Supporting Online Material for

Entrainment of a Population of Synthetic Genetic Oscillators

Octavio Mondragón-Palomino, Tal Danino, Jangir Selimkhanov, Lev Tsimring, Jeff Hasty*

*To whom correspondence should be addressed. E-mail: hasty@ucsd.edu

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 to S9
References

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/333/6047/1315/DC1)

Movies S1 to S6

Supporting Online Material for

Entrainment of a Population of Synthetic Genetic Oscillators

Octavio Mondragón-Palomino, Tal Danino, Jangir Selimkhanov,

Lev Tsimring, Jeff Hasty.

Correspondence to: hasty@ucsd.edu

Contents

Materials and Methods.

Oscillator strain.

The dual-feedback oscillator is described in full detail in (*1*). Briefly, a cycle of the oscillator has two stages, dominated by either protein production or enzymatic degradation. The first stage is a positive feedback driven burst of mRNA from the clock genes *lacI* and *araC*, which goes on until enough LacI tetramer repressor accumulates and transcription is turned off, setting the conditions for the second phase where the activator and repressor proteins are enzymatically degraded. Protein degradation brings the system back to a state where a new transcription burst can occur. The time necessary to complete a cycle depends on the magnitude of the activation burst, which is roughly proportional to the induction level of the promoter. Therefore, the natural period of oscillations T_n can be externally tuned through the concentrations of inducers Arabinose and IPTG.

For this work we used the double feedback oscillator strain JS011 (*1*). In this strain the genetic oscillator is contained in two plasmids. The activator and reporter modules are on a derivative of pZE24, a medium-copy ColE1 plasmid, and the repressor module is on a derivative of pZA14, a lower-copy p15A plasmid. These plasmids are transformed into $\Delta araC\Delta lacI$ *E. coli* to obtain JS011 cells.

Microfluidics and microscopy

In the context of synthetic and cell biology, microfluidic devices have been used to obtain abundant dynamic data by lengthening the *in chip* life of microbes (*2–6*). In these designs cells are constrained to microscopic cavities of variable geometry where they are continuously supplied with nutrients. In the case of bacterium *E. coli*, our group and others realized that the mechanical interactions among bacteria and between bacteria and the walls of the micro-cavities leads

to cell spatial ordering (*7, 8*) (Movie S1). Moreover, we previously developed a microfluidic switch that allows the generation of chemical signals of any shape (*5*).

We built on this work by developing a microfluidic device where bacterial colonies can grow for long periods of time in a dynamic and precisely controlled environment (Fig. S1). The device consists of two parts: a trapping region and a dynamic switch. In the trapping region there are 48 rectangular cell chambers distributed in 4 columns. A cell chamber is a rectangular cavity with dimensions 40x50x0.95 μm^3, with the long sides open to the flow of medium. Since *E. coli* cells have a $\sim$1μm diameter, a trap allows focused observation of the $\sim$400 confined cells. The striped arrangement of cells inside small traps facilitates the exchange of nutrients, inducers and waste between the colony and the surrounding flow. The design also allows continuous discharge of cells into the flow, which carries them outside the device (Movie S1). In the dynamic switch, media flowing in from ports I and B (for Inducer and Background) are mixed to generate the modulated signal of inducers. During experiments, the concentrations of small molecules in medium I were $[IPTG] = 2$ mM and $[Ara] = 0.3+A$ %, and in the background medium B $[IPTG] = 2$ mM and $[Ara] = 0.3-A$ %. Therefore, by changing the hydrostatic pressure at ports I and B sinusoidally we changed the proportion of media I and B in the total flow. Media mixed by diffusion in the long and narrow channel leading from the switch to the trapping region, obtaining the concentration profiles of the signal $[Ara](t) = 0.3 + A\sin(2\pi t/T_f)$ % and $[IPTG] = 2$ mM.

In preparation for experiments JS011 cells were cultured overnight in LB medium with antibiotics (kanamycin at 100 μg/mL, ampicillin at 100μg/mL). The next day cells are passed into 200 mL of fresh medium in a 1000-fold dilution with antibiotics and average concentrations of inducers ($[IPTG] = 2$ mM, $[Ara] = 0.3$ %). Cells were grown until they reached an OD$_{600}$ of 0.2-0.25. This culture was centrifuged and cells were resuspended in 1.5 mL of fresh medium supplemented with 0.075 % Tween 20 (Sigma-Aldrich), antibiotics and average concentrations

of inducers. Cells were introduced to the device from the cell port (C). The hydrostatic pressure at ports W1 and W2 was set so that most cells flowed through the trapping region into W2. No cells were allowed to flow into the dynamic switch area to avoid contamination of the media. Cells were loaded into the traps by gentle flicking of the fluidic line that delivered cell suspension from the reservoir to the chip. To supply loaded cells with media, the hydrostatic pressure at ports C, W1 and W2 was adjusted so that port C became the recipient of the trapping region flow-through and of water flowing from port W1. Water diluted media coming into port C, delaying the clogging of the port by cells. Finally, in order to guarantee oscillators were out of phase before delivering the modulated arabinose signal, cells were allowed to grow inside the traps for 2-3 hours in a constant environment with average inducer concentrations.

In each experiment, the microfluidic device was mounted to the stage of the microscope and primed using a solution of 0.1 % Tween 20 surfactant in water to reduce surface tension and improve wetting. Priming of the device was also necessary to remove all air from the microfluidic channels. This prevented bubble formation when fluid reservoirs were connected to the device inlets and cells were loaded. Culture temperature was maintained at 37° C using a plexiglas incubation chamber. Images were collected at 100x magnification at intervals of 30 s for phase contrast and 3 min for GFP and MCherry fluorescence. Exposure times were chosen to prevent photobleaching and phototoxicity of cells (100 ms phase contrast, 500 ms GFP, 80 ms MCherry).

Data analysis

Ordered bacterial monolayers facilitated the automated tracking of cells and lineage reconstruction. Starting with the original sequence of brightfield images, we cropped each image such that only the cells inside the chamber were visible. We corrected for the positional drift that tends to occur in longer experiments by tracking the location of the frame in the original

images. Any out-of-focus images were replaced with a copy of the preceding brightfield image (happens rarely). Potential cell objects were extracted from the cropped images by automatic segmentation by a modified Otsu filter, user preset thresholding, and morphological operations completed in MATLAB (2009b, The MathWorks, Natick, MA). The Otsu threshold filter was used to obtain the background image containing all of the cell objects. Next, morphological opening, erosion, and dilation were applied to the background image to obtain cell object seeds. Using the background and object seed images, CellProfiler(TM) subfunction was utilized as a propagation algorithm to extract individual cell objects. Regionprops command was then employed to collect area, centroid location, and pixel information about each cell object. The goal of cell tracking was to track a cell object from one time lapse image to the next.

Using MATLAB, this was accomplished through a three step process. First, under the assumption that cells did not shift considerably between consecutive time lapse images, two successive images were overlaid and cell objects were matched up based on pixel overlap and distance between centroids. Matches were satisfied based on a one-to-one correspondence. In case of cell division, we were able to match only one of the two daughter cell objects to the mother cell object. In the second step of the cell tracking procedure, division events were identified based on significant reduction in the size of an object from one image to the next. The second daughter cell object, which did not have a matching cell object after the first tracking step, was determined based on pixel overlap and distance from the mother cell object. In the final tracking step, all of the matched objects from the two successive images were removed and any still unmatched objects were reassessed based on pixel overlap and centroid distance. Single cell fluorescence was calculated by averaging the corresponding object pixel values in fluorescence images. We applied Savitzky-Golay smoothing filter to the fluorescence trajectories before extracting the periods. We defined the period length as the time distance between two successive peaks. Using image dilation we identified individual peaks and the respective

troughs for each peak. Amplitudes of each peak were taken as the fluorescence difference between the troughs and the peak. The peaks whose amplitudes fell below a predetermined cutoff value, were removed.

Each peak in single cell trajectories marks the increase of the phase by 2π (or one cycle). Using linear interpolation between peak times, we obtained phase trajectories for every cell. Using a similar approach we calculated the phase trajectory of the input signal. By taking the difference between the signal and cell phase trajectories, we obtained the phase difference between the input signal and each cell trajectory as a function of time. Single cell trajectories that did not start simultaneously with the signal where shifted to the origin of the time coordinate.

The single cell data obtained from automated tracking and lineage reconstruction is available online at http://biodynamics.ucsd.edu/downloads. Each folder contains three types of MATLAB mat files. The file **trajectories_∗Ara_∗Min_xy∗.mat** contains fluorescent trajectory data (fluorescence, area, position) with rows representing cells and columns representing time points (3min spacing). The name of the file specifies strength of the driving signal (A % (w/v)), period of the driving signal (T_f, 0Min refers to no free running data), and the number of the colony (up to 4). The file **MaxIntensityArray_0.15.mat** contains a matrix for the cutoff for the maximum fluorescence value with rows representing signals with different entrainment period (0min not included), and columns representing trap number. The file **SignalArray_0.15.mat** contains a matrix with columns representing time series of arabinose signals (0min not included) and rows representing time points (3min spacing).

Quantitative analysis

Entrainment regions

Any self-sustained oscillator can be characterized by three main parameters; the amplitude (A), natural frequency ($\omega_n = 2\pi/T_n$), and the phase (ϕ). In an autonomous regime, the phase grows linearly with time as $\phi_n = \omega_n t$. However, when an oscillator is driven by a periodic signal with frequency ω_f (or period $T_f = 2\pi/\omega_f$), and phase $\phi_f = \omega_f t$, its phase ϕ is perturbed according to the phase equation $d\Delta\phi/dt = \omega_n - \omega_f + \epsilon Q(\Delta\phi)$ (*9*). In this equation, $\Delta\phi = \phi - \phi_f$ is the difference between the phase of the oscillator and the phase of the driving force, and the term $\epsilon Q(\Delta\phi)$ represents the effect of external perturbation with amplitude ϵ. Entrainment occurs if $\epsilon Q_{min} < \omega_f - \omega_n < \epsilon Q_{max}$, where Q_{min} and Q_{max} are the extremes of the 2π-periodic function $Q(\Delta\phi)$ (*9*). This condition defines a wedge in the plane (T_f, ϵ) near $T_f/T_n = 1$ which is known as the main Arnold tongue. When the oscillator is nonlinear, or the driving signal is non-sinusoidal, the entrainment may also occur near other rational values of T/T_f, however the corresponding higher-order Arnold tongues are much narrower. If the natural frequency distribution occupies the interval $[\omega_n - \Delta\omega, \omega_n + \Delta\omega]$, the entrainment interval will broaden to $\epsilon Q_{min} - \Delta\omega < \omega_f - \omega_n < \epsilon Q_{max} + \Delta\omega$.

Definition of the entrainment index ρ

The entrainment index is defined as $\rho = 1 - S/S_{max}$, where $S = -\sum_{k=1}^{N} P_k \ln P_k$ is the entropy of the discrete period distribution, P_k is the normalized occupancy of the kth bin, N is the total number of bins, and $S_{max} = \ln N$ (*10*).

Computational modeling

Deterministic and stochastic computational models constructed in (*1*) were modified to include the periodic modulation of arabinose. The underlying set of biochemical reactions can be found in the supplementary information of (*1*). For all computational simulations the nominal values of kinetic constants where the same and as specified here: $b_a = b_r = 0.36$ min^{-1}, $\alpha = 35$, $k_{-a} = k_{-r} = 1.8$ min^{-1}, $t_a = 85$ min^{-1}, $t_r = 90$ min^{-1}, $d_a = d_r = 0.54$ min^{-1}, $k_{fa} = k_{fr} = 0.9$ min^{-1}, $k_{da} = k_{dr} = k_t = 0.018$ min^{-1} molecules^{-1}, $k_{-da} = k_{-dr} = k_{-t} = 0.00018$ min^{-1}, $k_l = 0.36$ min^{-1}, $k_{ul} = 0.18$ min^{-1}, $\gamma = 1080$ molecules/min, $c_e = 0.1$ molecules, $\lambda = 2.647$, $\epsilon = 0.2$, $Na = 50$, $Nr = 25$, $C_r^{max} = 0.2$ molecules^{-1}, $C_r^{min} = 0.01$ molecules^{-1}, $k_{r1} = 0.035$ mM, $b_1 = 2$, $C_a^{max} = 1$ molecules^{-1}, $C_a^{min} = 0$ molecules^{-1}, $k_{a1} = 3.5\%$, $k_{r2} = 1.8$ mM, $c1 = 2, b2 = 2$. [IPTG] is the IPTG concentration in mM, [ara] is the concentration of arabinose in % w/v.

We calculated the boundaries of the Arnold tongues numerically using the deterministic model of the oscillator. At every value of the driving amplitude we computed the time series of the concentration of AraC $a(t)$ and extracted the values $a_n = a(nT_f)$ separated by the driving period. They form iterations of the *return map* $a_{n+1} = F(a_n)$ associated with the continuous-time deterministic model. Within the main Arnold tongue, these iterates converge to a fixed point defined by condition $a_* = F(a_*)$. The 2:1 resonance tongue corresponds to a fixed point of the second iterate of the return map, $a_0 = F(F(a_0))$. The boundaries of the main Arnold tongue for a given value of the driving amplitude are manifested as bifurcation points at which the fixed point of the corresponding return map becomes unstable. Numerically, we calculated the coefficient of variation of the sequence of 100 iterations a_n (for the main tongue) and a_{2n} (for the 2:1 resonance) after the transients died out, and identified the bifurcation points by the values of the driving period at which the CV crossed the nominal threshold value 10^{-3}. In principle, there is a countable set of other higher-order Arnold tongues corresponding to

other $n : m$ resonances, however they are very narrow and difficult to identify, especially in experiment. Outside the Arnold tongues the phase of the oscillator is not locked to the phase of the driving signal. The continuous phase drift is the indication of the quasiperiodic behavior (9). Close to the boundaries of Arnold tongues, the phase drift is non-uniform in time: the intervals of almost constant phase difference are interrupted by periodic phase slips.

We used the stochastic model to simulate the forcing of the double-feedback oscillator when its components are subject to intrinsic noise only. At every value of the driving amplitude and period we computed time series of the number of AraC and LacI molecules. We smoothed them using a Savitsky-Golay filter and used a threshold value to distinguish true oscillation peaks from noise. From time series we obtained the period T and the relative phase $\Delta\phi$, which are defined in the same way as for the experimental fluorescence time series (see Fig. 1C). The resulting probability distributions of the period and relative-phase can be compared to our experimental results (Figs. S3-S6).

The deterministic model was also used to simulate the forced oscillations of individual oscillators in a colony of 550 cells where extrinsic noise is the only source of variability. To model this we assumed the copy-number of the activator and repressor plasmids (Na, Nr), the transcription and translation rates of AraC and LacI (b_a, t_a, b_r, t_r) and the rate of degradation (γ) of the ssrA tagged proteins by the ClpXP protease are distributed normally around their nominal values. The degradation of tagged proteins includes the monomer and oligomer forms of AraC and LacI in any folding state and in their free or DNA bound states. The coefficient of variation (CV$= \sigma/\mu$) of the distributions is set to CV$= 0.15$, which was chosen to best fit our experimental results and is close to the experimental probability distribution of the free-running period CV$= 0.18$.

At every value of the driving amplitude and period we computed time series of the concentration of AraC and LacI molecules. From time series we obtained the period T and the relative

9

phase $\Delta\phi$, which are defined in the same way as for the experimental fluorescence time series (see Fig. 1C). The resulting probability distributions of the period and relative-phase can be compared to our experimental results (Figs. S7-S10).

References and Notes

1. J. Stricker, *et al.*, *Nature* **456**, 516 (2008).

2. M. Bennett, J. Hasty, *Nature Reviews Genetics* **10**, 628 (2009).

3. A. Groisman, *et al.*, *Nature Methods* **2**, 685 (2005).

4. F. Balagaddé, L. You, C. Hansen, F. Arnold, S. Quake, *Science* **309**, 137 (2005).

5. M. Bennett, *et al.*, *Nature* **454**, 1119 (2008).

6. T. Danino, O. Mondragón-Palomino, L. Tsimring, J. Hasty, *Nature* **463**, 326 (2010).

7. H. Cho, *et al.*, *PLoS Biology* **5**, e302 (2007).

8. D. Volfson, S. Cookson, J. Hasty, L. Tsimring, *Proceedings of the National Academy of Sciences* **105**, 15346 (2008).

9. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A universal concept in nonlinear sciences* (Cambridge Univ Pr, 2003).

10. P. Tass, *et al.*, *Physical Review Letters* **81**, 3291 (1998).
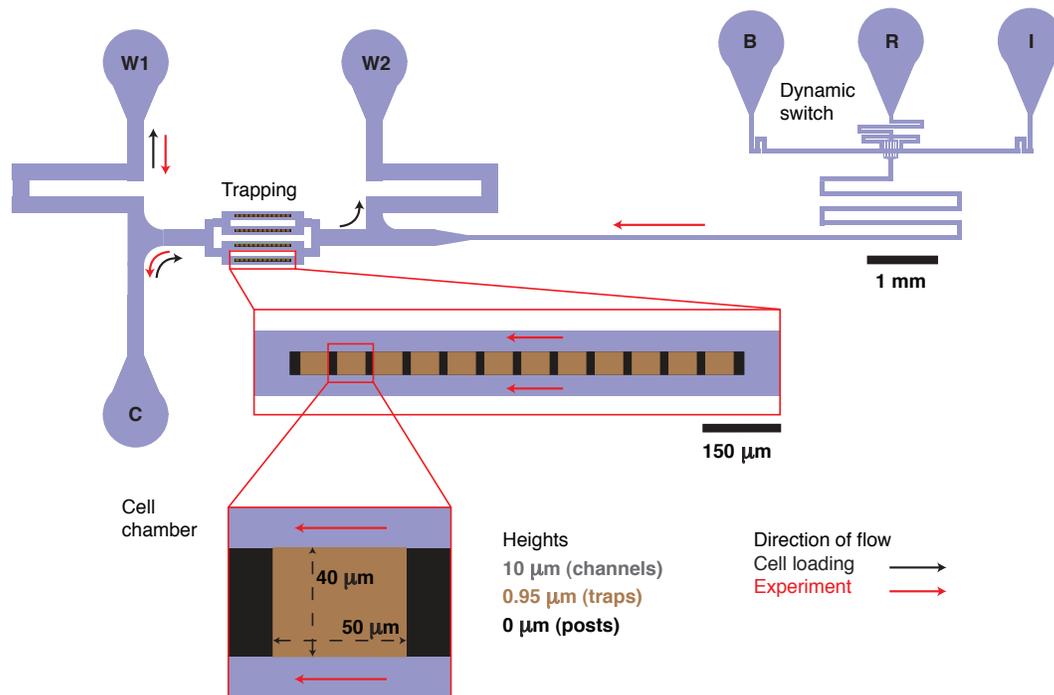
Fig. S1: Microfluidic device constructed for this study. The device consists of two parts: the signal generator (dynamic switch) and the trapping region, where 48 cell chambers host the same number of monolayer bacterial colonies.
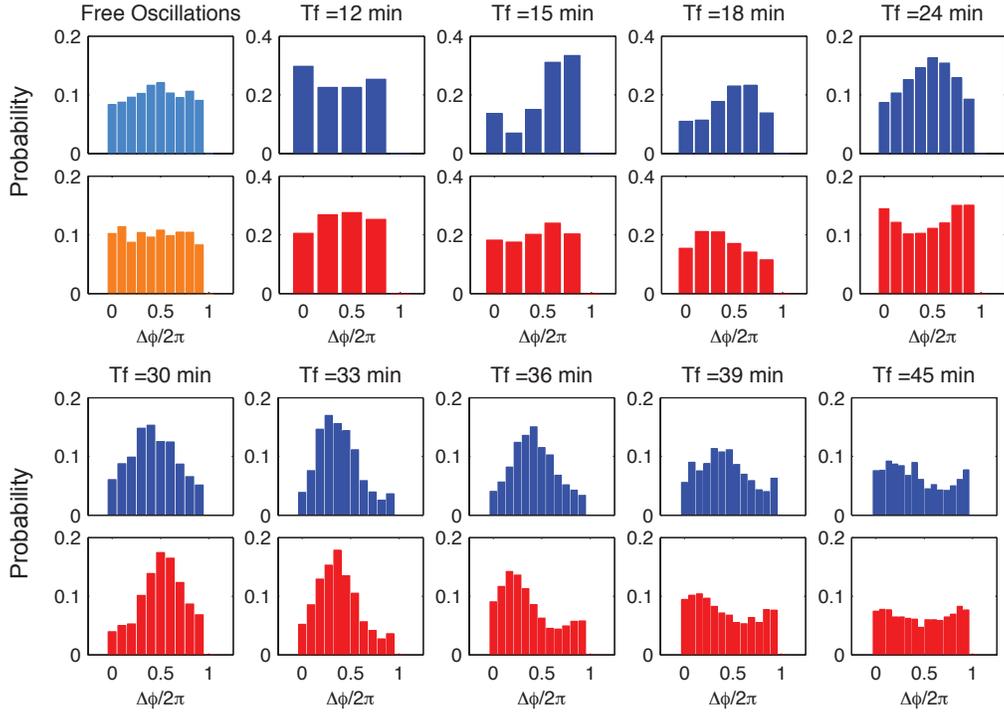
Fig. S2: Comparison of the experimental (blue) and stochastic model (red) probability distributions of the normalized relative phase $\Delta\phi/2\pi$ for a forcing amplitude $A =$0.075 %.
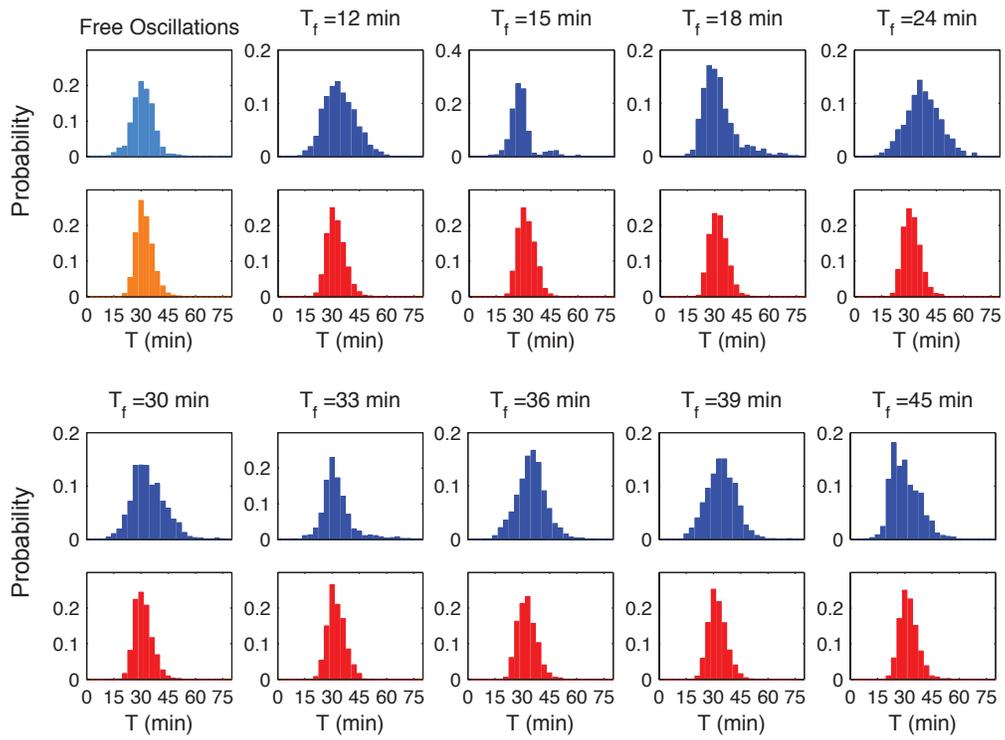
Fig. S3: Comparison of the experimental (blue) and stochastic model (red) probability distributions of the period T for a forcing amplitude $A = 0.075$ %.
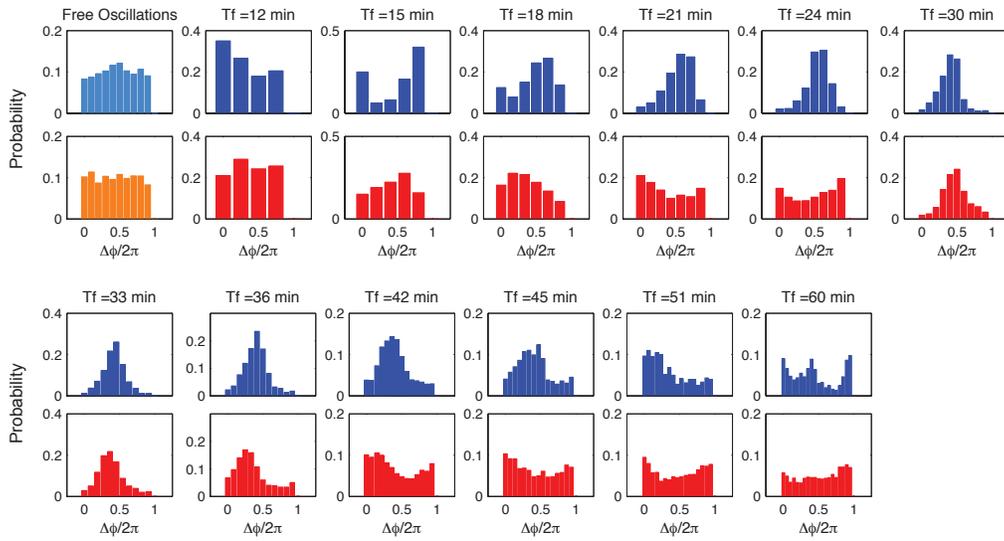
Fig. S4: Comparison of the experimental (blue) and stochastic model (red) probability distributions of the normalized relative phase $\Delta\phi/2\pi$ for a forcing amplitude $A = 0.15$ %.
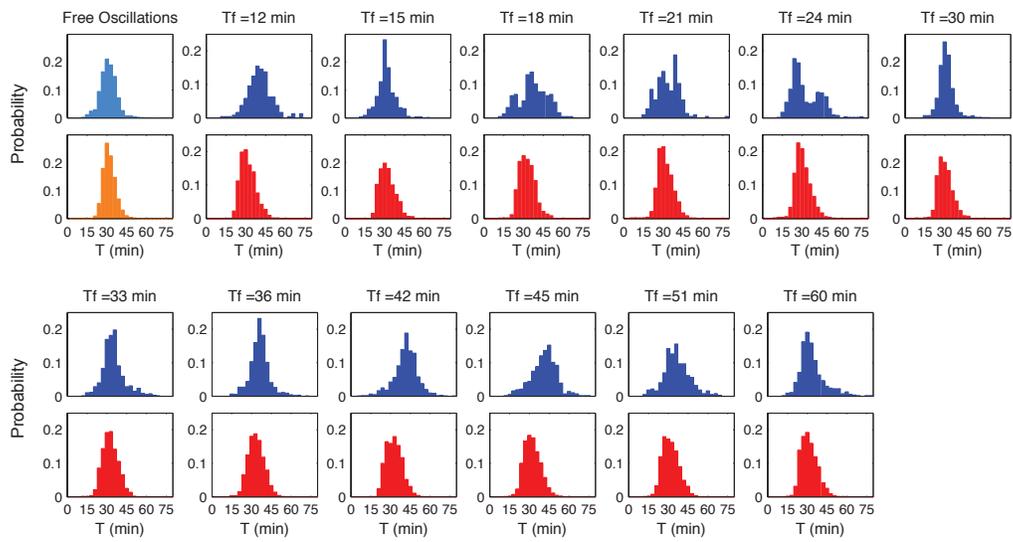
Fig. S5: Comparison of the experimental (blue) and stochastic model (red) probability distributions of the period T for a forcing amplitude $A = 0.15$ %.
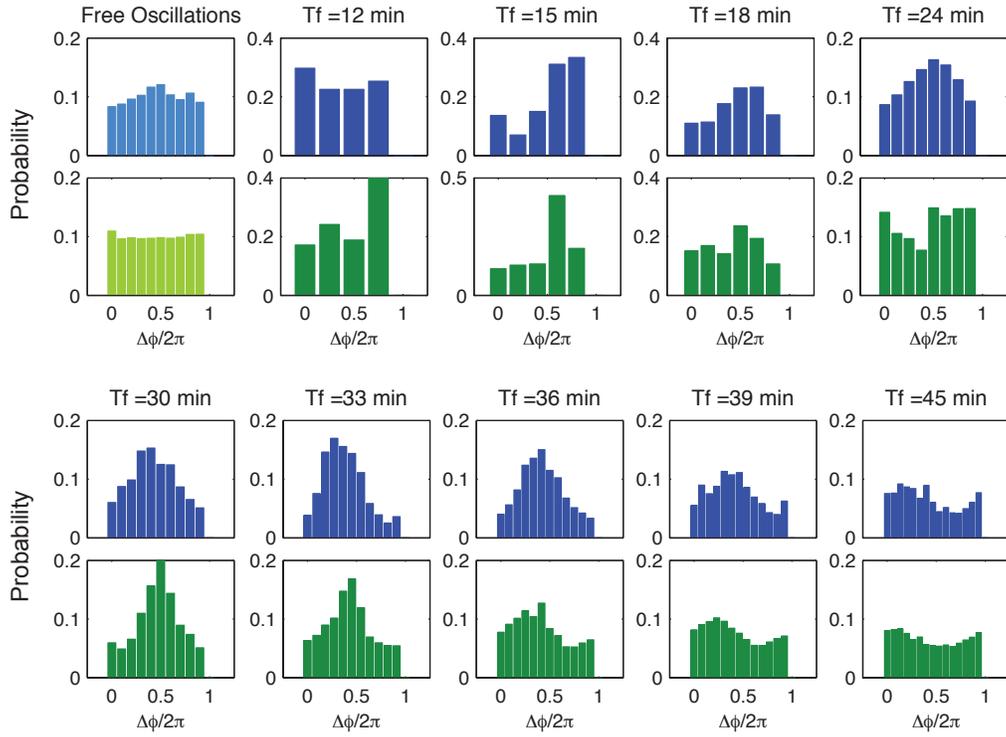
Fig. S6: Comparison of the experimental (blue) and deterministic (green) probability distributions of the normalized relative phase $\Delta\phi/2\pi$ for a forcing amplitude $A = 0.075$ %.
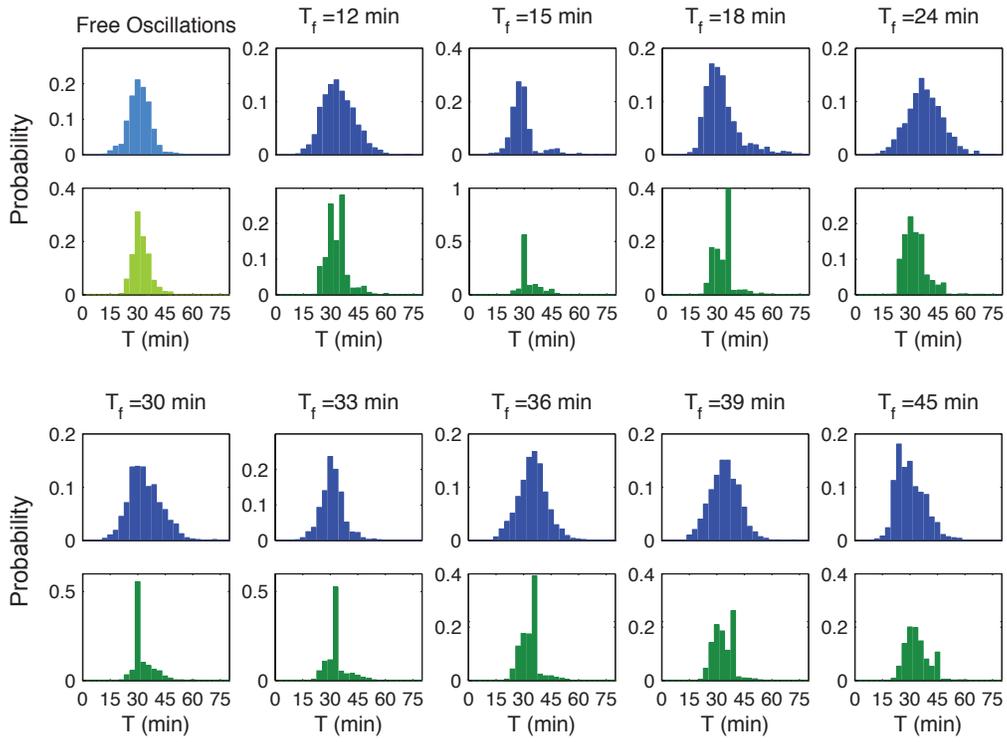
Fig. S7: Comparison of the experimental (blue) and deterministic (green) probability distributions of the period T for a forcing amplitude $A =0.075$ %.
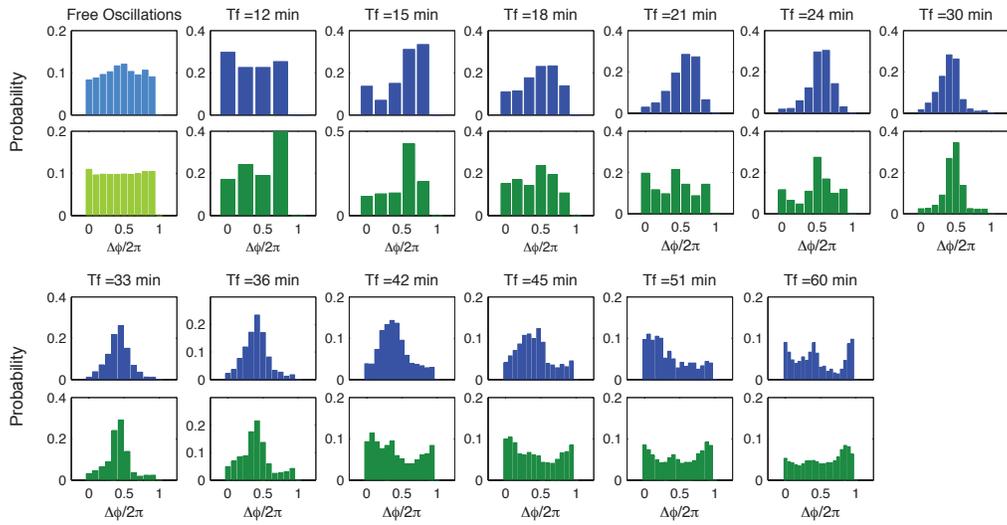
Fig. S8: Comparison of the experimental (blue) and deterministic (green) probability distributions of the normalized relative phase $\Delta\phi/2\pi$ for a forcing amplitude $A = 0.15$ %.
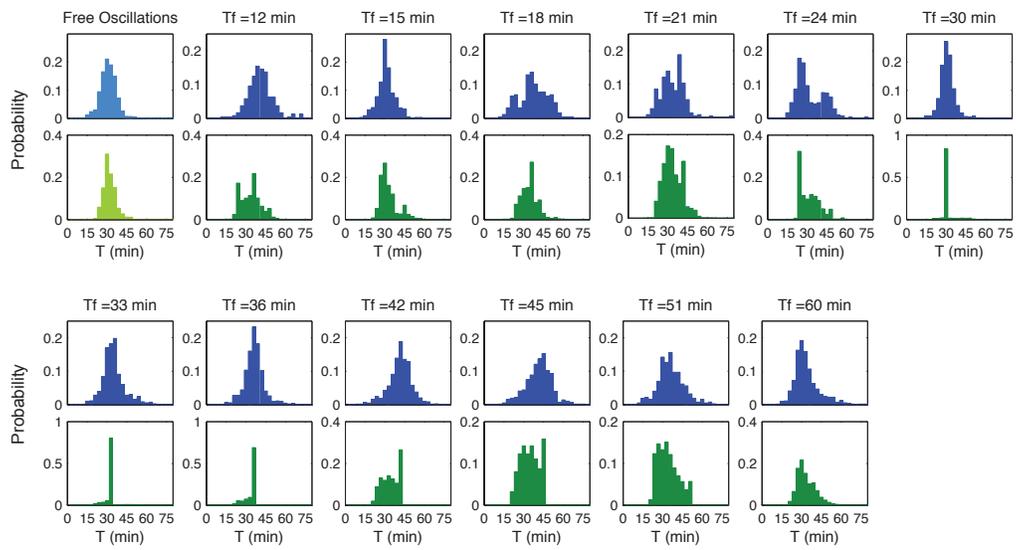
Fig. S9: Comparison of the experimental (blue) and deterministic (green) probability distributions of the period T for a forcing amplitude $A = 0.15$ %.

Movie S1. Time lapse fluorescence microscopy of JS011 cells in a microfluidic chamber subject to a sinusoidally changing concentration of arabinose with average [ara] = 0.3 %, amplitude A = 0.15 % and period Tf = 30 min, and constant IPTG concentration 2 mM. The GFP and MCherry fluorescence are shown in green and red respectively, the phase contrast image is shown in grey. Total time of the movie is 300 min with a sampling rate of one image every 3 min.

Movie S2. Time lapse fluorescence microscopy of JS011 cells in a microfluidic trap at constant inducer concentrations [ara] = 0.3 % and [IPTG] = 2 mM. The GFP and MCherry fluorescence are shown in green and red respectively, the phase contrast image is shown in grey. Total time of the movie is 360 min with a sampling rate of one image every 3 min.

Movie S3. Time lapse fluorescence microscopy of JS011 cells in a microfluidic chamber subject to a sinusoidally changing concentration of arabinose with average [ara] = 0.3 %, amplitude A = 0.075 % and period Tf =30 min, and constant IPTG concentration 2 mM. The GFP and MCherry fluorescence are shown in green and red respectively, the phase contrast image is shown in grey. Total time of the movie is 372 min with a sampling rate of one image every 3 min.

Movie S4. Time lapse fluorescence microscopy of JS011 cells in a microfluidic chamber subject to a sinusoidally changing concentration of arabinose with average [ara] = 0.3 %, amplitude A = 0.075 % and period Tf = 15 min, and constant IPTG concentration 2 mM. The GFP and MCherry fluorescence are shown in green and red respectively, the phase contrast image is shown in grey. Total time of the movie is 300 min with a sampling rate of one image every 3 min.

Movie S5. Time lapse fluorescence microscopy of JS011 cells in a microfluidic chamber subject to a sinusoidally changing concentration of arabinose with average [ara] = 0.3 %, amplitude A = 0.15 % and period Tf = 15 min, and constant IPTG concentration 2 mM. The GFP and MCherry fluorescence are shown in green and red respectively, the phase contrast image is shown in grey. Total time of the movie is 300 min with a sampling rate of one image every 3 min.

Movie S6. Example of the cell tracking algorithm applied to E. coli bacteria in a microfluidic chamber described in the supporting online text. Top left panel shows the capability of our algorithm to track three individual cells using bright field images. The top right panel shows fluorescence profiles extracted from the fluorescence images. The bottom plot shows the fluorescent trajectories corresponding to the average fluorescence value for each cell.