A synthetic oscillatory network of transcriptional regulators

Michael B. Elowitz & Stanislas Leibler

Departments of Molecular Biology and Physics, Princeton University, Princeton, New Jersey 08544, USA

Networks of interacting biomolecules carry out many essential functions in living cells1, but the ‘design principles’ underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems2. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional regulator systems that are not part of any natural biological clock3,4 to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its activity in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such ‘rational network design’ may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

In the network shown in Fig. 1a, the first repressor protein, LacI from E. coli, inhibits the transcription of the second repressor gene, tetR from the tetracycline-resistance transposon Tn10, whose protein product in turn inhibits the expression of a third gene, cI from λ phage. Finally, CI inhibits lac expression, completing the cycle.

That such a negative feedback loop can lead to temporal oscillations in the concentrations of each of its components can be seen from a simple model of transcriptional regulation, which we used to design the repressilator and study its possible behaviours (Box 1). In this model, the action of the network depends on several factors, including the dependence of transcription rate on repressor concentration, the translation rate, and the decay rates of the protein and messenger RNA. Depending on the values of these parameters, at least two types of solutions are possible: the system may converge to a stable steady state, or the steady state may become unstable, leading to sustained limit-cycle oscillations (Fig. 1b, c).

We found that oscillations are favoured by strong promoters coupled to efficient ribosome-binding sites, tight transcriptional repression (low ‘leakiness’), cooperative repression characteristics, and comparable protein and mRNA decay rates (Box 1, Fig. 1b). A general obstacle to the design of biochemical networks is uncertainty about the values of parameters that characterize the interactions between different components. In our network, estimates of the order of magnitude of the relevant parameters seem to be compatible with the possibility of oscillations. Nevertheless, to increase the chances that the artificial network would function in the oscillatory regime, we made two alterations to natural components. First, to address transcriptional strength and tightness, we used strong, yet tightly repressible hybrid promoters, developed previously, which combine the λ P1 promoter with lac and tet operator sequences5. Second, to bring the effective repressor protein lifetimes closer to that of mRNA (about 2 min, on average, in E. coli), we inserted a carboxy-terminal tag, based on the sRA RNA sequence6, at the 3’ end of each repressor gene. Proteases in E. coli recognize this tag and target the attached protein for destruction7,8. Such tags have been shown to reduce the half-life of the λ repressor DNA-binding domain from more than 60 min to around 4 min (ref. 8) and diminish the half-life of green fluorescent protein (GFP) to about 30–40 min (ref. 11).

With these considerations in mind, we used standard molecular biology techniques to construct a low-copy plasmid encoding the repressilator and a compatible, higher-copy reporter plasmid containing the tet-repressible promoter P1tetO1 (ref. 6) fused to an intermediate stability variant of gfp11 (Fig. 1a). Because the inducer IPTG interferes with repression by LacI, we expected that a transient pulse of IPTG might be capable of synchronizing a population of repressor-containing cells. A culture of E. coli MC4100 containing the two plasmids and grown in media containing IPTG displayed what appeared to be a single damped oscillation of GFP fluorescence per cell after transfer to media lacking IPTG (results not shown). Because individual cells have no apparent means of maintaining synchronization, we studied the repressilator by isolating single cells under the microscope and monitoring their fluorescence intensity as they grew into small two-dimensional microcolonies consisting of hundreds of progeny cells. In these experiments, total observation time was limited by the colony entering a stationary phase after about 10 hours of growth at
letters to nature

Figure 1 Construction, design and simulation of the repressilator. a. The repressilator network. The repressilator is a cyclic negative-feedback loop composed of three repressor genes and their corresponding promoters, as shown schematically in the centre of the left-hand plasmid. It uses pLlacO1 and pLtetO1, which are strong, tightly repressible genes and their corresponding promoters, as shown schematically in the centre of the network. The repressilator is a cyclic negative-feedback loop composed of three repressor genes that are expressed in an alternating manner: repressor 1 represses itself, repressor 2 represses repressor 1, and repressor 3 represses repressor 2. b. Stability diagram for a continuous symmetric repressilator model (Box 1). The parameter space is divided into two regions in which the steady state is stable (top left) or unstable (bottom right). Curves A, B and C mark the boundaries between the two regions for different parameter values: A, \( n = 2.1 \); B, \( n = 2 \); C, \( n = 2, \alpha_2/\alpha_3 = 10^{-3} \). The unstable region (A), which includes unstable regions (B) and (C), is shaded. c. Oscillations in the levels of the three repressor proteins, as obtained by numerical integration. Left, a set of typical parameter values, marked by the ‘X’ in b, were used to solve the continuous model. Right, a similar set of parameters was used to solve a stochastic version of the model (Box 1). Colour coding is as in a. Insets show the normalized autocorrelation function of the first repressor species.

Figure 2 Repressilator in living bacteria. a, b. The growth and timecourse of GFP expression for a single cell of E. coli host strain MC4100 containing the repressilator plasmids (Fig. 1a). Snapshots of a growing microcolony were taken periodically both in fluorescence (a) and bright-field (b). c. The pictures in a and b correspond to peaks and troughs in the timecourse of GFP fluorescence density of the selected cell. Scale bar, 4 \( \mu \text{m} \). Bars at the bottom of c indicate the timing of septation events, as estimated from bright-field images.

30 °C. At least 100 individual cell lineages in each of three microcolonies were tracked manually, and their fluorescence intensity was quantified.

The timecourse of the fluorescence of one such cell is shown in Fig. 2. Temporal oscillations (in this case superimposed on an overall increase in fluorescence) occur with a period of around 150 minutes, roughly threefold longer than the typical cell-division time. The amplitude of oscillations is large compared with baseline levels of GFP fluorescence. At least 40% of cells were found to exhibit oscillatory behaviour in each of the three movies, as determined by a Fourier analysis criterion (see Methods). The range of periods, as estimated by the distribution of peak-to-peak intervals, is 160 ± 40 min (mean ± s.d., \( n = 63 \)). After septation, GFP levels in the two sibling cells often remained correlated with one another for long periods of time (Fig. 3a–c). Based on the analysis of 179 septation events in the 3 movies, we measured an average half-time for sibling decorrelation of 95 ± 10 min, which is longer than the typical cell-division times of 50–70 min under these conditions. This indicates that the state of the network is transmitted to the progeny cells, despite a strong noise component. We observed significant variations in the period and amplitude of the oscillator output both from cell to cell (Fig. 3d), and over time in a single cell and its descendants (Fig. 3a–c). In some individuals, periods were omitted or phase delayed in one cell relative to its sibling (Fig. 3a, c).

Recent theoretical work has shown that stochastic effects may be responsible for noisy operation in natural gene-expression networks12. Simulations of the repressilator that take into account the stochastic nature of reaction events and discreteness of network components also exhibit significant variability, reducing the correlation time for oscillations from infinity (in the continuous model) to about two periods (Box 1, Fig. 1c, insets). In general, we would like to distinguish such stochastic effects from possible intrinsically complex dynamics (such as intermittence or chaotic behaviour). Further studies are needed to identify and characterize the sources of fluctuations in the repressilator and other designed networks. In particular, longer experiments performed under chemostatic conditions should enable more complete statistical characterization of
Figure 3 Examples of oscillatory behaviour and of negative controls. a–c, Comparison of the repressilator dynamics exhibited by sibling cells. In each case, the fluorescence timecourse of the cell depicted in Fig. 2 is redrawn in red as a reference, and two of its siblings are shown in blue and green. a, Siblings exhibiting post-division phase delays relative to the reference cell. b, Examples where phase is approximately maintained but amplitude varies significantly after division. c, Examples of reduced period (green) and long delay (blue). d, Two other examples of oscillatory cells from data obtained in different experiments, under conditions similar to those of a–c. There is a large variability in period and amplitude of oscillations. e, f, Examples of negative control experiments. e, Cells containing the repressilator were disrupted by growth in media containing 50 μM IPTG. f, Cells containing only the reporter plasmid.

Box 1
Network design

Design of the repressilator started with a simple mathematical model of transcriptional regulation. We did not set out to describe precisely the behaviour of the system, as not enough is known about the molecular interactions inside the cell to make such a description realistic. Instead, we hoped to identify possible classes of dynamic behaviour and determine which experimental parameters should be adjusted to obtain sustained oscillations.

Deterministic, continuous approximation

Three repressor-protein concentrations, $p_i$, and their corresponding mRNA concentrations, $m_i$ (where $i$ is lacI, tetR or $cI$) were treated as continuous dynamical variables. Each of these six molecular species participates in transcription, translation and degradation reactions. Here we consider only the symmetrical case in which all three repressors are identical except for their DNA-binding specificities. The kinetics of the system are determined by six coupled first-order differential equations:

$$\frac{dm_i}{dt} = -m_i + \alpha \left( \frac{1}{1 + p_i^n} \right) + \alpha_0 \left( i = lacI, tetR, cI \right)$$

$$\frac{dp_i}{dt} = -\beta (p_i - m_i)$$

where the number of protein copies per cell produced from a given promoter type during continuous growth is $\alpha_0$ in the presence of saturating amounts of repressor (owing to the ’leakiness’ of the promoter), and $\alpha + \alpha_0$ in its absence; $\beta$ denotes the ratio of the protein decay rate to the mRNA decay rate; and $n$ is a Hill coefficient. Time is rescaled in units of the mRNA lifetime; protein concentrations are written in units of $K_{m}$, the number of repressors necessary to half-maximally repress a promoter; and mRNA concentrations are rescaled by their half-life, yielding the average number of proteins produced per mRNA molecule. The numerical solution of the model shown in Fig. 1c used the following parameter values: promoter strength, $5 \times 10^{-4}$ (repressed) to 0.5 (fully induced) transcripts per s; average translation efficiency, 20 proteins per transcript; Hill coefficient, $n = 2$; protein half-life, 10 min; mRNA half-life, 2 min; $K_{m}$, 40 monomers per cell.

This system of equations has a unique steady state, which becomes unstable when

$$\frac{(2+1)^2}{\beta \alpha} < \frac{3X^2}{4+2X}$$

where $X = -\alpha p^{-n-1}$. The solution to $\rho = \frac{1}{(1 + \rho^n)^2} + \alpha_0$, the boundary between the stable and unstable domains can therefore be plotted (Fig. 1b). The unstable domain becomes much larger when the Hill coefficient increases, removing any limitation on $\beta$ for sufficiently large $\alpha$ (compare curve B, for which $n = 2$, to curve A, for which $n = 2$). The effect of leakiness, $\alpha_0$, can be seen by plotting the stability boundary for a constant ratio of $\alpha_0/\alpha$, as in curve C. When $\alpha_0$ becomes comparable to $K_{m}$ (1 in our units), the unstable domain shrinks (compare curve B, for which $\alpha_0 = 0$, to curve C, for which $\alpha_0/\alpha = 10^{-5}$). Similar analysis of the stability of the steady state can be performed for generalized models of cyclic transcriptional feedback loops. The simplest such networks supporting limit-cycle oscillations are those containing a single repressor and a single activator, or an odd number of repressors exceeding 3. In general, the period of oscillations in such networks is determined mainly by the protein stability and more detailed calculations, with non-Hill-function repression curves, or using thermodynamic binding energies to predict equilibrium operator occupancies, and taking repressor dimerization into account, yield similar stability results19. It is possible that, in addition to simple oscillations, this and more realistic models may exhibit other complex types of dynamic behaviour.

Stochastic, discrete approximation

The preceding analysis neglects the discrete nature of the molecular components and the stochastic character of their interactions, however. Such effects are believed to be important in biochemical and genetic networks12. We therefore adapted the above equations to perform stochastic simulations, as described20. To obtain cooperativity in repression analogous to the continuous case, we assumed the presence of two operator sites on each promoter and the following reactions: binding of proteins to each operator site (1 nM$^{-1}$s$^{-1}$); unbinding of protein from the first-occupied (224 s$^{-1}$) and the second-occupied (9 s$^{-1}$) operator; transcription from occupied (5 × 10$^{-4}$ s$^{-1}$) and from unoccupied (0.5 s$^{-1}$) promoters; translation (0.167 mRNA$^{-1}$s$^{-1}$); protein decay (10 min half-life); and mRNA decay (2 min half-life). These parameters were chosen to correspond as closely as possible to the continuous model described above, assuming that 1 molecule per cell corresponds to a concentration of ~1 nM. Oscillations persist for these parameter values (Fig. 1c) but with a large variability, resulting in a finite autocorrelation time (compare insets in Fig. 1c).
the noisy repressilator dynamics. In addition, varying the host species and genetic background would allow us to check for, and minimize, spurious interactions with endogenous cellular subsystems, and to investigate how the network is embedded in the cell. For instance, in the repressilator network, the cell-division cycle does not seem to be coupled with the repressilator, as the timing of oscillations is uncorrelated with cell septation events (Fig. 2). However, entry into the stationary phase causes the repressilator to halt, indicating that the network is coupled to the global regulation of cell growth.

The levels of many cellular components vary over time in growing cells, and even strains that constitutively express GFP exhibit significant heterogeneity, so we performed several control experiments to check that the observed oscillations are indeed due to the repressilator (Fig. 3e, f). These included deliberate disruption of the network (by adding sufficient IPTG to interfere with LacI) and observation of GFP expression in the absence of the repressilator (from plasmids with different promoters and origins of replication).

Our results show that it is possible to design and construct an artificial genetic network with new functional properties from generic components that naturally occur in other contexts. Such work is analogous to the rational design of functional proteins from well-characterized motifs. Further characterization of components and alteration of network connectivity may reveal general features of this and related networks, and provide a basis for improved design and possible use in biotechnological applications. Moreover, comparing designed networks with their evolved counterparts may also help us to understand the ‘design principles’ underlying the latter. For instance, circadian clocks are found in many organisms, including cyanobacteria in which the cell-division time may be shorter than the period.

However, the reliable performance of such circadian oscillators can be contrasted with the noisy, variable behaviour of the repressilator. Instead of three repressors, it seems that circadian oscillators use both positive and negative control elements. Does this design lead to improved reliability? Recent theoretical analysis suggests that, in the presence of interactions between positive and negative control elements that lead to bistable, hysteretic behaviour, an oscillating circuit does indeed exhibit high noise-resistance. It would be interesting to see whether one could build an artificial analogue of the circadian clock, and, if so, whether such an analogue would display the noise resistance and temperature compensation of its natural counterpart.

Methods

Network construction

The three repressors (lacI, tetR and cl), the β promoter Ps and the unstable variants of gfp were all cloned by the polymerase chain reaction (PCR) and verified by sequencing. In the case of lacI, the GTG start codon was changed to ATG with the 5’ primer. 3’ PCR primers were used to add the coding sequence for the 11-amino-acid signal peptide to the PCR product.

GFP was fused to the N-terminus of the lacI gene by PCR amplification using the primers described above and the lacI300 primer. 3’ PCR primers were then used to add the coding sequence for the 11-amino-acid signal peptide to the PCR product. The resulting fusion, with a shorter half-life than the wild-type lacI, was then cloned into pTEV2 (with or without transcriptional initiation from a 11-amino-acid signal peptide).

Data acquisition and analysis

Cells of E. coli lacI strain MC4100 (unless otherwise specified) transformed with appropriate plasmids were grown in minimal media (7.6 mM [NH₄]₂SO₄, 2 mM MgSO₄, 30 μM FeSO₄, 1 mM EDTA, 60 mM potassium phosphate, pH 6.8) supplemented with 0.5% glycerol, 0.1% casamino acids and appropriate antibiotics (20 μg ml⁻¹ kanamycin or 20 μg ml⁻¹ ampicillin), when needed. Time-lapse microscopy was conducted on a Zeiss Axiovert 135TV microscope equipped with a 512 × 512-pixel cooled CCD camera (Princeton Instruments). Cultures were grown for at least 10 hours to optical density of 0.1 at 600 nm, diluted into fresh media, spotted between a coverslip and 1 ml of liquid 2% SeaPlaque low-melt agarose (FMC) in media, and sealed. The temperature of the samples was maintained at 30–32°C by using Peltier devices (Melcor). Bright-field (0.3 s) and epifluorescence (0.05–0.5 s) exposures were taken periodically (every 5 or 10 min). All light sources (standard 100 W Hg and halogen lamps) were shuttered between exposures. Images were flat-field corrected with custom software. For the synchronization experiment, overnight cultures were diluted back 1:100 in media containing 100 μM IPTG, grown to mid-log, washed several times, and diluted again and grown in 96-well plates at 30 with shaking in a Wallac Victor 2 plate reader, while fluorescence and absorbance measurements were taken every 5 min.

For analysis, cells were selected from the final frames of bright-field movies, without regard to their fluorescence signals, and tracked manually backward in time until the first frame. At each time point, the cell position was identified on the bright-field image, and fluorescence intensity data were averaged over a 28-pixel region, similar in size to the cell diameter, in the corresponding location on the fluorescence image. A fast Fourier transform was applied to the temporal fluorescence signal from each analysed cell lineage and divided by the transform of a decaying exponential with a time constant of 90 min, the mean fluorescence spectra exhibiting peaks of more than four times the background at frequencies of 0.2–0.5 per hour were classified as oscillatory (the choice of threshold alters the fraction of oscillatory cells defined by this criterion). The sibling ‘decorrelation’ half-time was defined as the time necessary for the fluorescence signals from pairs of daughter cells to reach half of its asymptotic value. Here, I(1) and I(2) are the fluorescence intensities of the two sibling cells at time, t, starting from the moment of septation (±10 min). In this analysis, only cells that were observed by the fluorescent criterion were considered.

Various negative control experiments were performed. First, 50 μM IPTG was added to the media to disrupt the functioning of the network (Fig. 3e, f). Second, we used a version of the repressilator plasmid containing a transcriptional unit (in lacI strain, MJ109). We thus varied GFP expression of the reporter plasmid by controlling TetR levels in JM109. Third, we examined cells containing only the reporter plasmid (Fig. 3f). Fourth, we measured GFP expression from reporter plasmids modified in several ways either by replacing the PtetO1 promoter with each of the two other promoters, and gfp fused with gfp-lite (the suffix ‘lite’ indicates the presence of a C-terminal ssrA tag), or by replacing the ColE1 origin of replication with the lower-copy pSC101 origin normally used on the repressor plasmid. In none of these control experiments did we observe oscillations similar to those produced by the repressilator.

Received 6 July; accepted 9 November 1999.


Acknowledgements

We thank H. Bujard, S. Freundlieb, A. Hochschild, R. Lutz and C. Sternberg for plasmids and advice on discussions; M. S. Sauer for advice and encouragement; L. Hartwell and C. Weitz for comments on the manuscript; and K. Kafatos and E. Karsenti for hospitality and support at the European Molecular Biology Laboratory (EMBL), where part of this work was done. This work was partly supported by the US National Institutes of Health and the von Humboldt Foundation.

Correspondence and requests for materials should be addressed to M.B.E. (e-mail: melowitz@princeton.edu).

© 2000 Macmillan Magazines Ltd

LETTERS TO NATURE 338 | NATURE | VOL. 403 | 20 JANUARY 2000 | www.nature.com