_{\text{g}}\), and \(c\), refer to Eqs. (2) and (5). The cost reduction due to the time-specific or phase-specific r(t) is defined as \((c_{\text{g}} - c)/c_{\text{g}}\). We here assume constant k in Eq. (18). We treat x(t) as dimensionless through the normalization of x(t) by its peak value (Figs. 2a, 3a and 4a), and thus units of \(r_{\text{min}}, \ c_{\text{g}}, \) and \(c\) in the table are h\(^{-1}\). The cost reduction itself is not a quantity affected by the normalization of x(t), and hence there is no loss of generality in its values.

Enigmatic elements of animal circadian systems. Thus far, we have investigated circadian dynamics driven by protein synthesis and degradation in Eq. (1). We now discuss another class of circadian dynamics with Eq. (8) below, which is crucial for...
mammals and insects, but does not follow the underlying mechanism of Eq. (1).

The core part of the mammalian clock harbors a transcriptional/post-translational negative feedback loop\cite{10,35,36}, which involves transcription factors, CLOCK and BMAL1 proteins. CLOCK–BMAL1 heterodimers activate the transcription of Per and Cryptochrome (Cry) genes, and the encoded PER and CRY proteins form PER–CRY complexes that are translocated to the nucleus. In the nucleus, they interact with CLOCK–BMAL1 complexes to inhibit the CLOCK–BMAL1 transcriptional activities. These positive (CLOCK and BMAL1) and negative (PER and CRY) arms constitute a negative feedback loop.

In the following equations, $x_A(t)$ and $x(t)$ represent the concentrations of active and inactive CLOCK–BMAL1 complexes in the nucleus, respectively, and $y(t)$ represents the concentration of nuclear PER–CRY complexes that are not binding to CLOCK–BMAL1 complexes:

$$\frac{dx_A(t)}{dt} = \tilde{a}(t) + k_1 x(t) - k y(t)x_A(t) - r_1 x_A(t), \quad (6)$$

$$\frac{dx(t)}{dt} = k y(t)x_A(t) - k_1 x(t) - r_2 x(t). \quad (7)$$

Here, $\tilde{a}(t)$ is a rate of CLOCK–BMAL1 translocation from the cytoplasm to the nucleus, $k_1$ and $k$ are, respectively, dissociation and association rate constants of two complexes, CLOCK–BMAL1 and PER–CRY, and $r_1$ and $r_2$ correspond to the sums of degradation rates and the rates of translocation to the cytoplasm. Employing another variable $x_A(t) \equiv x_A(t) + x(t)$ for the total CLOCK–BMAL1 concentration, the upper equation can
be rewritten as
\[
\frac{dx_A(t)}{dt} = g_A(t) - [r_0 + ky(t)]x_A(t),
\]
where \( g_A(t) \equiv k_1x_A(t) + \bar{a}(t) \) and \( r_0 \equiv k_1 + r_1 \).

Equation (8) represents a class of circadian dynamics distinguished from our previous case, Eq. (1). Equation (8) for the core mammalian clock captures the dynamics of active CLOCK–BMAL1 complexes (i.e., CLOCK–BMAL1 that is not binding to PER–CRY) in the nucleus, as depicted in Fig. 5a, and is applied to the insect clock as well. A fundamental difference between Eqs. (1) and (8) is as follows: in Eq. (1), \( g(t) \) exhibits high-amplitude oscillation as evident from the transcript profiles of many plant clock genes and mammalian Per genes, and hence \( g(t) \) is a main driving force of \( x(t) \)’s oscillation. In contrast, in Eq. (8), \( x_A(t) \)’s oscillation is largely driven by \( y(t) \)’s oscillation, rather than by \( g_A(t) \)’s. Compared to PER2 levels (see t; Fig. 4a), BMAL1 levels are only weakly oscillating over time, and correspondingly, \( g_A(t) \) would be only weakly oscillating. In fact, cyclic BMAL1 expression is not even required for mammalian circadian rhythmicity, as the mutant with constitutive BMAL1 expression still exhibits circadian rhythms. Given the apparently minor role of BMAL1’s abundance oscillation in circadian rhythmicity, what beneficial effects on the clock might follow from this BMAL1 oscillation?

Diverse phase differences between clock components. To the above enigmatic presence of BMAL1’s abundance oscillation in the mammalian clock, the effect of protein production cost \( c \) in our previous analysis is not straightforwardly relevant. Unlike \( x(t) \) in Eq. (1), \( x_A(t) \) involves the only active, not the total, molecules. In other words, \( x_A(t) \) is mainly driven by the relatively costless, post-translational conversion of inactive to active molecular forms, devoid of severe biosynthetic cost problems in the previous analysis. Again, we suggest that the clue for the effect of the BMAL1 oscillation can be found from waveforms, especially \( x_A(t) \) and \( y(t) \) from the CLOCK–BMAL1 and PER–CRY complexes. As will be shown later, such BMAL1 oscillation confers at least two advantages on the circadian system: one is a wide range of a peak time difference between the two clock components, active CLOCK–BMAL1 and PER–CRY that is not binding to CLOCK–BMAL1. The other advantage is the symmetry of the waveforms of these components. For the sake of convenience, we will henceforth drop subscript \( A \) from \( x_A(t) \) and \( x’(t) \), and simply write them as \( x(t) \) and \( x’(t) \). Equation (8) can be rewritten as
\[
ky(t) = \frac{g_A(t) - x'(t)}{x(t)} - r_0.
\]

If we assume that \( g_A(t) = \text{const} \) (i.e., \( g_A(t) = g \)) waveforms \( y(t) \) and \( x(t) \) in Eq. (9) are substantially constrained by each other. Specifically,
\[
ky(t) = \frac{g - x'(t)}{x(t)} - r_0,
\]
where the inequality of the lower relation comes from \( x(t) \geq 0 \) and \( r_0 + ky(t) \geq 0 \). We will use a notation \( t_y \equiv \arg \max_t f(t) (0 < t_y \leq T) \) for any given periodic function \( f(t) (f(t) = f(t + T)) \) when \( t_y \) is uniquely determined by \( 0 < t_y \leq T \). For example, \( t_y \) denotes the peak time of \( x(t) \) during \( 0 < t \leq T \). From Eq. (10),
\[
t_y \leq t_y \leq t_1.
\]

Note that \( t_1 \) is identical to the hotspot \( t_0 \). In other words, \( t_y \) is even closer to \( x(t) \)'s trough time than \( t_0 \). This range of \( t_y \) is illustrated in Fig. 5b. Generally, a peak time difference between \( y(t) \) and \( x(t) \) takes such a narrow range that \( y(t) \)'s peak time is almost the same as \( x(t) \)'s trough time. Therefore, if \( g_A(t) \) stays constant over time, waveforms \( x(t) \) and \( y(t) \) are only allowed to have a near anti-phase relationship.

To exemplify the above point, we consider the case with a sinusoidal wave \( x'(t) = L \sin(\omega t) \) where \( L \) is a constant and \( \omega = 2\pi/T \). In this case,
\[
x(t) = -\frac{L}{\omega} \cos(\omega t) + \frac{L}{\omega} + h_0,
\]
with an additional constant \( h_0 \). In the subsequent analyses, we treat \( x(t) \) in Eq. (13) as dimensionless, without loss of generality. We define a phase difference between \( x(t) \) and \( y(t) \) in Eq. (10) as \( \phi \equiv \arg \omega (t_y - t_0) \). Using Eq. (10), \( \phi = \pi - 2 \tan^{-1} [(\sqrt{C^2 - L^2 + g^2 - g}/(C + L)) \), where \( C = h_0 + L \) and \( g \geq g_{\min} = |L| \) from Eq. (11). This exact solution of \( \phi \) is indeed very close to \( \pi \), as shown in Fig. 5c. This tendency is consistent with our generic result that constant \( g_A(t) \) forces \( x(t) \) and \( y(t) \) into a near anti-phase relationship.

In contrast, if \( g_A(t) \) is no longer constant but cycles over time, as observed with cyclic BMAL1 expression in nature, then waveforms \( y(t) \) and \( x(t) \) in Eq. (9) are not much constrained by each other, and their peak time difference (or phase difference) can be flexible depending on \( g_A(t) \)'s oscillatory form. Because there is a lack of compelling experimental data on the waveform of \( g_A(t) \), we start with the following assumption:
\[
g_A(t) = a + \beta x(t + \tau),
\]
where \( a, \beta, \) and \( \tau \) are constants, and \( \beta \geq 0 \) and \( \tau \geq 0 \). From Eq. (9), \( r_0 + ky(t) \geq 0 \), and \( g_A(t) \geq 0 \), \( a \) should satisfy
\[
a \geq a_{\min} \equiv \max \left\{ \max_i [x'(t) - \beta x(t + \tau)], \max_i [-\beta x(t)] \right\}.
\]

Next, we show that \( x(t) \) and \( y(t) \) can have almost any in-phase to anti-phase relationship, covering a wide range of the phase difference. If \( \tau \ll T \),
\[
x(t + \tau) \approx x(t) + \tau x'(t).
\]

Combined with Eq. (9), it leads to
\[
ky(t) \approx \frac{a - (1 - \beta \tau)x'(t)}{x(t)} + \beta - r_0.
\]

Depending on signs of \( a \) and \( 1 - \beta \tau \) in Eq. (17), \( y(t) \) is now allowed to peak anytime of a day relative to \( x(t) \)'s peak time, as proven in Methods.

This result is illustrated in Fig. 5d. Together, if BMAL1 level (\( g_A(t) \) in Eqs. (8) and (9)) is not constant but varies over time, it confers much freedom on the waveform \( y(t) \) of PER–CRY’ that is not binding to CLOCK–BMAL1, and thus allows various phase differences between those unbinding CLOCK–BMAL1 and PER–CRY complexes (\( x(t) \) and \( y(t) \)) through the adjustment of parameters \( a, \beta, \) and \( \tau \). As in Fig. 5d, the unbinding CLOCK–BMAL1 and PER–CRY complexes can take almost any in-phase to anti-phase relationship. This result is in sharp contrast to the case with constant \( g_A(t) \), where the unbinding CLOCK–BMAL1 and PER–CRY complexes have a
predominantly anti-phase-like relationship. Our predictions can be verified by experimental techniques, such as co-immunoprecipitation assays, measuring the time series of CLOCK–BMAL1 and PER–CRY levels across different tissues or developmental stages, while excluding the levels of inactive CLOCK–BMAL1.

These potentially diverse phase differences, conferred by cyclic expression of positive arm components, may help in the coordination of tissue-specific or developmental stage-specific clock events in complex multicellular organisms, such as mammals and insects. Interestingly, the fungus Neurospora crassa, a relatively simple species, shows almost constant levels of white collar-1 (WC-1) expression, and thus would have almost constant $g_s(t)$. Therefore, we expect that the fungal clock may have an only anti-phase-like relationship between its core components, nuclear WC-1 and FREQUENCY (FRQ) proteins (see Supplementary Discussion).

To illustrate the above diverse phase differences conferred by BMAL1 cycling, we revisit the case with a sinusoidal wave $x(t)$ in Eq. (13) and consider the oscillation of $g_s(t)$ in Eq. (14). From Eq. (9), we obtain the exact solution of the phase difference between $x(t)$ and $y(t)$, as plotted in Fig. 5e. This exact solution is in good agreement with our generic results based on the approximation Eqs. (16) and (17).

**Symmetries of waveforms.** Another advantage of cyclic expression of positive arm components in mammalian and insect clocks is in the symmetry of waveforms about the peak phases. Previous experimental data from the mammalian clock indicate the existence of such symmetry that ascending and descending phases span almost the same time intervals, while its phenotypic significance still remains unknown.

To demonstrate the effect of BMAL1 cycling on the waveform symmetry, we first assume the contradictory scenario that $g_s(t)$ is constant over time with $g_s(t) = g$. Therefore, $ky(t) = [g - x'(t)]/x(t) - r_0$ from Eq. (10). In this case, waveforms $x(t)$ and $y(t)$ from CLOCK–BMAL1 and PER–CRY complexes cannot easily satisfy both symmetric relations $x(t_2 - t_1) = x(t_1 + t_2)$ and $y(t_2 - t_1) = y(t_1 + t_2)$ at the same time, because $x'(t)$ term in $ky(t) = [g - x'(t)]/x(t) - r_0$ breaks the symmetry of either $x(t)$ or $y(t)$ waveform unless $g \gg \max|x'(t)|$ to diminish the effect of $x'(t)$.

In contrast, if BMAL1 level ($g_s(t)$) is not constant but varies over time, both unbinding CLOCK–BMAL1 and PER–CRY profiles ($x(t)$ and $y(t)$) are allowed to have symmetric waveforms relatively easily. For example, if $g_s(t) = a + bx(t + \tau) + x(t + r)\tau - x(t) + x'(t)$ with $\tau \approx \tau_0$, then $ky(t) = [a - (1 - \beta)r]x(t')/x(t) + \beta - r_0$ in Eq. (17). Therefore, both $x(t)$ and $y(t)$ waveforms can be approximately symmetric at the same time, as long as $|a/|1 - \beta r|\max x'(t)|$ for the diminished effect of $x'(t)$. This condition can be satisfied more easily than the previous one.

This waveform symmetry, along with the above phase difference between two core components (unbinding CLOCK–BMAL1 and PER–CRY complexes), shows that the waveforms are useful to understand the effect of the enigmatic oscillation in BMAL1 expression.

**Discussion**

In this study, we have revealed that protein waveforms are informative about the underlying mechanisms of circadian clockwork.

A sharp waveform at the hotspot time point (i.e., with large $r_{max}$) implies rhythmic post-translational regulation that yields a phase-specific protein half-life; otherwise, too large costs of protein syntheses may be incurred for those waveforms. Such rhythmic degradation rates are observed in plant and mammalian circadian clocks, and can substantially reduce the protein production costs, as demonstrated in Table 1. If more experimental data become available, our waveform-cost analysis can be extended to other clock proteins. For example, the orphan nuclear receptor REV-ERBα in the mammalian clock may have a phase-specific half-life, driven by the rhythmic activity of glycogen synthase kinase-3β that regulates the REV-ERBα stability. Hence, if the half-lives measured at multiple specific time points become available, REV-ERBα will be a good target candidate for our cost analysis, aided by the existing REV-ERBα expression profiles.

On the other hand, regarding any possible extra costs that may be incurred by rhythmic degradation rates, we note that the cost $c$ of a given protein is not the concept to include the production cost of its proteolytic factor. Yet, the half-life can exhibit a rhythmic pattern by the proteolytic factor’s oscillation, and thus one may suggest that the cost $c$ should be extended to the proteolytic factor’s production cost. This extra cost from the proteolytic factor production, however, is not always relevant and needs cautious analyses in the future. For example, if the proteolytic factor has not only evolved for the degradation of a particular protein but also for other functions, then the cost of the proteolytic factor production shall not be covered by the cost $c$ in question. This is because such a proteolytic factor continues to be produced for multiple purposes, not exclusively for the degradation of that particular protein.

In this study, we also suggest that seemingly dispensable, cyclic expression of certain clock proteins in mammals and insects may allow both a broad range of phase differences between clock components and the symmetries of the waveforms. The various phase differences may be important for tissue-specific or developmental stage-specific clock coordination in complex multicellular organisms, such as mammals and insects. As previously mentioned, fungi do not show such cyclic expression of the corresponding components, and their relatively simple organismal forms may not necessitate as widely ranging phase differences as in the cases of mammals and insects.

Our waveform-guided approach is well supported by experimental data (Figs. 2b, 3b, and 4b), and provides insights into circadian mechanisms of evolutionarily distant organisms. Furthermore, we envisage that the concepts presented in this study can be applied beyond circadian dynamics, such as to time-course data from cell cycle systems and synthetic genetic oscillators.

**Methods**

**Experimental measurement of the PRR7 half-life.** We describe the details of our experimental methods for the measurement of the PRR7 degradation rate, of which data are available in Fig. 2c, and Supplementary Figs. 1 and 2. For CHX assays, PRR7pro::FLAG-PRR7-GFP seedlings were grown on MS media with 3% sucrose and 1% agar under 12L:12D cycles (white fluorescent light; 30–40 μmol m−2 s−1) at 22 °C for 14 days. Seedlings were transferred to MS liquid media with 100 μM CHX or mock (ethanol) at ZT17 in darkness. The tissues were kept in the dark under slow shaking and collected at 0, 1, 3, 5, and 7 h post treatment.

For immunoblots, the tissue was ground in liquid nitrogen and extracted in protein extraction buffer (50 mM Tris–Cl, pH 7.5, 150 mM NaCl, 0.5% Nonidet P-40, 1 mM EDTA, 3 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 5 μg ml−1 leupeptin, 1 μg ml−1 aprotinin, 1 μg ml−1 pepstatin, 5 μ g ml−1 antipain, 5 μg ml−1 chymostatin, 50 μg MG132, 50 μg MG115, 50 μl ALLN). Total proteins were separated using an 8% SDS–PAGE gel (acylamide:bisacrylamide, 37.5:1), immunoblotted and probed with anti-GFP antibody (Abcam, ab6556) and polyclonal anti-ADK antibody (gift from Dr. David Bisaro) diluted to 1:4000 and 1:15,000, respectively, followed by anti-rabbit IgG conjugated with horseradish peroxidase (GE healthcare, NA934). Chemiluminescent detection was performed using SuperSignal® West Pico Chemiluminescent Substrate (Thermo Scientific, 34080). The FLAG-PRR7-GFP protein signals were calculated by ImageJ software (NIH, version 1.8.0) from three biological repeats, and were normalized to their corresponding ADK (adenosine kinase) signal intensities individually.

8 COMMUNICATIONS BIOLOGY | DOI: 10.1038/s42003-018-0217-1 | www.nature.com/commsbio
All unique biological materials used in this study (PRR7pre:FLAG-PRR7-GFP) are available from the authors upon request.

Analysis of data from the plant and mammalian clock. By writing the protein synthesis rate \( g(t) \) as \( g(t) = k'(\rho(t)) \) and by assuming the roughly constant \( k' \), i.e., \( k(t) = k \), Eq. (1) can be written as

\[
\frac{dx(t)}{dt} \approx \kappa_0 g(t) - r(t)x(t),
\]

which leads to

\[
k \approx \frac{\kappa_0 - r(t)x(t)}{\kappa_0 g(t)},
\]

where \( t \) corresponds to each time point \( t = t \) with experimentally available degradation rate \( r(t) \). In the case of PRR7, we used the protein degradation rates at \( t = 4 \), 12, and 18 h (Fig. 2 and Supplementary Fig. 3). The first two degradation rates were obtained from the protein abundance data in Fig. 7b of Farre et al.28, while the last degradation rate was from our own experimental data in Fig. 3c. We also obtain experimental data of the mRNA and protein profiles from Fig. 5d of Flis et al.27 and Fig. 5a of Nakamichi et al.29, respectively. Both datasets have 2-h sampling intervals under 12L:12D cycles. These mRNA and protein levels were normalized by the peak levels of their splines, and adopted for \( g(t) \) and \( x(t) \) in Eq. (18), respectively. From Eq. (5), \( c_0 = 0.40 \) h⁻¹. Using \( r(t) \), \( x(t) \), and \( g(t) \), we obtained \( k(t) \) from Eq. (19). To be precise, although the small \( t_0 \) can have different values at \( t_0 = 4 \) h, \( t_0 = 12 \) h, \( t_0 = 18 \) h, the overall difference between \( k(t) \) is unlikely to be larger than \( 1.02 \) h⁻¹ every 2 h, except for \( 10 \) h. The former peak is a natural consequence of \( r(t) = \min{1, 21, 42.5, 11.6} \) molecules cell⁻¹. Therefore, we used the experimental protein levels for \( x(t) \) in Supplementary Fig. 3c. On the other hand, without a correction \( x(t) \) in Eq. (18), respectively. Following a similar procedure to the case \( x(t) \) whose values at \( t = 19, 22, 24 \) h are replaced by that of \( x(t) \) max20h. Using \( k(t) \) and \( r(t) \), we estimated \( \epsilon_0 = 0.32 \) h⁻¹ and \( \epsilon = 0.12 \) h⁻¹. Here, \( \epsilon \) is calculated from the newly estimated \( r(t) \) in Supplementary Fig. 4c. In the other hand, without a correction \( r(t) \) at 9.6 h to 33 h for one circadian period \( T = 23.4 \) h, and smoothened them with a moving window average (3-h window). These data were normalized by their peak level, and the resulting \( x(t) \) appears in Fig. 4a. \( R(t) \) derived from this \( x(t) \) is very noisy, and therefore smoothened with a moving window average (1-h window). For experimental protein degradation rates, we used the instantaneous half-lives after 0.5 h since CHX treatment at \( t = 19, 22, 25, 28, \) and 30 h in Supplementary Fig. 1a of Zhou et al.25.

Full details of the PRR7, PRR5, and PER2 data collection are provided in Supplementary Methods and Supplementary Table 1.

Analysis of data from the algal clock. In the case of CCA1 and TOC1 proteins in the Ostreococcus circadian system, we obtained the full time-course degradation rate \( r(t) \) and protein level \( x(t) \) data from Fig. 1a, b of van Oojen et al.30, respectively (12L:12D-cycle condition). We did not perform any normalization of \( x(t) \), and the unit of \( x(t) \) here follows that of \( x(t) \) molecules cell⁻¹. For the calculation of \( r(t) \), we estimated \( \epsilon_0 = 0.28 \) h⁻¹, \( \epsilon_0 = 0.17 \) h⁻¹, and \( \epsilon = 42.5 \) and 11.6 molecules cell⁻¹ respectively. Because \( x(t) \)‘s sampling resolution was rather low (4-h sampling interval), we did not apply \( r(t) \) and \( x(t) \) to Eq. (3) wherein the specific form of \( R(t) \) could be sensitive to the \( x(t) \)’s sampling resolution. For the calculation of \( c_0 \), we estimated \( r_{\text{min}} = \min{(\epsilon(t), \max{\phi(t), r(t)}), \epsilon(t)} \) with regards to possibly inaccurate \( r(t) \) from the low sampling resolution of \( x(t) \). For the calculation of \( \epsilon \), we adopted \( r(t) \) and \( \epsilon \) in Fig. 1c of van Oojen et al.30. As a result, for CCA1 and TOC1, \( r_{\text{min}} = 0.28 \) h⁻¹, and \( \epsilon = 42.5 \) and 11.6 molecules cell⁻¹ respectively. In other words, the cost of CCA1 and TOC1 production is about two-thirds of the assumed cost in the case of constant degradation rates.

Effects of oscillating BMAL1 expression. If \( - \beta r < T \) in Eqs. (14) and (17) can be used to calculate a phase difference between \( x(t) \) and \( r(t) \). Without loss of generality, let \( x(t) \) be the lowest at \( t = T', i.e., \( t' \leq T \). Depending on signs of \( \alpha \) and \( - \beta r \) in Eq. (17), we consider the following four cases:

1. If \( \alpha > 0 \) and \( - \beta r < T \), \( y(t) \) in Eq. (17) is described essentially in the same way as Eq. (10), while extra constants in Eq. (17) do not affect the way to determine a phase difference between \( x(t) \) and \( y(t) \).

\[
T' \leq t' \leq T,
\]

and the phase difference between \( x(t) \) and \( y(t) \) is determined in a similar way to the case with constant \( g(t) \) (i.e., \( g(t) = g_0 \)). In this case \( y(t) \) will be called \( y^1(t) \).
2. If $a > 0$ and $\beta r > 1$, $y(t)$ is determined in a similar way to $y_1(t)$, but with the flipped sign of $x(t)$. Therefore,

$$0 \leq t_x \leq \frac{t_2 - t_1}{T}$$

(22)

In this case, $y(t)$ will be called $y_2(t)$.

3. If $a < 0$ and $\beta r < 1$, $y(t)$ is described in a similar way to $-y_3(t)$. Therefore,

$$t_2 \leq t_x \leq t_1$$

(23)

4. If $a < 0$ and $\beta r > 1$, $y(t)$ is described in a similar way to $-y_3(t)$. Therefore,

$$t_2 \leq t_x \leq t_1$$

(24)

In addition, both $x(t)$ and $y(t)$ can have symmetric waveforms as long as $|a/(1-\beta r)| \geq \max|x(t)|$ (for example, this condition can be satisfied when $\beta r = 1$).

We further assume the waveform of $x(t+T/2)$ to be $x(t)$, where $x$ is a constant satisfying $1 \approx (2/T)\int x(t)dt$. From Eq. (9),

$$k(t) = x(t) = \frac{\alpha + \beta - (1 + \beta)e^{i\omega t}}{\alpha + \beta}$$

(25)

This equation takes a similar form to Eq. (17). By dividing four different categories of $y(t)$ depending on signs of $a + \beta r$ and $1 + \beta e^i$, it is straightforward to obtain similar results to our previous analysis of a phase difference between $x(t)$ and $y(t)$ when $\tau \ll T$.

In addition, both $x(t)$ and $y(t)$ can have symmetric waveforms as long as $|a + \beta r|/(1 + \beta e^i) \geq \max|x(t)|$ (for example, this condition can be satisfied when $\beta r = -1$).

To illustrate the diverse phase differences conferred by BMAL1 cycling, we study the case with a sinusoidal wave $x(t)$ in Eq. (13) and consider the oscillation of $g_x(t)$ in Eq. (14). From Eq. (9), $t_y$ is obtained as

$$\omega t_y = 2\pi n + 2\pi t_1$$

(26)

with $C = h_2 + L$ and a constant $n$ that satisfies $0 < \omega t_y \leq 2\pi$. From Eq. (15), $\alpha$ satisfies

$$\alpha \geq \alpha_{min} = \max \left\{ L \left( \frac{1}{2} - 2\beta \omega \sin(\omega t) + \frac{\beta^2}{\omega^2} - \frac{\beta}{\omega} - 0 \right) - \beta h_0 \right\}$$

(27)

Equation (27) and $t = T/2$ give rise to the exact solution of the phase difference $\phi$ between $x(t)$ and $y(t)$ ($\phi = |x(t) - y(t)|$), as plotted in Fig. 5e. This exact solution is in good agreement with our generic results based on the approximations Eq. (16) and (17).

**Code availability.** Source codes for analyzing data in the manuscript have been deposited in the public repositories GitHub and Zenodo.

**Data availability**

All relevant data are available in Methods, Figs. 2–4, Supplementary Methods, Supplementary Figs. 1–5, and Supplementary Data.

Received: 22 May 2018 Accepted: 1 November 2018

Published online: 26 November 2018

**References**

1. Nagel, D. H. & Kay, S. A. Complexity in the wiring and regulation of plant circadian networks. *Curr. Biol.* 23, R648–R657 (2013).

2. Schlegel, A., Price, J. L., Man, B. & Young, M. W. Loss of circadian behavioral rhythms and per RNA oscillations in the *Drosophila* mutant timeless. *Science* 263, 1603–1606 (1994).
Acknowledgements
This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) Grant 2015R1D1A1A01059598 funded by the Ministry of Education (H.-H.J.) and by the National Research Foundation of Korea (NRF) Grants N01160447 funded by the Ministry of Science, ICT and Future Planning (J.K.K.). This work was also supported by National Institutes of Health grant R01GM093285 and by a grant from the Next-Generation BioGreen21 Program (Systems and Synthetic Agrobiotech Center, Project PB1327305), Rural Development Administration, Republic of Korea (D.E.S.). P.-J.K. acknowledges the support by Hong Kong Baptist University, Research Committee, Start-up Grant for New Academics.

Author contributions
H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. designed the research. H.-H.J., Y.J.K., M.F., and P.-J.K. performed the research. H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. analyzed the data. H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. wrote the manuscript.

Additional information
Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-018-0217-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.