VERSITA

Central European Journal of Biology
DOI: 10.2478/s11535-007-0016-z
Research article
CEJB 2(2) 2007 233–253

Evolutionary mechanisms of circadian clocks

Guillermo Rodrigo[1], Javier Carrera[1], Alfonso Jaramillo*[2]

[1] *Depto. de Matematica Aplicada, Universidad Politecnica de Valencia, 46022 Valencia, Spain*

[2] *Laboratoire de Biochimie, CNRS - UMR 7654, Ecole Polytechnique, 91128 Palaiseau, France*

Received 19 February 2007; accepted 26 March 2007

Abstract: An intriguing question in biology is to know how circadian molecular networks could have evolved their particular topologies to adjust to a daily period. We analyze the mechanism of the evolution of such networks by using a computational design strategy that allows for the generation of synthetic networks with a targeted 24 hours oscillation. We have performed a systematic analysis of all possible two-gene network topologies based on a core activator-repressor frequently found in circadian mechanisms. We have considered transcriptional and post-translational regulations to implement this core. We have applied our analysis to both, eukaryotic and prokaryotic circadian machinery. Finally, we conjecture a possible mechanism for the evolution of circadian clocks.

Keywords: Circadian clocks, systems biology, biological networks

1 Introduction

Living organisms are immersed in cyclic processes and evolution has provided some of them with molecular clock systems. Some of these clocks are composed of genetic circuits having a rhythm of about 24 h period. Interestingly, circadian clocks can work even in a constant environment (e.g., constant darkness). Circadian clocks control many important activities, such as the sleep-wake cycle or the migration of birds [1].

Traditionally, it was thought that bacteria did not have circadian clocks and that these types of circuits were specific to eukaryotic cells [2]. However, it has been found

* E-mail: Alfonso.Jaramillo@polytechnique.edu

that prokaryotic cells also exhibit an oscillatory behavior [3]. The first studied organisms were *Drosophila* [4] and *Neurospora* [5], but later work on circadian clocks have extended these circuits to cyanobacteria [6], plants [7], and mammals [8]. In mammals, clock genes are expressed in the suprachiasmatic nuclei (SCN) [9], where each SCN cell produces a periodic signal.

In the studied models of circadian clocks, the interactions between proteins and genes were found to give complex networks [10]. To analyze them and obtain their system response, it is necessary to introduce mathematical and computational tools together with the use of quantitative models to get an accurate explanation of their dynamical mechanism. From the perspective of the emerging field of synthetic biology, the *de novo* design of proteins and gene networks constitutes a new approach in biology [11, 12] that allows obtaining biological understanding by following a model-driven approach. By introducing new synthetic proteins [13] and biological networks [14] within the cell we can test our current biological knowledge and iteratively correct our assumptions. Rational design techniques have been found to be very fruitful to design new genetic circuits with a targeted behavior [15], and even computational methods are also starting to be used with this goal [16]. The success of new designed synthetic oscillators [11, 17] gives a strong boost to the understanding of the molecular mechanism underlying some circadian clocks.

Here, we have analyzed simplified models of circadian clocks present in various organisms. Principally, these are reduced genetic networks containing protein-protein interactions. Following the work of Dunlap [1], these models are based on two elements: a positive element (activator) and a negative element (repressor). The activator is a transcription factor which induces the transcription of the repressor element. On the other hand, the repressor binds to the activator through a protein-protein interaction [18]. In addition, we have proposed equivalent transcriptional models of them by changing the protein sequestration interaction by a transcriptional repression [17]. We apply our analysis to eukaryotic (*Drosophila*, *Neurospora* and mammals) and prokaryotic (*Synechococcus*) clock networks. We have developed an algorithm to design transcriptional clocks. In addition, we have performed a systematic study of reduced gene models to know their oscillatory ability.

2 Materials and methods

2.1 Mathematical model

For this work, we have used a deterministic model based on differential equations taking the form

$$\begin{aligned} \frac{dy}{dt} &= f(t, y, k), \\ y(0) &= 0, \end{aligned} \tag{1}$$

where y is a vector containing the protein concentrations, f is a vector function, and k represents the various parameters of the model. To simplify, the initial conditions for

all species are considered to be 0. We have studied two different types of models. In the first case, the mRNA degradation is high and its dynamics are fast (with respect to protein dynamics). We then use an effective transcriptional regulatory function and the time-evolution of our transcription factor concentration (Y_i) a given by

$$\frac{d}{dt}[Y_i] = \alpha_i F_i([Y_j], K_{ij}, n_{ij}) - \beta_i[Y_i] + \gamma_i, \tag{2}$$

where α_i is the transcription-translation rate of gene i, β_i is the corresponding degradation rate, and γ_i is the basal synthesis rate. The function $F_i([Y_j], K_{ij}, n_{ij})$ is the regulatory factor, which is implemented with a standard Hill function, K_{ij} is the regulatory coefficient and n_{ij} is the Hill coefficient.

In our second type of models, we suppose that the mRNA (M_i) dynamics is slow (i.e., compared with the time scale of protein dynamics) and we explicitly model its dynamics. In this case, the Hill function

$$\begin{aligned} \frac{d}{dt}[M_i] &= \alpha_i F_i([Y_j], K_{ij}, n_{ij}) - \delta_i[M_i] + \gamma_i, \\ \frac{d}{dt}[Y_i] &= \lambda_i[M_i] - \beta_i[Y_i], \end{aligned} \tag{3}$$

where α_i is now the transcription rate of gene i, λ_i the translation rate, and δ_i is the mRNA degradation rate. In our study we did not need to go to more accurate descriptions by also modeling the DNA (e.g., its binding equilibrium between promoters and transcription factors) [19]. It has been shown that the real effect of explicitly modeling mRNA is that it introduces a delay in the build up of protein concentration [20] and it is possible to obtain similar oscillatory dynamics without considering mRNA if such delay is explicitly modeled in the Hill function. We did not need to consider this simplification either.

Stochastic models are also used to simulate the molecular noise [21], and in circadian clocks noise introduces constraints on the oscillatory mechanism [18]. Results can differ if we use a model based on differential equations or via Gillespie [22]. For instance, we could consider a continuous model for a circuit and predict a steady behavior, but by modeling stochastic reactions the system could be capable of oscillating. In this work, we have chosen to discard this type of behavior and concentrate on analyzing the general oscillatory mechanisms arising from the network topology. Therefore, we have not considered molecular noise.

2.2 Computational design procedure

We have implemented an algorithm in C to engineer transcriptional regulatory networks with a specified behavior following the methodology used by Francois and Hakim [16]. We have used simulated annealing [23] as our optimization procedure, which is implemented though the use of Monte Carlo algorithm with Metropolis updating. The algorithm starts from a random circuit and performs an evolutionary procedure towards the optimized circuit. For this purpose, at each step of our optimization, we modify the circuit by

randomly changing a parameter or by adding/removing an interaction. We then solve the dynamics of the corresponding genetic network and we compute a scoring function (Ψ)

$$\Psi = \int |y - z| \chi dt, \tag{4}$$

where y is the species concentration at a given time, z is the targeted concentration, which could be implemented as an oscillatory function with a period of 24 h for circadian circuits, and χ is a binary weighting factor that scores only in the neighborhoods of maxima and minima. We have also developed a MATLAB algorithm to train parameters for a given network following a different score, based on the eigenvalues of the Jacobian matrix of Eq. 1,

$$\Psi = \prod \frac{Re(\lambda_i)}{1 - 0.99e^{-|Im(\lambda_i)|}}. \tag{5}$$

This score function allows us to find the Hopf bifurcation of a given circuit [24]. Particularly, we have applied our methodology to obtain oscillatory circuits. This procedure does not consider protein-protein interactions and it also assumes a fast mRNA dynamics. Our optimization algorithm does not guarantee to provide unique answers and further work is needed to estimate alternative solutions.

2.3 Core activator-repressor

In principle, to accurately study the whole mechanism of the cellular circadian clocks would require knowing all the interactions and parameters between the species involved in the process, but recently there have been many advances in this respect [25–27]. Instead, we will consider simplified models that will allow us to understand the design principles of circadian clocks [18, 28]. In particular, we have focused in the oscillatory behaviors that could arise from a transcriptional core of activator-repressor [18].

To study the dynamical behavior of the system, we analyze the linear stability of the system around a fixed point [29], defined by $f(t, y^*, k) = 0$. For this, we calculate the Jacobian matrix and its corresponding eigenvalues (λ) in that fixed point ($J = \frac{\partial f}{\partial y}(y^*)$). When the eigenvalues are complex (with positive real part), the system has sustained oscillations in a neighborhood of the fixed point.

For the particular case of having a system composed of two genes and neglecting the mRNA dynamics (assumed to be fast enough), we can solve exactly the conditions required for oscillations. We can calculate the eigenvalues according to

$$\lambda = \frac{1}{2}\tau \pm \frac{1}{2}\sqrt{\tau^2 - 4\delta}, \tag{6}$$

where τ is the trace of the Jacobian matrix, and δ its determinant. If λ is a real negative number then the system is stable, whereas if it is positive then the system is unstable. When λ is a complex number ($\tau^2 - 4\delta < 0$) the system has an oscillatory behavior. If the real part is negative, the oscillations are damped and a stable steady state is achieved, but if it is positive, then the system becomes a permanent oscillator.

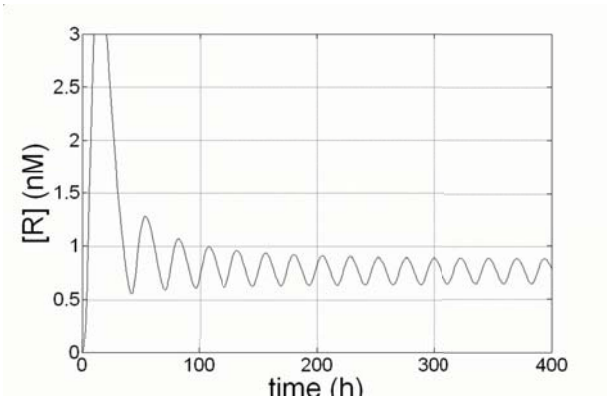


Fig. 1 Gene A dynamics corresponding to the oscillatory core with negative feedback (first case in Table 1). Parameter values (Eq. 7) are as follows: $\alpha_A = 2.083$ nM/h, $K = 0.5$ nM, $n = 9$, $\delta = 0.17$ h^{-1}, $\gamma_A = 0$ nM/h, $\lambda = 1.04$ h^{-1}, $\beta_A = 0.104$ h^{-1}, $\beta_R = 0.146$ h^{-1}, $\nu_1 = 0.04$ nM^{-1}h^{-1}, and $\nu_2 = 0.01$ h^{-1}.

Hence, we have used this fact to systematically characterize the circuits made of our two-genes core. If the eigenvalues are real, they will have to be negative as the system is bounded. If they are complex, the real part could be positive or negative. To know which circuits could be susceptible of showing an oscillatory behavior, once their parameter values are conveniently chosen, we have looked for eigenvalues with a positive real part (i.e. with a positive trace of the Jacobian matrix, $\tau > 0$).

On the other hand, when the model considers the dynamics of mRNA the analytic study is harder. Still, systems modeled with three equations (considering mRNA for one gene) can be analytically analyzed without much trouble using the fact that a given third degree polynomial equation ($x^3 + ax^2 + bx + c = 0$, with $a, b, c > 0$) has all roots with the negative real part if only if $ab - c > 0$. Following a similar reasoning as above, the oscillatory condition becomes $c > ab$. For a $3x3$ matrix J, we can calculate the characteristic polynomial coefficients as $a = -\tau = -tr(J)$ and $b = \eta = tr(adj(J))$ ($adj(J)$ is the adjoint of J), and $c = -\delta = -det(J)$ where for our problem $\delta, \tau < 0$ and $\eta > 0$.

3 Results

3.1 Circadian clocks core

We have started our systematic analysis of genetic networks of two genes by considering a circuit based on a transcriptional repression (negative feedback) and a protein activation (see first case in Table 1).

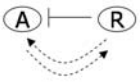
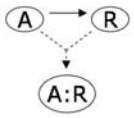
Circuit	Oscillatory Condition	Possible
	$\lambda\nu_1\alpha_A \frac{n}{K}\frac{(R/K)^n}{1+(R/K)^n} > (\beta_A+\delta)[(\beta_R+\nu_2)(\beta_R+\nu_2+\beta_A+\delta)+\delta\beta_A]$ with $\alpha_A \frac{1}{1+(R/K)^n}+\gamma_A = \frac{\delta\beta_A(\nu_2+\beta_R)R-\alpha_R}{\lambda\nu_1}$	Yes
	$\lambda\theta_1\alpha_R A \frac{n}{K}\frac{(A/K)^n}{1+(R/K)^n} > (\beta_A+\theta_1 A+\beta_R+\theta_1 R)$ $[\delta(\delta+\beta_A+\theta_1 A+\beta_R+\theta_1 R)+\beta_A\beta_R+\theta_1(\beta_A A+\beta_R R)]$ with $\alpha_R \frac{(A/K)^n}{1+(A/K)^n}+\gamma_R = \frac{\delta}{\lambda}\left(\beta_R R+\theta_1 AR\frac{\beta_{A:R}}{\theta_2+\beta_{A:R}}\right)$ $\alpha_A+\frac{\theta_1\theta_2 AR}{\theta_2+\beta_{A:R}} = \beta_A A+\theta_1 AR$	Yes

Table 1 Oscillatory conditions for two activation-repression cores. If the condition is satisfied the corresponding genetic network can oscillate under a suitable choice of parameters. Bold arrows denote activation, and bold blunt lines repression. Dashed lines mean protein interaction. We have modeled the mRNA dynamics. We have used the following kinetic rates: α_i as the transcription of gene i, γ_i as its Basal expression, n as the Hill coefficient, K as the regulatory coefficient, δ_i mRNA degradation, β_i protein degradation, λ_i translation, ν_1 as the production rate of R from A (e.g., phosphorylation), ν_2 as the inverse rate (e.g., dephosphorylation), θ_1 as the protein binding between A and R, and θ_2 as the unbinding (which is assumed 0). All the variables are considered to be positive.

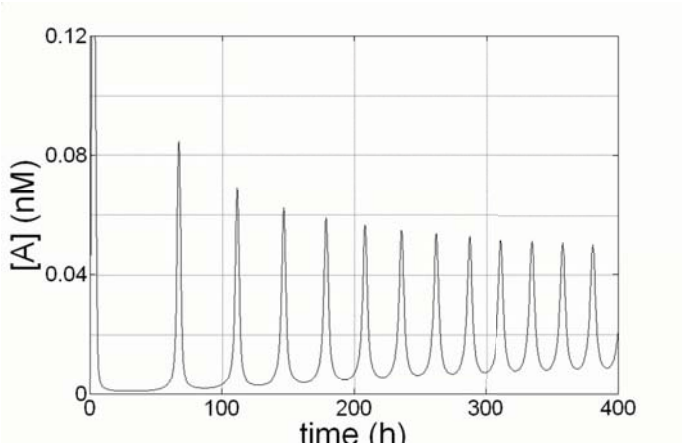


Fig. 2 Gene A dynamics corresponding to the oscillatory core with positive feedback (second case in Table 1). Parameter values (Eq. 8) are as follows: $\alpha_R = 0.0166$ nM/h, $K = 0.05$ nM, $n = 6$, $\delta = 0.0304$ h^{-1}, $\gamma_R = 0.0001$ nM/h, $\lambda = 2.7645$ h^{-1}, $\alpha_R = 0.0923$ nM/h, $\beta_R = 0.0009$ h^{-1}, $\beta_{A:R} = 0.0092$ h^{-1}, $\theta_1 = 92.15$ nM^{-1}h^{-1}, and $\theta_2 = 0$ h^{-1}.

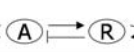
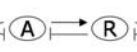
Type	Circuit	Oscillatory Condition	Possible
1		$-\beta_A-\beta_R>0$	No
2		$-\beta_A-\beta_R+\alpha_A G_{AA}>0$	Yes
3		$-\beta_A-\beta_R+\alpha_R G_{RR}>0$	Yes
4		$-\beta_A-\beta_R-\alpha_A G_{AA}>0$	No
5		$-\beta_A-\beta_R-\alpha_R G_{RR}>0$	No
6		$-\beta_A-\beta_R+\alpha_A G_{AA}+\alpha_R G_{RR}>0$	Yes
7		$-\beta_A-\beta_R-\alpha_A G_{AA}-\alpha_R G_{RR}>0$	No
8		$-\beta_A-\beta_R+\alpha_A G_{AA}-\alpha_R G_{RR}>0$	Yes
9		$-\beta_A-\beta_R-\alpha_A G_{AA}+\alpha_R G_{RR}>0$	Yes

Table 2 Oscillatory conditions for all possible two-gene topologies corresponding to an activator-repressor transcriptional circuit. Arrows mean activation, and blunt lines repression. The oscillatory condition corresponds to the Jacobian trace, where β denotes the degradation rate, α the transcription-translation rate, F the regulatory function, and G the derivate of the regulatory function (e.g., $G_{AR} = \partial F_A/\partial R$).

The model is given by

$$\begin{aligned} \frac{d}{dt}[M_A] &= \alpha_A \frac{1}{1+(\frac{[R]}{K})^n} - \delta[M_A] + \gamma_A \\ \frac{d}{dt}[A] &= \lambda[M_A] - \beta_A[A] - \nu_1[A] + \nu_2[R] \\ \frac{d}{dt}[R] &= \nu_1[A] - \nu_2[R] - \beta_R[R], \end{aligned} \tag{7}$$

where ν_1 is the binding rate of A with another molecule (such a phosphorus or a protein) and ν_2 the corresponding unbinding. On the other hand, the second circuit in Table 1 is based on a transcriptional activation (positive feedback) coupled with a post-translational repression. Its dynamics is given by

$$\begin{aligned} \frac{d}{dt}[M_R] &= \alpha_R \frac{(\frac{[A]}{K})^n}{1+(\frac{[A]}{K})^n} - \delta[M_R] + \gamma_R, \\ \frac{d}{dt}[R] &= \lambda[M_R] - \beta_R[R], \\ \frac{d}{dt}[A] &= \alpha_A - \theta_1[A][R] + \theta_2[A:R] - \beta_A[A], \\ \frac{d}{dt}[A:R] &= \theta_1[A][R] - \theta_2[A:R] - \beta_{A:R}[A:R]. \end{aligned} \tag{8}$$

where θ_1 and θ_2 correspond to the binding and unbinding rates respectively between the

proteins A and R. In Table 1 we show the required conditions for having a oscillatory dynamics for these two models of circadian cores found in natural organisms. Furthermore, in Fig. 1 and 2 we show the dynamics for the first circuit and second circuits, respectively. Both circuits have a period of 24 h. We notice that those circuits appear coupled in living systems. In addition, if we restrict ourselves to the pure transcriptional core, those circuits are not capable to sustain oscillations when the mRNA degradation rate is high (see Table 2). On the other hand, if the mRNA degradation rate is comparable to the protein rates then we can obtain an oscillatory behavior. We can easily obtain this result by considering the eigenvalues equation $(x+\beta_A)(x+\beta_R)(x+\delta_A)(x+\delta_R)+\Gamma=0$, with $\Gamma=\lambda_A\lambda_R\alpha_A\alpha_R G_{RA}G_{AR}$ where $G_{AR}=\partial F_A/\partial R$. If $\beta_A=\beta_R=\delta_A=\delta_R=0.1$ and $\Gamma=1$ the eigenvalues are $-0.8\pm0.7i$ and $0.6\pm0.7i$. Thus, we concluded that the circuit will oscillate.

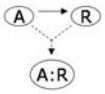
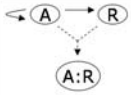
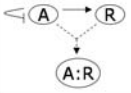
Type	Circuit	Oscillatory Condition	Possible
10		$-\beta_R-\beta_{AR}-\frac{\theta_1 AG_{RA}+\beta_A(\theta_1 A+\theta_2)}{\theta_1 R+\beta_A}>0$	No
11		$-\beta_R-\beta_{AR}-\frac{\theta_1 AG_{RA}+(\beta_A-\alpha_A G_{AA})(\theta_1 A+\theta_2)}{\theta_1 R+\beta_A-\alpha_A G_{AA}}>0$	Yes
12		$-\beta_R-\beta_{AR}-\frac{\theta_1 AG_{RA}+(\beta_A+\alpha_A G_{AA})(\theta_1 A+\theta_2)}{\theta_1 R+\beta_A+\alpha_A G_{AA}}>0$	No

Table 3 Oscillatory conditions for several activator-repressor gene networks involving protein-protein interactions. We assume that A has a faster dynamics than R. β denotes the degradation rate, α the transcription-translation rate, F the regulatory function, and G the derivate of the regulatory function (e.g., $G_{AR}=\partial F_A/\partial R$).

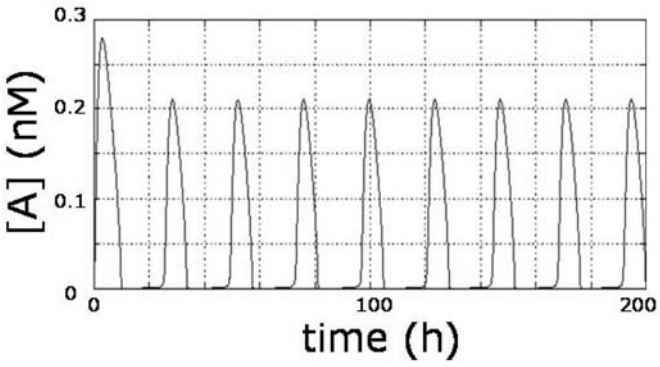


Fig. 3 Protein A dynamics corresponding to the oscillator of type 11 (see Fig. 3) [18]. Parameter values are as follows: $\alpha_A=0.416$ nM/h, $K_{AA}=0.05$ nM, $n_{AA}=1$, $\beta_A=1$ h^{-1}, $\gamma_A=0.04$ nM/h, $\alpha_R=0.045$ nM/h, $K_{RA}=0.1$ nM, $n_{RA}=1$, $\beta_R=0.2$ h^{-1}, $\gamma_R=0.01$ nM/h, and $\theta_1=2000$ nM^{-1}h^{-1} [30].

In addition, we have modeled all possible two-gene transcriptional circuits (assuming a fast mRNA degradation rate), where we have established the corresponding oscillatory conditions (based on the trace of the Jacobian matrix) (see Table 2). We have also performed the same study for circuits containing a post-translational repression, where we have assumed a fast mRNA degradation, and we show in Table 3 the resulting dynamics with 24 h period.

In Table 2, we show an analysis of the oscillatory capability for all the two-gene transcriptional core activator-repressor circuits (transcriptional repression). We have also extended the analysis to the most relevant two-gene genetic circuits involving protein-protein interactions (i.e., post-translational repression) having the activator-repressor mechanism, which is shown in Table 3. In the latter case, we have considered an equilibrium between A, B and $A:B$, where θ_1 and θ_2 would be the binding and unbinding rates respectively, and we assume that A has a fast degradation (i.e., $\beta_A > \beta_R$). Thus, we consider $d[A]/dt = 0$ and we calculate A as a function of R and $A:R$. Actually, this analysis is also useful for more complex models of circadian cores (e.g., those composed of a double repression or a double activation). However, our analysis just signals the possibility of oscillating for a certain range of kinetic parameters (which could end up being unrealistic).

It is interesting to analyze the circadian clock system proposed by Barkai and Leibler [18], assuming $d[M]/dt = 0$ (see Eq. 9) and taking the parameter values from [30], and we show in Fig. 3 the corresponding dynamics. This circuit (which corresponds to the second case in Table 3) is also called VKBL oscillator. It composed of two genes, an activator (A) and a repressor (R). The activator regulates both its own transcription together and the repressor levels. On the other hand, the repressor only regulates the activator transcription. Additionally, the activator and the repressor bind to obtain a new complex ($A:R$). This binding reaction is not reversible, although in that complex the activator protein can also be degraded which then releases the repressor protein. The model is given by

$$\begin{aligned}
\frac{d}{dt}[A] &= \alpha_A \frac{(\frac{[A]}{K_{AA}})^{n_{AA}}}{1+(\frac{[A]}{K_{AA}})^{n_{AA}}} - \beta_A[A] + \gamma_A - \theta_1[A][R], \\
\frac{d}{dt}[R] &= \alpha_R \frac{(\frac{[A]}{K_{RA}})^{n_{RA}}}{1+(\frac{[A]}{K_{RA}})^{n_{RA}}} - \beta_R[R] + \gamma_R - \theta_1[A][R], \\
\frac{d}{dt}[A:R] &= \theta_1[A][R] - \beta_{A:R}[A:R].
\end{aligned} \tag{9}$$

Using the parameter set shown in caption of Fig. 3, this system has oscillations of 24 h period.

It is important to remark that our study is only valid under the stated assumptions. For instance, the circuit type 10 (see Table 3) has been catalogued as a non-oscillator when mRNA degradation rate is high. However, as we see in Table 1, if we assume moderate mRNA degradation rates then we find that the system is capable of oscillating. In addition, this circuit was shown to either behave as an oscillator or as a bistable

switch depending on the kinetic parameters [19]. Therefore, a more accurate systematic study of circadian genetic networks, requires to take into account intermediate mRNA degradation rates (we could even consider the DNA dynamics by modeling the binding between transcription factors and promoters).

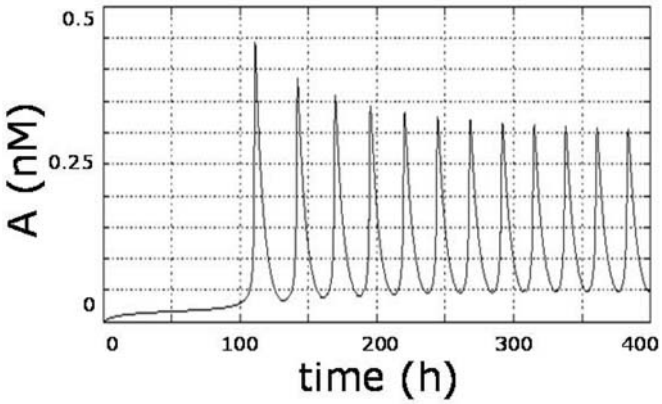


Fig. 4 Protein A dynamics corresponding to the circuit of type 2 (see Fig. 2) [17]. Parameter values are as follows: $\alpha_A = 5.78$ nM/h, $K_{AA} = 0.70$ nM, $n_{AA} = 2.26$, $K_{AR} = 0.85$ nM, $n_{AR} = 3.54$, $\beta_A = 0.17$ h^{-1}, $\gamma_A = 0.0017$ nM/h, $\alpha_R = 8.16$ nM/h, $K_{RA} = 0.51$ nM, $n_{RA} = 2.58$, $\beta_R = 0.17$ h^{-1}, $\gamma_R = 0.0017$ nM/h.

On the other hand, we have also simplified the VKBL oscillator to a transcriptional-only model (type 2 in Table 2). We have added on gene A a transcriptional repression (by R), instead of the previous protein interaction between these two proteins [17]. The model is given by

$$\begin{aligned} \frac{d}{dt}[A] &= \alpha_A \frac{(\frac{[A]}{K_{AA}})^{n_{AA}}}{1+(\frac{[A]}{K_{AA}})^{n_{AA}}} \frac{1}{1+(\frac{[R]}{K_{AR}})^{n_{AR}}} - \beta_A[A] + \gamma_A, \\ \frac{d}{dt}[R] &= \alpha_R \frac{(\frac{[A]}{K_{RA}})^{n_{RA}}}{1+(\frac{[A]}{K_{RA}})^{n_{RA}}} - \beta_R[R] + \gamma_R. \end{aligned} \tag{10}$$

If we consider the previous VKBL oscillator parameter values and evolve it by introducing an additional regulation, the new system does not have an oscillatory behavior. However, if we optimize the parameters for this network (see caption of Fig. 4), we can recover the cyclic dynamics with a 24 h period (see Fig. 4) (although after a long transient phase of 100 h).

3.2 Natural clocks

We can apply our results to analyze the eukaryotic circadian systems. Previous studies [4, 5, 8] have contributed to elucidate the structure for these clocks. Here, we have studied mammalian circadian clocks involving genes *Clock*, *BmalI*, *Rev-Erbα*, *Per*, and *Cry* [26] (see Fig. 5). The model uses the complex *BmalI:Clock* as the transcriptional regulator,

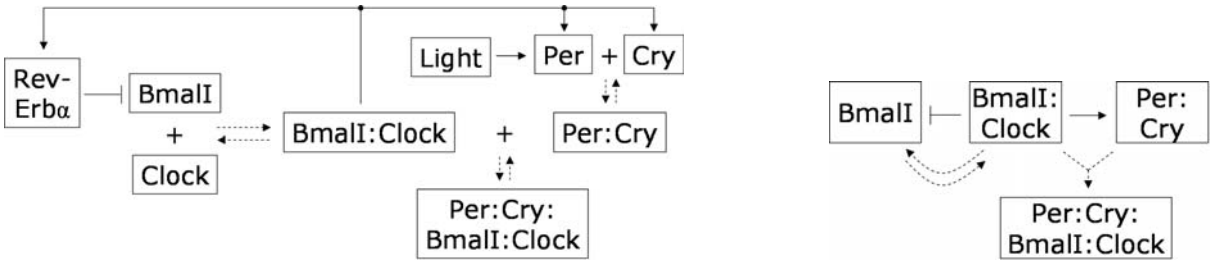


Fig. 5 Model of gene network for mammalian circadian clocks containing both, transcriptional and post-translational regulation. On the right we show its corresponding reduced model.

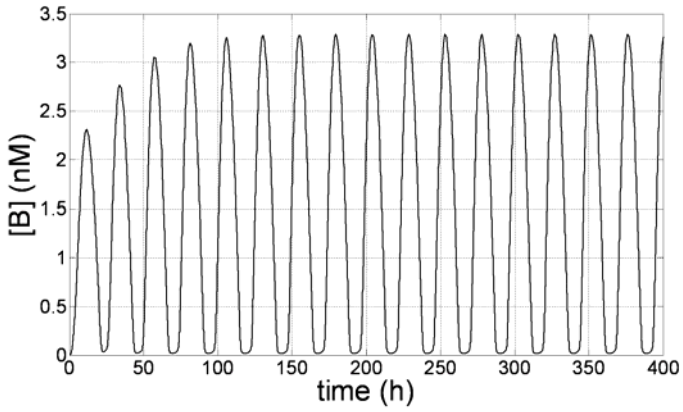


Fig. 6 Protein B dynamics corresponding to a natural circadian clock (see circuit (d) in Fig. 9). Parameter values are as follows: $\alpha_A = 2.083$ nM/h, $K_{AB} = 0.5$ nM, $n_{AB} = 9$, $\delta_A = 0.17$ h^{-1}, $\gamma_A = 0$ nM/h, $\lambda_A = 1.04$ h^{-1}, $\beta_A = 0.104$ h^{-1}, $\beta_B = 0.146$ h^{-1}, $\nu_1 = 0.04$ nM^{-1}h^{-1}, $\nu_2 = 0.01$ h^{-1}, $\alpha_C = 0.0166$ nM/h, $K_{CB} = 0.05$ nM, $n_{CB} = 6$, $\delta_C = 0.0304$ h^{-1}, $\gamma_C = 0.0001$ nM/h, $\lambda_C = 2.7645$ h^{-1}, $\beta_C = 0.0009$ h^{-1}, $\beta_{B:C} = 0.0092$ h^{-1}, $\theta_1 = 92.15$ nM^{-1}h^{-1}, and $\theta_2 = 0$ h^{-1}

acting positively on genes *Rev-Erbα*, *Per*, and *Cry*. Moreover, *Rev-Erbα* represses the transcription of gene *BmalI*). Light also activates gene *Per* transcription. In addition, this model considers the protein interactions between *BmalI* and *Clock*, between *Per* and *Cry*, and between *Per:Cry* and *BmalI:Clock*. For simplification purposes, we do not consider the duplication of genes such as *Per* or *Cry*. The previous studies have also unveiled that this model could have more than one possible circadian mechanism. Indeed, if we either remove the negative feedback on *BmalI* or stop the *Per* synthesis, the system is able to restore the oscillations with the same parameter values (solely in the second case increasing the degree of cooperativity of *BmalI:Clock*) [26]. On the other hand, in *Drosophila* the complex *Per-Tim* is the analogous of *BmalI:Clock* and it represses the transcription of genes *Per* and *Tim*. In *Neurospora* the repressive feedback is implemented by gene *Frq* [31].

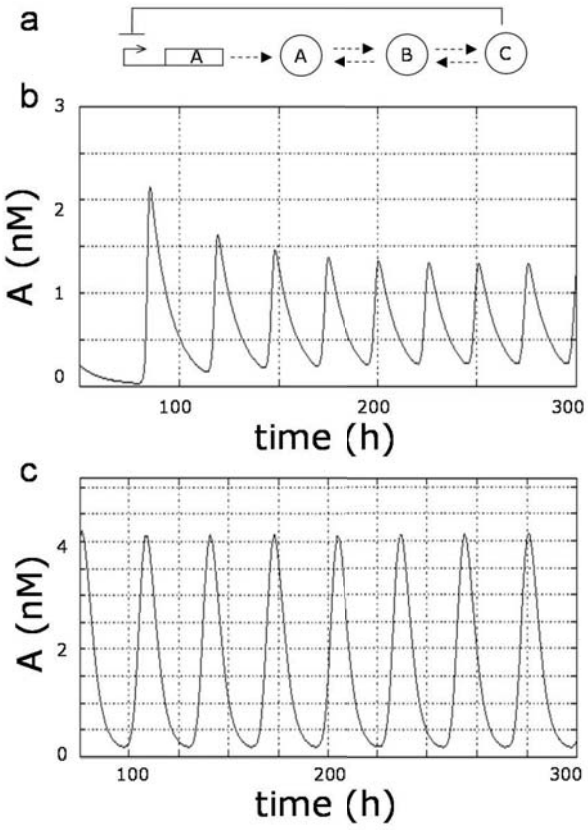


Fig. 7 (a) Genetic network for a reduced model of mammalian circadian clocks. (b) Dynamics for gene A. Parameter values are as follows: $\alpha_A = 8.66$ nM/h, $K_{AC} = 20$ nM, $n_{AC} = 40$, $\gamma_A = 0.0003$ nM/h, $\theta = 0.11$ h^{-1}, $\mu = 0.053$ h^{-1}, $\delta = 0.11$ h^{-1}, $\sigma = 0.053$ h^{-1}, $\beta_B = 0.053$ h^{-1}, ratio cytoplasmic / nuclear volume $W = 100$. (c) Dynamics for gene A considering modeling two phosphorylation steps together with Michalis-Menten rates. Parameter values are as follows: $\alpha_A = 3$ nM/h, $K_{AC} = 2$ nM, $n_{AC} = 10$, $\gamma_A = 0.001$ nM/h, $\theta_1 = 1.6$ nM/h, $K_{\theta_1} = 1.5$ nM, $\mu_1 = 0.8$ nM/h, $K_{\mu_1} = 2$ nM, $\theta_2 = 1.6$ nM/h, $K_{\theta_2} = 1.5$ nM, $\mu_2 = 0.8$ nM/h, $K_{\mu_2} = 2$ nM, $\beta = 1.6$ nM/h, $K_\beta = 1.5$ nM, $\delta = 0.6$ h^{-1}, $\sigma = 0.3$ h^{-1}.

We can reduce the previous mammalian model with the aim of identifying the core responsible for the oscillatory behavior. *BmalI* together with *Clock* (which is assumed constitutive) form the complex *BmalI:Clock*, which indirectly represses the *BmalI* synthesis (via *Rev-Erbα*). In addition, this complex also activates the transcription of genes *Per* and *Cry*. Those two proteins bind to form a complex *Per:Cry*, which binds to *BmalI:Clock* to obtain the macro-complex *Per:Cry:BmalI:Clock*. Therefore, we could represent our model by activator-repressor circuit d in Fig. 9, where the labels A, B and C would correspond to *BmalI, BmalI:Clock* and *Per:Cry* respectively (see dynamics in Fig. 6).

In Fig. 7a we have further reduced the model by taking into account the equilibrium reactions between the double phosphorylated, phosphorylated and non-phosphorylated species. In addition, for eukaryotic cells, when the complex goes into the nucleus its

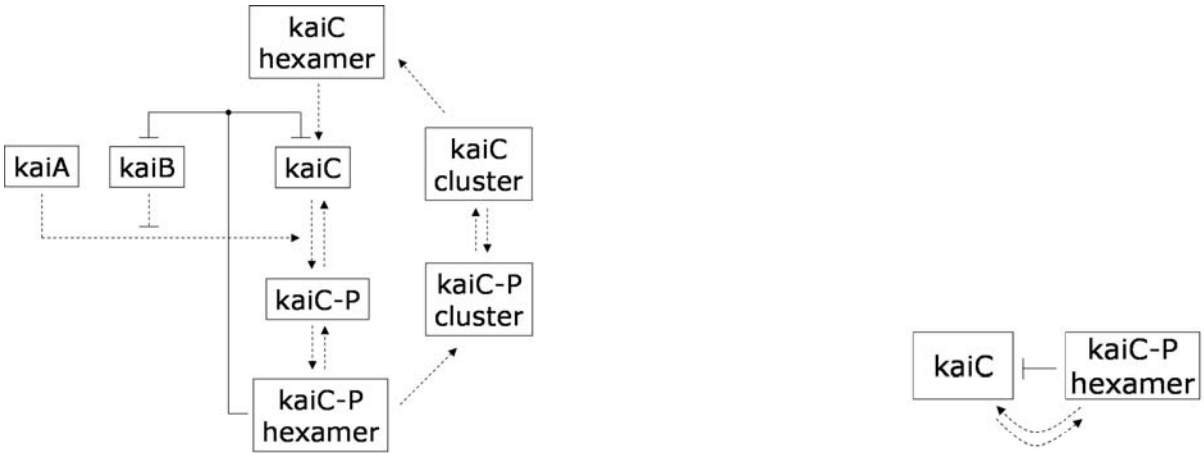


Fig. 8 Corresponding gene network model for cyanobacterial circadian clocks. We show its reduced model on the right.

concentration is automatically increased due to the reduced volume effect. Only the fully phosphorylated complex is degraded. According to that, we can proposed a simplified model following the equations

$$\begin{aligned} \frac{d}{dt}[A] &= \alpha_A \frac{1}{1+(\frac{[C]}{K_{AC}})^{n_{AC}}} + \gamma_A - \theta[A] + \mu[B], \\ \frac{d}{dt}[B] &= \theta[A] - \mu[B] - \beta_B[B] - \delta[B] + \sigma\frac{1}{W}[C], \\ \frac{d}{dt}[C] &= \delta W[B] - \sigma[C], \end{aligned} \tag{11}$$

where θ and μ are the kinetic rates of phosphorylation and dephosphorylation, δ and σ the kinetic rates in the cytoplasmic-nuclear equilibrium. W is the volume ratio between the cytoplasm and the nucleus which is assumed equal to 100. Thus, with this model the system is capable of exhibiting a 24 h rhythm (see Fig. 7b). The model could be improved by considering the two phosphorylation steps. Furthermore, this could also be approached by using Michaelis-Menten functions instead of kinetic rates (i.e., using $\frac{\theta[A]}{K+[A]}$ instead of $\theta[A]$). We show the corresponding results in Fig. 7c), where similar results are found. One remarkable thing is that the value of the degree of cooperativity (the Hill coefficient n) is considerably decreased when we perform these two improvements. We could reduce more this value by taken into account the mRNA dynamics, but this goes beyond the scope of this work.

On the other hand, it has been shown that some kinds of bacteria, concretely cyanobacteria, have a certain circadian rhythm [3]. Importantly, this rhythm is associated to the photosynthesis process [32]. The system is formed by a cluster of three genes (*kaiA*, *kaiB*, *kaiC*) [33] (see Fig. 8), containing two promoters: one upstream of gene *kaiA*, and another upstream of polycistronic genes *kaiB* and *kaiC*. Recent works have demonstrated that the underlying oscillatory mechanism is only based on protein-protein interactions (i.e., possible transcriptional regulations are not the principal cause of the rhythm) [27]. As cyanobacteria are not very complex organisms (they are one of the oldest on the

Earth [34]), the oscillatory mechanism is not highly complicated and it involves fewer genes than in higher organisms, such as *Drosophila*, *Neurospora* or mammals. Some mathematical models have been proposed [35], with monomer exchange during the day and cluster formation during the night. Here, we propose a model for primitive cyanobacteria and even other unknown organisms with circadian rhythms. The model is based on a transcriptional self-repression by gene *kaiC* with intermediate phosphorylated proteins, similarly to the proposed above for the reduced mammalian clock (although in this case there is no nucleus).

3.3 Genetic evolution *in silico*

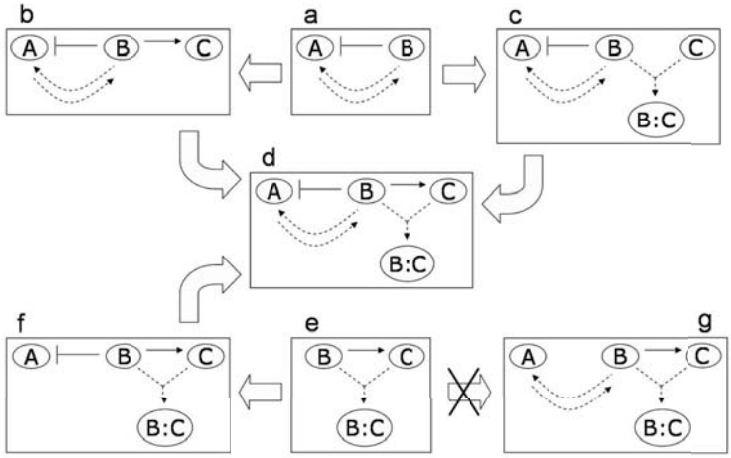


Fig. 9 Possible mechanism for the evolution of circadian clocks if we impose that a given network should not loose its 24 h dynamics during evolution. Circuit (*d*) is found in living systems, and could be modeled as the assembly of circuit *a* and *e* (both of them are able to oscillate separately). This show different evolutive pathways, where we have assumed that multiple mutations preserving the period will not occur simultaneously but in a gradual way. Circuit *g* is not able to sustain oscillations (see Fig. 11).

In Fig. 9 we analyze the different pathways to evolve towards a natural clock by starting from basic oscillatory circuits *a* and *e*. The resulting network is the circuit *d*, which corresponds to the coupling of both circuits. First, we have used our optimization procedure with both circuits (using Eq. 5 as fitness function) to obtain the required parameters for having a 24 h oscillation, (see Figs. 1 and 2). Using the parameters from circuits *a* and *e*, we have simulated the model *d* in Fig. 9 and we have obtained sustained oscillations of 24 h period. For instance, from circuit *a* to *d* there are two possible pathways. The first one consists of introducing a transcriptional activation to a new gene (that we call *c*), while maintaining all the kinetic parameters. This new network *b* maintains the same oscillatory behavior. The second path introduces firstly a post-translational sequestration (assuming a constitutive gene *c*). The dynamics of the

corresponding circuit c (maintaining all the kinetic parameters) has a oscillation but with different period of around 30 h (see Fig. 10). On the other hand, we can do the same analysis to explore the evolution of circuit e to d. Introducing a transcriptional repression on gene A (case f in Fig. 9), it maintains the same oscillatory behavior. However, when we add the equilibrium reaction between proteins A and B the new circuit g loses its oscillatory behavior (see Fig. 11).

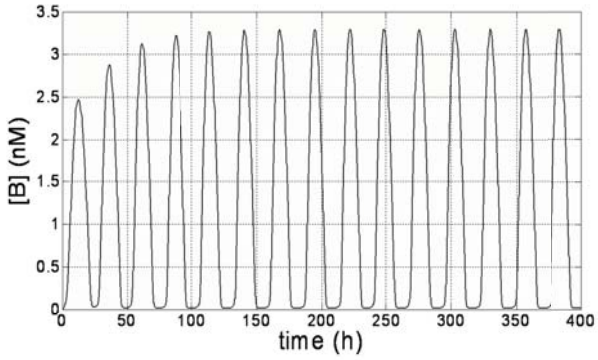


Fig. 10 Protein B dynamics corresponding to circuit (c) in Fig. 9. This corresponds to circuit c once we have removed the transcriptional activation on gene C from B while maintaining all other kinetic parameters (see caption of Fig. 6). The oscillatory behavior is maintained.

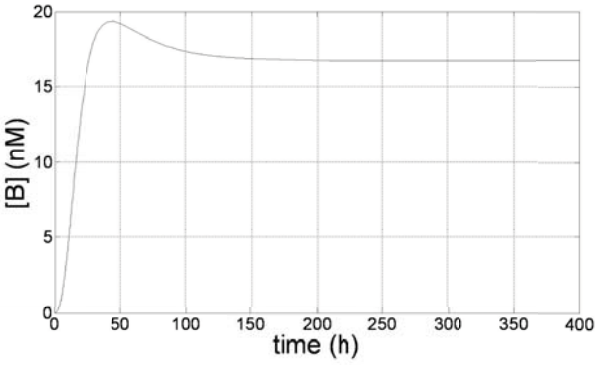


Fig. 11 Protein B dynamics corresponding to circuit (g) in Fig. 9. This corresponds to circuit e once we have removed the transcriptional repression on gene A from B while maintaining all other kinetic parameters (see caption of Fig. 6). We observe the lost of oscillatory behavior when we evolve circuit e to circuit g.

We have studied the possible network evolution of circuit type 2 (VKBL oscillator) to type 6, and vice-versa, in Table 2. With a certain set of parameters (Ω_1, see caption of Fig. 4) the circuit type 2 has a circadian period [30]. If some mutations occur that would change the parameters of the circuit (Ω_2, see caption of Fig. 13) but maintaining its topology, the system would remain oscillatory for only a certain time, particularly up to 400 h (see Fig. 12a). During that time, the system has the possibility of introducing a self-activation on the repressor gene in order to recover a sustainable periodic behavior.

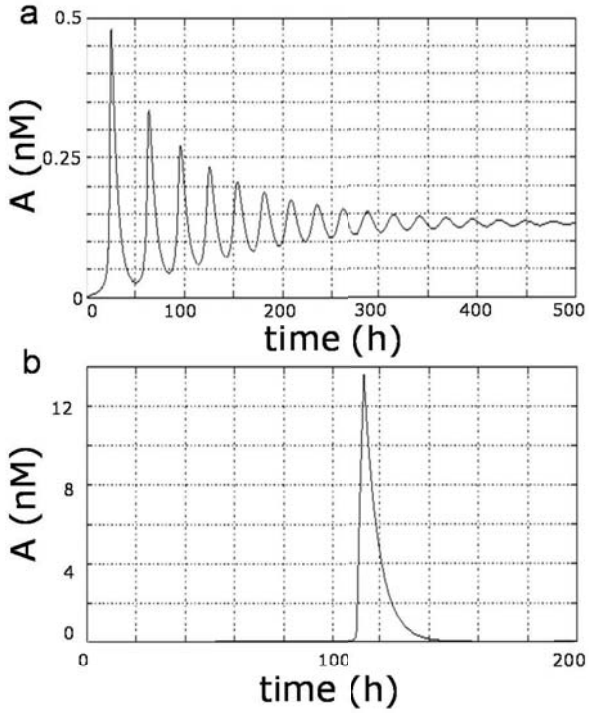


Fig. 12 Gene A dynamics for two types of transcriptional networks: (a) Type 2 see Table 2), where the kinetic parameter are shown in the caption of Fig. 13. (b) Type 6, where the kinetic parameter are shown in the caption of Fig. 4 (for the self-activation loop on gene R see the caption of Fig. 13)

Like that, the system has evolved to a type 6-topology with the set of parameters Ω_2 which is a circadian clock (see Fig. 13). Following that, we have constructed an evolutionary scenario whose mechanism comprises two evolution steps. Nevertheless, we could think in inverting these two steps. If we firstly mutate the topology the circuit does not behave as a clock, not even for a short time (see Fig. 12b).

It is interesting to see what kind of two-gene circadian circuit we could obtain by using our computational procedure. This resulting circuit is classified as type 6 in Table 2. Its dynamics is shown in Fig. 13. We notice that it has the same topology as the transcriptional VKBL oscillator, except for an additional self-activation of gene R. The corresponding model is given by

$$\begin{aligned}\frac{d}{dt}[A] &= \alpha_A \frac{(\frac{[A]}{K_{AA}})^{n_{AA}}}{1+(\frac{[A]}{K_{AA}})^{n_{AA}}} \frac{1}{1+(\frac{[R]}{K_{AR}})^{n_{AR}}} - \beta_A[A] + \gamma_A,\\ \frac{d}{dt}[R] &= \alpha_R \frac{(\frac{[A]}{K_{RA}})^{n_{RA}}}{1+(\frac{[A]}{K_{RA}})^{n_{RA}}} \frac{(\frac{[R]}{K_{RA}})^{n_{RA}}}{1+(\frac{[R]}{K_{RA}})^{n_{RA}}} - \beta_R[R] + \gamma_R.\end{aligned} \tag{12}$$

On the other hand, to study the evolvability of circuits, we simulated the transcriptional VKBL oscillator using the parameter values from our optimized circuit (see caption of Fig. 13). We found that although the system had a transient oscillatory behavior, it reached a steady state (see Fig. 12a).

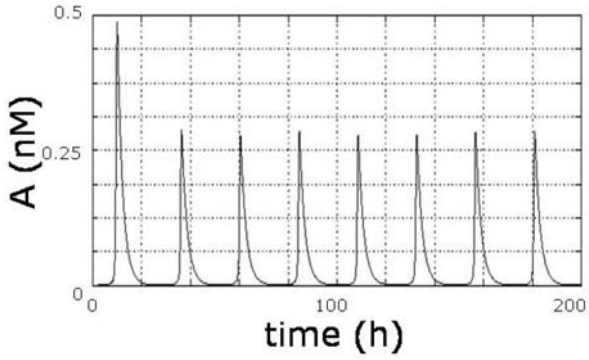


Fig. 13 Gene A dynamics corresponding to the transcriptional network of type 6 in Table 2. We have obtained this circuit by evolution *in silico*, where we have obtained the following parameter values: $\alpha_A = 9.18$ nM/h, $K_{AA} = 1.1$ nM, $n_{AA} = 1.6$, $K_{AR} = 0.52$ nM, $n_{AR} = 2.75$, $\beta_A = 0.459$ h^{-1}, $\gamma_A = 0.0027$ nM/h, $\alpha_R = 13.07$ nM/h, $K_{RA} = 1$ nM, $n_{RA} = 2$, $K_{RR} = 1.5$ nM, $n_{RR} = 1.2$, $\beta_R = 0.27$ h^{-1}, $\gamma_R = 0.0108$ nM/h.

4 Discussion

We have analyzed possible models for the origin of circadian clocks. They are composed of a core of two molecular elements, one assuming the role of activation and the other of repression [1]. We have aimed to determine possible mechanisms for their evolution by considering their dynamics and assuming a selective pressure for 24 h oscillations. We should point out that our study assumes constant conditions, where the oscillations would arise endogenously. Although circadian clocks are coupled with daily oscillations, our study focuses in the endogenous mechanisms.

We have considered transcriptional and post-translational regulations. We have distinguished between circuits with fast and slow mRNA dynamics, according to the ratio between mRNA and protein degradation (i.e., $r = \delta/\beta$). For the first case ($r \gg 1$), we have found that only the circuits having genes with self-activation are able to behave rhythmically (see Figs. 2 and 3). Thus, a transcriptional core (type 1 in Table 2) is not capable of maintaining oscillations by itself. However, if we introduce a self-activation on the activator gene (type 2 in Table 2), we obtain an oscillatory behavior. The circuit behaves as an oscillator for an appropriate choice of kinetic parameters [17]. We have obtained the same conduct when we considered protein sequestration instead of transcriptional repression (types 10 and 11 in Fig. 3) [18]. In addition, we have applied our combinatorial optimization procedure, which samples the transcriptional network space, to obtain clock circuits having a 24h-period (see Fig. 13). Our computational procedure did not consider mRNA dynamics (assumed to be fast) and the resulting circuits had indeed self-activations. In fact, we obtained the circuit of type 6 in Table 2, which was more robust to keep the circadian rhythmicity than the circuit considered in [17]. We notice that its Jacobian trace expression incorporates a new positive term due to the self-activation of the repressor element. This suggests that the self-activation incorporation is a possible evolutionary mechanism to get an oscillatory behavior.

On the other hand, when we consider $r \sim 1$, circuits that could not have a rhythm when $r \gg 1$ can now show oscillations (see Table 1). Now, we do not need to have self-activations to have a periodic behavior. We could ask whether organisms had evolved by incorporating a new self-activation or by stabilizing mRNA (making $r \sim 1$). Certainly, mRNA is degraded more quickly than proteins. In this case, the organism may get a ratio $r \sim 1$, because transcriptional activations are difficult to obtain spontaneously as the outcome of a random process. Furthermore, self-activated genes are not usually found in natural clocks models. Nevertheless, additional studies in other organisms are required to clarify this point.

In addition, we have modeled a system based on the first circuit in Fig. 1, where we have introduced a new intermediate specie B (see Fig. 7*a*) and we have assumed $r \gg 1$. This circuit is able to oscillate thanks to the delay introduced by the new specie. It could be understood as a double phosphorylation of protein A ($B = A : P$ and $C = A : PP$, which is the active form), or as a single phosphorylation and an insertion in the cellular nucleus (for eukaryotes) [28]. We have simulated this by using mass-action (see Fig. 7*b*) and Michaelis-Menten (see Fig. 7*c*) kinetics. This phosphorylation process is key to get circadian behavior because it introduces the required delay to produce oscillations.

We are interested in revealing the evolutionary mechanism by which circadian clocks have obtained their networks. Our results let us to conjecture that natural circadian circuits could have evolved by starting from a minimal two-gene activation-repression network, after which they would have acquired new interactions through evolution. It is difficult to evaluate the *in vivo* relevance of our kinetic parameters due to the fact that we use an effective model with a reduced number of species. Still, it provides a possible mechanism for the dynamical behavior although it is not necessary the only one. In all cases investigated so far, the circadian mechanism relies on a negative self-regulation of gene expression [1] (see Figs. 5 and 8). Previous studies in *Drosophila*, *Neurospora* and mammals have shown similar circadian clocks [26]. In mammals, the networks are more complex due to gene duplication, although the feedback structure is the same [25]. It has been shown that this gene duplication improves the rhythmic robustness [38]. We have seen how we could use a targeted dynamics to construct a fitness function that allowed us to analyze the possible networks showing an oscillatory behavior. We used it to explore the possible evolutionary pathways that would give us a circadian clock as found in living organisms.

Acknowledgments

We thank J. Salgado for discussions. This work was supported by the EU grant FP6-NEST 043340.

References

[1] J.C. Dunlap: "Molecular bases for circadian clocks", *Cell*, Vol. 96, (1999), pp. 271–290.

[2] C.F. Ehret and J.J. Wille: "Photobiology of microorganisms", Halldal, P. (Ed.), Wiley, New York (USA), 1970, pp. 369–416.

[3] C.H. Johnson, S.S. Golden, M. Ishiura and T. Kondo: "Circadian clocks in prokaryotes", *Mol. Microbiol.*, Vol. 21, (1996), pp. 5–11.

[4] M.W. Young and S.A. Kay: "Time zones: a comparative genetics of circadian clocks", *Nat. Rev. Genet.*, Vol. 2, (2001), pp. 702–715.

[5] J.J. Loros and J.C. Dunlap: "Genetic and molecular analysis of circadian rhythms in Neurospora", *Annu. Rev. Physiol.*, Vol. 63, (2001), pp. 757–794.

[6] T. Mori and C.H. Johnson: "Circadian programming in cyanobacteria", *Semin. Cell Dev. Biol.*, Vol. 12, (2001), pp. 271–278.

[7] L.C. Roden and I.A. Carre: "The molecular genetics of circadian rhythms in Arabidopsis", *Semin. Cell Dev. Biol.*, Vol. 12, (2001), pp. 305–315.

[8] S. Reppert and D. Weaver: "Coordination of circadian timing in mammals", *Nature*, Vol. 418, (2002), pp. 935–941.

[9] E. Herzog, J. Takahashi and G. Block: "Clock controls circadian period in isolated suprachiasmatic nucleus neurons", *Nature Neurosci.*, Vol. 1, (1998), pp. 708–713.

[10] A. Millar: "Molecular intrigue between phototransduction and the circadian clock", *Ann. Bot.*, Vol. 81, (1998), pp. 581–587.

[11] M.B. Elowitz and S. Leibler: "A synthetic oscillatory network of transcriptional regulators", *Nature*, Vol. 403, (2000), pp. 335–338.

[12] T.S. Gardner, C.R. Cantor and J.J. Collins: "Construction of a genetic toggle switch in E. Coli", *Nature*, Vol. 403, (2000), pp. 339–342.

[13] A. Jaramillo, L. Wernisch, S. Hery and S.J. Wodak: "Folding free energy function selects native-like protein sequences in the core but not on the surface", *Proc. Natl. Acad. Sci. USA*, Vol. 99, (2002), pp. 13554–13559.

[14] S. Basu, Y. Gerchman, C.H. Collins, F.H. Arnald and R. Weiss: "A synthetic multicellular system for programmed pattern formation", *Nature*, Vol. 434, (2005), pp. 1130–1134.

[15] J. Hasty, D. McMillen and J.J. Collins: "Engineered gene circuits", *Nature*, Vol. 420, (2002), pp. 224–230.

[16] P. Francois and V. Hakim: "Design of genetic networks with specified functions by evolution in silico", *Proc. Natl. Acad. Sci. USA*, Vol. 101, (2004), pp. 580–585.

[17] M.R. Atkinson, M.A. Savageau, J.T. Myers and A.J. Ninfa: "Development of genetic circuit exhibiting toggle switch or oscillatory behavior in Escherichia Coli", *Cell*, Vol. 113, (2003), pp. 597–607.

[18] N. Barkai and S. Liebler: "Circadian clocks limited by noise", *Nature*, Vol. 403, (1999), pp. 267–268.

[19] P. Francois and V. Hakim: "Core genetic module: the mixed feedback loop", *Phys. Rev. E*, Vol. 72, (2005), art. 31908.

[20] N. Monk: "Oscillatory expression of Hes1, p53, and NF-kB driven by transcriptional time delays", *Curr. Biol.*, Vol. 13, (2003), pp. 1409–1413.

[21] H. El-Samad, M. Khammash, L. Petzold and D. Gillespie: "Stochastic modelling of gene regulatory networks", *Int. J. Robust Nonlinear Control*, Vol. 15, (2005), pp. 691–711.

[22] D. Gillespie: "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", *J. Comput. Phys.*, Vol. 22, (1976), pp. 403–434.

[23] S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi: "Optimization by simulated annealing", *Science*, Vol. 220, (1983), pp. 671–680.

[24] V. Chickarmane, S.R. Paladugu, F. Bergmann and H.M. Sauro: "Bifurcation discovery tool", *Bioinformatics*, Vol. 21, (2005), pp. 3688–3690.

[25] D.B. Forger and C.S. Peskin: "A detailed predictive model of the mammalian circadian clock", *Proc. Natl. Acad. Sci. USA*, Vol. 100, (2003), pp. 14806–14811.

[26] J.C. Leloup and A. Goldbeter: "Toward a detailed computational model for the mammalian circadian clock", *Proc. Natl. Acad. Sci. USA*, Vol. 100, (2003), pp. 7051–7056.

[27] Y. Xu, T. Mori and C.H. Johnson: "Cyanobacterial circadian clockwork: roles of KaiA, KaiB and the kaiBC promoter in regulating KaiC", *EMBO J.*, Vol. 22, (2003), pp. 2117–2126.

[28] D. Gonze, J. Halloy and A. Goldbeter: "Robustness of circadian rhythms with respect to molecular noise", *Proc. Natl. Acad. Sci. USA*, Vol. 99, (2002), pp. 673–678.

[29] W.P. Dayawansa and C.F. Martin: "Stability of phase locked oscillations in the circadian clock", *Proc. Am. Control Conf.*, Vol. Arlington (USA), (2001), pp. 252–256.

[30] J.M.G. Vilar, H.Y. Kueh, N. Barkai and S. Leibler: "Mechanisms of noise-resistance in genetic oscillators", *Proc. Natl. Acad. Sci. USA*, Vol. 99, (2002), pp. 5988–5992.

[31] U. Albretch: "Regulation of mammalian circadian clock genes", *J. Appl. Physiol.*, Vol. 92, (2002), pp. 1348–1355.

[32] S. Golden, M. Ishiura, C.H. Johnson and T. Kondo: "Cyanobacterial circadian rhythms", *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, Vol. 48, (1997), pp. 327–354.

[33] M. Ishiura, S. Kutsuna, S. Aoki, H. Iwasaki, C.R. Andersson, A. Tanabe, S.S. Golden, C.H. Johnson and T. Kondo: "Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria", *Science*, Vol. 281, (1998), pp. 1519–1523.

[34] B.A. Whitton: "Survival and dormancy of microorganisms", Hennis, Y. (Ed.), Wiley, New York (USA), 1987, pp. 109–167.

[35] E. Emberly and N.S. Wingreen: "Hourglass model for a protein-based circadian oscillator", *Phys. Rev. Lett.*, Vol. 96, (2006), pp. 38303 1–4.

[36] H. Iwasaki, T. Nishiwaki, Y. Kitayama, M Nakajima and T. Kondo: "KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria", *Proc. Natl. Acad. Sci. USA*, Vol. 99, (2002), pp. 15788–15793.

[37] S. Ohno: "Evolution by Gene Duplication", (Ed.), Springer, Berlin (Germany), 1970, pp. 1–160.

[38] D.B. Forger and C.S. Peskin: "Stochastic simulation of the mammalian circadian clock", *Proc. Natl. Acad. Sci. USA*, Vol. 102, (2005), pp. 321–324.