Modelling Circadian Rhythms in *Drosophila* and Investigation of VRI and PDP1 Feedback Loops Using a New Mathematical Model

D. Kulasiri[1] and Z. Xie[2]

Centre for Advanced Computational Solutions (C-fACS)
Cell Biology Group, Lincoln University, New Zealand

Abstract. We present a brief review of molecular biological basis and mathematical modelling of circadian rhythms in *Drosophila*. We discuss pertinent aspects of a new model that incorporates the transcriptional feedback loops revealed so far in the network of the circadian clock (PER/TIM and VRI/PDP1 loops). Conventional Hill functions are not used to describe the regulation of genes, instead the explicit reactions of binding and unbinding processes of transcription factors to promoters are probabilistically modelled. The model is described by a set of ordinary differential equations, and the parameters are estimated from the in vitro experimental data of the clocks' components. The model is robust over a wide range of parameter variations. Through the sensitivity analysis of the model, roles of VRI and PDP1 feedback loops are investigated, and it is proposed that they increase the robustness of the clock.

1 Introduction

Mathematical models in molecular biology enable us to understand functionalities of the underlying metabolic, genetic and signal transduction networks, particularly by validating the models with experimental data, and when discrepancies are found in the comparisons, by reformulating the models to capture the missing biological knowledge that may have

---

[1]Corresponding author. E-mail: kulasird@lincoln.ac.nz
[2]Co-corresponding author. Email: zhi.xie@jhmi.edu

contributed to the discrepancies. This is an iterative process which will allow us to expand our knowledge of the underlying networks systematically [6]. Careful and realistic interpretation of the results of the computational experiments of mathematical models can suggest novel wet lab experiments for testing the hypotheses on which the models are based [56] and may aid in the study and analysis of systemic properties that are not accessible through *in vitro* experiments[40]. In addition, mathematical modelling can also be used for designing desirable products based on existing biological networks [2].

Many physiological processes in living beings follow a daily periodicity. All eukaryotes and some prokaryotes are capable of maintaining sustained oscillations in terms of gene activity, metabolism, physiology and behaviour with a period close to 24 h ([38], [37],[53],[36]). These oscillations are known as circadian rhythms.

Circadian systems in most organisms have input pathways that transmit molecular signals from external stimuli such as light and temperature to the internal clock; these signals synchronise the internal biological clock with the environment; the output pathways carry information from the internal clock to the relevant metabolic and genetic networks. These networks drive the daily biochemical and physiological changes in the cell. Examples of behaviours controlled by the clock are sleep-wake cycles, photosynthesis, and hormonal control [43]. The internal clock comprises a number of clock molecules which autonomously produce their oscillations, with or without external stimuli. This description of circadian systems is oversimplified: there are numerous overlaps where the different components and pathways can utilise the same molecules for different roles.

In recent decades, many components and molecular mechanisms comprising circadian systems have been unravelled, mainly due to the advances in molecular biology experimental techniques such as forward mutagenesis screens. Specifically, mutations in each of the genes responsible for time keeping produce circadian rhythm abnormalities, ranging from alterations in the circadian period to complete arrhythmicity. Such experimentation and analysis have produced a wealth of information and insight [10] and the internal clocks among the different organisms share a common theme [53]. At the core of all the circadian clocks there is a network of positive and negative elements. The positive elements (i.e. activators) activate the transcription of the negative elements, whereas the negative elements (i.e. repressors) block their own transcription by eliminating the positive elements. The genes and proteins of the activators and the repressors form transcriptional regulatory networks in the circadian clocks with feedback loops.

Drosophila (fruit fly) is one of the most extensively researched organism because of its status as a central model in eukaryote biology. *Drosophila* research has contributed most to the timing mechanism studies of the central circadian clock due to relative ease of genetic manipulation and the property of being suited to large-scale mutant screening [53].

The objectives of this paper are: (1) to present a <u>very brief review</u> of pertinent molecular biological knowledge related to circadian models and relevant mathematical models; (2) to introduce some relevant aspects of a new mathematical model which captures the behaviours of recently discovered VRI and PDP1 transcriptional factors and promoter binding sites; and, (3) to investigate the possible functionalities of VRI and PDP1 feedback loops using

sensitivity analysis of the model. As this is not a comprehensive treatment of the subject and the readers are encouraged to consult the references at the end of the paper.

2   Molecular basis of the *Drosophila* circadian clock

In this section, we briefly review the components of the *Drosophila* circadian system to facilitate the explanation of the modelling research. Hardin (e.g. [18]) has written excellent reviews of this subject, and readers are referred to them for further elucidations. (Notation: names of genes are in lower case italics; mRNA of gene, x, is identified as x mRNA, and proteins are in upper case.) The *period* (*per*) was the first gene identified as a clock component in *Drosophila* by the isolation of its mutants [24]. Three types of circadian rhythms were found based upon the period of rhythmicity under free running conditions in constant darkness: arrhythmicity in the *per* null allele (per^0), a shortened daily rhythm in the "short" period allele (per^s) and an elongated rhythm in the 'long' period allele (per^l) [24]. PER protein shares a domain with a family of proteins including ARNT (Arhl hydrocarbon receptor nuclear translocator) and SIM (Single Minded Protein) ([7], [21]). This region is named the PAS domain (PER ARNT SIM) and is involved in protein-protein interactions [22]. PER forms heterodimers with TIM, the protein product of the second clock gene identified in *Drosophila*, *timeless* (*tim*) ([12],[35]). *tim* is vital for rhythmic behaviour [44]: a null mutation of *tim* results arrhythmicity. Although the early genetic evidence suggested that the *per* and *tim* genes play crucial roles in the circadian cycling machinery, no insight had been given into how these genes participated in the clock mechanism until their discovery by Hardin et al.[19]. They showed that PER was feedback to regulate its own mRNA levels by observing that the mutants in *per*, affected the quantity and quality of *per* mRNA cycling. Moreover, the experiments showed that the molecular levels of *per* mRNA fell when PER levels rose, and *per* mRNA levels rose when PER levels fell [19]. These findings suggested that PER directly or indirectly repressed its own gene expression and the oscillating PER and TIM levels resulted from the oscillating levels of *per* and *tim* mRNAs [45].

However, the mechanism of feedback regulation of *per* and *tim* transcription remained poorly understood because neither PER nor TIM was found to have a DNA-binding domain. The breakthrough came with promoter dissection studies which identified a $\sim$ 70-basepair (bp) enhancer sequence, found $\sim$ 500 bp upstream of the *per* transcription initiation site, as a circadian regulatory sequence. This fragment, which contains a consensus E-box element (CACGTG), was required for transcriptional activation [17]. Similarly, a consensus CACGTG E-box was found in the *tim* upstream sequence that was also necessary for *tim* transcription [32]. Subsequently, the mechanisms underlying PER and TIM transcription feedbacks became clear when genetic screening for mutations identified that two genes, *clock* (*clk*) and *cycle* (*cyc*), played critical roles in circadian rhythmicity. In flies with *clk* and *cyc* mutations, *per* and *tim* expression is arrhythmic and low, suggesting that *clk* and *cyc* act as positive regulators for *per* and *tim* transcription [1]. Since E-box elements are known targets for basic helix-loop-helix domain (bHLH) transcriptional factors (TFs), and the products of *clk* and *cyc* expression, CLK and CYC, are bHLH-PAS TFs, these suggest a model for a

transcription feedback loop: CLK and CYC activate *per* and *tim* transcription by binding to E-box elements within their promoters, and PER and TIM inhibit their transcription by binding to CLK and CYC through their PAS domains [9].

The biology discussed above explains the generation of PER and TIM oscillation, but it does not explain the oscillation of *clk* mRNA and CLK proteins. In *Drosophila*, *clk* mRNA levels peak just after dawn, roughly in anti-phase with *per* and *tim* mRNA levels [9]. Moreover, *clk* mRNA levels are constitutively low in per^{01} and tim^{01} mutants which produce non-functional PER and TIM, suggesting that PER and TIM are required for activating *clk* expression [3]. However, *clk* mRNA levels are surprisingly high in mutant clk^{Jrk} which produces non-functional CLK, suggesting that CLK represses its own expression [13]. Since CLK is known to be a transcription activator and there are no consensus E-boxes in or around the *clk* promoter, it is unlikely that CLK directly represses its transcription. In addition, the high levels of *clk* mRNA in the per^{01}; clk^{Jrk} double mutant indicate that a separate *clk* activator is present [13]. All these findings prompted the discovery of a CLK repressor VRILLE (VRI), which is a rhythmically expressed PAS domain transcription factor and is activated by CLK/CYC. It was found that a consensus "VRI box" is in the *clk* promoter and VRI over-expression represses *clk* mRNA levels, suggesting that VRI directly represses *clk* expression [14]. Subsequently, a second PAR domain factor was identified by its homology with VRI, the PAR Domain Protein 1 (PDP1) [8]. Like the *per* and *tim* genes, E-boxes were also found in the promoters of *vri* and *pdp1* genes and CLK/CYC dimers had been shown to activate *vri* and *pdp1* expression *in vitro* in an E-box-dependent manner ([8], [14]). Both VRI and PDP1 belong to basic zipper TFs with highly conserved basic DNA binding domains, suggesting that they bind to the same set of target genes. Indeed, *in vitro* experiments showed that PDP1 can bind to the VRI box consensus sequence (henceforth referred to as a V/P box), and compete with VRI to regulate the *clk* mRNA [8].

In summary, six genes have been identified as necessary for the circadian clock functions in *Drosophila*. These genes can be divided into two categories according to the molecular nature of their protein products. These proteins include (1) transcriptional activators: CLK, CYC and PDP1; and (2) transcriptional repressors: PER, TIM and VRI. These components appear to be organised into a transcriptional regulatory network where the protein products of one or more clock genes indirectly regulate expression of their own genes.

3   A summary of Transcriptional feedback loops

The *Drosophila* circadian clock is composed of two interlocked feedback loops in gene expression ([13], [8], [18]). The first loop, named the PER/TIM loop, starts with activation of the *per* and *tim* expression from mid day. Activation of the *per* and *tim* transcription is mediated by two TFs, CLK and CYC. CLK and CYC form dimers that target E-boxes in the *per* and *tim* promoters [1]. After initial activation of the *per* and *tim* expression, there is a 4 h – 6 h delay between the peak concentrations of *per* and *tim* mRNAs and that of PER and TIM proteins ([58], [57]). As a result, CLK/CYC can continue to activate transcription of the *per* and *tim* genes, while PER and TIM proteins accumulate in the cyto-

plasm. PER and TIM also form PER/TIM dimers while accumulating. In the middle of the night PER/TIM dimers are transported into the nucleus. After entering the nucleus, they can bind to CLK/CYC dimers effectively inhibiting CLK/CYC binding ability to E-boxes without disrupting the dimeric structure of CLK/CYC [25]. This inhibition lasts until PER and TIM proteins are degraded. Then the expressions of *per* and *tim* are reactivated by CLK/CYC dimers in the following mid day.

The second loop, named the VRI/PDP1 loop, consists of a VRI-mediated negative feedback loop and a PDP1- mediated positive feedback loop. This loop starts with activation of *vri* and *pdp1* transcription by CLK/CYC during the late day and early night. VRI accumulates first in phase with its mRNA then PDP1 accumulates during the mid to late evening. VRI binds the V/P box in the *clk* regulatory elements to inhibit the *clk* transcription and PDP1 can compete with VRI for binding to the V/P box and activate the *clk* transcription [8]. The effects from the initial VRI-dependent repression in the early night and the subsequent PDP1-dependent activation in the middle to late night determine the rhythmic expression of *clk*. However, the newly produced CLK at the end of night and early morning is inactive temporarily due to high levels of PER/TIM dimers induced by the previously produced CLK. Once PER/TIM dimers are degraded, CLK/CYC reactivates the gene expression of *per*, *tim*, *vri* and *pdp1* and starts a new cycle.
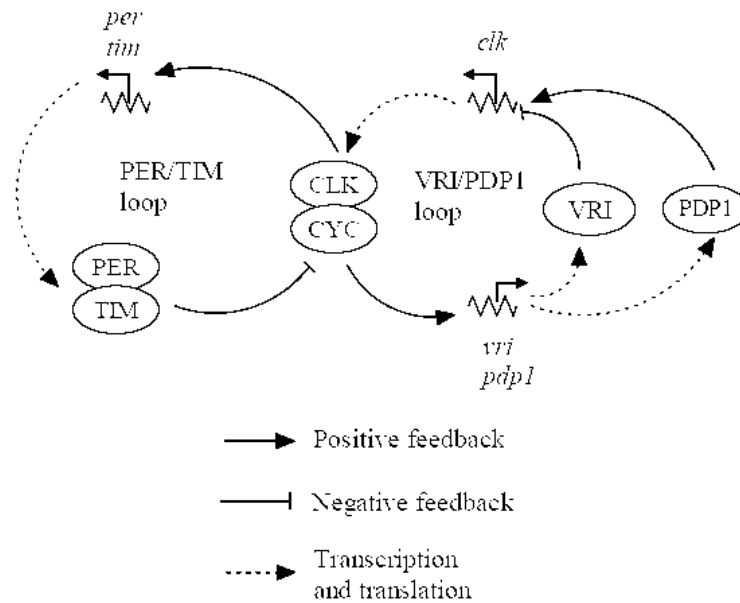

Figure 1: Interactions in the two loop model of Cyran et al. [8].

In addition to regulation at the transcriptional level, many clock components in *Drosophila* are also regulated post-transcriptionally and post-translationally. For example, Doubletime (DBT) destabilises PER. Casein Kinase 2 (CK2) destabilises PER and also affects its nuclear localisation. Shaggy (SGG) phosphorylates TIM to promote nuclear localisation of

PER/TIM dimers. Slimb (SLMB) targets phosphorylated PER for degradation [18]. These processes might provide time delays between mRNAs and proteins. For example, a 4 h – 6 h delay between accumulation of *per* mRNA in the cytoplasm and PER in the nucleus results from the initial destabilisation of PER by DBT dependent phosphorylation, and possibly also CK2 dependent phosphorylation, followed by the stabilisation of PER by dimerisation with TIM before nuclear entry [39].

4   Mathematical Models of the circadian clock in *Drosophila*

A range of mathematical models for the circadian clocks in different organisms have been proposed in the literature, including *Arabidopsis thaliana*, *Neurospora*, *Drosophila* and mammals ([15], [47], [26], [30]). The common character of these models is that they contain at least one negative feedback loop and have a capability to produce sustained oscillations for at least one protein in such a loop in the appropriate parameter regimes [16]. The first circadian clock model of *Drosophila* was proposed by Goldbeter [15] which is based on the negative feedback exerted by PER on the transcription of the *per* gene. This simple model containing five variables described the multiple phosphorylation of PER. Numerical simulations showed that a single negative feedback of PER alone can produce limit-cycle oscillations for appropriate parameter values. However, the early model did not account for the effect of light on the circadian system because the light receptor TIM was not taken into account in the model. Leloup et al. presented an extended model based on the auto-regulatory negative feedback exerted by a complex between PER and TIM proteins on the expression of *per* and *tim* genes ([27], [28]). The model produced essentially the same results as the first model; in addition, it explicitly incorporated the effect of light on the TIM degradation rate. A closely related model incorporating the formation of a PER-TIM complex has been proposed for *Drosophila* circadian rhythms [51]. The difference from the model proposed by Leloup et al. [27] is that the model [51] contains an additional positive feedback loop based on stabilisation of PER upon dimerisation. The model proposed by Tyson et al. [51] accounted for several properties of circadian rhythms, including temperature compensation and the per^L mutant. With the discovery of an additional regulatory loop related to the *clk* gene and CLK [13], more detailed mathematical models involving two interlocked *per-tim* and *clk* loops were created and examined ([47], [52]). The simulations showed that the model can produce sustained oscillations of *clk* mRNA and CLK which were not explained in previous models, and the analysis revealed that the interlocked feedback model provided a possible explanation for the robust oscillation of *Drosophila* circadian rhythms.

More recently, while two additional proteins, VRI and PDP1, were identified to be involved into the regulation of *clk* gene, two new models that reproduced CLK expression regulation by VRI and PDP1 were proposed ([49], [41]). The model proposed by Smolen et al. [49] contained feedback loops based on transcriptional regulation of *per*, *clk*, *pdp1*, and *vri*, in particular, *pdp1* expression was modelled with time delay. The role of PER was

described in detail in the model in which PER protein underwent a two-step phosphorylation in the cytoplasm and nucleus. Simulations suggested that *vri* and *pdp1* feedback loops were not essential for oscillations, however, the negative feedback loop in which PER represses *per* expression was critical for producing oscillations. In addition, the model simulated a range of behaviour of the circadian clock in *Drosophila*, including null mutations of *per, vri, pdp1* and *clk*, photic phase-response curves resembling experimental curves, and the entrainment to light-dark cycles. The incompleteness of Smolen's model is that the regulation of TIM was not included.

The core of the model proposed by Ruoff et al. [41] is that CLK is subjected to positive and negative regulations by PDP1 and VRI, whose transcriptions are activated by CLK . The model did not differentiate the *per* and *tim* gene expressions, instead, treating PER/TIM complex as a whole whose expression is activated by CLK. The results of simulations suggested that the positive feedback loop and negative feedback loop of *pdp1* and *vri* were essential for the overall oscillations. The PER/TIM complex only played a role in amplification and stabilisation of the oscillations. This conclusion contradicts the one drawn by Smolen in which the PER feedback loop was found to be vital to produce oscillations for the circadian clock system. The other contribution of this model is the calculation of the phase resetting of temperature compensation and losses of temperature compensation in per^S and per^L mutants. However, the model showed poor entrainment under light/darkness cycles.

The mathematical models discussed above produce oscillations of some mRNAs or proteins in the clock system although they are based on incomplete conceptual models of the regulatory networks of the circadian clock in *Drosophila*. For example, these models have not incorporated (1) the six experimentally found TFs (PER, TIM, CLK, CYC, VRI and PDP1) and their roles, (2) functionalities of E-boxes of *tim, per, vri* and *pdp1* genes and V/P box in *clk* gene in transcriptional and translational stages of gene expression, and (3) interlocked nature of the PER/TIM and VRI/PDP1 feedback loops. We have developed a new model which includes these missing aspects of biological reality extending the previous work [55], and explains some of the experimental observations in meaningful manner.

5 A new model

We summarise the pertinent aspects of our model in this section and for details refer to [55]; the mathematical model follows the conceptual model shown in Figure 2 . The core structure of the model is similar to the model proposed by Cyran et al. [8], as schematised in Figure 1. The development of the mathematical model was also relied upon a number of assumptions that were used to simplify the model development, and the assumptions in brief are given in the appendix (for details see [55]). We have reproduced the equations in the appendix along with the estimated parameters and initial conditions.

In our model, we modelled the binding probabilities of TFs to E-boxes, and incorporated them in the reaction dynamics. There are a number of E-boxes in the promoter region of *per, tim, vri* and *pdp1* genes. Here we make an important assumption that CLK/CYC
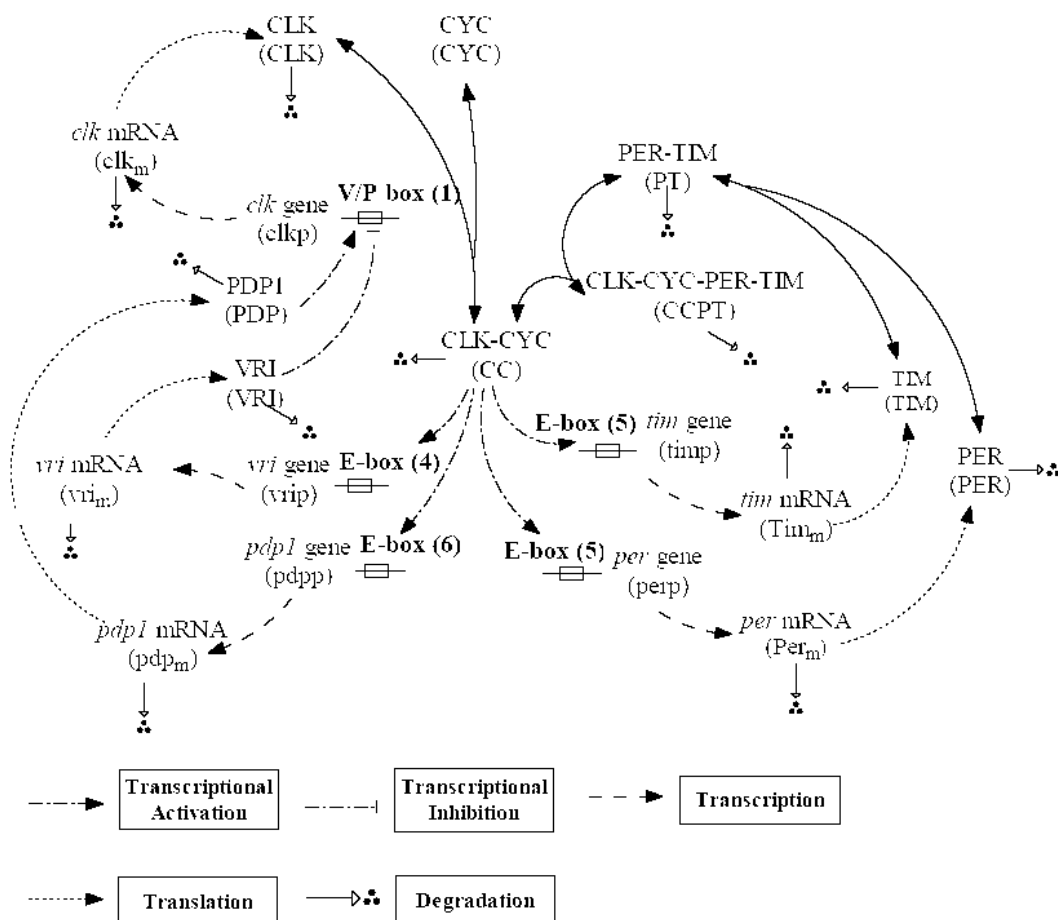
Figure 2: The schematic diagram of the model proposed by Xie and Kulasiri [55]. The model shows the regulatory relationships among genes, mRNAs and proteins in the negative and positive transcriptional feedback loops. Transcription of *per, tim, vri* and *pdp1* genes are activated by CLK/CYC dimers binding to E-boxes in their promoter regions. In one loop, *per* and *tim* mRNAs are translated to PER and TIM proteins which form PER/TIM dimers. PER/TIM binds to CLK/CYC to form PER/TIM/CLK/CYC complex. In another loop, *vri* and *pdp1* mRNAs are translated to VRI and PDP1 proteins. They compete to bind the V/P box in the promoter in *clk* gene. Transcription of clk gene is repressed by VRI and activated by PDP1. *clk* mRNA is translated to CLK which forms CLK/CYC dimers with CYC. Proteins, mRNAs, dimers and complexes are degraded at certain kinetic rates. CYC is assumed to be constant, therefore, there is no degradation process for CYC. Variable names used in the model are indicated in parentheses. The number of the E-boxes and the V/P box in the promoters is also shown here.

8

dimers independently bind to an individual E-box in a promoter. In the functional analysis of E-boxes in the mouse *mPer1* promoter, the levels of *mPer1* transcriptional expression activated by CLK/BMAL1 were roughly proportional to the number of conserved E-boxes [20]. The result suggests that there is no or negligible cooperative interaction in the E-box binding activities of CLK/BMAL1. Since no information is available about cooperativity in the E-box binding activities by CLK/CYC in *Drosophila*, we treat it as the case in *mPer1* promoter. For the same reason, we also assume that if CLK/CYC is bound to just one E-box for a given gene, the transcription of that gene is activated and the effect of binding additional E-boxes on transcription activation is additive [20].

Suppose there are n E-boxes in a promoter where CLK/CYC dimers can bind. Since CLK/CYC dimers bind to individual E-boxes independently, we can consider each E-box separately. The binding and unbinding processes of CLK/CYC to an E-box can be formulated as

$$B + T \underset{ub_{bt}}{\overset{b_{bt}}{\rightleftharpoons}} BT, \tag{5.1}$$
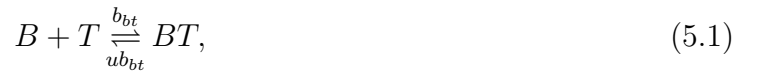
where B is an available binding site, i.e., an unbound E-box; T is the TF, CLK/CYC; and BT denotes CLK/CYC bound to the E-box; b_{bt} is the rate of CLK/CYC binding to the E-box and ub_{bt} is the rate of CLK/CYC releasing from the E-box. Using mass-action kinetics,

$$d[BT]/dt = [B]\,[T]\,b_{bt} - [BT]\,ub_{bt}, \tag{5.2}$$

where [B] is the concentration of unbound binding sites, [T] is the concentration of the TFs and [BT] is the concentration of bound binding sites. Suppose the volume of the cell is V. The number of B and BT in the cell are $[B]\,V$ and $[BT]\,V$. If the total number of B and BT is n then,

$$d[BT]/dt = ((n/V) - [BT])\,[T]\,b_{bt} - [BT]\,ub_{bt}. \tag{5.3}$$

Let Pr_{bt} be the ratio of the number of bound binding sites over the total number of binding sites. $[BT] = $ total number of binding sites$/V \times \mathrm{Pr}_{bt}$. Since the total number of binding sites is n, Eq. (5.3) becomes

$$d((n/V)\,\mathrm{Pr}_{bt})/dt = ((n/V) - (n/V)\,\mathrm{Pr}_{bt})\,[T]\,b_{bt} - (n/V)\,\mathrm{Pr}_{bt}\,ub_{bt}, \tag{5.4}$$

which simplifies to

$$d\mathrm{Pr}_{bt}/dt = (1 - \mathrm{Pr}_{bt})\,[T]\,b_{bt} - \mathrm{Pr}_{bt}\,ub_{bt}. \tag{5.5}$$

Now we can calculate probabilities for CLK/CYC binding to the whole promoter in a gene. Assuming that CLK/CYC can bind independently to any of n binding sites (E-boxes) and if one or more E-boxes are bound, transcription of that gene is activated at a rate tc_{av} otherwise at a deactivated rate tc_{dvpmt}. The probability of none of the E-boxes being bound is $(1 - \mathrm{Pr}_{bt})^n$. The rate of transcription would then be

$$tc_{av}\left(1-(1-\Pr_{bt})^n\right)\ +\ tc_{dvpmt}\left(1-\Pr_{bt}\right)^n. \tag{5.6}$$

6 Modelling competition of PDP1 activation and VRI repression

Here we consider the *clk* expression activated by PDP1 and repressed by VRI. We first calculate probabilities of VRI and PDP1 binding to the V/P box in the *clk* promoter. Assuming there is only one binding site B, i.e. the V/P box, in the *clk* promoter and an activator PDP1, denoted by A, and a repressor VRI, denoted by R, competing to bind that site. We write the reactions as below,

$$B + A \underset{ub_{ba}}{\overset{b_{ba}}{\rightleftharpoons}} BA, \tag{6.7}$$

and

$$B + R \underset{ub_{br}}{\overset{b_{br}}{\rightleftharpoons}} BR, \tag{6.8}$$

where b_{ba} is the rate of PDP1 binding to the V/P box, and ub_{ba} is the rate of PDP1 releasing from the V/P box; b_{br} is the rate of VRI binding to the V/P box, and ub_{br} is the rate of VRI releasing from the V/P box. We can obtain Eqs. (6.9) and (6.10) by using the mass action rate law:

$$d[BA]/dt = [B]\,[A]\,b_{ba} - [BA]\,ub_{ba}, \tag{6.9}$$

and

$$d[BR]/dt = [B]\,[R]\,b_{br} - [BR]\,ub_{br}. \tag{6.10}$$

The number of B, BA and BR in the cell are [B] V, [BA] V and [BR] V. As the total number of B, BA and BR is one, we eliminate [B] from the above two equations and obtain,

$$d[BA]/dt = (1/V - [BA] - [BR])\,[A]\,b_{ba} - [BA]\,ub_{ba}, \tag{6.11}$$

and

$$d[BR]/dt = (1/V - [BR] - [BA])\,[R]\,b_{br} - [BR]\,ub_{br}. \tag{6.12}$$

Let $\Pr_{ba}$ and $\Pr_{br}$ be the probabilities of A bound to B and R bound to B. We can write Eqs. (6.11) and (6.12) in the form of probabilities,

$$d\Pr_{ba}/dt = (1 - \Pr_{ba} - \Pr_{br})\,[A]\,b_{ba} - \Pr_{ba}\,ub_{ba}, \tag{6.13}$$

and

$$d\Pr_{br}/dt = (1 - \Pr_{ba} - \Pr_{br})\,[R]\,b_{br} - \Pr_{br}\,ub_{br}. \tag{6.14}$$

Assuming that if a PDP1 is bound to the V/P box, transcription of *clk* genes occurs at a rate of tc_{pc}; if a VRI is bound to the V/P box, transcription rate is tc_{vc} and if neither

PDP1 nor VRI binds the V/P box, transcription occurs at a deactivated rate tc_{dvpmt}. The transcription rate of *clk* gene would then be

$$tc_{pc}\mathrm{Pr}_{ba} + tc_{vc}\mathrm{Pr}_{br} + tc_{dvpmt}\left(1 - \mathrm{Pr}_{ba} - \mathrm{Pr}_{br}\right). \tag{6.15}$$

7 In Silico experiments

A range of *in silico* experiments were conducted using the model developed in [55] including the tests of the circadian rhythms under the condition of constant darkness, the robustness of the system, and the response of the system to light and to a number of mutants. All of the *in silico* experimental data were thoroughly and rigorously compared with the *in vitro* experimental data. (For details see [55].) Using a set of parameters, the model produced autonomous sustained oscillations under conditions corresponding to constant darkness [55]. The simulated results showed correct phases for all the components in the system, correct phase and anti-phase relationships of mRNAs and proteins, as well as appropriate lags between mRNAs and proteins. The model also accounted for the disappearance of the oscillations in the condition of constant light.

Robustness is an important characteristic of the circadian clock, which should produce close to 24 hours periodic oscillations regardless of modest variations in parameters under certain conditions. We measured variations of period by increasing and decreasing each parameter in turn by 20%. The oscillatory patterns remained in all the cases with the largest period of variation being around 0.8 hours for 20% parameter perturbations [55]. Parameter sensitivity analysis suggested that some of the most sensitive parameters were binding rate of PDP1 to *clk* promoter, and binding rates of CLK/CYC to *pdp1* and *per* promoters. These are all positive elements (transcriptional activators) in the network.

It is also essential that the circadian clock should have the ability to reset phases in response to Zeitgeber (time giver), where light is the most important. It has been experimentally shown that light enhances degradation of TIM [57], and as TIM stabilizes PER in the cytoplasm, the indirect effect of light is to regulate the localisation of PER and in turn, to decrease the PER levels in the nucleus. Therefore, we simulated the effect of light by increasing the degradation rates of TIM and PER. Simulations have shown the entrainment of the system by light dark cycles and the induction of phase shifts by light pulses. In the entrainment by light dark cycles, the phase relationship in mRNAs and proteins were well maintained with a period of 24 hours and the phase of oscillations was delayed depending on the particular degradation rates we chose. We also constructed a phase-response curve (PRC) to represent the phase shifts induced by temporal promotion of TIM and PER degradation. When normalising the simulated PRC by advancing it for 5 hours, the agreement between the normalised and experimental PRCs appeared very good. Both data showed a dead zone in the middle of a day, a phase delay during early night, and a phase advance during late night. The time lag between the simulated and the experimental data suggested that some unpresented mechanisms in the model, such as phosphorylation and nuclear entry of TIM and PER, are important to providing a time delay in response to light.

We also carried out a number of tests for simulating mutations. Mathematical mutants were simulated by setting an appropriate parameter value according to the functionality of mutants. The simulated short and long mutants, per^S and per^L, resembled their phenotypes where 19 h and 29 h of period were found, respectively. In arrhythmic mutants, oscillations of all the mRNAs and proteins were blocked in per^{01}, tim^{01} and clk^{Jrk}, as shown in the experiments. However, some mRNA levels significantly differed from the experimental data. Particularly, simulated data have shown high levels of *per* and *tim* mRNA in per^{01} and tim^{01} and low level of *clk* mRNA in clk^{Jrk}, which were opposite to that in experiments. This deficiency obviously came from the structure of the model. In the model we assumed that the *per*, *tim*, *vri*, and *pdp1* promoters were all strongly activated by CLK/CYC. The low levels of *per* and *tim* mRNAs in per^{01} and tim^{01} can not be explained with this model because the loss of PER/TIM directly resulted in a high level of CLK and, consequently, high levels of *per* and *tim* mRNAs. Furthermore, although the assumption of strong binding ability of VRI to CLK gave a reasonable low level of *clk* mRNA in per^{01} and tim^{01}, this assumption, nevertheless, makes a low level of *clk* mRNA in clk^{Jrk}, which was again different from the experimental observations in which a high level of *clk* mRNA was found. The deficiency of the model could indicate the possibility of an unknown part in the gene-protein network.

8   Possible function of VRI and PDP1 feedback loops

The mutant analysis revealed that the expression of *per* and *tim* is critical for maintaining the oscillations of all the components in the circadian clock [55]. A question arising here is the nature of functionality of the newly found VRI and PDP1 feedback loops. Previous research on other organisms, such as *Neurospora*, suggested that the interlocked feedback loops may contribute to the robustness of the circadian clock [5]. Do the VRI and PDP1 feedback loops play a role in increasing the robustness of the circadian clock in *Drosophila*? To answer this question, we compared the robustness of the system, with and without the two feedback loops, towards parameter changes. To get a quantitative perspective of the robustness, sensitivity analysis of the model was performed.

We first removed the VRI negative feedback loop and/or the PDP1 feedback loop in the model by fixing their expression. It was found that the rhythmicity of the mRNAs and the proteins was preserved in all the cases. We plotted the time evolution of proteins when VRI and PDP1 loops were removed one at a time in Figure 3 . For comparison, we also plotted the protein concentrations when both VRI and PDP1 feedback loops were removed simultaneously in Figure 3. When *per* and *tim* gene expressions are fixed the oscillations of all proteins disappeared as one would expect from the model.

9   Sensitivity Analysis

Sensitivity analysis is a general purpose technique that is often used to analyse how sensitive a system is with respect to the change of a set of parameter values. Mathematical theories
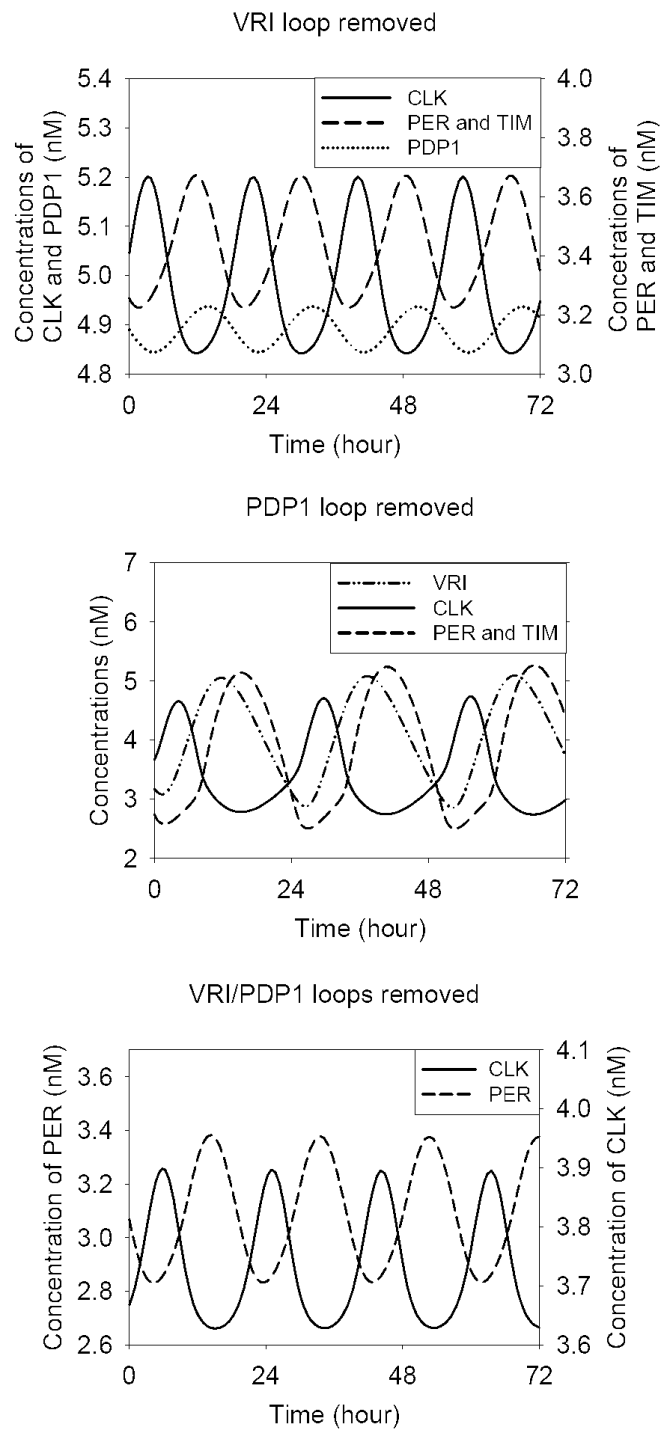
Figure 3: Time evolutions of proteins. The feedback loop was removed by making its corresponding gene expression constant.

about sensitivity analysis can be found in [54]. In the dynamics of biological networks, the state variables mostly correspond to molecule concentrations, and the parameter set consists of biochemical constants related to the system dynamics and initial conditions. The purposes of parameter sensitivity analysis in the biological networks are twofold: (1) to validate information about insensitivities and critical system parameters and to provide a basis for system identification, and (2) to observe intrinsic system properties like stability and robustness of the system behaviour with respect to parameter fluctuations in a systematic way [11].

For a model whose dynamics is described by a system of ODEs,

$$\frac{dx}{dt} = f(x, p, t), \quad \text{with} \quad x(t_0) = x_0, \tag{9.16}$$

where x is the $n_S \times 1$ vector of state variables, t is the time where $t \geq t_0$, and p is the $n_P \times 1$ vector containing the parameters of interest for the system. Suppose the solution of the system is $x = x(t, p)$. The parameter sensitivity with respect to the system's states along a specific trajectory $s(x; p_j)$ is defined by $\frac{\partial x(t, p_j)}{\partial p_j}$. This is also defined as the first-order local sensitivity, or simply local sensitivity of the dependent variable, x, with respect to the input parameter, pj. Although higher-order local sensitivities can be defined in a similar fashion, we limit the treatment to first-order local sensitivities, since most applications are based on linear sensitivity analysis [54]. Sometimes the local sensitivity is also called the absolute sensitivity. It is noticeable that the local parameter sensitivities are valid only in the neighbourhood of a specific parameter set. Thus, they provide information on the resilience of a model with respect to a particular parameter. Another quantity related to local sensitivity, commonly used in sensitivity analysis, is the normalised sensitivity, or relative sensitivity, defined as

$$S(x; p_j) = \frac{p_j}{x} \cdot \frac{\partial x}{\partial p_j} = \frac{\partial \, In \, x}{\partial \, In \, p_j} = \frac{p_j}{x} \cdot s(x; p_j). \tag{9.17}$$

In most cases, the relative sensitivities are more meaningful and therefore we have used this measure in our analysis.

In oscillating systems, the primary interest of parameter sensitivity is generally period sensitivities which capture the change of period length upon changes in parameters. Suppose $\tau(p)$ defines the period of the system for a given parameter p; the absolute period sensitivity to the parameter pj and the normalised periodic sensitivity are given by $\frac{\partial \tau(t, p_j)}{\partial p_j}$, and $\frac{p_j}{\tau(p_j)} \frac{\partial \tau(t, p_j)}{\partial p_j}$, respectively.

These measures have been both used ([54], [50]) for analysing period sensitivity of oscillations in chemical and biochemical systems.

We have calculated the periodic sensitivity indices of the models without the VRI and/or PDP1 feedback loops and compared those of the model with complete feedback loops by computing the percent ratios of the respective indices. Five percent perturbation in parameters was used; smaller the S value, greater the robustness of the system to the perturbations for a particular parameter.

The simulation results are plotted in Figure 4, which shows five parameters whose sensitivity increases more than 10 times whereas the sensitivity of the other 39 parameters only yielded small changes when removing the VRI or/and PDP1 feedback loops. (Table 2 provides the corresponding parameter for each index, and Table 1 explains the parameters.) All five parameters were related to the *clk* gene, *clk* mRNA or CLK protein. In particular, the sensitivities of the transcription and translation rates of CLK increased more than fifty times when removing the VRI feedback loop. Because of the critical roles CLK plays in regulating all the rhythmically expressed genes, and even some non-rhythmically expressed genes, we propose that the VRI and PDP1 feedback loops decrease the sensitivity of CLK to parameter variations and therefore increase the robustness of the circadian clock.
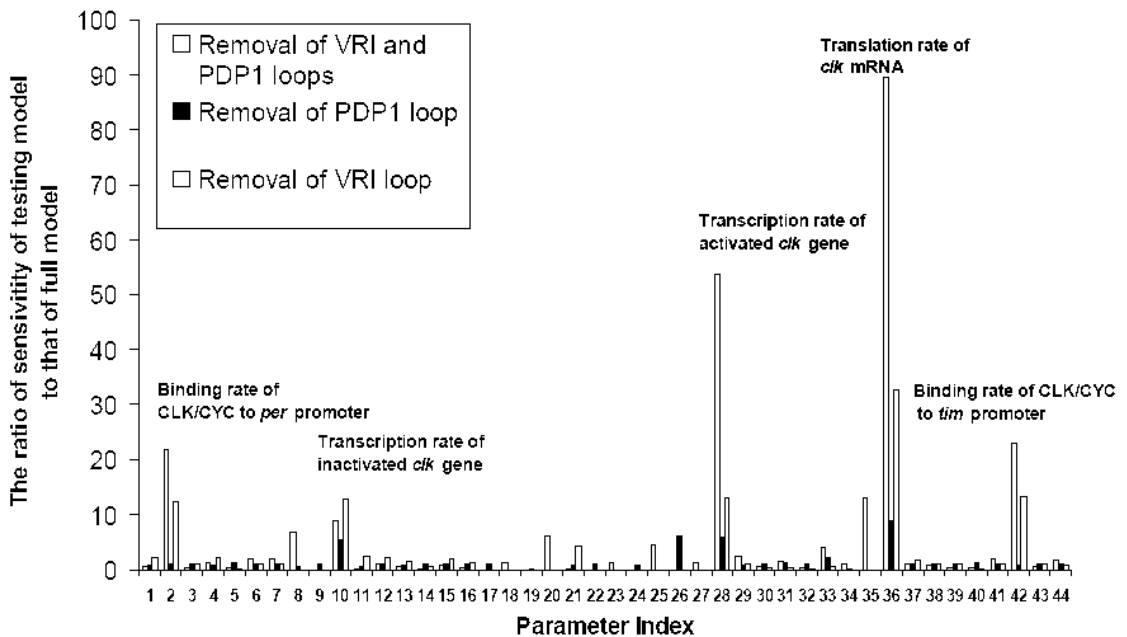

Figure 4: Parametric sensitivity results. The parameter indices are as indicated in Table 2 in the appendix.

10   Discussion

The main difference in terms of model representation between Smolen's and the present model is that different assumptions are used to capture the essence of various interactions. Smolen's model used the Hill function and Michaelis-Menten kinetics describing transcriptional activation and phosophorylation processes, and discrete time delay terms were included in the equations to describe the time lags between proteins. Our model took account of binding and unbinding processes of TFs to promoters but ignored the nuclear entry of

proteins and phosophorylation of PER. However, the simulated results of the two models were very similar regarding oscillations in constant darkness, photic entrainment of oscillations, the PRC and null mutations of *per* and *clk*. Nevertheless, different predictions were made by the two models. For example, E-box mutations were readily simulated in our model whereas some short and long period mutants were observed in Smolen's model. This shows that different mathematical models might be able to create similar behaviours, and they are mainly distinguished by the different predictions they suggest and how close they are to the real biology [34]. As both models were simplified to some extent from the real network, we expect that a more sophisticated model should be developed in future as more data emerge from experiments.

The current model required the assumption that gene regulation is primarily accomplished through transcription regulation. Although the reconstructed model has a good predictive performance for some phenotypes, there are still a large number of phenotypes to be explained. A more complete model could include more detailed post-transcriptional and post-translational regulations, such as phosphorylation of PER by DBT and CK2 [42] and phosphorylation of TIM by SGG [31]. Also, inclusion of separate compartments is necessary in an extended model.

The differential equations are assumed to follow the mass action rate law. This assumption has been used simply because it can simplify the equations greatly. However, it is generally considered that for reactions of the model that are catalysed – which is always the case in biochemical reactions – it is more justifiable to think of them as Michaelian. The most striking problem with using the mass action law is that mass action kinetics is unbounded. A comparative study of the current system described by the mass action rate law and by Michaelis-Menten kinetics could direct the choice of the kinetic approach in future research.

Parameter estimation is a bottleneck in any model development of this nature. The initial parameters are estimated by trial and error, still a common practice [23]. We have used stochastic sampling to aid in our parameter estimation; however, no 'optimal' parameters can be guaranteed. So we can only claim that our model equations and this particular parameter set are sufficient to account for many properties of the system studied.

Nearly all the organisms which have developed circadian rhythms share a great similarity in their circadian clocks. Among them, *Drosophila* is one of the most well- researched species. A comparative study between the circadian clock in *Drosophila* and other species could facilitate the understanding of the structures of the circadian clocks for the different species. Indeed, there has already been a successful model developed in which gene *gigantea* was found to be a candidate in a new feedback loop in the *Arabidopsis* circadian clock by comparing it with the circadian clocks in *Drosophila* and mammals [30].

References

[1] R. Allada, N. E. White, W. V. So, J. C. Hall, M. Rosbash. *A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless.* Cell, 93 (1998), 791-804.

[2] A. P. Arkin. *Synthetic cell biology.* Curr. Opin. Biotechnol., 12 (2001), 638-44.

[3] K. Bae, C. Lee, D. Sidote, K.-y. Chuang, I. Edery. *Circadian regulation of a Drosophila homolog of the mammalian Clock gene: PER and TIM function as positive regulators.* Mol. Cell. Biol., 18 (1998), 6142-51.

[4] Y. Cao, D. T. Gillespie, L. R. Petzold. *The slow-scale stochastic simulation algorithm.* J. Chem. Phys., 122 (2005), 14116.

[5] P. Cheng, Y. Yang, Y. Liu. *Interlocked feedback loops contribute to the robustness of the Neurospora circadian clock.* Proc. Natl. Acad. Sci. USA, 98 (2001), 7408-13.

[6] M. W. Covert, B. O. Palsson. *Transcriptional regulation in constraints-based metabolic models of Escherichia coli.* J. Biol. Chem., 277 (2002), 28058-64.

[7] S. T. Crews, J. B. Thomas, C. S. Goodman. *The Drosophila single-minded gene encodes a nuclear protein with sequence similarity to the per gene product.* Cell, 52 (1988), 143-51.

[8] S. A. Cyran, A. M. Buchsbaum, K. L. Reddy, M. C. Lin, N. R. Glossop, P. E. Hardin, M. W. Young, R. V. Storti, J. Blau. *vrille, Pdp1, and dClock form a second feedback loop in the Drosophila circadian clock.* Cell, 112 (2003), 329-41.

[9] T. K. Darlington, K. Wager-Smith, M. F. Ceriani, D. Staknis, N. Gekakis, T. D. Steeves, C. J. Weitz, J. S. Takahashi, S. A. Kay. *Closing the Circadian Loop: CLOCK-Induced Transcription of Its Own Inhibitors per and tim.* Science, 280 (1998), 1599-603.

[10] J. C. Dunlap. *Molecular bases for circadian clocks.* Cell, 96 (1999), 271-90.

[11] V. V. Fedorov, P. Hackl. Model-oriented design of experiments. Springer, New York, 1997, pp. viii, 117p.

[12] N. Gekakis, L. Saez, A. M. Delahaye-Brown, M. P. Myers, A. Sehgal, M. W. Young, C. J. Weitz. *Isolation of timeless by PER protein interaction: defective interaction between timeless protein and long-period mutant PERL.* Science, 270 (1995), 811-5.

[13] N. R. Glossop, L. C. Lyons, P. E. Hardin. *Interlocked feedback loops within the Drosophila circadian oscillator.* Science, 286 (1999), 766-8.

[14] N. R. J. Glossop, J. H. Houl, H. Zheng, F. S. Ng, S. M. Dudek, P. E. Hardin. *Vrille feeds back to control circadian transcription of Clock in the Drosophila circadian oscillator.* Neuron, 37 (2003), 249-61.

[15] A. Goldbeter. *A model for circadian oscillations in the Drosophila period protein (PER).* Proc. R. Soc. Lond. B, 261 (1995), 319-24.

[16] A. Goldbeter. *Computational approaches to cellular rhythms.* Nature, 420 (2002), 238-45.

[17] H. Hao, D. L. Allen, P. E. Hardin. *A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster.* Mol. Cell. Biol., 17 (1997), 3687-93.

[18] P. E. Hardin. *The circadian timekeeping system of Drosophila.* Curr. Biol., 15 (2005), 714-22.

[19] P. E. Hardin, J. C. Hall, M. Rosbash. *Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels.* Nature, 343 (1990), 536-40.

[20] A. Hida, N. Koike, M. Hirose, M. Hattori, Y. Sakaki, H. Tei. *The human and mouse Period1 genes: five well-conserved E-boxes additively contribute to the enhancement of mPer1 transcription.* Genomics, 65 (2000), 224-33.

[21] E. C. Hoffman, H. Reyes, F. F. Chu, F. Sander, L. H. Conley, B. A. Brooks, O. Hankinson. *Cloning of a factor required for activity of the Ah (dioxin) receptor.* Science, 252 (1991), 954-8.

[22] Z. J. Huang, I. Edery, M. Rosbash. *PAS is a dimerization domain common to Drosophila period and several transcription factors.* Nature, 364 (1993), 259-62.

[23] F. Hynne, S. Dano, P. G. Sorensen. *Full-scale model of glycolysis in Saccharomyces cerevisiae.* Biophys. Chem., 94 (2001), 121-63.

[24] R. J. Konopka, S. Benzer. *Clock mutants of Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA, 68 (1971), 2112-6.

[25] C. Lee, K. Bae, I. Edery. *The Drosophila CLOCK protein undergoes daily rhythms in abundance, phosphorylation, and interactions with the PER-TIM complex.* Neuron, 21 (1998), 857-67.

[26] J. C. Leloup, A. Goldbeter. *Toward a detailed computational model for the mammalian circadian clock.* Proc. Natl. Acad. Sci. USA, 100 (2003), 7051-56.

[27] J. C. Leloup, A. Goldbeter. *A model for circadian rhythms in Drosophila incorporating the formation of a complex between the PER and TIM proteins.* J. Biol. Rhythms., 13 (1998), 70-87.

[28] J. C. Leloup, A. Goldbeter. *Modeling the molecular regulatory mechanism of circadian rhythms in Drosophila.* Bioessays, 22 (2000), 84-93.

[29] J. C. Leloup, D. Gonze, A. Goldbeter. *Limit cycle models for circadian rhythms based on transcriptional regulation in Drosophila and Neurospora.* J. Biol. Rhythms, 14 (1999), 433-48.

[30] J. C. W. Locke, M. M. Southern, L. Kozma-Bognar, V. Hibberd, P. E. Brown, M. S. Turner, A. J. Millar. *Extension of a genetic network model by iterative experimentation and mathematical analysis.* Mol. Syst. Biol., 1 (2005), 13.

[31] S. Martinek, S. Inonog, A. S. Manoukian, M. W. Young. *A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock.* Cell, 105 (2001), 769-79.

[32] M. J. McDonald, M. Rosbash. *Microarray analysis and organization of circadian gene expression in Drosophila.* Cell, 107 (2001), 567-78.

[33] D. A. McQuarrie. *Stochastic approach to chemical kinetics.* J. Appl. Prob., 4 (1967), 417-78.

[34] J. D. Murray. Mathematical biology. Springer, New York, 2002, pp. 2 v.

[35] M. P. Myers, K. Wager-Smith, C. S. Wesley, M. W. Young, A. Sehgal. *Positional cloning and sequence analysis of the Drosophila clock gene, timeless.* Science, 270 (1995), 805-8.

[36] M. N. Nitabach. *Circadian rhythms: clock coordination.* Nature, 438 (2005), 173-5.

[37] S. Panda, J. B. Hogenesch, S. A. Kay. *Circadian rhythms from flies to human.* Nature, 417 (2002), 329-35.

[38] C. S. Pittendrigh. *Circadian rhythms and the circadian organization of living systems.* Cold Spring Harb. Symp. Quant. Biol., 25 (1960), 159-84.

[39] J. L. Price, J. Blau, A. Rothenfluh, M. Abodeely, B. Kloss, M. W. Young. *double-time is a novel Drosophila Clock gene that regulates PERIOD protein accumulation.* Cell, 94 (1998), 83-95.

[40] L. Pritchard, D. B. Kell. *Schemes of flux control in a model of Saccharomyces cerevisiae glycolysis.* Eur. J. Biochem., 269 (2002), 3894-904.

[41] P. Ruoff, M. K. Christensen, V. K. Sharma. *PER/TIM-mediated amplification, gene dosage effects and temperature compensation in an interlocking-feedback loop model of the Drosophila circadian clock.* J. Theor. Biol., 237 (2005), 41-57.

[42] S. Sathyanarayanan, X. Zheng, R. Xiao, A. Sehgal. *Posttranslational regulation of Drosophila PERIOD protein by protein phosphatase 2A.* Cell, 116 (2004), 603-15.

[43] J. C. Schoning, D. Staiger. *At the pulse of time: protein interactions determine the pace of circadian clocks.* FEBS Lett., 579 (2005), 3246-52.

[44] A. Sehgal, J. L. Price, B. Man, M. W. Young. *Loss of circadian behavioral rhythms and per RNA oscillations in the Drosophila mutant timeless.* Science, 263 (1994), 1603-6.

[45] A. Sehgal, A. Rothenfluh-Hilfiker, M. Hunter-Ensor, Y. Chen, M. P. Myers, M. W. Young. *Rhythmic expression of timeless: a basis for promoting circadian cycles in period gene autoregulation.* Science, 270 (1995), 808-10.

[46] P. Smolen, D. Baxter, J. Byrne, *Modeling transcriptional control in gene networks– methods, recent results, and future directions.* Bull. Math. Biol., 62 (2000), 247-92.

[47] P. Smolen, D. A. Baxter, J. H. Byrne. *Mathematical Modeling of Gene Networks.* Neuron, 26 (2000), 567-80.

[48] P. Smolen, D. A. Baxter, J. H. Byrne. *Modeling circadian oscillations with interlocking positive and negative feedback loops.* J. Neurosci., 21 (2001), 6644-56.

[49] P. Smolen, P. E. Hardin, B. S. Lo, D. A. Baxter, J. H. Byrne. *Simulation of Drosophila circadian oscillations, mutations, and light responses by a model with VRI, PDP-1, and CLK.* Biophys. J., 86 (2004), 2786-802.

[50] J. Stelling, E. D. Gilles, F. J. Doyle, 3rd. *Robustness properties of circadian clock architectures.* Proc. Natl. Acad. Sci. USA, 101 (2004), 13210-5.

[51] J. J. Tyson, C. I. Hong, C. D. Thron, B. Novak. *A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM.* Biophys. J., 77 (1999), 2411-17.

[52] H. R. Ueda, M. Hagiwara, H. Kitano. *Robust oscillations within the interlocked feedback model of Drosophila circadian rhythm.* J. Theor. Biol., 210 (2001), 401-6.

[53] R. N. van Gelder, E. D. Herzog, W. J. Schwartz, P. H. Taghert. *Circadian rhythms: in the loop at last.* Science, 300 (2003), 1534-5.

[54] A. Varma, M. Morbidelli, H. Wu. Parametric sensitivity in chemical systems. Cambridge Univ. Press, Cambridge, 1999, pp. xvi, 342 p.

[55] Z. Xie, D. Kulasiri. *Modelling of circadian rhythms in Drosophila incorporating the interlocked PER/TIM and VRI/PDP1 feedback loops.* J. Theor. Biol., 245 (2007), 290-304.

[56] C. H. Yuh, H. Bolouri, E. H. Davidson. *Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control.* Development, 128 (2001), 617-29.

[57] H. Zeng, Z. Qian, M. P. Myers, M. Rosbash. *A light-entrainment mechanism for the Drosophila circadian clock.* Nature, 380 (1996), 129-35.

[58] D. M. Zerr, J. C. Hall, M. Rosbash, K. K. Siwicki. *Circadian fluctuations of period protein immunoreactivity in the CNS and the visual system of Drosophila.* J. Neurosci., 10 (1990), 2749-62.

A   Assumptions (for biological justification see [55])

1. The separation between the nuclear and cytoplasmic compartments is ignored in the model; instead we assume that all the reactions take place over a whole cell.

2. Phosphorylation of proteins is not considered. Although we are aware that phosphorylation is important to provide the time delay between mRNAs and proteins, the focus of the current model is on the transcriptional regulation, and phosphorylation of proteins is not included at this stage for the sake of simplification.

3. Gene expression of *per*, *tim*, *vri* and *pdp1* is activated by binding of CLK/CYC dimers to E-boxes in their promoter regions. Five binding sites are assumed, including E-boxes and E-box-like binding sites, in the *tim* and *per* promoter regions.

4. PER/TIM dimers are assumed not to bind to CLK/CYC dimers if the latter are bound to promoters.

5. It is assumed that the concentration of CYC in the system is constant (1.00 nM is assumed) so that there is always enough CYC bound to CLK to form dimers.

B   Kinetic equations

The model described by a system of ODEs is outlined below. For clarity, these equations are grouped into four categories. The names of mRNAs are written in lower case with a subscript 'm' denoting mRNA. The names of proteins and complexes are written in upper case. Abbreviations used for variable names are: PDP for PDP1, CC for CLK/CYC dimer, PT for PER/TIM dimer and CCPT for CLK/CYC/PER/TIM complex. The biochemical meaning of the parameters is explained in Table1.

B.1   Probabilities of TFs binding to a binding site (an E-box or a V/P box) in promoters:

The binding probabilities defined in the model are CLK/CYC binding to an E-box element in *per* promoter (Pr_{cper}), in *tim* promoter (Pr_{ct}), in *vri* promoter (Pr_{cv}), and to *pdp1* promoter (Pr_{cpdp}); VRI binding to the V/P box in *clk* promoter (Pr_{vc}), and PDP1 binding to that in *clk* promoter (Pr_{pc}).

$$d(\mathrm{Pr}_{cper})/dt = (1 - \mathrm{Pr}_{cper}) \times bccper_p \times CC - \mathrm{Pr}_{cper} \times ubccper_p \tag{B.1}$$

$$d(\mathrm{Pr}_{ct})/dt = (1 - \mathrm{Pr}_{ct}) \times bcctim_p \times CC - \mathrm{Pr}_{ct} \times ubcctim_p \tag{B.2}$$

$$d(\mathrm{Pr}_{cpdp})/dt = (1 - \mathrm{Pr}_{cpdp}) \times bccpdp_p \times CC - \mathrm{Pr}_{cpdp} \times ubccpdp_p \tag{B.3}$$

$$d(\mathrm{Pr}_{cv})/dt = (1 - \mathrm{Pr}_{cv}) \times bccvri_p \times CC - \mathrm{Pr}_{cv} \times ubccvri_p \tag{B.4}$$

$$d(\mathrm{Pr}_{pc})/dt = (1 - \mathrm{Pr}_{vc} - \mathrm{Pr}_{pc}) \times bpdpclk_p \times PDP - \mathrm{Pr}_{pc} \times ubpdpclk_p \tag{B.5}$$

$$d(\mathrm{Pr}_{vc})/dt = (1 - \mathrm{Pr}_{vc} - \mathrm{Pr}_{pc}) \times bvriclk_p \times VRI - \mathrm{Pr}_{vc} \times ubvriclk_p \tag{B.6}$$

B.2   Time evolution of mRNAs of *per, tim, clk, vri* and *pdp1*:

$$
\begin{aligned}
d(clk_m)/dt =\ & (\mathrm{Pr}_{vc} \times tcvriclk_p + \mathrm{Pr}_{pc} \times tcpdpclk_p + (1 - \mathrm{Pr}_{vc} - \mathrm{Pr}_{pc}) \times tcclk_p) \\
& \times clk_p - dclkm \times clk_m
\end{aligned}
\tag{B.7}
$$

$$
\begin{aligned}
d(per_m)/dt =\ & ((1 - (1 - \mathrm{Pr}_{cper})^{npt}) \times tcccper_p + (1 - \mathrm{Pr}_{cper})^{npt} \times tcdvpmt) \\
& \times per_p - dperm \times per_m
\end{aligned}
\tag{B.8}
$$

$$
\begin{aligned}
d(tim_m)/dt =\ & ((1 - (1 - \mathrm{Pr}_{ct})^{npt}) \times tccctim_p + (1 - \mathrm{Pr}_{ct})^{npt} \times tcdvpmt) \\
& \times tim_p - dtimm \times tim_m
\end{aligned}
\tag{B.9}
$$

$$
\begin{aligned}
d(vri_m)/dt =\ & ((1 - (1 - \mathrm{Pr}_{cv})^{nvri}) \times tcccvri_p + (1 - \mathrm{Pr}_{cv})^{nvri} \times tcdvpmt) \\
& \times vri_p - dvrim \times vri_m
\end{aligned}
\tag{B.10}
$$

$$
\begin{aligned}
d(pdp_m)/dt =\ & ((1 - (1 - \mathrm{Pr}_{cp})^{npdp}) \times tcccpdp_p + (1 - \mathrm{Pr}_{cp})^{npdp} \times tcdvpmt) \\
& \times pdp_p - dpdpm \times pdp_m
\end{aligned}
\tag{B.11}
$$

B.3   Time evolution of PER, TIM, CLK, VRI and PDP1 proteins:

$$d(PER)/dt = tlper \times per_m - bpt \times PER \times TIM + ubpt \times PT - dper \times PER \tag{B.12}$$

$$d(TIM)/dt = tltim \times tim_m - bpt \times PER \times TIM + ubpt \times PT - dtim \times TIM \tag{B.13}$$

$$d(CLK)/dt = tlclk \times clk_m - bcc \times CLK \times CYC + ubcc \times CC - dclk \times CLK \tag{B.14}$$

$$d(VRI)/dt = tlvri \times vri_m - dvri \times VRI \tag{B.15}$$

$$d(PDP)/dt = tlpdp \times pdp_m - dpdp \times PDP \tag{B.16}$$

B.4   PER/TIM, CLK/CYC and PER/TIM/CLK/CYC complexes:

$$
\begin{aligned}
d(PT)/dt \;=\; & bpt \times PER \times TIM - ubpt \times PT - bccpt \times PT \times CC \\
& +ubccpt \times CCPT - dpt \times PT
\end{aligned}
\tag{B.17}
$$

$$
\begin{aligned}
d(CC)/dt \;=\; & bcc \times CLK \times CYC - ubcc \times CC - bccpt \times PT \times CC \\
& +ubccpt \times CCPT - dcc \times CC
\end{aligned}
\tag{B.18}
$$

$$
d(CCPT)/dt = bccpt \times PT \times CC - ubccpt \times CCPT - dccpt \times CCPT
\tag{B.19}
$$

In summary, the circadian clock model is described by a system of 19 ODEs with 47 parameters.

Parameter	Biochemical meaning
bccpdpp	binding rate of CLK/CYC to an E-box in *pdp1* promoter
bccperp	binding rate of CLK/CYC to an E-box in *per* promoter
bcctimp	binding rate of CLK/CYC to an E-box in *tim* promoter
bccvrip	binding rate of CLK/CYC to an E-box in *vri* promoter
bpdpclkp	binding rate of PDP1 to the V/P box in *clk* promoter
bvriclkp	binding rate of VRI to the V/P box in *clk* promoter
ubccpdpp	unbinding rate of CLK/CYC to an E-box in *pdp1* promoter
ubccperp	unbinding rate of CLK/CYC to an E-box in *per* promoter
ubcctimp	unbinding rate of CLK/CYC to an E-box in *tim* promoter
ubccvrip	unbinding rate of CLK/CYC to an E-box in *vri* promoter
ubpdpclkp	unbinding rate of PDP1 to the V/P box in *clk* promoter
ubvriclkp	unbinding rate of VRI to the V/P box in *clk* promoter
bcc	association rate of CLK/CYC dimer
bpt	association rate of PER/TIM dimer
bccpt	association rate of CLK/CYC/PER/TIM complex
ubcc	dissociation rate of CLK/CYC dimer
ubpt	dissociation rate of PER/TIM dimer
ubccpt	dissociation rate of CLK/CYC/PER/TIM complex
tcccpdpp	transcription rate of CLK/CYC-activated *pdp1* gene
tcccperp	transcription rate of CLK/CYC-activated *per* gene
tccctimp	transcription rate of CLK/CYC-activated *tim* gene
tcccvrip	transcription rate of CLK/CYC-activated *vri* gene
tcpdpclkp	transcription rate of PDP1-activated *clk* gene
tcvriclkp	transcription rate of VRI-repressed *clk* gene
tcclkp	transcription rate of *clk* gene binding neither PDP1 nor VRI
tcdvpmt	transcription rate of deactivated *per, tim, vri* or *pdp1* gene
tlclk	translation rate of *clk* mRNA
tlpdp	translation rate of *pdp1* mRNA
tlper	translation rate of *per* mRNA
tltim	translation rate of *tim* mRNA
tlvri	translation rate of *vri* mRNA
dclkm	degradation rate of *clk* mRNA
dpdpm	degradation rate of *pdp1* mRNA
dperm	degradation rate of *per* mRNA
dtimm	degradation rate of *tim* mRNA
dvrim	degradation rate of *vri* mRNA
dclk	degradation rate of CLK protein
dpdp	degradation rate of PDP1 protein
dper	degradation rate of PER protein
dtim	degradation rate of TIM protein
dvri	degradation rate of VRI protein
dpt	degradation rate of PER/TIM dimer
dcc	degradation rate of CLK/CYC dimer
dccpt	degradation rate of CLK/CYC/PER/TIM complex
npt	number of E-boxes in *per* or *tim* promoter
nvri	number of E-boxes in *vri* promoter
npdp	number of E-boxes in *pdp1* promoter

Table 1: Biochemical meaning of the parameters

Index	Parameter	Value	Index	Parameter	Value
1	bccpdpp	0.062	25	tcclkp	1.42
2	bccperp	0.069	26	tcdvpmt	0.053
3	bcctimp	0.069	27	tlclk	35
4	bccvrip	0.1	28	tlpdp	1.87
5	bpdpclkp	1.155	29	tlper	36
6	bvriclkp	1.858	30	tltim	36
7	ubccpdpp	0.145	31	tlvri	14.68
8	ubccperp	0.262	32	dclkm	0.643
9	ubcctimp	0.262	33	dpdpm	0.06
10	ubccvrip	0.276	34	dperm	0.053
11	ubpdpclkp	0.952	35	dtimm	0.053
12	ubvriclkp	1.043	36	dvrim	0.07
13	bcc	2.349	37	dclk	0.2
14	bpt	1.1	38	dpdp	0.156
15	bccpt	51	39	dper	0.62
16	ubcc	0.89	40	dtim	0.62
17	ubpt	2.93	41	dvri	1.226
18	ubccpt	7.89	42	dpt	0.279
19	tcccpdpp	9.831	43	dcc	0.184
20	tcccperp	11	44	dccpt	15.122
21	tccctimp	11	45	npt	5
22	tcccvrip	16.86	46	nvri	4
23	tcpdpclkp	125.54	47	npdp	6
24	tcvriclkp	0.028			

Table 2: Parameters of the model: The units of binding rates and association rates are nM-1.3h-1.3 and the units of the other parameters are h^{-1}

Species	Concentration (nM)	Species	Concentration (nM)
CC	0.5566	clkp*	0.003185
CCPT	0.4982	pdpp*	0.003185
CLK	3.6628	perp*	0.003185
clkm	0.2583	timp*	0.003185
PDP	4.1953	vrip*	0.003185
pdpm	0.3175		
PER	2.7527		
perm	0.2395	Probability	Value
PT	0.4014	prcpdp	0.08
TIM	2.7527	prcper	0.0431
timm	0.2395	prct	0.043
VRI	3.175	prcv	0.0585
vrim	0.2571	prpc	0.426
CYC*	1.00	prvc	0.489

Table 3: Initial conditions. Abbreviations: CC – CLK/CYC, PT – PER/TIM, CCPT – CLK/CYC/PER/TIM. Constant values in the system are denoted by *