A minimal model for the rhythmic protein per expression in Drosophila

DEGLI AGOSTI, Robert


Available at:
http://archive-ouverte.unige.ch/unige:42875

Disclaimer: layout of this document may differ from the published version.
A minimal model for the rhythmic protein *per* expression in *Drosophila*

R. DEGLI AGOSTI

Laboratory of Plant Biochemistry and Physiology, Plant Physiomatics, University of Geneva, 3 Place de l'Université, CH-1211 Genève 4, Switzerland. E-mail: Robert.DegliAgosti@bota.unige.ch

Introduction

Metabolic models with negative feedback regulation have been since a long time proposed to explain some rhythmic phenomena's [4,12]. They were temporarily discarded from consideration in circadian rhythms on the basis of theoretical grounds [see 2, for a discussion of this aspect]. However, it clearly appears now from experimental results obtained in the insect *Drosophila*, the fungus *Neurospora* and man that such kind of feedback's are involved in circadian rhythms or even constituting the core of biological clocks [1,2,5,8]. Among the current most studied systems is the *Drosophila* one. An important gene has been identified early and found to be closely involved in the timing mechanisms: the *per* gene. Its mRNA and the corresponding *per* protein oscillate in a circadian (~24 h) manner. Since then, other genes have also been identified [7,9,11]. We use here the concept of minimal modeling approach to model the *per* rhythmic circadian expression with the minimum number of variables (elements or state variables). We will use the method of metabolic network modeling with a negative feedback regulation [2].

The purpose of this paper is to illustrate how we can construct such models, starting from experimental data, to finally obtain coherent simulations on computers. This is not so trivial since from the experimental model (Fig.1) it is not intuitively clear how this can effectively lead to a correct simulation of the observed (measured) variables. In other words, how many people when observing Figure 1 could positively answer to the question: is this model oscillating with a circadian (~24 h) rhythmicity? To answer this question one must construct a mathematical model, obtain the evaluation of parameters and then simulate it. To be precise, we should mention here that methods do exist that extract values of parameters from raw data with just the mathematical model only (e.g. by non linear multiple least squares regression). This method will allow to estimate parameters where often not enough experimental data is available and is therefore very useful. However, the drawback of this is that the simulation will be a little biased to necessarily work with artificially estimated parameters. In this
paper we will follow the first method, namely obtain successively each parameters estimation from experimental data available.

1. The model

From an early "experimental" biological model presented in the literature for the *per* gene regulation in *Drosophila* the following network is presented in Figure 1.

![Figure 1. Early model of the *per* protein regulation involved in circadian rhythms in *Drosophila*. Very simplified metabolic network of the of protein *per* regulation according to [5]. RNA and *per* represent the relative concentrations of mRNA for the *per* protein and the protein *per* itself, respectively. RNA is synthesized with a basic flux (*ν₀*) controlled by the inhibition with the *per* protein (negative feedback *f(per)*). The RNA is degraded with a first order kinetic law (*k₁*). The synthesis of the protein *per* depends on the RNA concentration according to a first order kinetics reaction law (*k₂*), a delay has been experimentally observed in this reaction (~8 h). Finally the *per* protein is destroyed with a first order kinetics reaction law (*k₃*).](image)

The corresponding differential equations that allow the modeling and the simulation of the dynamics of the concentrations of the 2 substances may be written as follows:

\[
\begin{align*}
\frac{dRNA(t)}{dt} & = \nu_0 f(per(t)) - k_1 RNA(t) \\
\frac{dper(t)}{dt} & = k_2 RNA(t - delay) - k_3 per(t)
\end{align*}
\]

(1) (2)

2. Evaluation of the model parameters

The data which will allow us to evaluate the parameters of the models are available in the literature. We will describe this important aspect in details here for illustrating purposes. Data for RNA is found in [5] and for *per* protein in [13]. We have normalized the original values to "normalized units of concentrations (n.u.c)", by eliminating baselines and dividing each protein or mRNA *per* value by their respective maximal values. They are presented in Figure 2.
Minimal model for the rhythmic protein *per* expression in *Drosophila*

![Graph showing concentrations of per protein and per mRNA over time](image)

**Figure 2.** Experimental data expressed in normalized units of concentrations (n.u.c) of the *per* protein and *per* mRNA according to [13] and [6]. One might observe that the synthesis and degradation of *per* mRNA are triggered around a threshold of 0.2 n.u.c. of *per* protein.

### 2.1 The type of regulatory function

The regulatory function (negative feedback) of *per* mRNA synthesis may be identified directly by a closer look to these experimental results. Indeed, when a concentration higher of 0.2 n.u.c. of *per* is present the synthesis is switched off. Whereas, when a concentration lower than 0.2 n.u.c is present the synthesis is switched on (degradation is always active). This corresponds typically to one of the most simple threshold regulatory function as follows:

\[
f(\text{per}(t)) = \begin{cases} 
1 & \text{if } \text{per}(t) \leq 0.2 \\
0 & \text{if } \text{per}(t) > 0.2 
\end{cases}
\]

### 2.2 Estimating \( k_1 \)

When the synthesis of mRNA is inhibited (when \( \text{per} > 0.2 \) n.u.c), then using (3), the first term in equation (1) becomes 0, and we have:

\[
\frac{d\text{ARN}(t)}{dt} = -k_1\text{ARN}(t)
\]

Which after integration gives the well known solution:

\[
\text{ARN}(t) = \text{ARN}(t_0)e^{-kt}
\]

This is a first order decreasing kinetics. From the original data one plots the natural logarithm as function of the time and estimates \( k_1 \) from the slope (Fig. 3) as 0.402 ± 0.049 h\(^{-1}\) (correlation = 0.95).
2.3 Estimating \( v_0 \)

When the synthesis of \( per \) mRNA is switched on (i.e., \( f(per(t) = 1 \)), the equation (1) becomes:

\[
\frac{dRNA(t)}{dt} = v_0 - 0.402RNA(t)
\]  

(6)

Which after integration gives:

\[
RNA(t) = v_0 \left( \frac{1 - e^{-0.402t}}{0.402} \right)
\]  

(7)

This could be linearized in:

\[
RNA(t) = v_0 B(t)
\]  

(8)

Thus, the plot of experimentally measured RNA as a function of calculated \( B(t) \) should be a line and its slope will be \( v_0 \). It equals to 0.358 ± 0.015 n.u.c. h\(^{-1} \) (Fig. 4).
2.4 Estimating the delay

The delay value is an experimentally observed feature. It represents a time latency between the observed maxima of RNA and per protein (see [5]), delay = -8 h.

2.5 Evaluation of k_2 and k_3

These evaluations are more difficult to obtain, since data is far less precise (see Fig. 2). It is possible to obtain an estimation of k_3 (per protein degradation) from Figure 2 by estimating that the half-life (t_{1/2}) of the degradation kinetics is -2 h. This represents a k_3 of 0.35 h^{-1} (k_3 = ln(2)/t_{1/2}). A rough estimation of k_2 is obtained if we consider that a steady state is observed at 1.0 n.u.c. Indeed, at per steady state (per(t_{eq})), equation (2) becomes:

0 = \frac{dRNA(t)}{dt} = k_2RNA(t - 8) - k_3per(t_{eq}) \tag{9}

per(t_{eq}) = 1.0 n.u.c, k_3 = 0.35 h^{-1}, the value of RNA(t_{eq}-8) can be estimated directly from Figure 2, since per(t_{eq}) is observed at ~22 to 24 h, then RNA(t_{eq}-8) at 14 to 16 h is 1.0 n.u.c. Substituting these numbers in (9), it comes that k_2 = 0.35 h^{-1}.

3. Model simulation

Once each parameter has been estimated, the dynamics of the system is described by the following equations:

\frac{dRNA(t)}{dt} = 0.358f(per(t)) - 0.402RNA(t) \tag{10}

\frac{dper(t)}{dt} = 0.35RNA(t - 8) - 0.35per(t) \tag{11}

The regulation function f(per(t)) is defined by (3). It is clear that, from these equations, it is rather difficult to intuitively see if they oscillate and, more difficult even, is at which period. One way is to simulate their resolution by numerical methods (numerical integration by a runge-kutta 4th order method, e.g.) in order to obtain the RNA and per dynamics.

The simulation is presented in Figure 5: the system oscillates with a periodicity of ~26 h, which is very near the one experimentally observed (~24 h). A rough sensitivity study has been realized by varying some parameters to check the stability of the model. With a delay of 6 h instead of 8 h, the period is still circadian (21.6 h). Changing k_2 = k_3 (actually at 0.35 h^{-1}) to 0.175 h^{-1} or to 0.70 h^{-1} gives a periodicity of 29.8 h and 23.2 h respectively. The model is thus robust with respect to changes in critical parameters (particularly the ones that were the most difficult to evaluate).
Figure 5. Simulation of the *per* protein and its mRNA dynamics in the *Drosophila* model. The period of oscillation is circadian (28 h). The conceptual and mathematical (before and after evaluation of the parameters from the experimental data) are presented in Figure 1, equations (1 & 2) and equations (3,10 & 11) respectively.

4. Comparison between real and simulated data

Figure 6 shows the comparison between the original and simulated data for *per* mRNA (see also Figure 2). From visual inspection the fitting seems to be rather correct.

Figure 6. Comparison between the original experimental data of *per* mRNA and the results of the simulated *per* mRNA according to equations (3,10,11). Units of mRNA are in n.u.c.

Conclusion

More recent works have shown that the oscillation of the *per* protein is also dependent of another essential element: the protein *tim* [7,9,11], that possess a specific gene distinct from *per*. As it is possible to observe in this work, the model we have elaborated doesn't contain these informations. However, it seems to rather correctly
describe the endogenous oscillation. The model is systemic, it possess a great power of simplification (aggregation) of a complex phenomenon. This choice is deliberate. More precisely, the existence of a delay factor is actually corresponding to a lot of molecular / physiological / biochemical reactions that were introduced in this kind of model only by their final simplified result, but with a correct dynamical equivalent. Indeed, as some authors have suggested, this delay might be explained by the necessity for per to be stabilized by tim for the transport of this complex from the cytoplasm to the nucleus [9].

In other words, in order to elaborate a functional (imperfect, but useful) systemic model, it is not necessary to know all the detailed molecular mechanisms of the phenomenon, but only some essential dynamical characteristics, only some aggregating parameters are needed. It is clear that the identification of such key parameters is of the utmost importance for modeling complex phenomena's with minimal systemic modeling. Other important domains were complexity is the rule like ecological modeling are using this kind of approach with success since a long time now [10].

There are little doubts that once any system has been somehow reduced in some of its components, one needs again someway to integrate them in order to test the modeled "behavior" with the experimental observations. One method, among others, to achieve integrated approaches is by modeling and simulating techniques.

References

