A Simple Negative Interaction in the Positive Transcriptional Feedback of a Single Gene Is Sufficient to Produce Reliable Oscillations

Jesús M. Miró-Bueno*, Alfonso Rodríguez-Patón*
Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid, Madrid, Spain

Abstract

Negative and positive transcriptional feedback loops are present in natural and synthetic genetic oscillators. A single gene with negative transcriptional feedback needs a time delay and sufficiently strong nonlinearity in the transmission of the feedback signal in order to produce biochemical rhythms. A single gene with only positive transcriptional feedback does not produce oscillations. Here, we demonstrate that this single-gene network in conjunction with a simple negative interaction can also easily produce rhythms. We examine a model comprised of two well-differentiated parts. The first is a positive feedback created by a protein that binds to the promoter of its own gene and activates the transcription. The second is a negative interaction in which a repressor molecule prevents this protein from binding to its promoter. A stochastic study shows that the system is robust to noise. A deterministic study identifies that the dynamics of the oscillator are mainly driven by two types of biomolecules: the protein, and the complex formed by the repressor and this protein. The main conclusion of this paper is that a simple and usual negative interaction, such as degradation, sequestration or inhibition, acting on the positive transcriptional feedback of a single gene is a sufficient condition to produce reliable oscillations. One gene is enough and the positive transcriptional feedback signal does not need to activate a second repressor gene. This means that at the genetic level an explicit negative feedback loop is not necessary. The model needs neither cooperative binding reactions nor the formation of protein multimers. Therefore, our findings could help to clarify the design principles of cellular clocks and constitute a new efficient tool for engineering synthetic genetic oscillators.

Introduction

Cellular clocks control important functions of the cell, such as circadian (24-hour) rhythms, cell cycle, metabolism and signaling. Clock operation appears to involve the coupling of two different types of oscillators. The first are oscillators based on cytoplasmic reactions, such as phosphorylation [1] and oxidation [2,3]. The second are genetic oscillators depending on gene expression regulation [4,5]. In the last decade several synthetic genetic oscillators have been implemented in the laboratory [6–12]. The first mathematical model of a genetic oscillator was developed by Goodwin for periodic enzyme production [13]. This model was the groundwork for subsequent theoretical research on genetic oscillators in living systems, such as fungi and flies [14–19]. In these models, the rhythms are generated by a gene with a negative transcriptional feedback (NTF) (Fig. 1A). This NTF needs time delay and sufficiently strong nonlinearity in the transmission of the feedback signal for preventing the steady-state stabilization of the system [20,21]. It has also been analyzed variants, involving two genes, of the model presented in the Fig. 1A [22].

Positive transcriptional feedbacks (PTFs) are also present in many celular clocks [23–25]. Models with two or more genes involving PTFs have been studied in genetic oscillators [26–34]. In these models the PTFs increase the expression of repressor genes. It has been shown how PTFs can cause robustness of cellular clocks [37,38] and could provide robust adaptation to environmental cycles [39]. Previously, it has been demonstrated that a single gene with only PTF does not produce oscillations [40]. Here we study a model with a simple condition to produce biochemical rhythms in a single gene with PTF (Fig. 1B).

We chose a circadian period for the oscillator due to its relevance in biological systems. This model is based on two common features of genetic oscillators [4,21,26,28,38]. The first is a PTF created by a protein that activates the transcription of its own gene. The second is a negative interaction in which a repressor inhibits the transcription of its own gene. These interactions are mainly driven by two types of biomolecules: the protein, and the complex formed by the repressor and this protein. The main conclusion of this paper is that a simple and usual negative interaction, such as degradation, sequestration or inhibition, acting on the positive transcriptional feedback of a single gene is a sufficient condition to produce reliable oscillations. One gene is enough and the positive transcriptional feedback signal does not need to activate a second repressor gene. This means that at the genetic level an explicit negative feedback loop is not necessary. The model needs neither cooperative binding reactions nor the formation of protein multimers. Therefore, our findings could help to clarify the design principles of cellular clocks and constitute a new efficient tool for engineering synthetic genetic oscillators.
The model is instead of over $R$ as $R$ or any other molecule that binds to and inhibits the function of $A$. Can be thought of as a protease, a protein that sequesters and can be used several times. $A$ and forms the is translated into protein ($G$).

### Methods: Deterministic and stochastic

Negative transcriptional $A$ prevents ($Fig. 3G$) than in the other molecules from binding to its promoter. The molecule $G$ inhibits function of $A$. Specifically, $R$ binds to $A$ and forms the complex $C$. In this complex, $A$ is not able to bind to its promoter. $R$ is not degraded together with $A$ and can be used several times. Therefore, $R$ can be thought of as a protease, a protein that sequesters $A$ or any other molecule that binds to and inhibits the function of $A$ as explained above. The zigzag arrows stand for degradations. A different version of the model can be formulated with the negative interaction acting over $M$ instead of over $A$.

doi:10.1371/journal.pone.0027414.g002

diagram

---

oscillator comprises fast and slow oscillation creation stages. In our model these oscillations are characterized by sawtooth waveforms. Finally, we explain how the negative interaction works through a comparison with the dynamics of the typical enzymatic reaction. We show that the rate of the negative interaction is amplified by the PTF and has a saturation point.

### Results

#### Model and simulations

The model is a simple one-gene network with two well-differentiated parts (Fig. 2). The first is a PTF created by a protein $A$, which is a transcription factor of its own gene. When this protein binds to its promoter the transcription rate increases. The second part is a negative interaction in which a repressor molecule $R$ prevents $A$ from binding to its promoter. The molecule $R$ can be thought of as a protease, as a protein that sequesters $A$, or as any other molecule that inhibits the function of $A$ as shown in Fig. 2. A different version of the model can be formulated in which the negative interaction acts on the mRNA molecules instead of on protein $A$.

Eleven biochemical reactions provide a full description of the model (see (3) in the section Methods: Biochemical reactions and rates). The system is assumed to have a uniform mixture of biomolecules. For this reason, we did not take into account diffusion processes. In this approach, the dynamics of the biochemical reactions (3) can be described by two different formalisms known as stochastic and deterministic approaches (see Methods: Deterministic and stochastic simulations for more details). These two approaches can lead to different behaviors. The stochastic dynamics of the reactions (3) were simulated using the Gillespie algorithm [46] and the deterministic dynamics using the following ordinary differential equations:

\[
\begin{align*}
dG/dt &= -k_1GA + k_{-1}G_a \\
dG_a/dt &= k_1GA - k_{-1}G_a \\
dM/dt &= k_2G + k_3G_a - k_4M \\
dA/dt &= -k_7AR + k_9C + k_9 - k_{10}R \\
dR/dt &= -k_7AR + k_9C + k_9 - k_{10}R \\
dC/dt &= k_7AR - k_8C,
\end{align*}
\]

where the variables and rates are described in the section Methods: Biochemical reactions and rates. We used standard values within the diffusion limit for the rates [18,38,47].

The stochastic approach is more realistic than the deterministic simulation because it takes into account the randomness of the chemical reactions. This randomness produces fluctuations in the number of molecules. We fitted the reaction rates to obtain circadian oscillations in the stochastic simulation. Then, we compared the results with the deterministic simulation (Fig. 3). For both simulations the time evolution of the protein ($A$), repressor ($R$), protein-repressor complex ($C$) and mRNA ($M$) are very similar. The main difference is the appearance of fluctuations in the stochastic case around the number of molecules predicted by the deterministic approach. The fluctuations are more evident in the time evolution of $M$ (Fig. 3G) than in the other biomolecules. This is because the number of $M$ molecules oscillates in a lower range than $A$, $R$ and $C$. The oscillations in $C$ are characterized by sawtooth waveforms. On the other hand, there are differences between the stochastic and deterministic time evolution of the gene. There is a single gene in the model, which can be deactivated ($G$) or activated ($G_a$). Therefore, $G + G_a = 1$ molecule. The stochastic simulation shows realistic discrete transitions between 0 and 1 molecules (Figs. 3I and 3K). By

---

**Figure 2. Model of a one-gene oscillator with PTF.** The model is composed of two well-differentiated parts. The first part is a positive feedback loop in which a gene ($G$) is transcribed into mRNA ($M$). In turn, $M$ is translated into protein ($A$). This protein is a transcription factor of its own gene and increases the transcription rate when it binds to the promoter. The positive feedback needs a second part, consisting of a negative interaction in order to obtain reliable oscillations. In this part repressor molecules ($R$) enter the system at a constant rate. $R$ inhibits the function of $A$. Specifically, $R$ binds to $A$ and forms the complex $C$. In this complex, $A$ is not able to bind to its promoter. $R$ is not degraded together with $A$ and can be used several times. Therefore, $R$ can be thought of as a protease, a protein that sequesters $A$ or any other molecule that binds to and inhibits the function of $A$ as explained above. The zigzag arrows stand for degradations. A different version of the model can be formulated with the negative interaction acting over $M$ instead of over $A$.

doi:10.1371/journal.pone.0027414.g002
The stochastic simulation shows discrete transitions between 0 and 1 molecules. By contrast, the deterministic simulation shows unrealistic continuous transitions. The time evolution of Ga shows sawtooth waveforms. There is a single gene in the model; hence the oscillations is low. This is reflected in an increase of the amplitude and period standard deviations to 2,132 molecules and 5.2 hours, respectively (Fig. S2B).

In cells, there are also fluctuations in the number (or activity) of molecules such as polymerases, ribosomes and degradation machinery. These fluctuations are the source of so-called extrinsic noise [41,42]. We performed stochastic simulations varying the parameters in order to account for some aspect of extrinsic noise in the robustness study of the model. The results show that this oscillator is robust to small parameter variations [Fig. 4I] like more other complex models of genetic oscillators [27]. The largest amplitude and period changes occurred for variations in $k_1$ (see Table S1). The changes in the mean period and amplitude were always less than 15% and 31%, respectively. Particularly, variations in the rates $k_1, k_{-1}, k_2, k_6, k_7$ and $k_{10}$ produced changes of less than 3% and 8% in the mean period and amplitude, respectively. The changes in the standard deviation of the period and the amplitude were always less than 13% and 27%, respectively.

Reduced deterministic model

To identify the types of biomolecules mainly responsible for oscillations, it is useful to reduce the deterministic model by means of the quasi-steady-state assumption (QSSA) [43,48]. This approximation differentiates between fast and slow variables. The greater the time-scale separation between the variables the more accurate the approximation is. In this approach it is assumed that fast variables quickly reach the equilibrium, i.e., their derivatives are zero. This assumption means that slow variables are responsible for the system dynamics. In this model, we assumed that the fast variables are $G, G_a, M$ and $R$, and the slow variables are $A$ and $C$. Then, the set of Eq. (1) can be simplified to

$$
\frac{dA}{dt} = \frac{z + \beta A - k_9 A k_8 C}{\gamma + A A + k_9 A} - k_6 A,
$$

$$
\frac{dC}{dt} = \frac{k_2 A - k_1 k_2 - k_3 C}{\delta + A},
$$

where $z = G_1 k_{-1} k_2 k_3 k_4$, $\beta = G_1 k_2 k_3 k_4, \gamma = k_{-1} k_1, \delta = k_{10} k_7$ and $G_i = G + G_a$. A good way to check if this
Figure 4. Model robustness to noise. A. Stochastic and deterministic phase plane defined by $C$ and $A$. The general shape of the curves is the same in both cases, but the deterministic curve (dashed line) is well defined because the oscillations are identical. In contrast, the stochastic curve (solid line) spreads around the deterministic one. B, C. Amplitude and period histograms of the stochastic simulation of $A$, respectively. D. Autocorrelation of the stochastic oscillations in the number of $A$ molecules. The half-life of the autocorrelation is about 120 hours (intersection of dashed lines). E–H. Model robustness to intrinsic noise when the number of molecules is low. The changed rates are $k_4 \sim 1000$ hour$^{-1}$, $k_5 \sim 5000$ hour$^{-1}$, $k_6 = 25.5$ molecules hour$^{-1}$, $k_7 = 132.6$ hour$^{-1}$, and $k_9 = 1$ molecules hour$^{-1}$. In particular, we multiplied the rates $k_4$ and $k_5$ by 100 to obtain a low number of $M$ molecules. Simultaneously, we multiplied the rates $k_7$ and $k_8$, and divided the rate $k_9$ by 51 to obtain a low number of $R$ and $C$ molecules. The initial conditions are $G_{00} = 1$ and $G_{00} = M_{00} = A_{00} = R_{00} = C_{00} = 0$ molecules. The mean value of $M$ is 0.48 molecules. I. Model robustness to extrinsic noise. Scatter plot of amplitude versus period that shows the robustness of the model to parameter variation (data is presented in Table S1). Two stochastic simulations were performed for each parameter in which the value was increased and decreased by 15%. The x and y coordinates of each data point correspond to the mean values of the period and amplitude, respectively. The horizontal and vertical error bars are the standard deviation of the period and amplitude, respectively. The intersection between dashed lines shows the point obtained without changing the value of any rate (Figs. 4B and 4C). (B, C, D and each data point in I were calculated for 1,000 successive cycles. We assumed that a cycle occurs if the number of proteins $A$ increases to 1,000 molecules and then decreases to 700 molecules. The amplitude was calculated as the greatest number of $A$ molecules in each cycle. The period was calculated as the time interval that it takes the number of proteins $A$ to reach 1,000 molecules for the first time in two successive cycles.)

doi:10.1371/journal.pone.0027414.g004
approximation is correct is to compare the numerical solution of the complete and the reduced systems. Both numerical solutions agree except for quantitative differences in the period and the amplitude (Figs. 5A and 5B). These differences are due to the fact that the time-scale separation between fast and slow variables is not large enough for QSSA to be more accurate. Despite these differences, we can conclude that $A$ and $C$ are mainly responsible for the system dynamics. The other types of biomolecules can be considered to be at equilibrium. The fluctuations in the fast variables do not significantly affect the system dynamics [38]. This explains the robustness of the model when the number of molecules is low (Figs. 4E–H). In fact, the system produces reliable oscillations even if the average of $M$ is less than one molecule (Fig. 4H), and, surprisingly, even when the driven $C$ molecules oscillate in a range of less than 30 molecules (Fig. 4F).

The oscillations in the reduced deterministic model exhibit limit-cycle behavior (thin solid line in Fig. 5C). Therefore, if an external disturbance is applied to the oscillator, the system will go back to oscillating with the period and amplitude of its limit cycle. The unstable fixed point of the system is $A_0 = 552.4$ and $A_0 = 56.3$ molecules (circle in Fig. 5C). For a bifurcation analysis of limit-cycle oscillations, see Methods: Bifurcation diagram.

This genetic clock belongs to the so-called relaxation oscillators [28, 43, 44]. The mechanism responsible for the oscillations is represented by the nullclines $A_N$ and $C_N$ (Fig. 5C). These nullclines are the solution of the equations $dA/dt = 0$ and $dC/dt = 0$, respectively. The nullcline $A_N$ is a straight line and the nullcline $A_N$ has the characteristic “Z” shape of relaxation oscillators [43–45]. The shape of the $A$ nullcline is the same as the hysteresis diagram obtained if $C$ is assumed constant (Fig. S6). Therefore, this genetic clock contains some features of hysteresis in its oscillatory mechanism. The $A$ nullcline has two branches that we can call “high” and “low” (Fig. 5C). These branches are steady states if the $C$ is a constant (Fig. S6). In each oscillation the system switches from one branch to the other using the number of $C$ molecules as a transient signal. This process can be explained following the limit-cycle trajectory. When $A$ and $C$ are about 1 and 200 molecules, respectively, their number increases until $A$ reaches its maximum of about 7,330 molecules and $C$ reaches about 650 molecules. This is the transient from the low to the high branch. Then, the number of $A$ molecules is reduced to about 0 molecules, whereas $C$ reaches its maximum of about 1,260 molecules. This is the transient from the high to the low branch. Finally, the number of $C$ molecules is quickly reduced and the trajectory moves along the nullcline $A_N$, returning to the starting point where a new cycle begins.

This genetic clock is characterized by containing fast and slow stages. The time evolution of $C$ shows these two well-differentiated stages (Fig. 5D). In the slow stage $A \gg \delta$ and $k_A A \gg k_C C$, then the second differential equation in (2) can be approximated by $dC/dt \approx k_A$. In this stage, therefore, the number of $C$ molecules

![Figure 5. Model of the genetic oscillator reduced by QSSA. A, B. Comparison between the reduced (solid line) and complete (dashed line) deterministic simulation of the time evolution of $A$ and $C$, respectively. C. Phase plane. Limit cycle (thin solid line) and nullclines $A_N$ (thick solid lines) and $C_N$ (thick dashed line). The unstable fixed point of the system (marked by circle „) is $A_0 = 552.4$ and $A_0 = 56.3$ molecules. The $A_N$ and $C_N$ nullclines are the solution of equations $dA/dt = 0$ and $dC/dt = 0$, respectively. The two branches in the nullcline $A_N$ are called “high” and “low”. D. Slow and fast stages in the reduced system. The solid line is $C$ and the dotted line is $A$. $C$ exhibits a sawtooth waveform. (The arrows in C and D represent the direction of the oscillations. One and two arrows mean slow and fast stages, respectively).](http://www.plosone.org)
Intuitively, this effective negative feedback loop can be explained by the complex situation, the enzymes are saturated because most are part of complex substrate-enzyme and product, respectively. The total number of enzymes (Et = E + D) is constant in the system. The rate of catalysis in this reaction is defined as \( v \equiv dP/dt = C \cdot D \). The value of this rate can be approximated by \( k_C \cdot D \). Therefore, the rate of the catalysis increases asymptotically as a function of \( S \). The rate \( v \) reaches a maximum value \( V_{\text{max}} \) when the amount of \( S \) is large compared with the constant \( K_M \). In this situation, the enzymes are saturated because most are part of complex \( D \), and adding more \( S \) does not increase the rate \( v \). Therefore, \( D \approx E_t \), and the rate of the catalysis \( v \) reaches the constant value \( c_2 E_t \).

In our model, the negative interaction is \( A + R_t \xrightarrow{k_{C}} C \cdot k_s \cdot R \), where we assumed \( k_{C} \cdot \gamma = 0 \) to simplify the model. We can think of \( A \), \( R_t \), and \( C \) as \( S \), \( E \), and \( D \), respectively. Therefore, the rate of the negative interaction can be defined as \( v \equiv k_S \cdot C \).

In this model there is not an explicit negative feedback loop at the genetic level. It has been conjectured that all biochemical oscillators involve some sort of negative feedback loop [21]. In this genetic clock, an effective negative feedback loop appears in the reduced model (see the section Methods: The Jacobian matrix). Intuitively, this effective negative feedback loop can be explained as follows: when \( C \) is rare, \( A \) is increased by the positive feedback. This rise in the production of \( A \) leads to the accumulation of \( C \), which in turn increases \( v \). This accumulation of \( C \) increases until the negative interaction is fast enough to reduce the number of \( A \) molecules.

The biochemical reactions that fully describe the model in the feedback signal is an alternative and simple way of generating reliable oscillations. Our study suggests that PTF, besides increasing robustness in cellular clocks, could be more directly used several times. Therefore, the negative interaction does not produce an effective high nonlinearity (see Supporting Information: Text S1).

How the negative interaction works

The negative interaction decreases the number of free \( A \) molecules and takes the system back to the start of a new cycle. The detailed explanation of how this interaction works is related to the dynamics of the typical enzymatic reaction \( S + E \xrightarrow{k_{C}} P + E \), where \( S \), \( E \), \( D \) and \( P \) are the substrate, enzyme, complex substrate-enzyme and product, respectively. The total number of enzymes (Et = E + D) is constant in the system. The rate of catalysis in this reaction is defined as \( v \equiv dP/dt = C \cdot D \). The value of this rate can be approximated by \( k_C \cdot D \). Therefore, the rate of the catalysis increases asymptotically as a function of \( S \). The rate \( v \) reaches a maximum value \( V_{\text{max}} \) when the amount of \( S \) is large compared with the constant \( K_M \). In this situation, the enzymes are saturated because most are part of complex \( D \), and adding more \( S \) does not increase the rate \( v \). Therefore, \( D \approx E_t \), and the rate of the catalysis \( v \) reaches the constant value \( c_2 E_t \).

In our model, the negative interaction is \( A + R_t \xrightarrow{k_{C}} C \cdot k_s \cdot R \), where we assumed \( k_{C} \cdot \gamma = 0 \) to simplify the model. We can think of \( A \), \( R_t \), and \( C \) as \( S \), \( E \), and \( D \), respectively. Therefore, the rate of the negative interaction can be defined as \( v \equiv k_S \cdot C \) (Fig. 6A). This rate represents the number of degraded \( A \) molecules per hour. The negative interaction works as follows. The number of \( A \) molecules increases quickly due to the positive feedback. This rise causes most of the \( R_t \) molecules to bind to \( A \) molecules forming the complex \( C \). At this point, the system reaches the saturation level (circle in Fig. 6A). The total number of repressor molecules in the system is \( R_t = R + C \). Therefore, at the saturation point, \( C \approx R_t \), and the rate \( v \) reaches the value \( k_S \cdot R_t \).

The negative interaction is not fast enough to decrease the growth of \( A \) molecules immediately after the saturation point is reached. This is because the number of \( R_t \) molecules is low at this point. Nevertheless, new \( R_t \) molecules enter the system at rate \( k_s \cdot R_t \). Therefore, \( R_t \) increases linearly over time \( R_t \approx k_s \cdot t \) compared with the enzymatic reaction in which \( E_t \) is constant. This means that the rate of the negative interaction increases linearly according to equation \( v \approx k_S \cdot k_s \cdot t \). The value of \( v \) increases until the negative interaction is fast enough to reduce the number of \( A \) molecules and take the system back to the start of a new cycle. The maximum rate reached by the negative interaction is \( v_{\text{max}} = 3,180 \) molecules/hour (square in Fig. 6A).

In this model there is not an explicit negative feedback loop at the genetic level. It has been conjectured that all biochemical oscillators involve some sort of negative feedback loop [21]. In this genetic clock, an effective negative feedback loop appears in the reduced model (see the section Methods: The Jacobian matrix). Intuitively, this effective negative feedback loop can be explained as follows: when \( C \) is rare, \( A \) is increased by the positive feedback. This rise in the production of \( A \) leads to the accumulation of \( C \), which in turn increases \( v \). This accumulation of \( C \) increases until the negative interaction is fast enough to reduce the number of \( A \) molecules. In this model, we assumed that \( C \) is not degraded. If this complex is degraded according to the reaction \( C \xrightarrow{k_s} \), \( v \) increases at a slower rate, and its maximum value \( v_{\text{max}} \) is reduced (Figs. 6B and S5). The oscillations stop when \( k_{C} = 0.6 \) hour\(^{-1} \) (Fig. S5G), because not enough \( C \) is accumulated in order to increase \( v \).

This genetic oscillator does not need cooperative binding reactions nor the formation of protein multimers, in contrast to the one-gene oscillator with TNF [Fig. 1A]. It has been demonstrated that protein sequestration produces an effective high nonlinearity [49,50]. But this high nonlinearity is not observed if the repressor molecule is recycled [49]. In our model the repressor \( C \) can be used several times. Therefore, the negative interaction does not produce an effective high nonlinearity (see Supporting Information: Text S1).

Discussion

Genetic networks with NTFs and PTFs play an important role in cellular clocks. In this paper, we provided a simple model illustrating that a single gene with PTF has also the potential to produce reliable oscillations. The sufficient additional requirement is a simple and usual negative interaction of degradation, sequestration or inhibition acting on the positive feedback signal.

The model presented in this article has a different oscillatory mechanism than the well-established NTF one-gene oscillator model. Our model can be classified as a relaxation oscillator. A two-gene model has been proposed as a different way of producing reliable circadian oscillations in cellular clocks [26], which also is a relaxation oscillator. This two-gene model is important because it is robust to noise [30]. The model introduced in this paper is a simpler way to produce relaxation oscillations than the previous two-gene oscillator. A comparison with our model reveals that the activation of the repressor gene is not a necessary condition to produce reliable circadian oscillations in the two-gene oscillator. We demonstrated that our model produces circadian oscillations that are just as robust to noise as the two-gene oscillator and other more complex models [18,27]. Similarly to the two-gene oscillator, our model produces good oscillations when the average number of mRNA molecules is less than one. In fact, the number of proteins oscillates satisfactorily even when the other types of molecules involved in the clock are less than 30. Therefore, this model is a simpler genetic relaxation oscillator than the current two-gene clocks [25]. Our model does not need the activation of a second repressor gene by the PTF, cooperative binding reactions nor the formation of protein multimers.

A single gene with PTF and a negative interaction in the feedback signal is an alternative and simple way of generating reliable oscillations. Our study suggests that PTF, besides increasing robustness in cellular clocks, could be more directly and deeply involved in the production of oscillations than at first thought. Further research is necessary to elucidate the presence and the role of this genetic oscillator in natural cellular clocks. On the other hand, thanks to its simplicity, this model has the potential to be a new tool for engineering synthetic genetic oscillators. In this case the period and amplitude of the oscillations could be possibly controlled by externally manipulating the entry rate of the repressor molecules.

Methods

Biochemical reactions and rates

The biochemical reactions that fully describe the model in the Fig. 2 are as follows:
where $G$ denotes the gene without $A$ bound to its promoter, $M$ denotes mRNA transcribed from $G$, $A$ denotes the activator protein translated from $M$, $G_a$ denotes the gene with $A$ bound to its promoter, $R$ denotes the repressor and $C$ denotes $R$ bound to $A$. All the biochemical species are measured in molecules. The description of the rates is as follows: $k_1$ is the binding rate of $A$ to the promoter of $G$, $k_{-1}$ is the unbinding rate of $A$ from the promoter of $G$, $k_2$ is the basal transcription rate, $k_3$ is the activated transcription rate, $k_4$ is the degradation rate of $M$, $k_5$ is the translation rate, $k_6$ is the degradation rate of $A$, $k_7$ is the binding rate of $R$ to $A$, $k_8$ is the decay rate of $C$ into $R$, $k_9$ is the creation (or entry) rate of $R$ and $k_{10}$ is the degradation (or exit) rate of $R$.

We used standard values within the diffusion limit for the rates [18,38,47]. They are as follows: $k_1 = 1$ molecules$^{-1}$ hour$^{-1}$, $k_{-1} = 50$ hour$^{-1}$, $k_2 = 50$ hour$^{-1}$, $k_3 = 500$ hour$^{-1}$, $k_4 = 10$ hour$^{-1}$, $k_5 = 50$ hour$^{-1}$, $k_6 = 0.1$ hour$^{-1}$, $k_7 = 0.5$ molecules$^{-1}$ hour$^{-1}$, $k_8 = 2.6$ hour$^{-1}$, $k_9 = 51$ molecules hour$^{-1}$ and $k_{10} = 1$ hour$^{-1}$. The cell has a single copy of the gene: $G_0 = G + G_a = 1$ molecule. The initial conditions are: $G_0 = 0$, $G_a = 1$, $M_0 = 5$, $A_0 = 1000$, $R_0 = 5$, and $C_0 = 1200$ molecules. The initial conditions have been chosen to obtain a first cycle with an amplitude similar to the limit-cycle oscillations. Note that the rates $k_1$ and $k_7$ include the volume of the system $V$. Hence, these rates can be written as $k_1 = k_4^f V$ and $k_7 = k_4^s V$, where the rates $k_4^f$ and $k_4^s$ are expressed in M$^{-1}$ hour$^{-1}$. In order to generate circadian oscillations, first, we varied all the reaction rates, according to the values used in [18,38] and [47], until we got oscillations with a period of around 24 hours in the stochastic simulation. Then we fine-tuned the oscillations varying rates $k_8$ and $k_9$ until a period closer to 24 hours was achieved.

**Deterministic and stochastic simulations**

Models based on chemical reactions in a well stirred system are usually described by two different formalisms from a mathematical point of view:

Deterministic: this formalism is suitable for large numbers of molecules. It is described by a set of coupled ordinary differential equations that follow the law of mass action. These equations are called reaction rate equations and they can only be solved analytically for simple systems. For more complex systems numerical methods are necessary. In this approach the amount of each chemical species and the time are continuous. The velocity at which reactions occur is given by the reaction rate constants $k$, or simply rate.

Stochastic: this formalism is suitable for small numbers of molecules because it takes into account the randomness of the chemical reactions. It is described by the so-called master equation, which is the time evolution of the probability that the system has a certain number of molecules of each chemical species at time $t$. Few systems can be solved analytically with the master equation. It is possible, however, to simulate the stochastic behaviour with the Gillespie algorithm [46]. In this approach the amount of each chemical species and the time are discrete, and the rates $k$ turn into probabilities.

**Bifurcation diagram**

We calculated the bifurcation diagram for parameters $k_8$ and $k_9$. These are key parameters for two reasons. First, the rate of the negative interaction $v$ is proportional to $k_8$ and $k_9$ when the saturation point is reached. Second, the fast and slow stages in the relaxation oscillations depend on $k_8$ and $k_9$ respectively.
Specifically, we studied the range values of $k_9$ that produce stable oscillations through a bifurcation diagram. Then we studied how this range changes when the parameter $k_8$ varies.

The bifurcation diagram of the reduced model depending on $k_9$ shows two Hopf bifurcation points (Fig. S3A). The first Hopf bifurcation appears at $k_9 = 4.78$ molecules hour$^{-1}$ and the second at $k_9 = 217.6$ molecules hour$^{-1}$. Most of the values of $k_9$ between these two points produce stable oscillations. Only for a short range of values around these points are the oscillations unstable (white circles in Fig. S3A). The oscillations have an amplitude of from 2,000 to 16,000 molecules, and a period of from 7 to 170 hours (Fig. S3B). The velocity of the reaction $\phi \to \infty$ in $R$ in (Eq. 3) does not depend on any biomolecule involved in the oscillator. Therefore, parameter $k_9$ can be interpreted as an external signal controlling the behaviour of the clock.

The variation of parameter $k_8$ changes the position of the two Hopf bifurcation points (white circles in Fig. S4). The different positions of these points define the regions with stable oscillations depending on the values of $k_8$ and $k_9$ (regions I and II in Fig. S4). If parameter $k_8$ is increased, the range of values of $k_9$ that produces stable oscillations decreases. This range shrinks faster if $k_8$ is greater than 20 hour$^{-1}$. We plotted an equivalent graph for the stochastic model because it is more realistic than the reduced model (black circles in Fig. S4). In particular, we assumed that oscillations occur in a region if the correlation in the first period is greater than 0.2. The stochastic model produces oscillations in the regions II and III (Fig. S4). The range of oscillations in the complete deterministic model is close to the region II.

The Jacobian matrix

The Jacobian matrix of the reduced system (2) is:

$$J = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \frac{\gamma}{(v + A)^2} - \frac{\delta (k_9 + k_6 C)}{(6 + A)^2} - k_6 - \frac{k_9 A}{\delta + A} \\ \frac{\delta (k_9 + k_6 C)}{(6 + A)^2} - \frac{k_9}{\delta + A} \end{pmatrix} \tag{4}$$

where the element $a_{12}$ and $a_{22}$ are always negative, the element $a_{21}$ is always positive and the element $a_{11}$ can be positive or negative depending on the values of the rates. With the rates given in the section Methods: Biochemical reactions and rates and the fixed point of the reduced system (Fig. 5C) the sign pattern for the Jacobian matrix is:

$$J = \begin{pmatrix} + & - \\ + & + \end{pmatrix} \tag{5}$$

A two-component negative feedback loop is created in the reduced model because $a_{12}a_{21} < 0$ (see Chapter 9 of the reference [48]). The Jacobian matrix (5) has a typically sign pattern that produces Hopf bifurcation in chemical systems with two variables [43,48].

The two-component systems with this sign pattern in the Jacobian matrix are called activator-inhibitor models [48].

Software

Code for stochastic and deterministic simulations was written in FORTRAN and XPPAUT (http://www.math.pitt.edu/bard/xpp/xpp.html), respectively. Simulations have been contrasted using CAIN software (http://cain.sourceforge.net/). The stability analysis to determine steady states and limit cycles was performed with XPPAUT. The histograms and autocorrelation function were plotted using FORTRAN and GNU Octave (http://www.gnu.org/software/octave/). The code for complete and reduced deterministic simulations in XPPAUT is available in File S1 and File S2. The code for stochastic and deterministic simulations in CAIN is available in File S3.

Supporting Information

Figure S1 Time evolution of $A$ with and without a low number molecules. Comparison between deterministic simulation of the time evolution of $A$ with (dashed line) and without (solid line) a low number of $M$, $R$, and $C$ molecules. (Solid line graph: the values of the parameters are as in the section Methods: Biochemical reactions and rates. Dashed line graph: the changed rates are $k_4 = 1000$ hour$^{-1}$, $k_5 = 5000$ hour$^{-1}$, $k_7 = 25.5$ molecules$^{-1}$ hour$^{-1}$, $k_8 = 132.6$ hour$^{-1}$ and $k_9 = 1$ molecules hour$^{-1}$.)

Figure S2 Amplitude and period histograms of the stochastic simulation of $A$, $A$, $B$. Amplitude and period histograms of the stochastic simulation of $A$, respectively. The values of the parameters are as in the section Methods: Biochemical reactions and rates but now we set $k_4 = 1000$ hour$^{-1}$, $k_5 = 5000$ hour$^{-1}$, $k_7 = 25.5$ molecules$^{-1}$ hour$^{-1}$, $k_8 = 132.6$ hour$^{-1}$ and $k_9 = 1$ molecules hour$^{-1}$. $A$ and $B$ were calculated for 1,000 successive cycles. We assumed that a cycle occurs if the number of proteins $A$ increases to 1,000 molecules and then decreases to 700 molecules. The amplitude was calculated as the greatest number of $A$ molecules in each cycle. The period was calculated as the time interval that takes the numbers of proteins $A$ to reach 1,000 molecules for the first time in two successive cycles.)

Figure S3 Bifurcation diagram of the reduced model. A. Bifurcation diagram depending on $k_9$. The solid/dashed line represents stable/unstable fixed points. Black/white circles are the maximum and minimum values of $A$ during unstable/stable oscillations. HB denotes a Hopf Bifurcation point. HB and HB appear when the value of $k_9$ is 4.78 and 217.6 molecules hour$^{-1}$, respectively. B. Period of the stable oscillations in $A$.

Figure S4 Oscillatory regions in the reduced and stochastic models depending on $k_9$ and $k_8$. Region I. Oscillations in reduced model. Region II. Oscillations in both reduced and stochastic model. Region III. Oscillations in the stochastic model. Region IV. No oscillations in any model. White circles represent the locus of Hopf bifurcations in the reduced model (data are presented in Table S2). Black circles represent the locus of oscillations in the stochastic simulation (data are presented in Table S3). We assumed in the stochastic case that oscillations occur in a region if the correlation in the first period is greater than 0.2. (The lines connecting circles are designed to clearly single out the different regions.)

Figure S5 Rate of the negative interaction for different values of $k_{11}$. Rate of the negative interaction $v = k_9 C$ for different values of $k_{11}$, where $k_{11}$ is the rate of reaction $C \longrightarrow \phi$. Deterministic simulations $A$, $B$, $C$, $D$, $E$, $F$ and $G$ correspond to $k_{11}$ equals 0.0, 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 hour$^{-1}$, respectively. The values of the other parameters are as in the section Methods: Biochemical reactions and rates. The oscillations stop when $k_{11} = 0.6$ hour$^{-1}$. If $k_{11}$ is increased, $v$ increases slower, and its maximum value ($v_{max}$) is lower. The value of $v_{max}$ corresponds to
the peak of the oscillations ($u_{max}$ is the value of the steady state in $G$).

**Figure S6** Hysteresis diagram. Hysteresis diagram depending on $C$. The curve is the solution of the equation $dA/dt = 0$, where $C$ is assumed constant. The two solid lines in the diagram are the two stable steady states “high” and “low” as a function of $C$. The dashed line represents the unstable points in the diagram.

**Table S1** Data points of Fig. 4I.

**Table S2** Data points of locus Hopf bifurcation in reduced model (Fig. S4).

**Table S3** Data points of locus of oscillations with less than 20% of correlation in the first period in the stochastic model (Fig. S4).

**Text S1** The negative interaction does not produce an effective high nonlinearity.

**References**


**File S1** Complete deterministic model (XPPAUT software).

**File S2** Reduced deterministic model (XPPAUT software).

**File S3** Stochastic and deterministic model (CAIN software).

**Acknowledgments**

Part of this work was carried out by JMMB during his stay at the Novel Computation Group led by Dr. Martyn Amos at MMU. We would also like to thank Rachel Elliott and Niall Murphy for polishing the English, and an anonymous reviewer for useful comments on the manuscript.

**Author Contributions**

Conceived and designed the experiments: JMMB ARP. Performed the experiments: JMMB. Analyzed the data: JMMB ARP. Contributed reagents/materials/analysis tools: JMMB. Wrote the paper: JMMB ARP.