Supporting Information

Delay-Managed tradeoff in molecular dynamics of segmentation clock

Henglin Song, Zhanjiang Yuan and Tianshou Zhou

1 Introduction

This supporting information contains three sections with details on the mathematical model for the pace-making circuit of segmentation clock, parameter value/range selection, numerical methods. In the section on the mathematical model, we provide details about modeling the dynamics of Hes7 mRNA and protein. In the section on parameter selection, we describe how we In the section on numerical methods, we give details on the

2 Mathematical models of pace-making circuit

2.1 Deterministic Models

We derive here the mathematical models used in the main text. Using the standard chemical kinetics and the assumption for the existence of two types of reactions: fast and slow, we introduce the following chemical species: \( P \), repressor monomer, \( P_2 \), repressor dimer, \( D_0 \), repressor promoter region, and \( D_1, D_2, D_3 \), the state of the three binding sites with the respective bound proteins. The system size is taken into account by introducing a parameter \( \Omega \) (including cell volume and Avogadro’s number).

Scenario I

Fast equilibrium reactions. We assume that DNA-binding reactions and dimerization are fast.

\[
\begin{align*}
D_0 + P_2 & \xrightarrow{k_{+1}} D_1 \\
D_1 + P_2 & \xrightarrow{k_{+2}} D_2 D_1 \\
D_2 D_1 + P_2 & \xrightarrow{k_{+3}} D_3 D_2 D_1 \\
P + P & \xrightarrow{k_{-4}} P_2
\end{align*}
\]

Slow reactions. Transcription, translation and degradation of the repressor are considered as slow reactions.
Fast reactions are assumed to be in equilibrium. We can thus introduce the corresponding dissociation constants for DNA-binding ($K_1 = [D_0][P_2]/[D_1] = k_{-1}/k_{+1}$, $K_2 = [D_1][P_2]/[D_2D_1] = k_{-2}/k_{+2}$, $K_3 = [D_2D_1][P_2]/[D_3D_2D_1] = k_{-3}/k_{+3}$) and dimerization ($K_4 = [P]_2/[P_2]_2 = k_{-4}/k_{+4}$). Defining concentrations as our dynamical variables, $m = [M]$, $p = [P]$, $p_2 = [P_2]$, $d_0 = [D_0]$, $d_1 = [D_1]$, $d_2 = [D_2D_1]$, and $d_3 = [D_3D_2D_1]$, we can write a 2-dimensional deterministic model as

$$
\frac{d [M]}{dt} = \alpha_0 [D_0] + \alpha_r [D_1] + \alpha_r [D_2D_1] + \alpha_r [D_3D_2D_1] - \gamma_m [M]
$$

$$
\frac{d [P]}{dt} = \alpha_p [M] - \gamma_p [P] - 2\sigma \gamma_p [P_2]
$$

Consider that the total concentration of promoter sites $d_t$ is a constant ($d_t = [D_0] + [D_1] + [D_2D_1] + [D_3D_2D_1]$), take the fast reactions to be in equilibrium, and making use of the above reactions and assumptions. We next simplify these equations by introducing dimensionless time and concentrations as follows:

$$
\tilde{t} = \gamma_m t
$$

$$
\tilde{m} = \frac{\alpha_p [M]}{\gamma_p}
$$

$$
\tilde{p} = [P]
$$

The time is then measured in units of mRNA decay. Upon substitution into Eq. (4), the previous 2-dimensional model becomes

$$
\frac{dm}{dt} = \alpha_0 + \frac{\alpha}{1 + (p/p_0)^2 + r_2 (p/p_0)^4 + r_2 r_3 (p/p_0)^6} - m
$$

$$
\frac{dp}{dt} = \beta m - \beta p - 2r_1 \beta p^2
$$
with the definitions

\[
\alpha_0 = \frac{\alpha_r \alpha_p d_T}{\gamma_m \gamma_p}, \quad \alpha = \frac{(\alpha_b - \alpha_r) \alpha_p d_T}{\gamma_m \gamma_p}
\]

\[
\beta = \frac{\gamma_p}{\gamma_m}, \quad r_1 = \frac{\sigma}{K_4}
\]

\[
r_2 = \frac{K_1}{K_2}, \quad r_3 = \frac{K_1}{K_3}
\]

\[
P_0 = \sqrt{K_1 K_4}
\]

(7)

Here we have suppressed the overbar on \( \tilde{t}, \tilde{m}, \) and \( \tilde{p} \). It can be proved from Bendixson’s Negative Criterion, however, that it is impossible for this pair of differential equations to generate sustained oscillations. This conclusion holds for any form of the function \( f(p) \), provided only that \( \beta \) is a positive number.

**Deterministic model I with delay**

The requirement of the somewhat mysterious third state variable \( z \) for sustained oscillations in the ODE model studied by Hirata et al. leads to the introduction of delay differential equation (DDE) models of the cellular oscillator. These models introduced biologically realistic transcription, translation, and transport delays into the production terms of protein and mRNA, which allowed sustained oscillations in a dynamical system with only two dependent variables and a reduced number of model parameters. Models with negative feedback and time delay have arguably become the most prominent ones of oscillatory gene expression in somitogenesis. We introduce a delayed model to track both mRNA and protein levels of a single clock gene. In this model, the clock protein is assumed to form a homodimer that represses its mRNA production after a delay. Three main sources of delay are the transcription and post-transcriptional processing of mRNA, the translation delay in protein production, and the transport of molecules between the nuclear and cytosolic compartments within a cell or between the cytosolic compartments of adjacent cells. This system is capable of autonomous, generating sustained oscillations of gene expression. The relative amounts of clock protein monomers, homodimers with the control protein are explicitly tracked, allowing different decay rates for each.

Absorbing all the delays (including transcription and translation) into one delay, \( \tau \), the model with delay is represented by the following DDE:

\[
\frac{dm}{dt} = \alpha_0 + \frac{\alpha}{1 + (p(t-\tau)/P_0)^2 + r_2 (p(t-\tau)/P_0)^4 + r_2 r_3 (p(t-\tau)/P_0)^6} - m(t)
\]

\[
\frac{dp}{dt} = \beta m(t) - \beta p(t) - 2r_1 \beta p(t)^2
\]

(8)

**Scenario II**

Since clock protein is degraded by the ubiquitin-proteasome pathway, we include Michaelis-Menten kinetics for the ubiquitination of protein.
We proceed similarly as in scenario I. The only difference lies in the protein decay mechanism. In this case, the enzyme $E$ binds to clock protein and favors catalyzed protein decay. The degradation of clock protein is caused by ubiquitination. If we pool the three enzymatic steps into one enzymatic reaction, ubiquitination can be displayed by the following reaction scheme:

**fast equilibrium reactions:**

$$
E + P \xrightarrow{k_{+5}} EP
$$

$$
E + P_2 \xrightarrow{k_{+6}} EP_2
$$

**slow reactions:**

$$
EP \xrightarrow{\delta_1 \gamma_p} E
$$

$$
EP_2 \xrightarrow{\delta_2 \gamma_p} E
$$

We formulate a model by using a rate equation approach, which neglects fluctuations and use concentrations as our dynamical variables. Taking the fast reactions to be in equilibrium, we eliminate some variables and obtain: $[P_2] = [P]^2 / K_4$, $[PE] = [P][E] / K_5$, and $[P_2E] = [P_2][E] / K_6$.

The rate equations for the concentration of protein and mRNA are given by

$$
\frac{d[M]}{dt} = \alpha_0 [D_0] + \alpha_r [D_1] + \alpha_r [D_2D_1] + \alpha_r [D_3D_2D_1] - \gamma_m [M]
$$

$$
\frac{d[P]}{dt} = \alpha_p [M] - \gamma_p [P] - \delta_1 \gamma_p [PE] - 2\delta_2 \gamma_p [P_2E]
$$

We eliminate the fast variables, and use the fact that the total number of DNA-binding sites is conserved ($d_t = [D_0] + [D_1] + [D_2D_1] + [D_3D_2D_1]$). Assume that the total amount of enzyme $E$ is constant ($E_{tot} = [E] + [PE] + [P_2E]$), and nondimensionalized by setting $t = \gamma_m t$, $m = \frac{\alpha \gamma_p K_4}{\gamma_p [M]}$ and $p = [P]$. The dimensionless equation is then

$$
\frac{dm}{dt} = \alpha_0 + \frac{\alpha}{1 + (p/P_0)^2 + r_2 (p/P_0)^4 + r_2 r_3 (p/P_0)^6} - m
$$

$$
\frac{dp}{dt} = \beta m - \beta p - \beta \mu \left( 1 - \frac{1}{1 + p/K + r_4 (p/K)^2} \right)
$$

where
\[ \mu = \delta_1 E_{tot} \]
\[ K = K_5 \]
\[ r_4 = \frac{K_5}{K_4 K_6} \]
\[ \delta_1 = 2 \delta_2 \]

**Deterministic model II with delay**

Absorbing all the delays (including transcription and translation) into one delay, \( \tau \), the model with delay is represented by the following DDE:

\[
\frac{dm}{dt} = \alpha_0 + \frac{\alpha}{1 + (p(t-\tau)/P_0)^2 + r_2 (p(t-\tau)/P_0)^4 + r_2 r_3 (p(t-\tau)/P_0)^6} - m(t)
\]
\[
\frac{dp}{dt} = \beta m(t) - \beta p(t) - \beta \mu \left(1 - \frac{1}{1 + p(t)/K + r_4 (p(t)/K)^2}\right)
\]

**3 Methods**

**3.1 Steady states and local stability**

The local stability can be characterized by jacobian matrix of mathematical model equations. We present local stability analysis for model I, and local stability analysis for model II can be performed in a similar way.

The steady states \( E^* = (p^*, m^*) \) are given by solving

\[
\begin{align*}
\alpha_0 + f(p) - m &= 0 \\
m - p - 2r_1 p^2 &= 0 \\
\end{align*}
\]

where

\[
f(p) = \frac{\alpha}{1 + (p/P_0)^2 + r_2 (p/P_0)^4 + r_2 r_3 (p/P_0)^6}
\]

and \( p^* \) satisfies \( f(p) - p - 2r_1 p^2 + \alpha_0 = 0 \), which clearly has a unique positive solution. The characteristic equation of the linearized equation of (6) around \( E^* \) is given by

\[
\begin{pmatrix}
-1 & \frac{f'(p^*)}{\beta} \\
\beta & -\beta (1 + 4r_1 p^*)
\end{pmatrix}
\begin{pmatrix}
\lambda + 1 \\
\lambda + \beta (1 + 4r_1 p^*)
\end{pmatrix} - \beta f'(p^*) = 0
\]
which clearly has eigenvalues with negative real parts. Thus, $E^*$ is asymptotically stable. In other words, without delays, the model will not generate sustained oscillations.

**Deterministic model with delay**

The characteristic (eigenvalue) equation associated with this system of equations depends on the delay $\tau$,

$$(\lambda + 1) (\lambda + \beta (1 + 4r_1p^*)) - \beta f'_\tau (p^*) e^{-\lambda \tau} = 0 \quad (18)$$

There is a critical value for the delay $\tau$, denoted by $\tau_{\text{crit}}$, at which the system is destabilized and undergoes a Hopf bifurcation. At the Hopf bifurcation point, a pair of eigenvalues $\lambda$ has a zero real part, i.e., $\lambda = \pm i\omega$. The value $\omega$ gives the frequency of oscillation at the Hopf bifurcation point.

Assuming that $s = 1 + 4r_1p^*, c = f'_\tau (p^*), \lambda = u + i\omega$, we have

$$u^2 - \omega^2 + (1 + \beta s) u + \beta s - \beta ce^{-u\tau} \cos (\omega \tau) = 0$$
$$2u\omega + (1 + \beta s) \omega + \beta ce^{-u\tau} \sin (\omega \tau) = 0 \quad (19)$$

When $u = 0, \omega \neq 0$,

$$-\omega^2 + \beta s - \beta c \cos (\omega \tau) = 0$$
$$(1 + \beta s) \omega + \beta c \sin (\omega \tau) = 0 \quad (20)$$

Using the fact that $\sin^2 (\omega \tau) + \cos^2 (\omega \tau) = 1$, we get

$$\omega^4 + (1 + \beta^2 s^2) \omega^2 + \beta^2 s^2 - \beta^2 c^2 = 0$$
$$\omega^2 = \frac{1}{2} \left( - (1 + \beta^2 s^2) + \sqrt{(1 - \beta^2 s^2)^2 + 4\beta^2 c^2} \right) \quad (21)$$

Unstable behavior, if present, should exist only for finite values of the delay. If a stable domain exists, then at its borders we have $\Re \lambda = 0$ (generically $\Re \lambda$ crosses zero and changes sign). Placing ourselves at this very border we assume that $\lambda = i\omega$ with real $\omega$. Substituting the value of $\lambda$ into (18), we obtain an eigenvalue equation and separate it into real and imaginary parts. We solve these two equations for $\cos \omega \tau$ and $\sin \omega \tau$ and implement the relation $\cos^2 \omega \tau + \sin^2 \omega \tau = 1$. The latter factorises into polynomials in $\omega^2$. It turns out that the polynomial in $\omega^2$ always has one real positive root provided that $s^2 < c^2$. Thus, one can satisfy the compatibility between the two equations resulting from (18). Substituting this root back into the equation for, say, $\cos \omega \tau$, one obtains a constraint to the parameters of the equation for the existence of the instability. We can numerically solve this constraint equation by using the function $f\text{zero}$ of Matlab.
3.2 Analytical power spectra of cellular oscillator

Let us first analyze briefly the dynamics of the macroscopic system. Defining $M(t)$ and $P(t)$ as the concentration of mRNA and clock protein, respectively, the model presented in our work is then represented by the following ODE:

\[
\frac{dm(t)}{dt} = \alpha_0 + f(p(t)) - m(t) \\
\frac{dp(t)}{dt} = \beta m(t) - \beta g(p)
\]  

where

\[
f(p) = \frac{\alpha}{1 + (p/P_0)^2 + r_2 (p/P_0)^4 + r_3 (p/P_0)^6}
\]

and for Model I

\[
g(p) = p + 2r_1 p^2
\]

for Model II

\[
g(p) = p + \mu \left(1 - \frac{1}{1 + p/K + r_4 (p/K)^2}\right)
\]

Its equilibria are obtained by solving the system

\[
m = \alpha_0 + f(p) \\
m = g(p)
\]  

and their stability is studied by means of a linear stability analysis. For the linear form of model equations

\[
\frac{dm}{dt} = a_{11} m + a_{12} p \\
\frac{dp}{dt} = a_{21} m + a_{22} p
\]  

with $a_{11} = -1, a_{12} = f'(p^*), a_{21} = \beta, a_{22} = -\beta g'(p^*)$ in which $p^*$ is the solution of (26). The stochastic dynamics are fully described by a multivariated master equation. A formal large $\Omega$ expansion of this equation gives rise to the Fokker-Planck equation. To look for oscillations in the fluctuation case, it is easier to work with the equivalent Langevin equations:
\[
\frac{d\xi_m}{dt} = a_{11}\xi_m + a_{12}\xi_p + \kappa_m \\
\frac{d\xi_p}{dt} = a_{21}\xi_m + a_{22}\xi_p + \kappa_p
\]

(28)

These are a pair of differential equations which describe the stochastic behavior of the model at a large but finite \(\Omega\). The variables \(\xi_m\) and \(\xi_p\) are stochastic corrections to the deterministic system with the relation

\[
M(t) = \Omega m(t) + \Omega^{1/2}\xi_m \\
P(t) = \Omega p(t) + \Omega^{1/2}\xi_p
\]

(29)

And, the \(\kappa_m, \kappa_p\) are Gaussian white noises with zero mean and a correlation function defined in terms of a noise covariance matrix \(D_{ij}\), satisfying \(\langle \kappa_m(t) \kappa_m(t') \rangle = D_m \delta(t-t')\), \(\langle \kappa_p(t) \kappa_p(t') \rangle = D_p \delta(t-t')\), and \(\langle \kappa_m(t) \kappa_p(t') \rangle = D_{mp} \delta(t-t')\). The noise intensities are given by:

\[
D_m = \alpha_0 + f(p) + m \\
D_p = \beta m + \beta g(p) \\
D_{mp} = D_{pm} = 0
\]

(30)

The constants \(a_{ij}\) are exactly the same as coefficients found from linear stability analysis about the non-trivial fixed point of Eq. (26). The matrix \(D_{ij}\), which is responsible for generating the large-scale oscillations, can not be determined from Eq. (22) and is derived from the master equation using the Van Kampen expansion.

Since we are interested in cycles, it is natural to work in terms of \(\tilde{\xi}_m(\omega)\) and \(\tilde{\xi}_p(\omega)\) which are the Fourier transforms of \(\xi_m\) and \(\xi_p\), respectively. Furthermore, taking the Fourier transform of the equations (28) allows us to solve for these variables very simply. We can now calculate the power spectra of the fluctuations about the mean-field values for the cellular oscillator to be

\[
S_m(\omega) = \left\langle \left| \tilde{\xi}_m(\omega) \right|^2 \right\rangle = \frac{F_m + D_m\omega^2}{(\Delta - \omega^2)^2 + \omega^2T^2} \\
S_p(\omega) = \left\langle \left| \tilde{\xi}_p(\omega) \right|^2 \right\rangle = \frac{F_p + D_p\omega^2}{(\Delta - \omega^2)^2 + \omega^2T^2}
\]

(31)

Where \(\Delta = a_{11}a_{22} - a_{12}a_{21}, T = a_{11} + a_{22}\), the constants \(F_m\) and \(F_p\) are defined as \(F_m = D_m a_{22}^2 - 2D_{mp}a_{12}a_{22} + D_p a_{12}^2\), \(F_p = D_p a_{11}^2 - 2D_{mp}a_{21}a_{11} + D_m a_{21}^2\), respectively.
3.3 Analytical power spectra with delay

In the fixed-point regime of the mean-field dynamics, the delay Langevin equations read

\[
\frac{d\xi_m}{dt} = a_{11}\xi_m(t) + a_{12}\xi_p(t - \tau) + \kappa_m(t)
\]
\[
\frac{d\xi_p}{dt} = a_{21}\xi_m(t) + a_{22}\xi_p(t) + \kappa_p(t)
\]

with \(a_{11} = -1, a_{12} = f'(p^*)\), \(a_{21} = \beta, a_{22} = -\beta g'(p^*)\) and \(p^*\) is the fixed point. The variables \(\xi_m\) and \(\xi_p\) are stochastic corrections to the deterministic system with the relation

\[
M(t) = \Omega_m(t) + \Omega_1^{1/2}\xi_m
\]
\[
P(t) = \Omega_p(t) + \Omega_1^{1/2}\xi_p
\]

And, the \(\kappa_m, \kappa_p\) are Gaussian white noises with zero mean and a correlation function that is defined in terms of noise covariance matrix \(D_m, D_p\) satisfying

\[
\langle \kappa_m(t) \kappa_m(t') \rangle = [\alpha_0 + f(p^*) + m^*] \delta(t - t')
\]
\[
\langle \kappa_p(t) \kappa_p(t') \rangle = [\beta m^* + \beta g(p^*)] \delta(t - t')
\]
\[
\langle \kappa_m(t) \kappa_p(t') \rangle = 0
\]

Inverting in Fourier space one then finds after some algebraic manipulations

\[
S_m(\omega) = \left\langle |\tilde{\xi}_m(\omega)|^2 \right\rangle
\]
\[
= \frac{[D_{11}a_{22}^2 - 2D_{12}a_{12}(a_{22}\cos(\omega\tau) + \omega\sin(\omega\tau)) + D_{22}a_{12}^2 + D_{11}^2\omega^2]}{(a_{11}a_{22} - a_{12}a_{21}\cos(\omega\tau) - \omega^2)^2 + (a_{12}a_{21}\sin(\omega\tau) - \omega T)^2}
\]
\[
S_p(\omega) = \left\langle |\tilde{\xi}_p(\omega)|^2 \right\rangle
\]
\[
= \frac{[D_{22}a_{11}^2 - 2D_{21}a_{21}a_{11} + D_{11}a_{21}^2 + D_{22}\omega^2]}{(a_{11}a_{22} - a_{12}a_{21}\cos(\omega\tau) - \omega^2)^2 + (a_{12}a_{21}\sin(\omega\tau) - \omega T)^2}
\]

with \(D_{11} = \alpha_0 + f(p^*) + m^*, D_{22} = \beta m^* + \beta g(p^*) p^*, D_{12} = D_{21} = 0\).

**Quality of noise-induced oscillations:**

The quality factor, \(Q\), is a dimensionless parameter, which characterizes an oscillator’s bandwidth relative to its peak frequency,

\[
Q = \frac{\omega_p}{\Delta\omega}
\]
where $\omega_p$ is the peak frequency and $\Delta \omega$ is the bandwidth. A high $Q$ corresponds to oscillations of better regularity. We calculate the $Q$ for the diagonal entries of the PSD matrix. The bandwidth $\omega_p$ satisfies $\Delta \omega = |\omega_1 - \omega_2|$, $S(\omega_1) = S(\omega_2) = [S(\omega_p) - S(0)]/2$.

3.4 Noise evaluation in the phase angular

Let $M(t)$ and $P(t)$ represent the number of mRNA and clock protein and $(M_{eq}, P_{eq})$ be the steady state. We have

$$r = \sqrt{(M(t) - M_{eq})^2 + (P(t) - P_{eq})^2}$$

$$\phi(t) = \text{atan} (P(t) - P_{eq}, M(t) - M_{eq})$$

(37)

where $\phi(t)$ is the phase angle of oscillator. We define $\varphi$ as the prephase satisfying $\varphi = \text{mod} (\phi, 2\pi)$. Then we have

$$\sigma_r(\varphi) = \sqrt{\langle (r - \langle r \rangle)^2 \rangle (\varphi)}$$

$$\sigma_m(\varphi) = \sigma_r(\varphi) \left| \cos \varphi \right|$$

$$\sigma_p(\varphi) = \sigma_r(\varphi) \left| \sin \varphi \right|$$

$$\eta_r(\varphi) = \frac{\sigma_r(\varphi)}{\langle r \rangle}$$

$$\eta_m(\varphi) = \frac{\sigma_r(\varphi) \left| \cos \varphi \right|}{\langle M \rangle}$$

$$\eta_p(\varphi) = \frac{\sigma_r(\varphi) \left| \sin \varphi \right|}{\langle P \rangle}$$

(38)

(39)

The Delayed Gillespie Direct Method.

1. Initialize. Set model parameters and the system size $\Omega$. Input values for initial state $X = (X_1, \cdots, X_N)$, set time $t = 0$ and list of scheduled delay reaction to an empty list.

2. Compute propensities of $M$ reactions $a_{\mu}$, $\mu = 1, \cdots, M$.

3. Compute $a_0(X(t)) = \sum_{j=1}^{M} a_j(X(t))$.

4. Generate uniform random numbers $u_1, u_2 \in [0, 1]$.

5. Compute time to next reaction, $\Delta t = \frac{1}{a_0(X(t))} \ln (1/u_1)$.

6. Check whether there are delayed reactions scheduled within time interval $[t, t + \Delta t]$. If delayed reactions are scheduled within $[t, t + \Delta t]$, then time $t$ advances to the time $t_d$, and $X$ states are updated according to the delayed reaction channel $k$.

$$X(t_d) = X(t) + \nu_k$$

$$t = t_d$$
\textbf{elseif} there is no delayed reaction scheduled for the interval \((t, t + \Delta t]\), then find the channel of the next reaction \(j\) such that

\[
\sum_{\mu=1}^{j-1} a_\mu(X(t)) < u_2a_0(X(t)) \leq \sum_{\mu=1}^{j} a_\mu(X(t))
\]

(7) \textbf{If} the selected reaction \(j\) is not delayed, update \(X\) according to the reaction channel \(j\), \(X(t + \Delta t) = X(t) + \nu_j\).

\textbf{else}

record time, \(t_d = t + \Delta t + \tau\), for delayed reaction \(j\).

(8) Update time \(t = t + \Delta t\) and go to step (2).