Oscillation, cooperativity, and intermediates in the self-repressing gene

David Lepzelter [a], Haidong Feng [b], Jin Wang [a,b,c,*]

[a] *Department of Physics and Astronomy, Stony Brook University, United States*
[b] *Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, United States*
[c] *State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, PR China*

ARTICLE INFO

ABSTRACT

Biological oscillators are vital to living organisms, which use them as clocks for time-sensitive processes. However, much is unknown about mechanisms which can give rise to coherent oscillatory behavior, with few exceptions (e.g., explicitly delayed self-repressors and simple models of specific organisms' circadian clocks). We present what may be the simplest possible reliable gene network oscillator, a self-repressing gene. We show that binding cooperativity, which has not been considered in detail in this context, can combine with small numbers of intermediate steps to create coherent oscillation. We also note that noise blurs the line between oscillatory and non-oscillatory behavior.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Biological oscillation and its mechanisms have recently become a subject of intense study by a number of experimental and theoretical groups. Its study is still in the early stages, and while the phenomenon itself is believed to be of great importance, knowledge of mechanisms by which it can occur is quite incomplete.

Oscillation is used in a number of biological systems, most notably in the circadian rhythms responsible for keeping organisms' biochemical processes in line with the day–night schedule of the Earth. A number of mechanisms can be used by organisms to give a reasonably coherent oscillation, including three-gene repressilator systems [1–3], self-repressors with maturation times or other explicit, deterministic time delays (as opposed to intermediate steps which would result from a full treatment of chemical intermediates) [4–7], more complex negative feedback loops tied to population dynamics [8] combined repression and activation loops [9], highly non-linear protein degradation [10], very large numbers (hundreds) of intermediate steps [11], and self-repressors whose production and gene repression involve diffusion through the nuclear membrane [12,13]. A self-activator with one stable and one metastable state, and stochastic switching between the two, can cause what may be called incoherent oscillation [14]. Additionally, certain kinds of self-repression can cause behavior which is not coherent oscillation when one considers the deterministic average protein and mRNA concentrations, but which still appear quite oscillatory and reasonably coherent when one considers any given system's stochastic trajectory through protein and mRNA concentration space over time [15].

This last example is of particular interest because in part of its simplicity; it is an uncomplicated system, easily modelled using straightforward Markov chain Monte Carlo methods.[1] There is, however, another reason for interest in the work. The binding mechanism suggested by the mathematics presented by McKane and Newman in Ref. [15] is in fact a combination of multiple n-mer bindings of regulatory proteins to the gene. In addition, the which n contributes most to the genetic activity here depends on the single parameter which McKane and Newman used to tweak their system and the average number of proteins present in the system. While this in itself may be non-physical, it inherently suggests that the number of bound proteins may be able to change the system from non-oscillatory to oscillatory with noise to (possibly) oscillatory even without noise.

We note that the simplest possible version of this last suggestion, a deterministic model with mRNA and with n regulatory proteins binding directly and instantaneously to the gene, cannot be oscillatory. A brief proof of this is included in the section on linear stability analysis (however, the stochastic version of the system can have oscillation-like behavior, and we explore this briefly).

2. Methods

We performed both deterministic and stochastic calculations. For ease of calculation and interpretation, the time units used were

---

* Corresponding author at: Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, United States.
E-mail address: jin.wang.1@stonybrook.edu (J. Wang).

[1] Deterministic time delays of the type $t - \tau$, the most common method used to get a simple self-repressor to oscillate, can reproduce most results from other methods. However, for stochastics they require non-Markov simulation. Furthermore, while the equations are straightforward, the authors know of no way to determine *a priori* the numerical values of time delays that substitute for the delaying steps. The methods outlined in this Letter, in contrast, are purely Markov and involve only the actual chemical rates of the represented processes.

mRNA lifetimes (also monomer protein lifetimes). The following were constant for all calculations: the degradation rate of mRNA $k_m = 1$ (hence, $\tau_m = 1$); the degradation rate of monomer protein $k_p = 1$ (hence, $\tau_p \geqslant 1$); the rates of gene binding and unbinding $\omega \gg 1$ (and so the system is very adiabatic); the ratio of the n-mer dissolution constant f to the formation constant h is $\frac{f}{h} = \left(\frac{g_1 g_p}{k_m k_p}\right)^n$; the rate of n-mer dissolution $f \gg 1$ in the simplest case, Fig. 1, and $f = 1$ in all other calculations; the number of complete n-mers which will cause the gene to be repressed by a factor of $1/e$, $X_{eq} = 10$; the protein synthesis rate from a single strand of mRNA $g_p = 3$; and the ratio of mRNA synthesis in the repressed versus the unrepressed gene state $g_0 = g_1/100$. It should be noted that g_p and X_{eq} can be rescaled in systems without internal noise; only when the actual number becomes important, as opposed to relative concentrations, does do these quantities have any significant effect other than a rescaling. It should also be noted that the ratio τ_p/τ_m is in rough agreement with the average for this sort of gene in yeast, given as approximately 2 in Ref. [16].

Deterministic calculations (involving the average numbers of mRNA, proteins, and protein aggregates) used the simple set of equations

$$\frac{dm}{dt} = -k_m m + g_0 P_{\text{off}} + g_1 P_{\text{on}} = -k_m m + g_0 + \frac{g_1 - g_0}{1 + c/X_{eq}},$$

$$\frac{dp}{dt} = -k_p p + g_p m - hnp^n + fnc,$$

$$\frac{dc}{dt} = hp^n - fc,$$

where m is the number of mRNA molecules, P_{off} (P_{on}) is the probability or fraction of the time that the gene is repressed (unrepressed), p is the number of monomer proteins, and c is the number of n-mer proteins in the system at time t. Simple modifications made additional intermediate steps possible (c becoming c_1, n becoming n_1, additional terms in the last equation for the $n_2 c_1 \rightarrow c_2$ step, the extra equations for c_2 in terms of c_1 and c_3, etc.).

We should also note that we assume all binding steps involved are fast compared to the other processes involved in the system, though binding events are relatively rare, and binding is highly cooperative. Slower binding, or binding that is less cooperative, can introduce very different elements to a system, as shown in Ref. [17].

The authors used time-series data to determine the oscillatory nature of the systems. A general solution to the linear stability analysis of the system would involve analytically solving an $n + 1$-degree polynomial with non-integer parameter-dependent coefficients, which is infeasible [18]. Even numerical solutions to the linear stability analysis problem alone could have glossed over complex behavior less useful to an organism than regular oscillation (i.e., chaos). However, individual points in different regions

identified by the time-series data have been checked numerically using linear stability analysis. In the regime noted as 'oscillatory' via the time-series data, the Jacobian at the equilibrium point has one real negative eigenvalue and two complex eigenvalues with positive real parts, confirming that the point can be oscillatory. Outside that region, all three eigenvalues' real parts are negative, confirming the non-oscillatory classification. We note that the line between the zones labeled 'decaying oscillation' and 'no oscillation' is much less well-defined than the line between those labeled 'oscillation' and 'decaying oscillation', as a system which displays decaying oscillation given one set of initial conditions can have no oscillations with another set of initial conditions. However, we believe the general trend is correct and useful, and so we indicate the distinction in the figure.

Stochastic calculations were straightforward, using the following equations:

$$R_{\text{mRNA synthesis}} = g_0 + \frac{g_1 - g_0}{1 + c/X_{eq}},$$

$$R_{\text{mRNA degradation}} = k_m m,$$

$$R_{\text{protein synthesis}} = g_p m,$$

$$R_{\text{monomer protein degradation}} = k_p p.$$

In the calculations for Fig. 1 (in which binding to the gene is coupled with n-mer formation), we used

$$c = \frac{h}{f} p(p-1) \cdots (p-n+1),$$

derived from the standard master equation rates of binding $hp(p-1) \cdots (p-n+1)$ and unbinding f. For other stochastic calculations, in which the proteins bind to each other before binding to the gene, we used instead

$$R_{n-\text{mer formation}} = hp(p-1) \cdots (p-n+1),$$

$$R_{n-\text{mer breaking}} = fc.$$

A time-step dt was then calculated using these rates, ensuring that events generally occurred one at a time by using $dt \sim 0.01/\sqrt{\sum R^2}$ (this is subtly different from the traditional Gillespie simulation in that multiple events were in theory allowed, if quite unlikely; however, the difference is very small and may reasonably be considered to be an advantage of our algorithm.) For practicality's sake, at very high levels of synthesis, multiple synthesis and decay events were allowed and occurred in the background of other events, but were kept to at most a 0.01% mean change in the number of molecules.

Each process was then treated as a Poisson process using the same dt, with the probability of a events of type b occurring being $\frac{e^{-R_b dt}(R_b dt)^a}{a!}$.
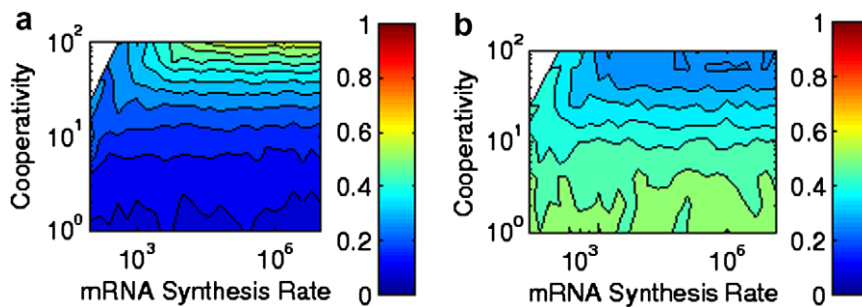
Fig. 1. (a) Stochastic calculation of coherence in a system which has mRNA and proteins cooperatively binding to the gene, with coherence given by $2\frac{\sum \Theta(d\phi)}{\sum |d\phi|} - 1$ where $d\phi$ is the difference in angle in mRNA–protein space. (b) Stochastic calculation of the standard deviation of the period distribution divided by its mean. Both colormaps are on the same scale as Fig. 3.

The programs used were written in C, and run using FEDORA 10 LINUX on a Dell desktop computer.

3. Linear stability analysis

We present, as an aside, a brief proof of the non-oscillation of the system described without delays. This system is described by the equations

$$\frac{dm}{dt} = -k_m m + g_0 + \frac{g_1 - g_0}{1 + p^n / X_{eq}},$$

$$\frac{dp}{dt} = -k_p p + g_p m.$$

The Jacobian of the system is given by

$$\begin{pmatrix} -k_m & -\frac{(g_1 - g_0) n p^{n-1} / X_{eq}}{(1 + p^n / X_{eq})^2} \\ g_p & -k_p \end{pmatrix},$$

with eigenvalues that add to a real negative number and multiply to a real positive number. If the eigenvalues are complex, due to the result of multiplication they have the same real part. Due to this and the result of the addition, both eigenvalues must have negative real parts, meaning oscillation is impossible. If the eigenvalues are real, both must be negative, again making oscillation impossible.

It should be noted that we made these conclusions based on the structure of the Jacobian without finding the fixed point, based on the fact that there are only two eigenvalues. The more complex cases described in this work would have three or more eigenvalues, which would make finding the fixed point necessary. They also add to the complexity of the equations for finding the fixed point, which are already of degree $n + 1$ with multiple terms and non-integer coefficients.

4. Results

We begin with the stochastic system described first, in which binding to the gene is coupled with n-mer formation (this could be either because the binding to the gene is itself cooperative or because it is very fast once the n-mer is formed). In the described regime, whose deterministic solutions yield at best decaying oscillation, we now explore the possibility of noise-induced oscillation. We note that the system in this case is between truly oscillatory and simply two-state with reasonably frequent switching. Simple two-state switching would lead to a 'period distribution' (time between maxima) with a normalized standard deviation of that period $\frac{\sigma_t}{t}$ of $\frac{1}{\sqrt{2}}$. Fig. 1b shows $\frac{\sigma_t}{t}$, and Fig. 1a shows the coherence (generalized from the definition used in Ref. [2]). Coherence is defined as $2 \frac{\sum_t \Theta(d\phi(t))}{\sum_t |d\phi(t)|} - 1$ where ϕ is the angle in mRNA–protein space from the average concentration, and $d\phi(t) = \phi(t) - \phi(t - 0.1)$. It is, in essence, a measure of how often a system goes forward (clockwise) instead of backward (counterclockwise) in its oscillation. Both the coherence and $\frac{\sigma_t}{t}$ strongly imply that increased cooperativity yields a steadier oscillatory behavior, but that in the regions explored there is no coherent oscillation (which would require a coherence close to 1 and a value of $\frac{\sigma_t}{t}$ significantly less than 1).

So far, the only pieces of the puzzle considered have been multimer binding of proteins to genes, mRNA, and noise. We now add the final piece, the possibility of proteins binding together before they bind to the gene. The number of steps involved here keeps oscillation from occurring at a cooperativity of 1, but increased n causes oscillatory behavior when combined with the rest of the system.

Specifically, Fig. 2 shows a stochastic system with $n = 16$ and the distribution of periods, defined here as time between local
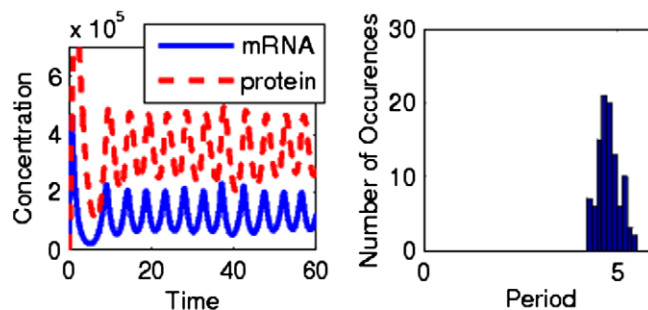

Fig. 2. Left graph: oscillation due to an intermediate step with a cooperativity of 16 and noise. Right graph: period distribution for the same system.

maxima in protein number which are at least $0.4 \cdot \tau_{mRNA}$ (in order to remove less significant fluctuations from consideration). This system is clearly oscillatory, with a reasonably sharp period distribution.

It should be noted that 16 is at best a marginally reasonable value for cooperativity in a simple genetic system. However, it is clear now that simple intermediate steps can interact with cooperativity, and between them can produce oscillations in which neither is the primary factor in the behavior. Furthermore, examining the parameter space in more detail, we find telling behavior in both deterministic and stochastic cases. Fig. 3a shows the deterministic phase diagram of the system, made using time-series calculations. Region I is oscillatory, region II displays decaying oscillation, and region III is non-oscillatory. High cooperativity and synthesis rate are both clearly necessary for this; also, it should again be noted that the additional intermediate step as compared with the system from Fig. 1 is necessary for deterministic oscillation.

In Fig. 3b, we see, in the same region, a graph of the coherence. It corresponds well with the deterministic calculations: the system reliably goes in the correct direction (associated with time going forward) in the stochastic case when its parameters would predict that it is deterministically oscillatory. Fig. 3c shows $\frac{\sigma_t}{t}$, which again corresponds well, although not perfectly; some loss of specificity in period at very high cooperativity and synthesis implies that the oscillation may be imperfect or require that other parameters be very specific at some level. However, the increase in $\frac{\sigma_t}{t}$ at these points is slight.

From these figures, it is apparent that coherence is possible with higher rates of synthesis, and that there is a tendency towards higher coherence with higher cooperativity. Additionally, we note that in the stochastic case there is no jump in coherence or $\frac{\sigma_t}{t}$ as we would find if the system exhibited the phase-transition behavior from the deterministic system, Fig. 3a; in the stochastic case (Fig. 3b and c), the line between oscillatory and non-oscillatory is blurred. It is important to note that noise alone is not sufficient for reliable or even semi-reliable oscillation. The deterministically oscillating region clearly has more stable oscillation in the stochastic regime, and systems close to it are more capable of reliable behavior than those far away from it.

For comparison, we now choose a point in the region of deterministic decaying oscillations, and make plots comparable to Fig. 2. In Fig. 4 we see a system whose behavior seems similar in many ways; the main difference in this case is a slightly widened period distribution. While this system is definitely in region II of Fig. 3a (with a cooperativity of 8), and the period distribution may be too wide for a truly accurate clock, it is obvious that the line between coherent oscillation and incoherent oscillation-like behavior is blurred in this case by noise.

We see in these figures that the optimal cooperativity for this number of intermediate steps in binding is still, at best, at the very
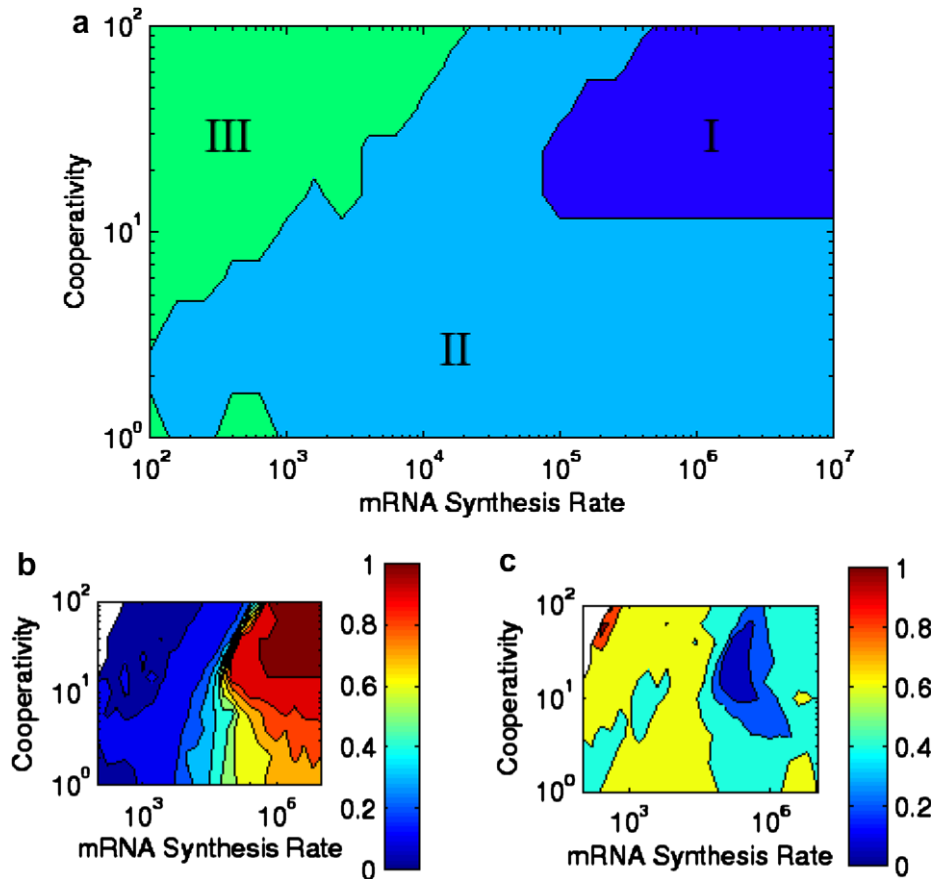
Fig. 3. (a) Deterministic calculation of oscillatory features in a system with mRNA, cooperative binding of proteins to each other, and a separate step binding to the gene; region I is oscillatory, region II has decaying oscillations, and region III is non-oscillatory. Due to the fact that the distinction betweeen II and III is somewhat dependent on initial conditions, there may be some overlap between those two regions as discussed in the Methods section. (b) Stochastic calculation of coherence, given by $2\frac{\sum \Theta(d\phi)}{\sum |d\phi|} - 1$ where $d\phi$ is the difference in angle in mRNA–protein space. (c) Stochastic calculation of the standard deviation of the period distribution divided by its mean.
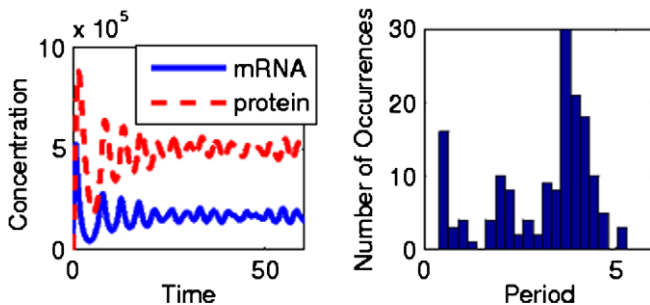
Fig. 4. Left graph: behavior with small oscillatory tendencies due to an intermediate step with a cooperativity of 8 and noise. Right graph: period distribution for the same system.

high end of cooperativity in reasonable biological systems. Therefore, we have attempted a less complete search of parameters considering additional steps (instead of monomer to n-mer, monomer to dimer to tetramer, etc.). While our searches were not exhaustive, we found at least one region in which such a system can oscillate at quite reasonable values for n, as low as 8 (octamer). For reference, this is not an uncommon amount of cooperativity for these sorts of proteins. SARS–CoV nsp10, for instance, forms a dodecameric structure ($n = 12$) [19]; *Rhizobium leguminosarum* NodD has been shown to bind to DNA preferentially in octamer form [20]; and λ phage CI, while it has a small autoregulatory effect

with tetramer and even dimer forms, only fully represses itself when an octamer form of the protein links distant sites on DNA together [21].

5. Conclusions

Our data show that oscillation occurs when cooperativity and intermediate steps are both present, when neither cooperativity nor intermediates would otherwise be sufficient. More, it implies that, with many intermediate steps, lower cooperativity can yield oscillation, while with higher cooperativity fewer intermediate steps are necessary. Oscillation behavior in deterministic calculations gives rise to more stable oscillation in the stochastic case.

Additionally, it should be noted that, even though we found no coherent oscillation for lower values of cooperativity, moderate coherence in an oscillation-like behavior may not necessarily be detrimental to a biological system which relies on it as a clock for functions which are not truly vital. Many biological clocks receive input of some kind from outside sources, whether light and food for 24-h clocks or some other form [22,23]. These inputs can reset a biological clock, and one may argue that they might be able to do so more easily in a fundamentally inexact clock than a single-mode oscillator.

In summary, we have discovered a relationship between generally realistic intermediate steps (as opposed to set-time time delays of the form $t - \tau$, which are realistic for a limited set of systems), cooperativity, and oscillation. Intermediate steps result-

ing from slow cooperative binding can cause oscillation at biologically relevant cooperativity. We have also explored the noise effect, in which the line between oscillatory and non-oscillatory (relatively clear in the deterministic case) is blurred in the stochastic case.

References

[1] M.B. Elowitz, S. Leibler, Nature 403 (2000) 335.
[2] M. Yoda, T. Ushikubo, W. Inoue, M. Sasai, J. Chem. Phys. 126 (2007) 115101.
[3] K.Y. Kim, D. Lepzelter, J. Wang, J. Chem. Phys. 126 (2007) 034702.
[4] J. Lewis, Curr. Biol. 13 (2003) 1398.
[5] W. Mather, M.R. Bennett, J. Hasty, L.S. Tsimring, Phys. Rev. Lett. 102 (2009) 068105.
[6] J. Stricker, S. Cookson, M.R. Bennett, W.H. Mather, L.S. Tsimring, J. Hasty, Nature 456 (2008) 516.
[7] P.-E. Morant, Q. Thommen, F. Lemaire, C. Vendermoëre, B. Parent, M. Lefranc, Phys. Rev. Lett. 102 (2009) 068104.
[8] F.K. Balagaddé, L. You, C.L. Hansen, F.H. Arnold, S.R. Quake, Science 309 (5731) (2005) 137.
[9] T.Y.-C. Tsai, Y.S. Choi, W. Ma, J.R. Pomerening, C. Tang, J.E.J. Ferrell, Science 231 (2008) 126.
[10] J.J. Tyson, C.I. Hong, D. Thron, B. Novak, Biophys. J. 77 (1999) 2411.
[11] L.G. Morelli, F. Jülicher, Phys. Rev. Lett. 98 (2007) 228101.
[12] A. Goldbeter, Proc. Roy. Soc. B 261 (1995) 319.
[13] J. Wang, L. Xu, E. Wang, PMC Biophys. 1 (2008) 7.
[14] D. Schultz, E.B. Jacob, J.N. Onuchic, P.G. Wolynes, PNAS 104 (2007) 17582.
[15] A.J. McKane, J.D. Nagy, T.J. Newman, M.O. Stefanini, J. Stat. Phys. 128 (2007) 165.
[16] V. Shahrezaei, P.S. Swain, PNAS 105 (45) (2008) 17256.
[17] M. Santillán, Math. Model. Nat. Phenom. 3 (2) (2008) 85.
[18] J.B. Fraleigh, A First Course in Abstract Algebra, fifth edn., Addison-Wesley, Chennai, 1994.
[19] D. Su et al., J. Virol. 80 (16) (2006) 7902.
[20] L. Songtao, L. Huafeng, H. Guofan, Sci. China C 41 (6) (1998) 592.
[21] I.B. Dodd, A.J. Perkins, D. Tsemitsidis, J.B. Egan, Genes Dev. 15 (22) (2001) 3013.
[22] J. Aschoff, S. Daan, The Entrainment of Circadian Systems, Handbook of Behavioral Neurobiology, vol. 12, Kluwer Academic/Plenum Publishers, 2001. p. 7.
[23] E. Challet, I. Caldelas, C. Graff, P. Pévet, Synchronization of the molecular clockwork by light- and food-related cues in mammals, Biol. Chem. 384 (2003) 711.