



## Supporting Online Material for

### **Circadian Gating of the Cell Cycle Revealed in Single Cyanobacterial Cells**

Qiong Yang, Bernardo F. Pando, Guogang Dong, Susan S. Golden,  
Alexander van Oudenaarden\*

\*To whom correspondence should be addressed. E-mail: [avano@mit.edu](mailto:avano@mit.edu)

Published 19 March 2010, *Science* **327**, 1522 (2010)  
DOI: 10.1126/science.1181759

#### **This PDF file includes:**

Materials and Methods

Figs. S1 to S11

Tables S1 to S4

References

# Circadian gating of the cell cycle revealed in single cyanobacterial cells

## SUPPLEMENTARY INFORMATION

Qiong Yang, Bernardo F. Pando, Guogang Dong, Susan S. Golden & Alexander van Oudenaarden

### Contents

I.	Parameterization of the coupling function	1
II.	Monte-Carlo simulations of the proposed model	4
III.	Fokker-Planck computation of the joint distribution $p_{\theta,d,\tau}(\theta, \tau)$	4
IV.	Distribution of division events across the circadian cycle	7
V.	Cell culture	7
VI.	Time-lapse microscopy	7
VII.	Single cell tracking method	8
VIII.	Alignment of <i>YFP</i> traces	10
IX.	Estimate of the expected distribution of timing of division events under the hypothesis of no coupling	11
X.	Estimate of the diffusion of circadian phases	12
XI.	Fits of $p_{\theta,d,\tau}$ according to the proposed model	12
XII.	Alternative parameterization of the coupling function and fit to the data	16
XIII.	Measurement of the phase delay in <i>YFP</i> relative to the day-night cycle	18

### I. Parameterization of the coupling function

The coupling function  $\gamma(\phi, \theta)$  is a non-negative periodic function of order 1 that describes how cell cycle speed is affected by the state of the circadian and cell cycle systems. One of the objectives of this work is to infer characteristics of this function from

quantitative single cell experiments. In order to do that we decided to parameterize this function with a handful of parameters and look for the set of parameters that best describes the data. Taking into account these restrictions that the coupling function must satisfy, we chose the following parameterization, which describes a single gate characterized by its position (given by the  $(\phi_0, \theta_0)$  coordinates), amplitude ( $A$ ) and widths along the  $\phi$  and  $\theta$  axes (quantified by the parameters  $\alpha$  and  $\beta$ ):

$$\gamma(\phi, \theta) = 1 + A \left[ \left| \cos\left(\frac{\phi - \phi_0}{2}\right) \right|^\alpha \left| \cos\left(\frac{\theta - \theta_0}{2}\right) \right|^\beta - \mathcal{N}_{\alpha, \beta} \right]. \quad (1)$$

The parameters  $\alpha$  and  $\beta$  regulate the extent of the gate along the cell-cycle-phase and circadian-phase coordinates respectively. The relative width of the gate along each direction with respect to the cycle period is correlated with the value of these parameters. Small values correspond to coupling functions that describe gates with a weak dependence on the corresponding variable. For instance, the limit  $\alpha \rightarrow 0$  corresponds to a coupling function that is completely independent of the cell-cycle coordinate.

The constant  $\mathcal{N}_{\alpha, \beta}$  is a normalization term that guarantees that this function is of order 1. We chose this normalization constant so that on average the coupling has no effect over a population of completely unsynchronized cells and therefore we implemented this restriction by imposing the condition

$$1 = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \gamma(\phi, \theta) d\phi d\theta, \quad (2)$$

which leads to the expression

$$\mathcal{N}_{\alpha, \beta} = \frac{1}{2^{\alpha+\beta}} \frac{\Gamma(1+\alpha)\Gamma(1+\beta)}{\Gamma(1+\frac{\alpha}{2})^2 \Gamma(1+\frac{\beta}{2})^2}. \quad (3)$$

In order for the coupling function to be non-negative the coefficient  $A$  must satisfy the condition  $\frac{-1}{1-\mathcal{N}_{\alpha,\beta}} \leq A \leq \frac{1}{\mathcal{N}_{\alpha,\beta}}$ . In this work we only considered maximal gates and so we fixed  $A = \frac{-1}{1-\mathcal{N}_{\alpha,\beta}}$ .

In Fig. S1 we show the shape of the coupling function for different choices of parameter values to illustrate the range of possibilities that we considered by using this approach. The first four panels illustrate the effect of the parameters  $\phi_0$ ,  $\theta_0$  and  $\beta$  and the last four show how as  $\alpha$  is taken to zero the coupling function becomes independent of the cell-cycle phase.

One peculiarity of this parameterization of the coupling function is that for any  $\alpha > 0$  and negative values of  $A$  (including those required for gates of maximal amplitude) the maximum of the coupling function is found at the loci defined by the conditions  $\phi = \phi_0 + \pi$  or  $\theta = \theta_0 + \pi$ , where the corresponding cosine terms become zero. This means that as a coupling function converges to one that lacks any dependence on  $\phi$  by taking the limit  $\alpha \rightarrow 0$ , for any value of  $\alpha > 0$  there is a region around  $\phi_0 + \pi$  where the coupling function takes values close to its maximum. In other words, the convergence is not uniform. However, the impact of the size of this region becomes negligible as  $\alpha \rightarrow 0$  and it does not contribute significantly to the dynamics of the underlying model. Moreover, alternative parameterizations that do not share this peculiarity, such as that presented in Section XII, yield the same inferences as those obtained with the parameterization introduced in this section. In Fig. S2 we illustrate this effect, comparing convergence to a coupling function that is independent of the cell cycle phase both in the parameterization described in this Section and that introduced in Section XII.

## II. Monte-Carlo simulations of the proposed model

Monte Carlo simulations of the proposed model were performed using an Euler integration scheme with a time interval of at most  $\Delta t = \frac{0.02}{v_0}$ . The noise terms were chosen independently at each time step from a normal distribution with standard deviation  $\sqrt{2D\Delta t}$  where  $D$  is the corresponding noise strength.

## III. Fokker-Planck computation of the joint distribution $p_{\theta,\tau}(\theta, \tau)$

In order to attack the problem of computing the joint distribution of circadian phases at division and cell cycle durations let's first consider a cell that is in state  $(\phi, \theta)$  at  $t = 0$  and let's ask what is the probability  $P(\phi, \theta, t)$  that it would not have completed its cell cycle by time  $t$ . This quantity satisfies the following Fokker-Planck-like equation (1)

$$\frac{\partial P}{\partial t} = D_\phi \frac{\partial^2 P}{\partial \phi^2} + D_\theta \frac{\partial^2 P}{\partial \theta^2} + v\gamma(\phi, \theta) \frac{\partial P}{\partial \phi} + v_0 \frac{\partial P}{\partial \theta} \quad (4)$$

subject to periodic boundary conditions along the  $\theta$  axis and an absorbing boundary at  $\phi = 2\pi$  (recall that we are associating cell cycle completion with the  $\phi$  variable crossing the  $2\pi$  boundary). We also need to indicate some boundary condition as  $\phi \rightarrow -\infty$ . Given that the speed of cell cycle progression is positive the probability that the cell will perform an excursion towards  $\phi = -\infty$  is negligible and we can approximate the situation by erecting a reflecting boundary at some finite value of  $\phi$ , which we have taken as 0 (for the sets of parameters explored in this study changing this value to smaller values does not change the results of the inferences significantly). The system becomes fully specified with the initial condition  $P(\phi, \theta, 0) = 1$ :

$$\begin{cases} \frac{\partial P}{\partial t} = D_\phi \frac{\partial^2 P}{\partial \phi^2} + D_\theta \frac{\partial^2 P}{\partial \theta^2} + v\gamma(\phi, \theta) \frac{\partial P}{\partial \phi} + v_0 \frac{\partial P}{\partial \theta} \\ P(\phi, \theta = 0, t) = P(\phi, \theta = 2\pi, t) \\ P(\phi = 2\pi, \theta, t) = 0 \\ \left. \frac{\partial P}{\partial \phi} \right|_{(\phi=0, \theta, t)} = 0 \\ \lim_{t \rightarrow 0^+} P(\phi, \theta, t) = 1. \end{cases} \quad (5)$$

We solved this PDE numerically using an implicit finite difference scheme over a grid of 24×24 points (for the sets of parameters explored in this study using a finer grid of 48×48 points did not alter the results of the inferences significantly).

Once this quantity is computed the probability density that a cell that was born at circadian phase  $\theta_b$  would have divided by time  $\tau$  can be computed as

$$p_{\tau|\theta_b}(\tau|\theta) = - \left. \frac{\partial P}{\partial t} \right|_{(\phi=0, \theta, \tau)}. \quad (6)$$

The joint distribution of birth phases and cell cycle durations satisfies

$$p_{\theta_b, \tau}(\theta, \tau) = p_{\tau|\theta_b}(\tau|\theta)p_{\theta_b}(\theta), \quad (7)$$

where  $p_{\theta_b}$ , the marginal distribution of birth phases, represents the probability of observing some birth phase irrespective of the lifetime associated with that cell:

$$p_{\theta_b}(\theta) = \int_0^\infty p_{\theta_b, \tau}(\theta, \tau) d\tau. \quad (8)$$

So far we don't have this quantity, but we can compute it in steady state by noting that once equilibrium is reached the marginal distribution of division phases should be equal to the distribution of birth phases ( $p_{\theta_b}(\theta) = p_{\theta_d}(\theta)$ ) and that there is a relationship between these quantities:  $\theta_d = \theta_b + v_0\tau \pmod{2\pi}$ , so that we can write

$$p_{\theta_d, \tau}(\theta, \tau) = p_{\theta_b, \tau}(\theta - v_0\tau, \tau) \quad (9)$$

where the difference in the first argument is to be understood modulo  $2\pi$ . We have that

$$\begin{aligned}
p_{\theta_b}(\theta) &= p_{\theta_d}(\theta) = \int_0^\infty p_{\theta_d,\tau}(\theta, \tau) d\tau = \int_0^\infty p_{\theta_b,\tau}(\theta - \nu_0\tau, \tau) d\tau \\
&= \int_0^\infty p_{\tau|\theta_b}(\tau|\theta - \nu_0\tau)p_{\theta_b}(\theta - \nu_0\tau) d\tau,
\end{aligned} \tag{10}$$

which defines an integral equation for the distribution of birth phases. Given that the kernel  $p_{\tau|\theta_b}$  had been already computed this equation can be solved numerically following an iterative procedure starting with, for example, a uniform distribution:

$$p_{\theta_b}(\theta) = \lim_{n \rightarrow \infty} p_{\theta_b}^n(\theta), \text{ where } p_{\theta_b}^{n+1}(\theta) = \int_0^\infty p_{\tau|\theta_b}(\tau|\theta - \nu_0\tau)p_{\theta_b}^n(\theta - \nu_0\tau) d\tau. \tag{11}$$

One possible way of gaining intuition about this equation is to think that it “evolves” a given distribution of birth phases for one cell cycle, *i.e.* if we have a population of cells that were born at different circadian phases this equation describes the distribution of birth phases of the next generation. The number of iterations necessary to achieve convergence depends on the strength of the diffusive terms. For the range of parameter sets explored during the fitting procedure (section XI) this procedure converged after about 10 iterations and in all cases we checked that the inferences were not modified if we doubled the number of iterations.

Once  $p_{\tau|\theta_b}$  and  $p_{\theta_b}$  are computed, the joint distribution of division phases and cell cycle durations can be calculated using equations (7) and (9).

This technique is equivalent to running Monte-Carlo simulations until a steady state is achieved and inferring the distribution from a sample but has the advantage of being a completely deterministic calculation, which eliminates uncertainties related to finite sampling and therefore reduces the computational cost of estimating the distribution.

#### **IV. Distribution of division events across the circadian cycle**

In order to illustrate some of the general properties of the proposed model in Fig. S3 we show the distribution of division events across the circadian cycle as a function of  $\nu/\nu_0$ , a parameter that quantifies the average speed of cell cycle progression with respect to the circadian speed. As the parameter  $\nu$  is increased the cell cycle gets synchronized to the circadian signal and division events get localized at different parts of the day. If the cell cycle phase moves forward too slowly compared to the speed of the circadian system the effect of the coupling becomes less important and division events tend to get distributed more uniformly across the day. Even though the effect of coupling is less important for slow cell-cycle speeds compared to speeds faster than that of the circadian signal it is possible to get significant entrainment for slow speeds if the noise strength is sufficiently small (Fig. S3E,F).

#### **V. Cell culture**

Cells corresponding to the WT strain with an *YFP* reporter were grown at 30 °C in BG11 media supplemented with 40  $\mu\text{g ml}^{-1}$  spectinomycin in a constant light environment maintained at an intensity of about 30  $\mu\text{E m}^{-2} \text{s}^{-1}$ .

#### **VI. Time-lapse microscopy**

Cell cultures were diluted three days before acquisition so that their optical densities at 750 nm were about 0.2 (HITACHI U-1800 Spectrophotometer). For each experiment, 1.5  $\mu\text{l}$  of cells were loaded in one of two wells of a glass chamber (LAB-TEK)

and covered with a square pad of 2% low-melting agarose. Next, a liquid 2% low-melting agarose was poured on top of the first solid pad and the chamber was loaded on the microscope for imaging. Phase contrast and *YFP* images were taken every 40 minutes for a total of 16 positions for each sample. The whole acquisition period lasted three days for a typical experiment, during which a constant cool-white fluorescent light source illuminated the sample with an intensity of either  $50 \mu\text{E m}^{-2} \text{s}^{-1}$  or  $25 \mu\text{E m}^{-2} \text{s}^{-1}$  and was only shut off during image acquisition. The acquired images were analyzed with custom-written MATLAB software to track each single cell as it divides and oscillates.

## **VII. Single cell tracking method**

A typical time-lapse microscopy video lasts for three days, starting with a single cyanobacterial cell and ending with a grown-up colony containing more than a hundred cells.

In order to quantitatively measure both the circadian signal and cell cycle progression of each single cell, we wrote a MATLAB program that is based on the basic Image Processing Toolbox™ functions. The program works on a time series of phase images segmenting and tracking each individual cell lineage, to obtain the cell cycle information. The segmented images are then overlapped with *YFP* images to extract the circadian signal corresponding to each cell.

The main program includes three parts and is mostly automated except for the last part which allows the users to check and interactively correct detected errors:

1) Before segmentation, a series of modifications are performed on the raw image data, preparing them to be segmented and tracked efficiently. First, the program detects and eliminates all dead pixels that come from imperfections in the camera. Then if one video includes more than one colony it separates them and crops away any non-cell regions for reducing the image size and speeding up the image processing. Last, it aligns all images of each growing colony in a time series based on their centers of mass. This last step overcomes stage drifting issues that are often associated with long-term time-lapse microscopy and is critical for correct lineage tracking later.

2) The aligned phase image series are now ready for segmentation and tracking. The basic algorithm for segmentation is based on a marker-controlled watershed transform. Once markers are extracted by a set of standard imaging processing tools including applying filters and morphological operations to the phase images, the watershed transform uses these markers to separate touching objects and detect the accurate contour of individual cells.

Correct marker detection is thus the key to the accurate segmentation. However, marker detection is not always successful, as it is very sensitive to image quality. To overcome the unpredictable detection problem, we took advantage of the fact that we are tracking a continuous growing culture. Instead of starting over and detecting markers every time from scratch, we let each frame (except the first frame) inherit the markers that have been detected and verified in the previous frame. This way the efficiency and robustness of the segmentation process increases, making the program automated to a large extent.

After complete the segmentation stage the program proceeds to track each cell lineage. It compares each newly segmented frame with the previous one and identifies cells based on the largest area overlapping between these two successive frames. If two regions in the current frame are found to match one of the regions in the previous frame, we consider these to be the daughter cells of the cell in the previous frame.

3) The automated segmentation and lineage tracking described above is considerably efficient and robust. However, sometimes when the image quality itself is not good enough, some errors may be generated. To increase the accuracy, we added a last step to check the results based on several criteria. For example, for wild type cells growing in reasonable conditions, one cell should not once divide into more than two daughters, the later frame should always have a larger number of cells than the previous frame, etc. Using these criteria, the program will inquire users to interactively check and correct potential errors. The correction is simple and straightforward. Based on what kind of error it may be, the users are allowed to make a click which can either delete or add a maker to the set of existing markers that had been automatically detected by the program (Fig. S4).

### **VIII. Alignment of *YFP* traces**

Given that the circadian signals of different colonies in each experiment were not necessarily synchronized to each other we synchronized them in-silico so that we could analyze cell cycle effects at different circadian phases over the population. For each colony we chose a time origin so that the average *YFP* signals from all colonies in one experiment would align properly. To do so we first resampled all *YFP* traces in the

colony with a resolution of 1 hour, we computed the average *YFP* signal over all traces, we smoothed it out with a 5-point running average filter and we detected the position of local maxima. We chose the time origin so as to maximize the alignment between both maxima and phases of the form  $2n\pi$  ( $n \in \mathbb{Z}$ ) and minima and phases of the form  $2n\pi + \overline{\Delta\theta}$  where  $\overline{\Delta\theta}$  represents the measured phase displacement between maxima and minima.

### **IX. Estimate of the expected distribution of timing of division events under the hypothesis of no coupling**

The fact that each colony is growing throughout the course of the experiment and that the experiment is of finite duration makes the expected distribution of the timing of division events under the hypothesis of no coupling non-trivial. Roughly speaking one would expect the number of division events to increase exponentially for early times due to cell divisions and to suffer from an exponential cutoff at late times due to the finite nature of the experiment. Under no coupling the circadian and cell cycle processes would be unsynchronized. Therefore, in order to estimate this distribution we followed a randomization procedure in which for each colony we considered 2000 random time offsets uniformly distributed over a time period corresponding to one circadian cycle. After doing this we pooled all realigned division events and constructed the histograms shown as gray traces in Fig. 4B.

## X. Estimate of the diffusion of circadian phases

In order to estimate the diffusion coefficient of circadian phases we computed the autocorrelation of *YFP* traces for each lineage and averaged them to obtain an average autocorrelation function  $\langle C_\theta(\tau) \rangle$  for each experiment. Fitting these curves by the expression  $e^{-D_\theta \tau} \cos(\nu_0 \tau)$  gave us estimates for the values of  $D$  associated with each experiment (Fig. S5). Averaging over the two experiments we obtained a value of  $D_\theta = (0.030 \pm 0.008)\nu_0$ . Taking into account that experimental errors as well as signal drift would make this number an overestimate we decided to use a value of  $D_\theta = 0.01\nu_0$  in our computations.

## XI. Fits of $p_{\theta_d, \tau}$ according to the proposed model

Given a set of parameters  $\psi$  that specifies the proposed model, the Fokker-Planck technique described in section III provides a prediction for the joint distribution of division phases and cell cycle durations,  $p_{\theta_d, \tau}(\theta_d, \tau | \psi)$ . To take into account the experimental uncertainty in the determination of the timing of division events we convolved this distribution with a Gaussian kernel with widths corresponding to 0.75 hours along each dimension, producing a new distribution,  $\bar{p}_{\theta_d, \tau}(\theta_d, \tau | \psi)$  that we could directly compare to the measured distribution of the pairs  $(\theta_d, \tau)$  over all cells in each experiment.

We determined the set of parameters that best described the data by confronting the measurements with the model predictions using a Bayesian inference approach (2). Briefly, by considering each pair  $(\theta_d, \tau)$  as an independent sample from the distribution

$\bar{p}_{\theta_d, \tau}$ , applying Bayes theorem and considering uniform priors over each parameter one can describe the *a posteriori* distribution of parameters as

$$P(\psi|\text{data}) \propto \prod_{j \in \text{data}} \bar{p}_{\theta_d, \tau}(\theta_d^{(j)}, \tau^{(j)}|\psi), \quad (12)$$

where  $(\theta_d^{(j)}, \tau^{(j)})$  represents the  $j$ th measurement. We sampled this distribution using a Metropolis-Hastings algorithm, which gave us estimates for the distribution of parameters with the highest likelihood of being good descriptors of the data.

The parameter set considered for the fit was  $\psi = (D_\phi, v_1, v_2, \phi_0, \theta_0, \alpha, \beta)$ , *i.e.*, we included the strength of the noise associated with the cell cycle system, the average speeds of cell cycle progression in each experiment and the location and shape parameters of the coupling function. The parameter  $D_\theta$  was kept fixed at a value of  $0.01v_0$  (see section X) and we only considered gates of maximal amplitude.

Parameter		Sampling parameters			
symbol	unit	starting point	jump width	lower bound	upper bound
$D_\phi$	$v_0$	0.4	0.1	0.1	0.8
$v_1$	$v_0$	1.3	0.1	0.1	4.0
$v_2$	$v_0$	2.2	0.1	0.1	4.0
$\phi_0$	-	0	0.2	-10	10
$\theta_0$	-	0	0.2	-10	10
$\alpha$	-	0.1	0.1	0	10
$\beta$	-	8.0	0.5	0	20

**Table S1.** Summary of the parameters used in the fitting procedure.

The sampling algorithm was run by following 25 independent Markov chains starting from the point indicated in Table S1. For each parameter, normal distributions with widths as indicated in Table S1 and centered in the previous point were used as jump distributions. We used lower and upper bounds for each parameter as indicated in

Table S1 but, except for the case of  $\alpha$ , the chains stayed away from the boundaries. Each chain was followed for 200 iterations and only the second half of the simulations was used to draw inferences.

Parameter		Inferences		
symbol	unit	best	mean	std
$D_\phi$	$\nu_0$	0.57	0.58	0.04
$\nu_1$	$\nu_0$	1.31	1.31	0.04
$\nu_2$	$\nu_0$	2.12	2.11	0.03
$\phi_0$	-	4	4	1
$\theta_0$	-	5.78	5.76	0.05
$\alpha$	-	0.12	0.09	0.07
$\beta$	-	8.2	8.4	0.8

**Table S2.** Summary of the inferred parameters statistics.

In Fig. S6 we show the inferred distributions for each parameter as well as the time-evolution of the *a posteriori* log-likelihood to illustrate the convergence of the procedure. First order statistics that describe the inferred values for each fit parameter can be found in Table S2. Note that the coupling function that we inferred is characterized by a value of  $\alpha$  very close to zero, where the parameter  $\phi_0$  loses meaning. This is consistent with the fact that our inferences on  $\phi_0$  are the ones with the largest uncertainty.

One of the main conclusions of the fit is that the best coupling function is one that represents a situation in which gating is independent of the phase of the cell cycle, which we infer from the fact that the sampled values of  $\alpha$  are distributed tightly close to  $\alpha = 0$ . To test this conclusion more extensively we performed another fit but imposing the condition  $\alpha > 2$ , which means that we restricted our attention only to coupling functions with a significant cell-cycle dependence. In Fig. S6J we show the distribution of sampled

likelihood values both for the original and the restricted fits and we observe that if we restrict our attention to coupling functions with  $\alpha > 2$  we obtain a significantly worse fit.

To further assess the sensitivity of the fit, in Figure S7A-D we show how the likelihood (eq. 12), which quantifies the quality of the fit, degrades as we change each parameter away from the values that characterize the best fit. In all cases the reduction in the likelihood score is consistent with the distribution of sampled parameters (Fig. S6).

In Fig. S8 we display both the coupling function corresponding to the best fit (Fig. S8A) and the expected coupling function (Fig. S8B), obtained by averaging the coupling across the sampled parameters. In Fig. S8C we exhibit the mean and standard deviation associated with the value of the inferred coupling function at different points in  $(\phi, \theta)$  space. The inferred values of the coupling function around  $\theta_0$  are significantly smaller than at the opposite circadian phase and are consistent with a uniform distribution across cell-cycle phase, from where we make the inference that the best description of our experiments is achieved by a coupling function that is essentially independent of the cell-cycle phase.

Finally, to show that the results presented in Fig. 1 are not necessarily restricted to specific coupling function shapes, in Fig. S9 we show the results of performing an analysis similar to that presented in Fig. 1 but using the coupling function that we inferred as the best fit to our data. The effect of synchronization can be still seen in the distribution at times at which divisions take place.

## XII. Alternative parameterization of the coupling function and fit to the data

To verify that the conclusions we arrived at are not an artifact of the specific parameterization of the coupling function we chose, we investigated an alternative parameterization and checked that the gating function inferred through the fitting procedure had features similar to the coupling function reported in section XI.

We based this parameterization on a family of functions related to Jacobi Elliptic Theta functions (3). Specifically, the building blocks are the functions

$$V_q(\varphi) = 1 + 2 \sum_{n=1}^{\infty} q^{n^2} \cos(n\varphi), \quad (13)$$

which are  $2\pi$ -periodic, with just one maximum (located at  $\varphi = 0$ ), just one minimum (located at  $\varphi = \pi$ ), and a typical width that is regulated by the parameter  $q$  ( $q = 1$  corresponding to an infinitely peaked function and  $q = 0$  corresponding to a situation in which there is no dependence on  $\varphi$  whatsoever). Using these functions we can construct a coupling function using the expression

$$\gamma(\phi, \theta) = 1 + A \left[ V_{q_\phi}(\phi - \phi_0) V_{q_\theta}(\theta - \theta_0) - 1 \right] \quad (14)$$

where we have already imposed the condition expressed by equation (2). In order to obtain a non negative coupling function the amplitude  $A$  must satisfy the constraints

$$\frac{-1}{V_{q_\phi}(0)V_{q_\theta}(0) - 1} \leq A \leq \frac{1}{1 - V_{q_\phi}(\pi)V_{q_\theta}(\pi)}. \quad (15)$$

As in the previous case, we only considered maximal gates and so we fixed the amplitude at  $A = \frac{-1}{V_{q_\phi}(0)V_{q_\theta}(0) - 1}$ .

This form of coupling function is characterized by parameters that describe the position of the gate ( $\phi_0$  and  $\theta_0$ ) and its widths along the cell-cycle and circadian phase

coordinates ( $q_\phi$  and  $q_\theta$  respectively), so this family of coupling functions is described in terms of parameters analogous to those used to describe the shape of the coupling function proposed in Section I. In Fig. S10A-D we exhibit the shape of coupling functions for some choice of parameters to illustrate the similarities with the coupling functions obtained from the parameterization described in Section I.

To analyze the effect of changing the parameterization of the coupling function on our conclusions we performed a fit to the data applying the same Bayesian inference algorithm described in Section XI but using the alternative description of the coupling function described in this section. The sampling parameters were kept as before, except for the new parameters  $q_\phi$  and  $q_\theta$ , for which the corresponding information is presented in Table S3.

In Table S4 we report the result of the fit and in Fig. S10E-H we report the best and average inferred coupling functions as well as a cut along the circadian phase of maximal gating to illustrate the uniformity of the coupling along the cell cycle coordinate.

We conclude that this alternative parameterization of the coupling function does not alter our conclusions significantly and so we infer that our statements are not an artifact of the specific parameterization initially chosen.

Parameter		Sampling parameters			
symbol	unit	starting point	jump width	lower bound	upper bound
$q_\phi$	-	0.5	0.05	0	1
$q_\theta$	-	0.5	0.05	0	1

**Table S3.** Summary of the parameters used in the fitting procedure for the new variables introduced in the parameterization described in this Section.

Parameter		Inferences		
symbol	unit	best	mean	std
$D_\phi$	$\nu_0$	0.59	0.57	0.04
$\nu_1$	$\nu_0$	1.32	1.31	0.04
$\nu_2$	$\nu_0$	2.10	2.11	0.03
$\phi_0$	-	4.4	4.2	0.6
$\theta_0$	-	5.75	5.74	0.05
$q_\phi$	-	0.02	0.02	0.02
$q_\phi$	-	0.80	0.80	0.01

**Table S4.** Summary of the inferred parameters statistics.

### XIII. Measurement of the phase delay in *YFP* relative to the day-night cycle

All experiments described in the main text were performed on cells that had not been synchronized and that had been cultured under constant light conditions for about one month by being periodically diluted. We did this to minimize the effect of a changing environment on the coupling between the cell cycle and circadian clocks. In these unsynchronized cultures a traditional population measurement is not sufficient to extract the correct circadian and cell cycle signals. Instead, we have utilized a single cell method.

In order to relate our data to the real day-night cycle, we have to take into account that our protein reporter, YFP-SsrA, even though forced to degrade faster than normal by way of the SsrA tag, still has non-negligible life time and this generates a phase delay in the fluorescent signal. To quantify the time lag, we performed an experiment on a synchronized culture in which cells had been alternatively transferred between constant light and constant dark environments every 12 hrs for a total of three days. Fig. S11 shows the *YFP* time course from a single colony of the synchronized

culture. Time 0 indicates the instant when the culture was transferred from dark to light. We saw that *YFP* peaked at  $(19 \pm 1)$  hours after the day start which is also consistent with a previous study (4).

## SUPPLEMENTARY INFORMATION REFERENCES

1. C. W. Gardiner, *Handbook of stochastic methods*. (Springer-Verlag, New York, NY, 2004).
2. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, *Bayesian data analysis*. (Chapman & Hall/CRC, Boca Raton, FL, 2004).
3. M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. (U.S. Govt. Print. Off., Washington, ed. 10, 1972).
4. J. R. Chabot, J. M. Pedraza, P. Luitel, A. v. Oudenaarden, *Nature* **450**, 1249 (2007).

## SUPPLEMENTARY FIGURE CAPTIONS

**Fig. S1. Sample of coupling function shapes.** Maximal coupling functions for different points in the space of parameters chosen for describing their shapes. **(A)**  $\alpha = \beta = 4$ ,  $\phi_0 = \theta_0 = \pi$ . **(B)**  $\alpha = \beta = 4$ ,  $\phi_0 = \pi$ ,  $\theta_0 = 3\pi/2$ . **(C)**  $\alpha = \beta = 4$ ,  $\phi_0 = 3\pi/2$ ,  $\theta_0 = \pi$ . **(D)**  $\alpha = 4$ ,  $\beta = 1$ ,  $\phi_0 = \theta_0 = \pi$ ; **(E-H)**  $\beta = 4$ ,  $\phi_0 = \theta_0 = \pi$  and  $\alpha = 6, 4, 2, 0$ ; note that in the last case the parameterization is degenerate and the parameter  $\phi_0$  loses meaning. This series illustrates how as  $\alpha \rightarrow 0$  one recovers a coupling function that is independent of the cell-cycle phase.

**Fig. S2. Convergence to coupling functions with no dependence on the cell cycle phase.** **(A)** Cuts along the circadian phase axis of coupling functions parameterized according to the model described in Section I. The parameters used were  $\alpha = 2, 1, 0.1, 0$  (from left to right panel),  $\beta = 4$ ,  $\phi_0 = \pi/2$ ,  $\theta_0 = 0$ . The different traces in each panel correspond, from bottom to top, to different values the circadian phase that go from 0 to  $2\pi$  in  $\pi/4$  intervals. **(B)** As in **A** but for the parameterization proposed in Section XII. The parameters used were  $q_\phi = 0.4, 0.2, 0.05, 0$  (from left to right panel),  $q_\phi = 0.5$ ,  $\phi_0 = \pi/2$ ,  $\theta_0 = 0$ . In this case the convergence to the a coupling function that is independent of cell cycle phase is uniform.

**Fig. S3. Synchronization of cell cycle phases to the circadian signal.** **(A)** Distribution of circadian phases at division,  $p(\theta_d)$ , as a function of  $v/v_0$  as computed using the Fokker-Planck calculation described in section III. The model parameters used were  $D_\phi = 0.05v_0$ ,  $D_\theta = 0.01v_0$ ,  $\theta_0 = \phi_0 = \pi$ ,  $\alpha = \beta = 4$ . **(B)**  $p(\theta_d)$  plotted as a function

of  $\theta_d$  for some values of  $\nu/\nu_0$  using the same parameters as in **A**. **(C)** As in **A** but with  $D_\phi = 0.2\nu_0$ . **(D)** As in **B** but with  $D_\phi = 0.2\nu_0$ . **(E)**  $p(\theta_d)$  at  $\nu/\nu_0 = 0.5$  for different values of cell-cycle progression noise strength. **(F)** As is **E** but for  $\nu/\nu_0 = 2.5$ .

**Fig. S4. Interactive corrections of automatically detected markers.** **(A)** Automatically detected markers (green dots) with one error reported. A white circle highlights the error. **(B)** Deleting the wrong marker by clicking. A star represents the click. **(C)** Adding correct markers by clicking.

**Fig. S5. Estimate of the diffusion of circadian phases.** Average autocorrelation function,  $\langle C(\tau) \rangle$ , and fits to the functional form  $e^{-D_\theta \tau} \cos(\nu_0 \tau)$  for the experiments performed under a light intensity of  $I \sim 25 \mu\text{E m}^{-2} \text{s}^{-1}$  **(A)** and  $I \sim 50 \mu\text{E m}^{-2} \text{s}^{-1}$  **(B)**. The values of  $D_\theta$  that corresponds to the best fits are  $(0.024 \pm 0.006)\nu_0$  **(A)** and  $(0.04 \pm 0.01)\nu_0$  **(B)**.

**Fig. S6. Inferences on the parameters that best describe the data.** **(A)-(G)**, Metropolis-Hastings samples of the parameter values that best describes the data, according to a Bayesian-inference approach (see section XI). **(H)** Evolution of the value of the *a posteriori* log likelihood as a function of the iteration number of the sampling algorithm. A constant was subtracted so that the value of 0 corresponds to the highest likelihood found, *i.e.* that corresponding to the parameter set that best describes the data. **(I)** Normalized correlations,  $C_{xy} = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\text{std}(x) \text{std}(y)}$ , between the sampled parameters. **(J)** Distribution of the log-likelihood estimates associated with the sampled parameters both for the case in which no constraint was imposed (light blue) and for the case in which the

gating functions were restricted to have a significant dependence on the cell-cycle phase (light red).

**Fig. S7. Sensitivity of the fit to changes in the parameters that shape the coupling function.** (A-D) Log-likelihood estimate as a function of each parameter when all the others are kept fixed at the value that correspond to the best fit found. The arrows indicate the values that correspond to the best fit.

**Fig. S8. Inferred coupling function.** (A) Coupling function corresponding to the set of parameters that best fits the experimental data. (B) Expected coupling function, obtained by averaging the coupling across parameters sampled according to their likelihood of being descriptors of the experimental data. (C) Confidence bands (mean  $\pm$  s.d.) for the inferred coupling function across cuts parallel to the  $\phi$  axis.

**Fig. S9. Synchronization in the case of the inferred coupling function.** (A) Coupling function corresponding to the set of parameters that best fits the experimental data. (B, C) Steady-state distribution of circadian phases at which division takes place according to Monte Carlo simulations using the parameters corresponding to the best fit; the solid line is the result of the Fokker-Planck computation. The plot in B is representative of the experiment performed at  $I \sim 25 \mu\text{E m}^{-2} \text{s}^{-1}$  whereas that in C corresponds to  $I \sim 25 \mu\text{E m}^{-2} \text{s}^{-1}$ . (D, E) As in B, C but considering lower noise strengths.

**Fig. S10. Alternative parameterization of the coupling function.** (A-D) Samples of coupling function for different choices of the shape parameters; A:  $q_\phi = q_\theta = 0.5$ ,

$\phi_0 = \theta_0 = \pi$ ; **B**:  $q_\phi = q_\theta = 0.5$ ,  $\phi_0 = \pi$ ,  $\theta_0 = 3\pi/2$ ; **C**:  $q_\phi = q_\theta = 0.5$ ,  $\phi_0 = 3\pi/2$ ,  $\theta_0 = \pi$ ,  
**D**:  $q_\phi = 0.5$ ,  $q_\theta = 0.2$ ,  $\phi_0 = \theta_0 = \pi$ . **(E)** Coupling function corresponding to the set of parameters that best fits the experimental data. **(F)** Expected coupling function, obtained by averaging the coupling across parameters sampled according to their likelihood of being descriptors of the experimental data. **(G)** Confidence bands (mean  $\pm$  s.d.) for the inferred coupling function across a cut parallel to the  $\phi$  axis at the circadian phase of maximal gating. **(H)** As in **G** but for a cut parallel to the  $\theta$  axis at  $\phi = \pi$ .

**Fig. S11. Measurement of the phase delay in YFP signal relative to the day-night cycle.** YFP intensity time course obtained from a colony that had been synchronized with three 12 hours - 12 hours light-dark cycles (boxes on the top left corner indicate the sequence during the last day of entraining). The orange curve represents a 6-hour running window average. Dashed gray vertical lines indicate the positions of local maxima in the YFP signal. Horizontal double arrows show the delay between the start of the subjective day and the following YFP peak.

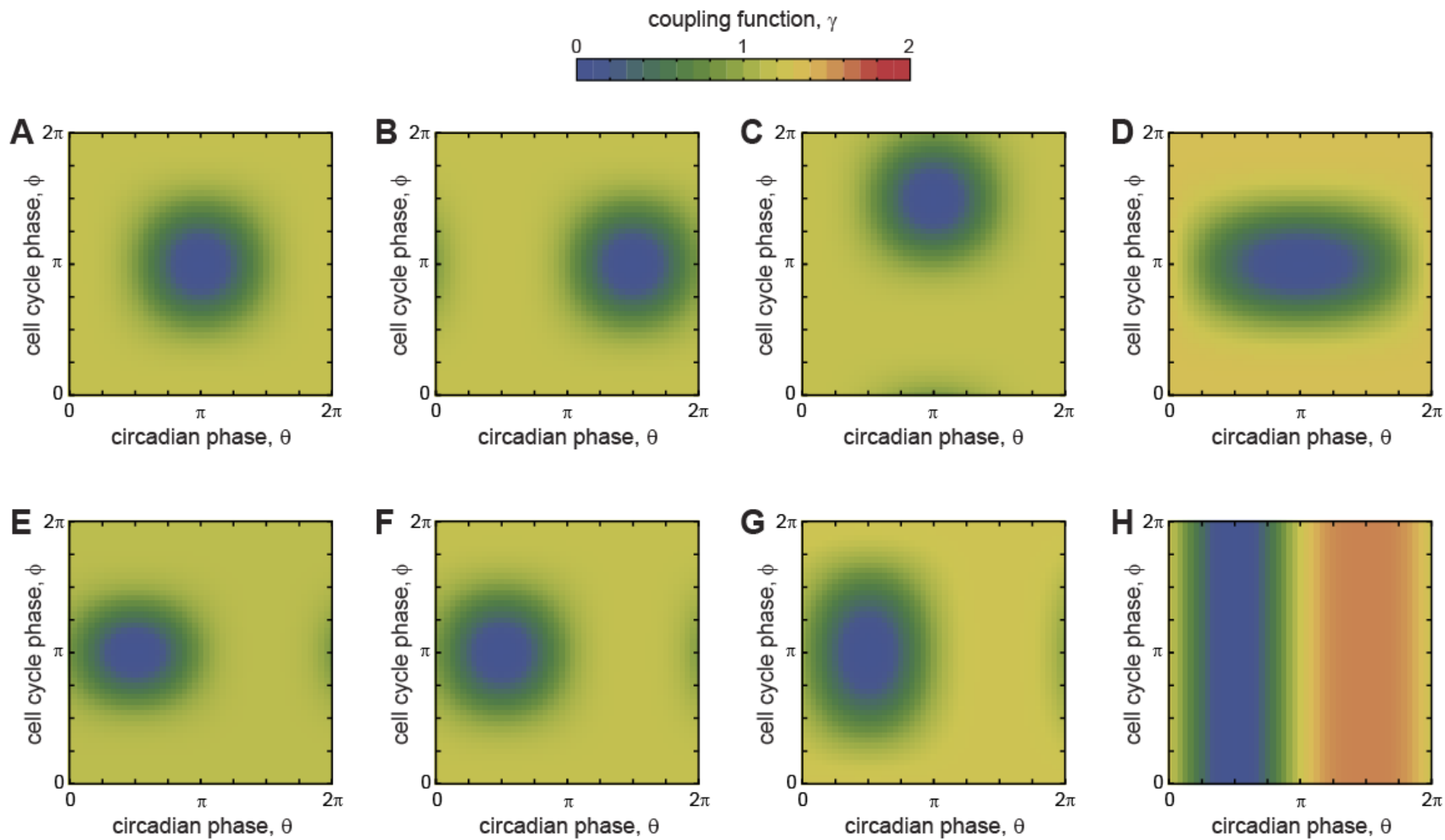


Figure S1

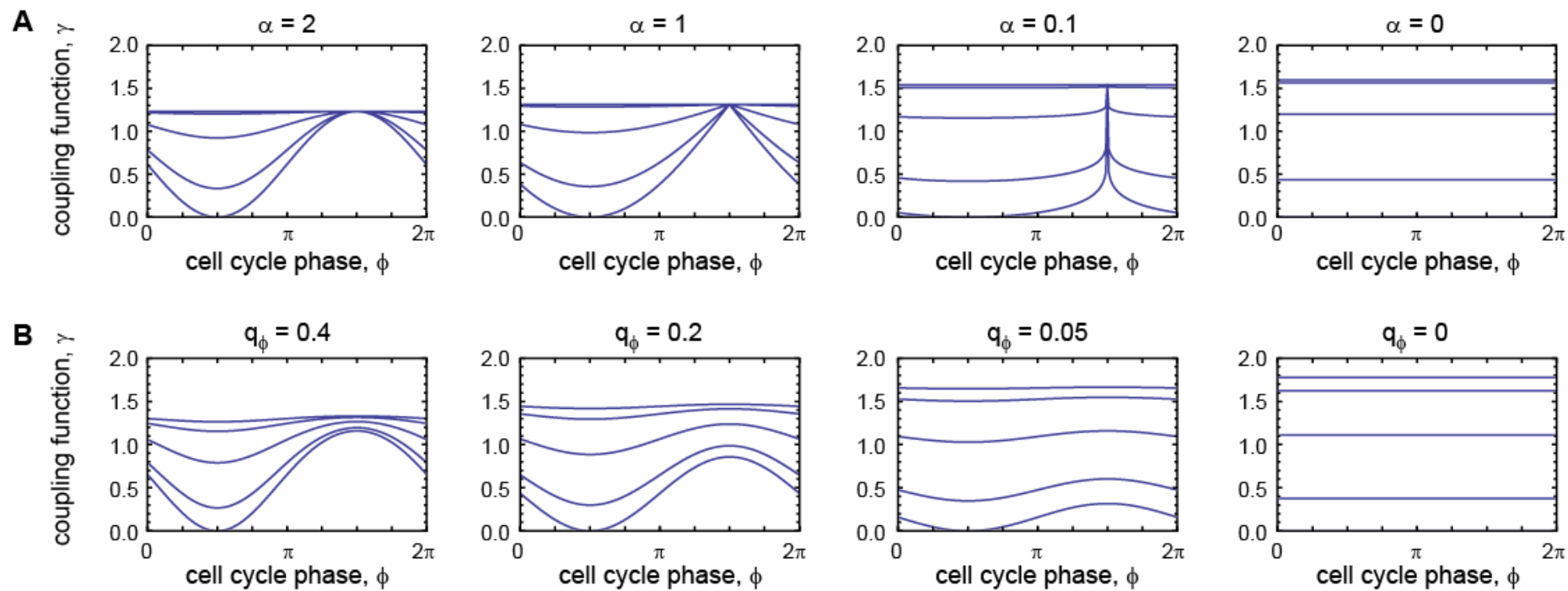


Figure S2

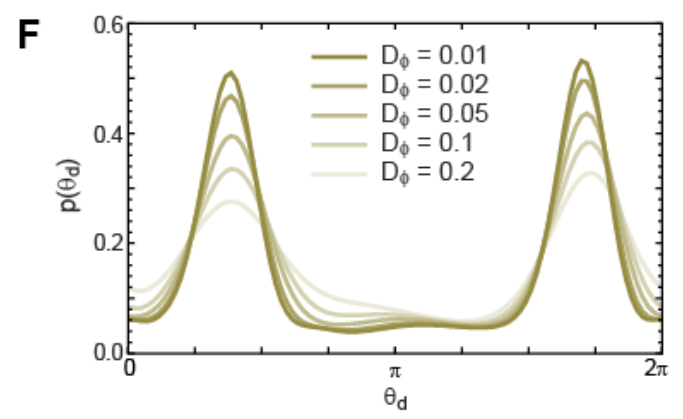
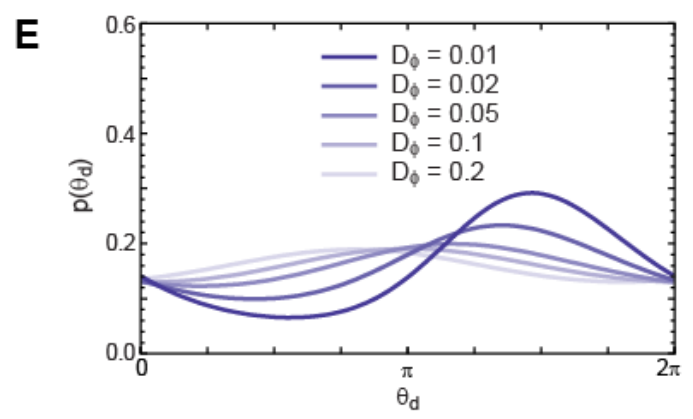
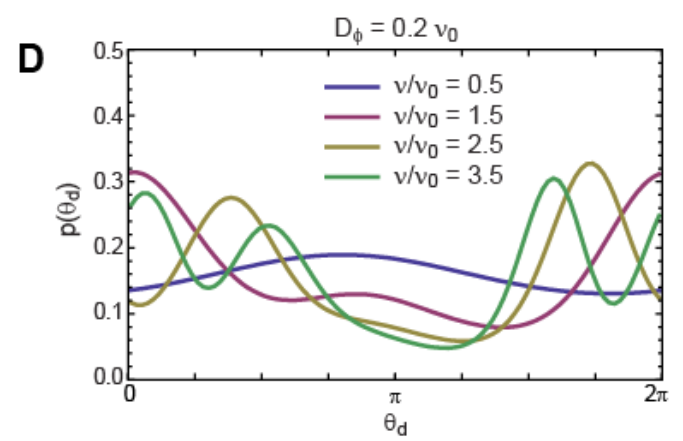
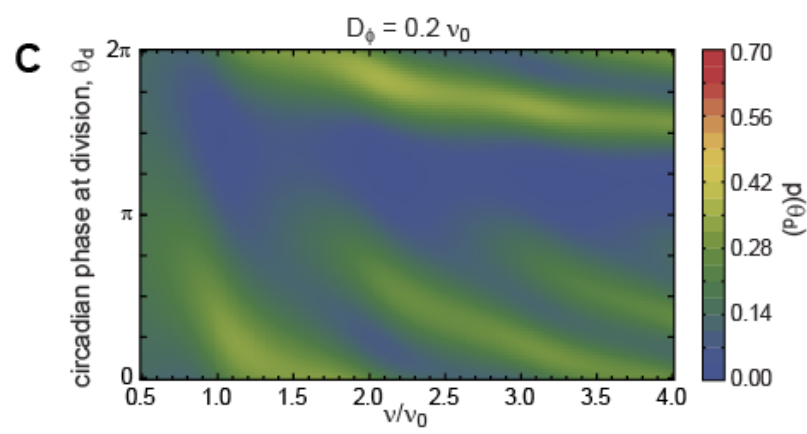
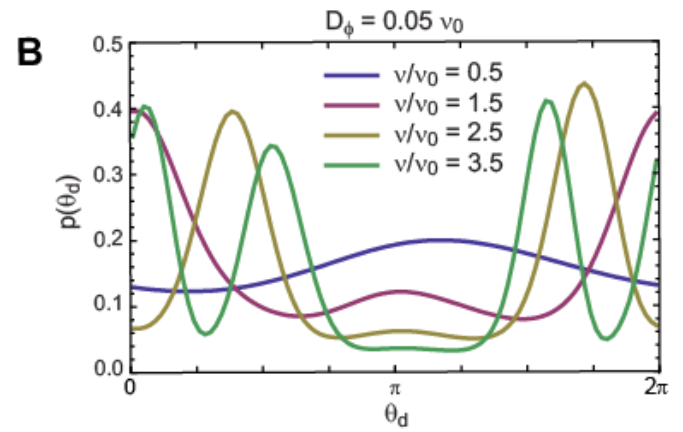
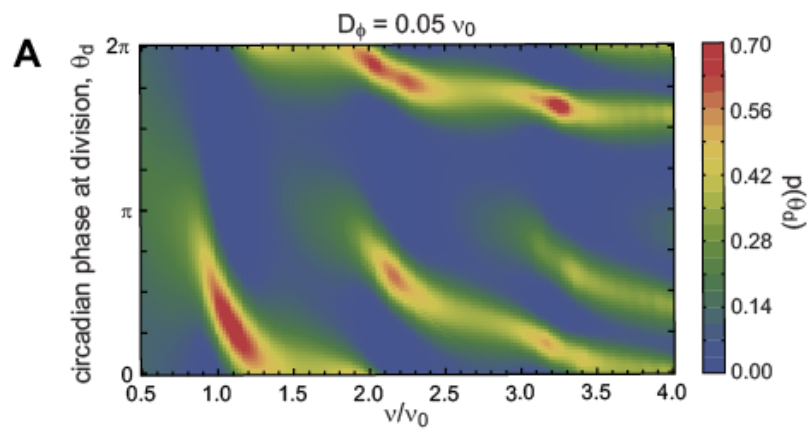


Figure S3

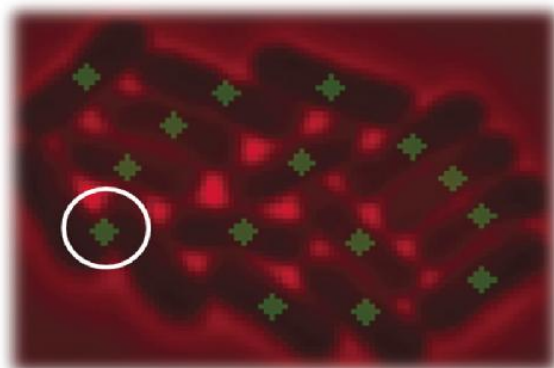
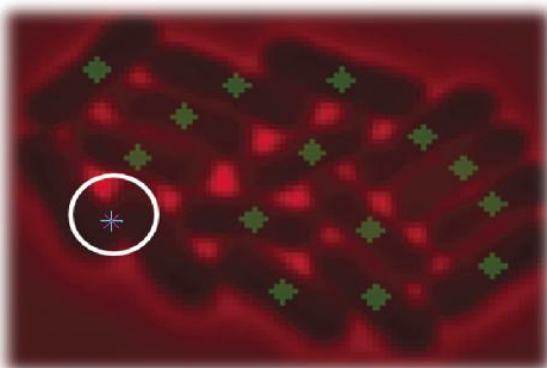
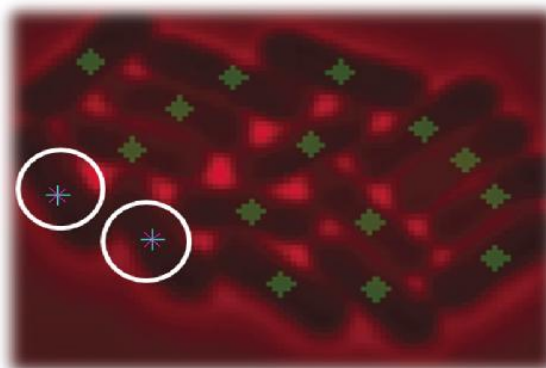
**A****B****C**

Figure S4

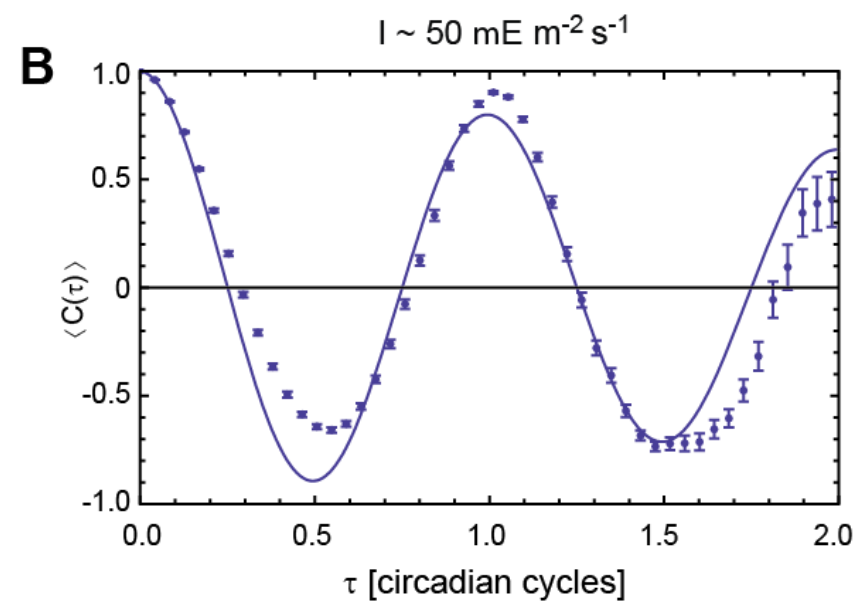
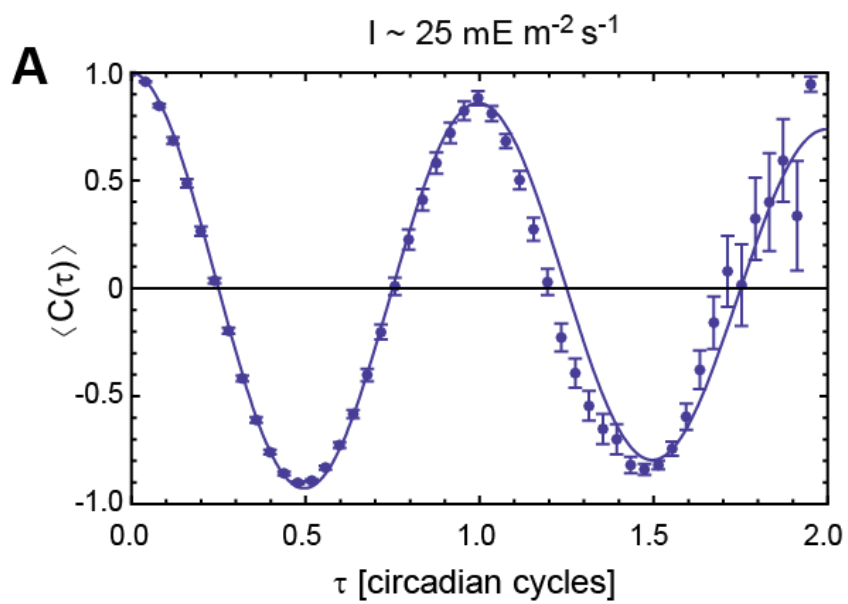


Figure S5

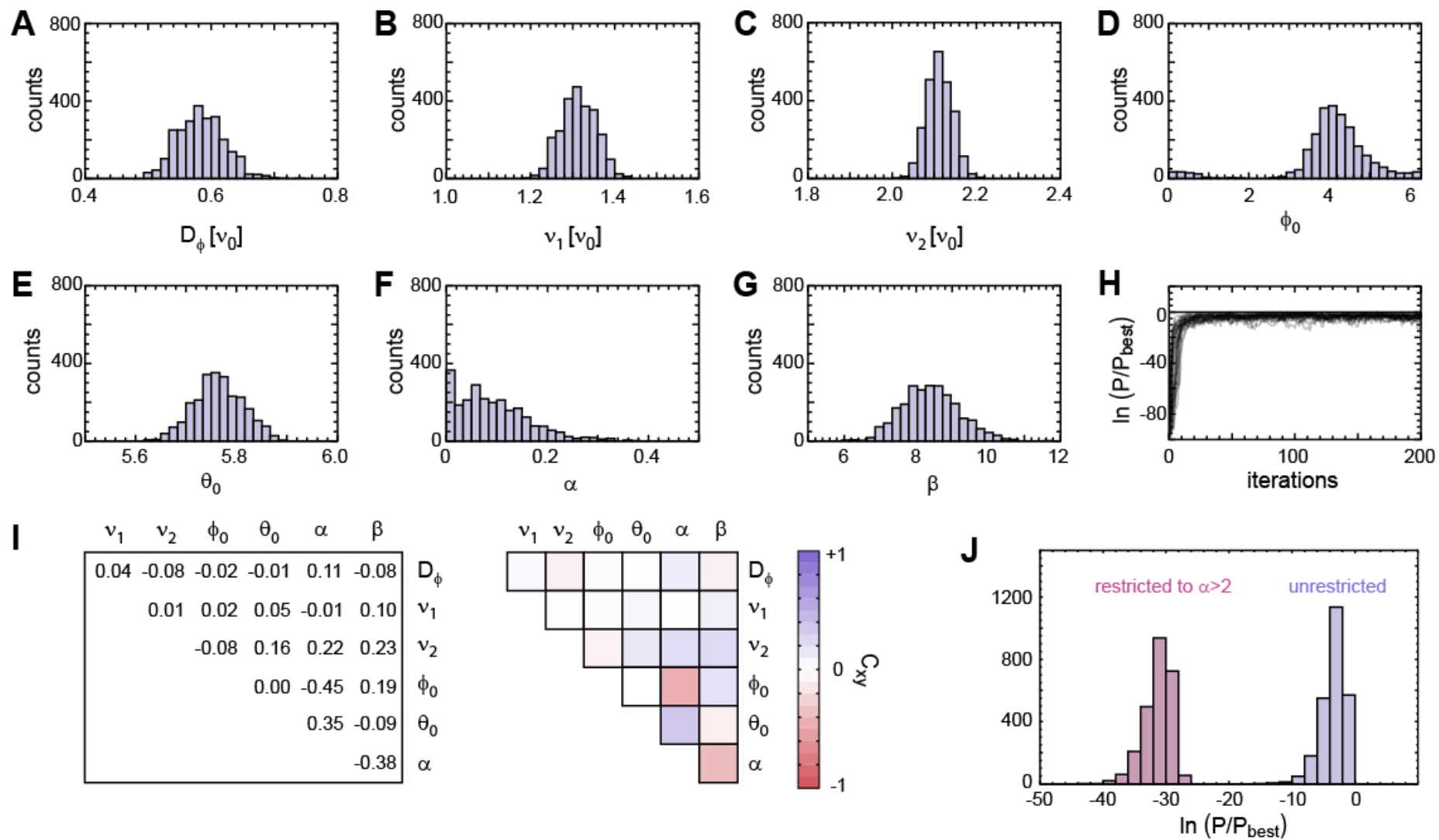


Figure S6

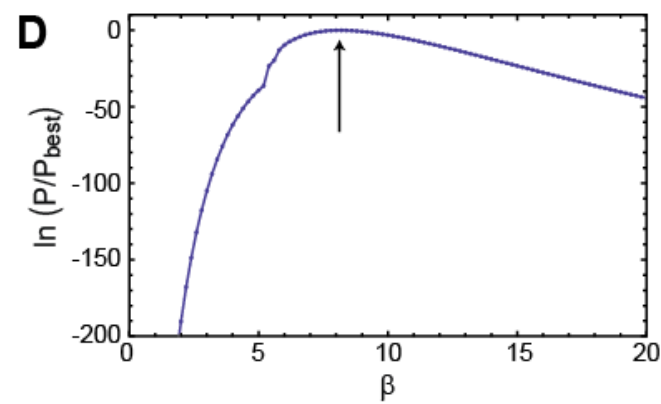
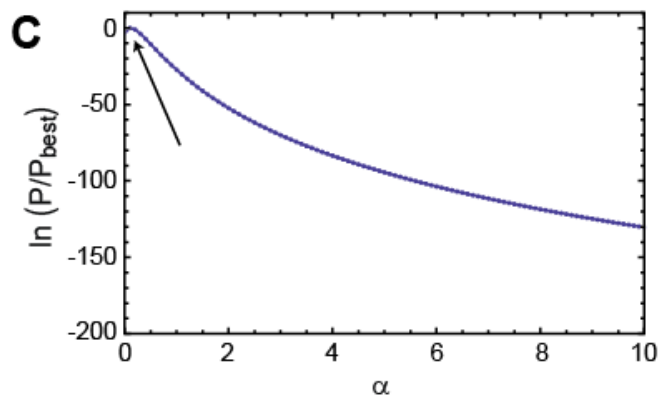
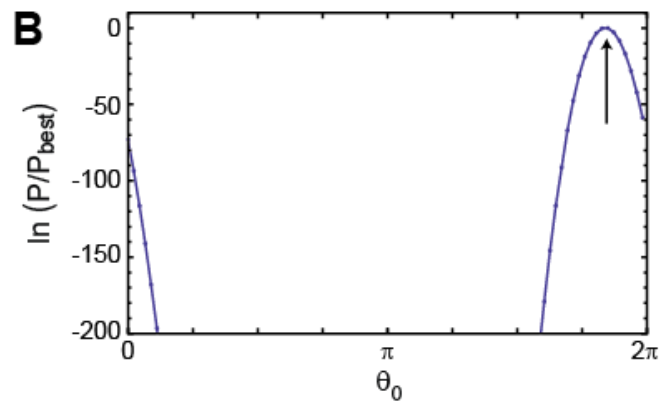
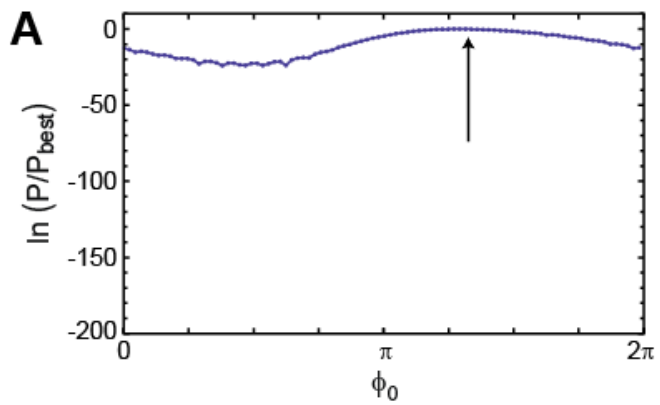


Figure S7

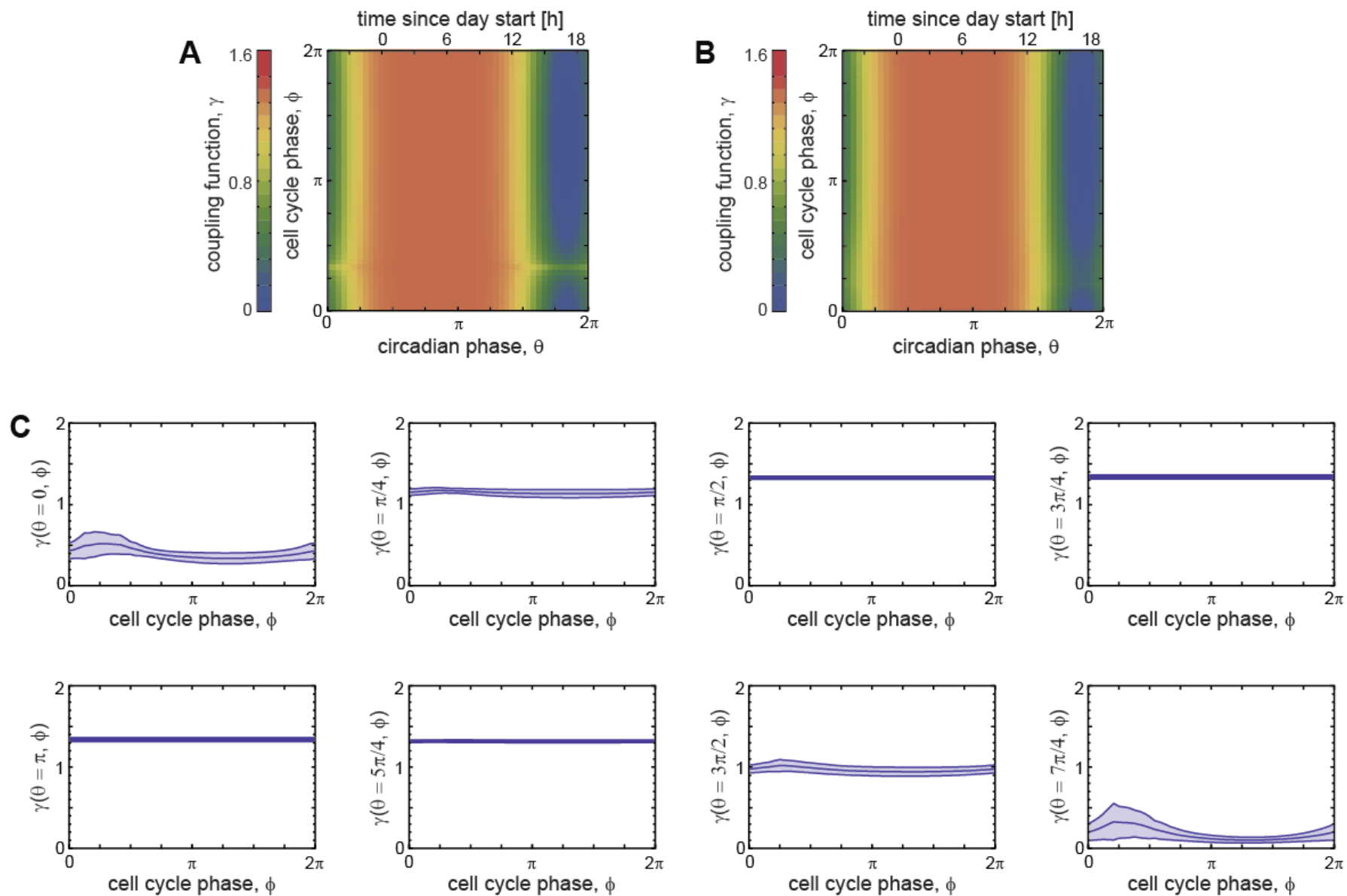


Figure S8

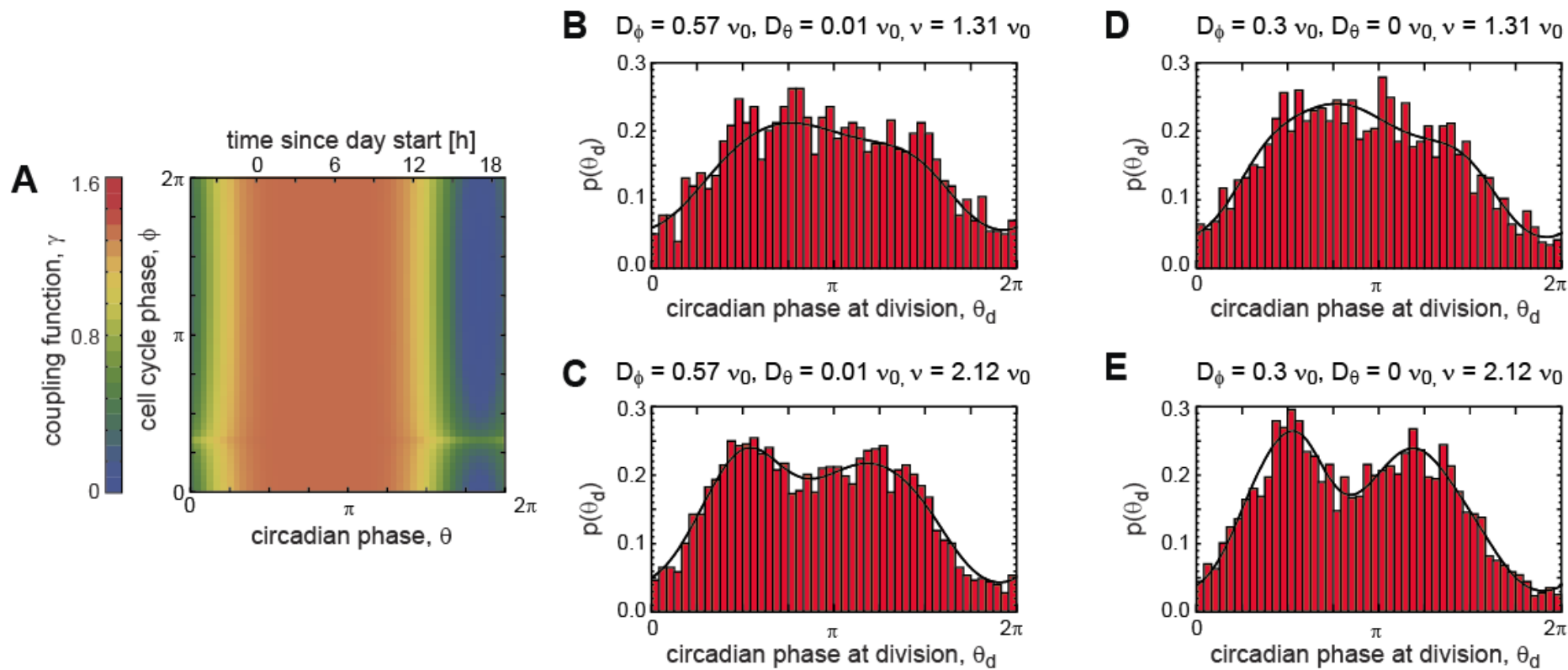


Figure S9

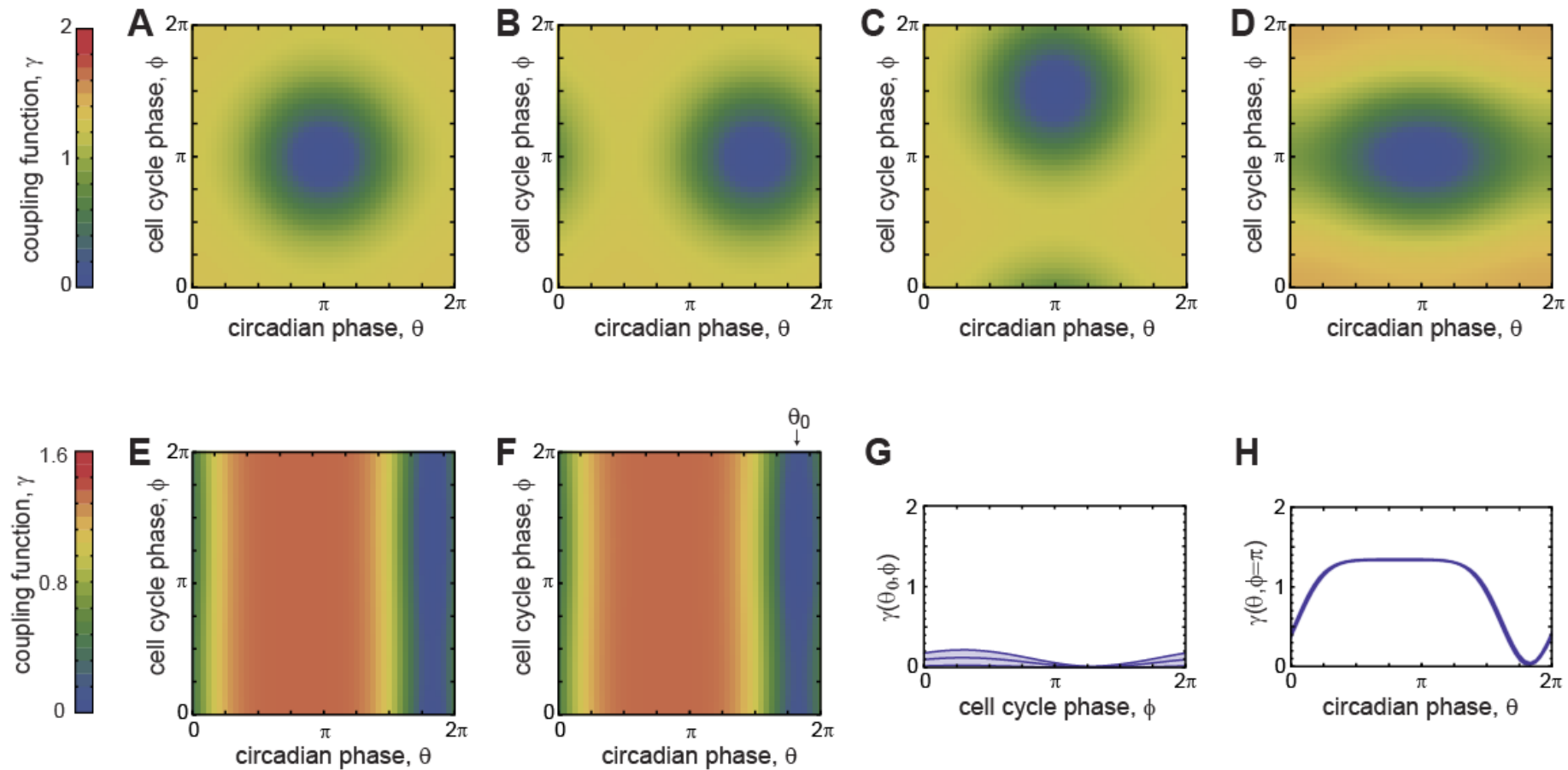


Figure S10

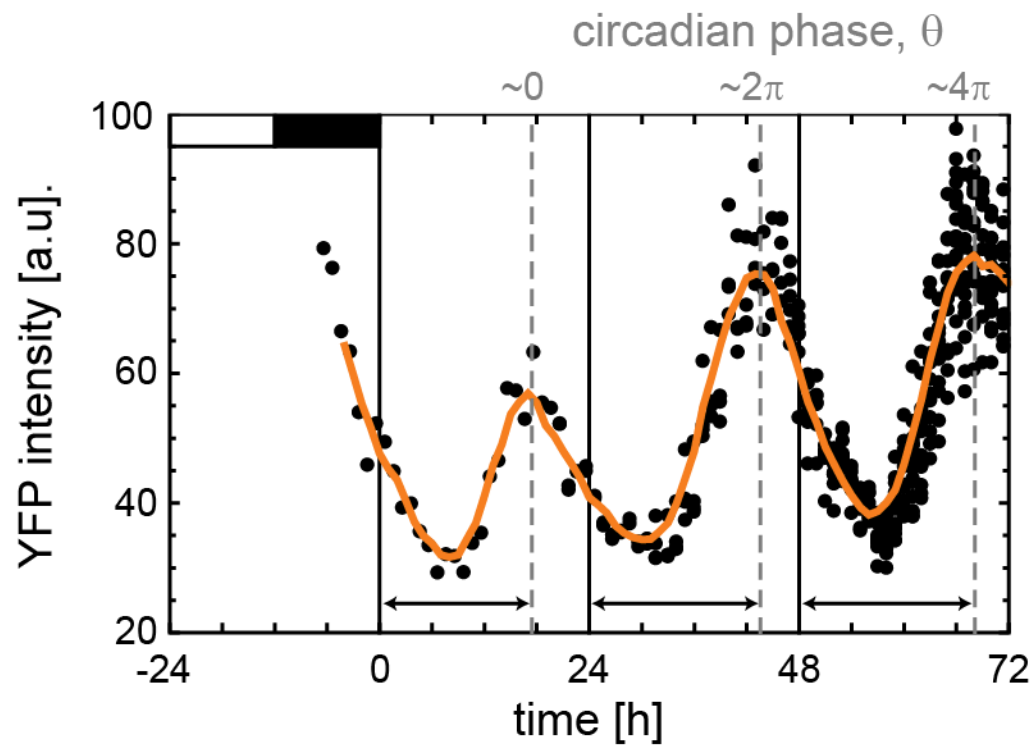


Figure S11