

# Oscillateurs génétiques simples. Application à l'horloge circadienne d'une algue unicellulaire

Pierre-Emmanuel Morant<sup>1</sup>, Constant Vandermoere<sup>1</sup>, Quentin Thommen<sup>1</sup>, Benjamin Parent<sup>2</sup>, François Lemaire<sup>3</sup>, Florence Corellou<sup>4</sup>, Christian Schwartz<sup>4</sup>, François-Yves Bouget<sup>4</sup>, Marc Lefranc<sup>1</sup>

<sup>1</sup> Laboratoire de Physique des Lasers, Atomes, Molécules, UMR CNRS 8523, UFR de Physique, Bât. P5, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France.

<sup>2</sup> Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS 8576, Bât. C9, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France.

<sup>3</sup> Laboratoire d'Informatique Fondamentale de Lille, UMR CNRS 8022, Bât. M3, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France.

<sup>4</sup> Laboratoire Modèles en Biologie Cellulaire et Evolutive, UMR CNRS-Paris 6 7628, Observatoire Océanologique de Banyuls sur mer, BP44, 66651 Banyuls sur Mer Cedex, France.

marc.lefranc@univ-lille1.fr

**Résumé.** Un gène réprimé par l'expression de sa propre protéine constitue l'exemple le plus simple de circuit génétique à boucle de rétroaction négative, et l'apparition d'oscillations dans ce système est un problème classique de la biologie théorique. Nous nous intéressons ici au cas où le taux de transcription ne suit pas instantanément la concentration en protéine, mais se comporte comme une variable dynamique indépendante. Nous observons que l'existence d'une dynamique transcriptionnelle favorise les oscillations, et que ces dernières apparaissent de manière systématique dans la limite où les dégradations de l'ARN et de la protéine sont totalement saturées. Nous considérons également la généralisation la plus directe du gène auto-régulé : une boucle à deux gènes, l'un activateur, l'autre répresseur, se régulant réciproquement, et nous comparons ses prédictions aux données expérimentales concernant les oscillations circadiennes d'une algue unicellulaire verte.

**Abstract.** A gene which is repressed by its own protein is the simplest example of a genetic circuit with a negative feedback, and the appearance of oscillations in this system is a classical problem in theoretical biology. Here we study the case where the transcription rate does not react instantaneously to changes in protein concentration but is an independent dynamical variable. We observe that the transcriptional dynamics favors oscillations, and that periodic regimes appear unconditionally in the limit where enzymatic degradations of ARN and protein are completely saturated. We also consider the simplest generalization of this oscillator, a circuit with two genes, an activator and a repressor, regulating each other, and compare its predictions to experimental data about circadian oscillations in a unicellular green alga.

## 1 Introduction

Les dizaines de milliers de gènes que porte la molécule d'ADN au cœur de chaque cellule contiennent l'information nécessaire à la synthèse des briques de la machinerie moléculaire de la Vie, les protéines. Cette synthèse s'effectue en deux étapes : "transcription" de la séquence codante en une molécule d'ARN messenger, puis "traduction" de cet ARN en une séquence d'acides aminés, c'est-à-dire une protéine. Or, les taux de production des ARN ne sont pas constants : l'activité des gènes est en effet régulée par des protéines produites par d'autres gènes, au travers de réseaux complexes. L'ensemble constitue donc un système dynamique fortement non linéaire, susceptible de présenter toute une gamme de comportements bien connus : bistabilité, mais aussi oscillations, comme par exemple celles intervenant dans la segmentation des somites lors de l'embryogénèse [1], ou dans les horloges circadiennes [2]. Ces dernières fournissent à un grand nombre d'organismes une mesure interne du temps leur permettant de faire varier de nombreuses grandeurs physiologiques sur une période de 24 heures, et de s'adapter ainsi à l'alternance jour-nuit. Leur caractère autonome est démontré par le fait qu'elles persistent en éclaircissement constant, avec une période naturelle légèrement différente de 24 heures.

L'oscillateur génétique le plus simple est a priori celui constitué d'un gène réprimé par la protéine qu'il produit, comme sans doute le gène *hes1* dans la segmentation des somites [1]. Il s'agit d'un problème ancien [3,4], pour lequel il est admis qu'on ne peut observer des oscillations que si on introduit soit une étape cinétique intermédiaire, par exemple une phosphorylation de la protéine [5] ou un transport entre cytoplasme et noyau [6], soit un terme explicite de délai dans les équations [7,8,9]. Nous avons revisité ce problème en tenant compte de deux effets complémentaires.

D'une part, des expériences récentes ont montré que le processus de transcription se caractérise par une cinétique complexe, et notamment par l'existence de "salves de transcription" [10] modulant l'activité transcriptionnelle sur des durées allant jusqu'à quelques dizaines de minutes. Comme François et Hakim [11], nous considérons donc le taux de transcription comme une variable dynamique à part entière, contrairement à l'immense majorité des études où on suppose qu'il réagit instantanément à la concentration en protéine. D'autre part, les analyses théoriques postulent généralement que les acteurs moléculaires sont dégradés par des mécanismes génériques, par exemple dégradation spontanée ou dirigée par une enzyme, avec une cinétique de type Michaelis-Menten. Or, l'importance de la cinétique de dégradation, et le pouvoir déstabilisant d'effets non linéaires, tels que la stabilisation de la forme dimère d'une protéine [12], ont été récemment soulignés [13].

Dans le cas du gène auto-régulé, nous avons constaté que l'existence d'une dynamique transcriptionnelle peut élargir considérablement le domaine de paramètres dans lequel un mécanisme de dégradation non linéaire induit des oscillations. Celles-ci sont observées de manière systématique dans la limite où les dégradations de l'ARN et de la protéine sont saturées mais peuvent apparaître bien avant.

## 2 Oscillations d'un gène réprimé par sa propre protéine

Comme François et Hakim [11], nous décrivons la dynamique transcriptionnelle par une simple équation cinétique décrivant des processus élémentaires d'association/dissociation entre la protéine et l'ADN, mais une modélisation plus complexe pourrait être envisagée. Dans ces conditions, la dynamique du circuit à un gène auto-régulé peut être modélisée par les trois équations adimensionnées suivantes :

$$\dot{g} = \theta [1 - g(1 + p^n)] \quad (1a)$$

$$\dot{p} = n\alpha [1 - g(1 + p^n)] + \delta[m - f(p)] \quad (1b)$$

$$\dot{m} = \mu + \lambda g - h(m) \quad (1c)$$

où  $g, p$  et  $m$  représentent respectivement l'activité du gène, et les quantités de protéines et d'ARN. L'entier  $n$  indique la coopérativité de la régulation, c'est-à-dire le nombre de protéines contenues dans le complexe protéique modulant l'activité du gène. L'unité de temps est le temps de demi-vie de l'ARN. Les coefficients  $\theta, \alpha$  contrôlent les échelles de temps des processus de dissociation et d'association à l'ADN, tandis que  $1/\delta$  est le temps de demi-vie de la protéine. Les paramètres  $\mu$  et  $\lambda$  déterminent l'activité du gène selon que celui-ci est libre et actif ( $g = 1$ ) ou lié et réprimé ( $g = 0$ ). Les fonctions  $f(p)$  et  $h(m)$ , qu'on suppose de pente unité à l'origine, décrivent respectivement les mécanismes de dégradation de la protéine et de l'ARN.

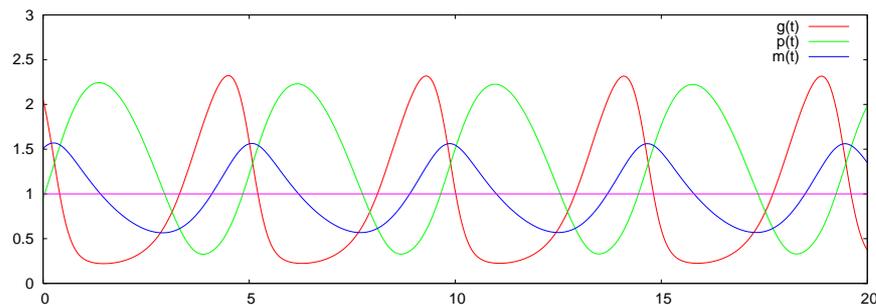
Pour étudier l'apparition d'oscillations dans ce système, nous n'envisagerons ici que la déstabilisation de l'état stationnaire des équations (1) via une bifurcation de Hopf menant à des oscillations périodiques. L'analyse de stabilité linéaire du système (1) montre que deux valeurs propres de la matrice jacobienne traversent l'axe imaginaire et acquièrent une partie réelle positive quand l'expression  $\mathcal{H}$  ci-dessous passe par zéro pour devenir négative (critère de Routh-Hurwitz) :

$$\begin{aligned} \mathcal{H} = & u h_0^2 (\alpha h_0 + \delta s \lambda) (\delta s \lambda + u \lambda + \alpha h_0) \\ & + \lambda^2 h_0 [h_0 (-\delta h_0 + 2u\alpha + \alpha \delta s) + \lambda(u + \delta s)^2] \theta \\ & + \lambda^4 (u + \delta s) \theta^2 \end{aligned} \quad (2)$$

où  $s$  et  $u$  sont les pentes des fonctions de dégradation  $f(p)$  et  $h(m)$  au point fixe et  $h_0$  est la valeur prise par la fonction de dégradation  $h(m)$  en ce point.

On voit facilement que lorsque  $T = u + \delta s \leq 0$ , l'expression (2) est négative pour toutes valeurs des constantes cinétiques  $\theta$  et  $\alpha$ . Cela indique que l'on observe alors systématiquement des oscillations, même pour des dynamiques transcriptionnelles extrêmement rapides, et en particulier dans le cas  $u, s \rightarrow 0$  où les dégradations enzymatiques sur l'ARN et la protéine sont saturées, un facteur d'instabilité bien connu [14]. Cela n'a rien de surprenant, car  $-T$  est la trace du jacobien du modèle à deux variables où l'activité du gène  $g$  est supposée être asservie à la concentration en protéine  $p$ , et l'on sait que pour un système à deux variables, la positivité de cette trace est synonyme d'instabilité [12,15].

L'expression (2) est plus intéressante si on adopte le point de vue que les constantes  $\theta$  et  $\alpha$  ne sont pas très grandes, comme on le suppose généralement, mais qu'elles doivent correspondre aux échelles de temps des "salves de transcription" observées expérimentalement. Ces dernières se caractérisent par des temps d'extinction allant jusqu'à quelques dizaines de minutes [10], soit  $\theta = O(1)$ . Nous avons observé que des oscillations peuvent alors apparaître pour des valeurs de  $T = u + \delta s$  nettement positives, ce qui correspond à des dégradations nettement moins saturées que lorsque  $\theta, \alpha \rightarrow \infty$ . La figure 1 montre ainsi des oscillations observées pour  $\theta \sim 0.25$ , ce qui correspond à des temps d'extinction d'environ 40 minutes pour une demi-vie de l'ARN de 10 minutes. Les pentes des fonctions de dégradation au point fixe sont alors  $u = 0.14$  et  $s = 0.56$ , à comparer à une valeur unité à faible concentration. On voit si la dégradation de la protéine est relativement saturée, celle de l'ARN ne l'est que modérément.



**Fig.1.** Oscillations du modèle (1) pour les valeurs des paramètres suivantes :  $\theta = 0.25$ ,  $\alpha = 8 \times 10^{-4}$ ,  $\delta = 0.76$ ,  $\lambda = 55.7$ ,  $\mu = 0.6$ ,  $n = 2$ . Les variables  $g$ ,  $p$ ,  $m$  sont normalisées par rapport à leur valeur au point fixe. La protéine est supposée être dégradée par une enzyme allostérique avec une cinétique d'ordre 2, tandis que la dégradation de l'ARN suit une cinétique de Michaelis-Menten classique. L'unité de temps est la demi-vie de l'ARN.

On peut légitimement se poser la question de la validité du modèle déterministe (1) si l'activité du gène  $g$  doit être considérée non comme une variable continue mais comme une variable stochastique alternant entre 0 et 1, et si les temps de commutation ne sont pas petits devant les temps d'évolution. A cela on peut répondre que les oscillations du modèle déterministe doivent se refléter de manière mesurable dans les propriétés statistiques du modèle stochastique, et entraîner par exemple une dispersion beaucoup moins importante des temps de commutation. D'autre part, il n'est pas exclu qu'une prise en compte plus fine des mécanismes de transcription montre la nécessité d'introduire certaines variables continues dans les description de ces mécanismes.

### 3 La boucle à deux gènes

Une généralisation naturelle du circuit à un gène auto-régulé est celui formé par une boucle de deux gènes, l'un activant le deuxième, le deuxième réprimant le premier. Nous utilisons dans ce qui suit un modèle semblable à (1), excepté que nous négligeons la dynamique transcriptionnelle. Nous nous intéressons ici à ce système en ce qu'il constitue un modèle minimal de l'horloge circadienne d'*Ostreococcus tauri*,

une algue verte unicellulaire dont la physiologie et l'appareil génétique se caractérisent par une compacité extrême, mais qui présente néanmoins de nombreux points communs avec les végétaux supérieurs. Deux gènes *TOC1* et *CCA1*, homologues de deux gènes centraux de l'horloge d'*Arabidopsis thaliana*, le modèle des végétaux supérieurs, ont pour l'instant été identifiés comme faisant partie de l'horloge circadienne de cette algue, qui est étudiée à l'Observatoire Océanologique de Banyuls.

En supposant des mécanismes de dégradation de type Michaelis-Menten, les équations réduites gouvernant la dynamique de la boucle à deux gènes peuvent s'écrire :

$$\frac{dm_T}{d\tau} = \mu_T + \frac{\lambda_T}{1 + p_C^{n_C}} - \delta \frac{\kappa_{m_T} m_T}{\kappa_{m_T} + m_T} \quad (3a)$$

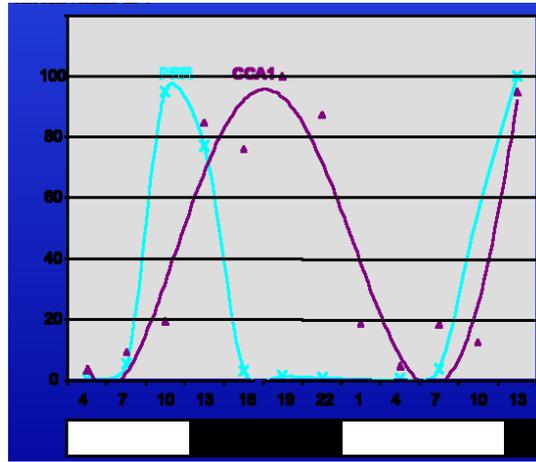
$$\frac{dp_T}{d\tau} = \delta_{p_T} \left( m_T - \frac{\kappa_{p_T} p_T}{\kappa_{p_T} + p_T} \right) \quad (3b)$$

$$\frac{dm_C}{d\tau} = \mu_C + \frac{\lambda_C p_T^{n_T}}{1 + p_T^{n_T}} - \frac{\kappa_{m_C} m_C}{\kappa_{m_C} + m_C} \quad (3c)$$

$$\frac{dp_C}{d\tau} = \delta_{p_C} \left( m_C - \frac{\kappa_{p_C} p_C}{\kappa_{p_C} + p_C} \right). \quad (3d)$$

où  $m_T$  et  $p_T$  ( $m_C$  et  $p_C$ ) représentent les quantités d'ARN et de protéine du gène *TOC1* (*CCA1*). Les paramètres  $n_{T,C}$ ,  $\lambda_{T,C}$ ,  $\mu_{T,C}$  et  $\delta$ ,  $\delta_{p_T,p_C}$  ont la même signification que dans (1). Les coefficients  $\kappa_i$  caractérisent la saturabilité des dégradations enzymatiques des différentes molécules en présence.

De même que pour le modèle (1), l'apparition d'oscillations dans le modèle (3) dépend de manière cruciale des mécanismes de dégradation. Plus précisément, il faut qu'au moins un certain nombre des quatre molécules impliquées dans la boucle soient dégradées de manière enzymatique, et que cette dégradation soit suffisamment saturée (à un moindre degré cependant que pour le circuit à un gène). Il est intéressant de noter au passage que le système (3) peut se ramener dans une certaine limite à la variante du célèbre oscillateur de Goodwin [3] donnée par Bliss *et al.* [17].

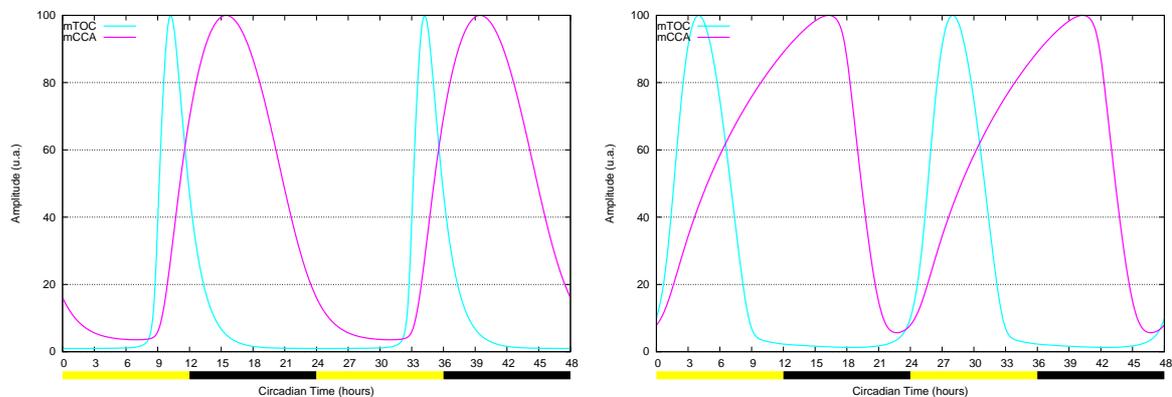


**Fig.2.** Niveaux d'expression en alternance jour/nuit des gènes *TOC1* (alias *PRR*) et *CCA1* d'O. tauri en fonction du temps circadien (CT : "circadian time"), CT0 correspondant au début du jour. Données expérimentales du groupe Horloge circadienne et cycle cellulaire de l'observatoire océanologique de Banyuls/mer. Malgré les incertitudes de mesure, on peut caractériser les deux courbes par des grandeurs relativement reproductibles. Ainsi, la quantité d'ARN de *TOC1* est maximale vers CT10.5, avec une largeur à mi-hauteur d'environ 6 heures, et un long passage à zéro entre CT17 et CT7. En ce qui concerne *CCA1*, la présence de l'ARN est beaucoup plus étalée dans le temps, avec un pic vers CT17, une largeur à mi-hauteur d'environ 12 heures et un point bas aux alentours de CT7.

Notre but est de comparer les prédictions du modèle (3) aux données expérimentales concernant les variations dans le temps des ARN et des protéines de l'horloge. Cette comparaison est d'autant plus intéressante que la boucle TOC1/CCA1 a été un temps évoquée comme modèle pour l'horloge d'*Arabidopsis* [16,18] avant d'être délaissée au profit de circuits plus sophistiqués à plusieurs boucles de rétroaction [19]. Or, comme on le voit sur la figure 1, qui montre les variations dans le temps des niveaux d'ARN des gènes TOC1 et CCA1 en alternance jour/nuit, l'horloge d'*Ostreococcus* présente une différence importante avec celle d'*Arabidopsis* : CCA1 est à son maximum d'expression en début de nuit plutôt qu'au petit matin. Etant donné qu'*Ostreococcus* se caractérise généralement par une relative simplicité, il était donc important de déterminer si la boucle à deux gènes pourrait être un meilleur modèle pour cette algue que pour *Arabidopsis*. Dans un premier temps, nous nous sommes attachés à reproduire les régimes en alternance jour/nuit, généralement plus reproductibles que les régimes en éclairage constant.

Le modèle (3) décrit la régulation réciproque des gènes TOC1 et CCA1, mais ne précise pas le mécanisme d'action de la lumière sur la boucle. En l'absence d'informations précises, il nous faut donc envisager plusieurs scénarios différents, associés à des modulations différentes des paramètres. L'horloge pourrait être ainsi entraînée et synchronisée au cycle jour/nuit par une dégradation accélérée d'une protéine ou d'une autre, et ce le jour ou plutôt la nuit, ou encore par une réduction de l'activité transcriptionnelle d'une des deux protéines dans l'une des deux périodes. On peut évidemment espérer que les tests de ces différents mécanismes nous fournissent des pistes sur le couplage effectivement présent.

La figure 2 montre ainsi deux simulations préliminaires du modèle (3). Ces profils temporels ont été obtenus en cherchant des jeux de paramètres pour lesquels ils se rapprochaient le plus des données expérimentales (fig. 2). On constate sur la partie gauche de la figure que l'hypothèse d'une dégradation accélérée de la protéine TOC1 la nuit permet au modèle à deux gènes d'ajuster relativement bien les données expérimentales : les caractéristiques des profils expérimentaux et théoriques coïncident avec une très bonne précision, si ce n'est un pic de CCA1 un peu en avance. On note toutefois sur la figure 2 qu'il n'est pas exclu que ce pic arrive en fait plus tôt que ne l'indique la ligne tracée pour guider l'oeil.



**Fig.3.** Simulations numériques du modèle (3) avec deux hypothèses différentes de coopérativité et de couplage de la lumière externe à la boucle génétique. Dans les deux cas, on teste un grand nombre de jeux de paramètres différents, et celui pour lequel les solutions s'approchent le plus de des courbes expérimentales de la figure 2 est retenue. (gauche) Dégradation accélérée de la protéine TOC1 la nuit, régulation par un monomère de TOC1 et un dimère de CCA1 ; (droite) Dégradation accélérée de la protéine CCA1 la nuit, régulation par des monomères de TOC1 et de CCA1. On constate que l'hypothèse de gauche est nettement plus vraisemblable que celle de droite.

Evidemment, des comparaisons plus précises impliquant également les profils temporels des protéines ainsi que les données en éclairage constant seront nécessaires avant de se prononcer définitivement sur la pertinence du système (3) en tant que modèle de l'horloge circadienne d'*Ostreococcus*. Les résultats préliminaires présentés ici sont cependant étonnamment encourageants.

## 4 Conclusion

Nous avons observé que la prise en compte d'une dynamique transcriptionnelle élargit les zones de paramètres où des mécanismes de dégradation non linéaires peuvent induire des oscillations dans l'expression d'un gène réprimé par sa propre protéine. Ces mécanismes de dégradation sont également importants pour comprendre l'apparition d'oscillations dans la boucle à deux gènes, qui est par ailleurs un modèle hypothétique de l'horloge circadienne de l'algue unicellulaire *Ostreococcus tauri*. Des calculs préliminaires montrent qu'à condition de supposer certains modes d'action de la lumière sur les acteurs moléculaires, ce système semble bien reproduire les observations expérimentales.

## Références

1. H. HIRATA *et al.*, Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop, *Science* **298**, 840–843 (2002).
2. C. A. STRAYER & S. A. KAY, The ins and outs of circadian regulated gene expression, *Curr. Opin. Plant Biol.* **2**, 114 (1999).
3. B. C. GOODWIN, Oscillatory behavior of enzymatic control processes, *Adv. Enzyme Regul.* **3**, 425–439 (1965).
4. J. S. GRIFFITH, Mathematics of cellular control processes I. Negative feedback to one gene, *J. Theor. Biol.* **20**, 202 (1968).
5. A. GOLDBETER, A model for circadian oscillations in the Drosophila period protein (PER), *Proc. R. Soc. Lond. B* **261**, 319 (1995).
6. J.-C. LELOUP, D. GONZE, AND A. GOLDBETER, Limit cycle models for circadian rhythms based on transcriptional regulation in Drosophila and Neurospora, *J. Biol. Rhythms* **14**, 433 (1999).
7. M. H. JENSEN, K. SNEPPEN & G. TIANA, Sustained oscillations and time delays in gene expression of protein Hes1, *FEBS Lett.* **541**, 176–177 (2003).
8. N. A. M. MONK, Oscillatory expression of Hes1, p53 and NK- $\kappa$ B driven by transcriptional time delays, *Curr. Biol.* **13**, 1409 (2003).
9. J. LEWIS, Autoinhibition with transcriptional delay : a simple mechanism for the zebrafish somitogenesis oscillator, *Curr. Biol.* **13**, 1398 (2003).
10. I. GOLDING, J. PAULSSON, S. M. ZAWILSKI, AND E. C. COX, Real-time kinetics of gene activity in individual bacteria, *Cell* **123**, 1025 (2005).
11. P. FRANÇOIS & V. HAKIM, Core genetic module : the mixed feedback loop, *Phys. Rev. E* **72**, 031908 (2005).
12. J. J. TYSON, C. I. HONG, D. THRON AND B. NOVAK, A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM, *Biophys. J.* **77**, 2411 (1999).
13. N. E. BUCHLER, U. GERLAND, AND T. HWA, Nonlinear protein degradation and the function of genetic circuits, *Proc. Natl. Acad. Sci. USA* **102**, 9559 (2005).
14. A. GOLDBETER, *Biochemical Oscillations and Cellular Rhythms : The molecular bases of periodic and chaotic behaviour* (Cambridge University Press, Cambridge, 1996).
15. C. P. FALL, E. S. MARLAND, J. M. WAGNER, AND J. J. TYSON, *Computational Cell Biology* (Springer, New York, 2002).
16. D. ALABADI, T. OYAMA, M. J. YANOVSKY, F. G. HARMON, P. MAS, S. A. KAY, Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock, *Science* **293**, 880 (2001).
17. R. D. BLISS, P. R. PAINTER, AND A. G. MARR, Role of feedback inhibition in stabilizing the classical operon, *J. Theor. Biol.* **97**, 177 (1982).
18. J. C. W LOCKE, A. J. MILLAR, AND M. S. TURNER, Modelling genetic networks with noisy and varied experimental data : the circadian clock in *Arabidopsis thaliana*, *J. Theor. Biol.* **234**, 383 (2005).
19. J. C. W LOCKE, M. M. SOUTHERN, L. KOZMA-BOGNAR, V. HIBBERD, P. E. BROWN, M. S. TURNER, AND A. J. MILLAR, Extension of a genetic network model by iterative experimentation and analysis, *Mol. Systems Biol.*, doi :10.138/msb4100018.