

Oscillations induites par une dynamique transcriptionnelle lente dans l'expression d'un gène réprimé par sa propre protéine

P.-E. Morant¹, Q. Thommen¹, F. Lemaire², C. Vandermoere¹, B. Parent³, & M. Lefranc¹

¹ Laboratoire de Physique des Lasers, Atomes, Molécules, UMR USTL/CNRS 8523, UFR de Physique, Université de Lille 1, F-59655 Villeneuve d'Ascq

² Laboratoire d'Informatique Fondamentale de Lille, UMR USTL/CNRS 8022, Université de Lille 1, F-59655 Villeneuve d'Ascq

³ Unité de Glycobiologie Structurale et Fonctionnelle, UMR USTL/CNRS 8576, Université de Lille 1, F-59655 Villeneuve d'Ascq

marc.lefranc@univ-lille1.fr

Résumé. Nous étudions la dynamique d'un gène réprimé par sa propre protéine dans le cas où le taux de transcription est une variable dynamique qui ne s'adapte pas instantanément à la concentration de protéines. Nous obtenons un critère analytique pour l'apparition d'oscillations spontanées qui montre que celles-ci peuvent alors se produire pour des mécanismes de dégradation bien moins non linéaires. Les prédictions déterministes sont aussi comparées à des simulations stochastiques.

Abstract. We revisit the dynamics of a gene repressed by its own protein in the case where the transcription rate does not adapt instantaneously to protein concentration but is a dynamical variable. We derive analytical criteria for the appearance of sustained oscillations and find that they require degradation mechanisms much less nonlinear than for infinitely fast regulation. Deterministic predictions are also compared with stochastic simulations of this minimal genetic oscillator.

1 Introduction

Les réseaux de gènes interagissant via des protéines qui régulent leurs activités sont des systèmes fortement non linéaires présentant une grande variété de comportements dynamiques [5,2]. Notre étude porte sur le circuit génétique élémentaire représenté schématiquement à la figure 1. Des molécules d'ARN

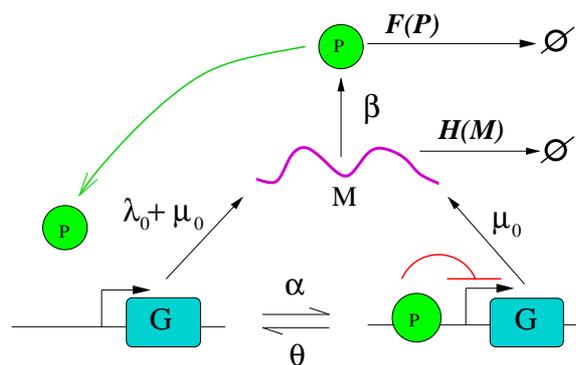


Fig.1. Schéma réactionnel du gène auto-régulé.

sont produites à partir d'un gène (*transcription*). Des protéines sont ensuite synthétisées à partir de ces ARN (*traduction*). Une des protéines produites peut se fixer sur le gène, entraînant une modification du taux de synthèse d'ARN (*taux de transcription*). Protéines et ARN peuvent être dégradés par des

processus externes. Enfin, le gène est présent en exemplaire unique et ne peut être dégradé. Trois variables dynamiques, G , M et P décrivent à chaque instant l'état du réseau. Les variables M et P représentent respectivement le nombre d'ARN et de protéines. Quant à la variable G , elle représente l'activité du gène.

Ce réseau simple décrit une boucle de rétroaction du gène sur lui-même. Comme nous nous intéressons aux comportements oscillatoires, nous nous limitons au cas d'une rétroaction négative, où la protéine réprime sa propre expression : le taux de production d'ARN est moindre lorsqu'une protéine est fixée à la zone régulatrice du gène.

Les études théoriques de réseaux génétiques ont généralement supposé que les processus d'activation d'un gène via son interaction avec la protéine régulatrice sont rapides par rapport à la dynamique des ARN et des protéines, et donc que le taux de transcription réagit instantanément à la concentration en protéines. Cependant, cette synthèse est un processus complexe [1], et cette hypothèse n'est pas toujours justifiée. Dans cet article, nous montrons comment la dynamique intrinsèque du gène peut modifier de manière spectaculaire le comportement du système dans le cas du gène auto-réprimé.

2 Dynamique déterministe

Nous modélisons le circuit de la figure 1 par le système d'équations différentielles (1a–1c)

$$\dot{G} = \theta_0(1 - G) - \alpha_0 GP \quad (1a)$$

$$\dot{P} = [\theta_0(1 - G) - \alpha_0 GP] + \beta_0 M - \delta_P F(P) \quad (1b)$$

$$\dot{M} = \mu_0 + \lambda_0 G - \delta_M H(M) \quad (1c)$$

L'équation (1a) décrit a priori la dynamique de l'interaction de la protéine avec le gène, les paramètres α_0 et θ_0 étant les taux d'accrochage et de relargage de la protéine. La variable G varie continûment entre 0 et 1 et peut être interprétée comme une moyenne temporelle de l'activité du gène lorsque le temps moyen de résidence de la protéine sur le gène est petit par rapport à la période d'oscillation [3]. De manière plus générale, l'équation (1a) décrit la dynamique d'une activité génique G relaxant lentement vers une valeur d'équilibre donnée par la fonction de régulation habituelle $G = 1/[1 + (P/P_0)]$, où $P_0 = (\theta_0/\alpha_0)$ est le seuil de demi-expression ($G = 1/2$ pour $P = P_0$). Une telle dynamique peut résulter de phénomènes coopératifs et de recrutement dans la machinerie de transcription. Les études précédentes ont généralement considéré θ_0 et α_0 comme très grands devant les autres paramètres.

Dans l'équation (1b), les deux premiers termes correspondent à la dynamique de fixation/relargage de la protéine; les troisième et quatrième décrivent la traduction et la dégradation. L'équation (1c) décrit la transcription avec un taux $\mu_0 + \lambda_0 G$ et la dégradation de l'ARN. Le critère d'oscillation est obtenu ci-dessous pour des fonctions de dégradation arbitraires $F(P)$ and $H(M)$, que nous choisissons de dérivée unitaire en zéro de sorte que δ_P et δ_M représentent les taux de dégradation à faible nombre de molécules.

En renormalisant le temps et les variables dynamiques comme suit :

$$t = \frac{t'}{\delta_M}, \quad G = g, \quad P = pP_0, \quad M = mM_0, \quad P_0 = \left(\frac{\theta_0}{\alpha_0}\right), \quad M_0 = \frac{\delta_P P_0}{\beta_0} \quad (2)$$

les équations (1a–1c) peuvent être réécrites sous la forme plus simple (on note $x' = dx/dt'$)

$$g' = \theta [1 - g(1 + p)], \quad (3a)$$

$$p' = \alpha [1 - g(1 + p)] + \delta [m - f(p)], \quad (3b)$$

$$m' = \mu + \lambda g - h(m). \quad (3c)$$

Les paramètres renormalisés et les fonctions de dégradation sont donnés par

$$\theta = \frac{\theta_0}{\delta_M}, \quad \alpha = \frac{\theta_0}{P_0 \delta_M}, \quad \delta = \frac{\delta_P}{\delta_M}, \quad \lambda = \frac{\lambda_0}{M_0 \delta_M}, \quad \mu = \frac{\mu_0}{M_0 \delta_M}, \quad f(p) = \frac{F(P)}{P_0}, \quad h(m) = \frac{H(M)}{M_0}. \quad (4)$$

Lorsque les fonctions f et g sont monotones, le modèle (3a–3c) a un unique point fixe (g_*, p_*, m_*) caractérisé par les équations :

$$g_* = \frac{1}{1 + p_*}, m_* = f(p_*), g_* = \frac{h(m_*) - \mu}{\lambda} \quad (5)$$

Remarquons que le point fixe ne dépend que des paramètres λ et μ et des fonctions de dégradations f et h , tandis que les paramètres θ , α , δ contrôlent les échelles de temps.

Le comportement des fonctions de dégradation au voisinage du point fixe est décrit par les pentes

$$s = \left. \frac{df(p)}{dp} \right|_{p=p_*} \quad \text{et} \quad u = \left. \frac{dh(m)}{dm} \right|_{m=m_*}$$

Dans le cas d'une dégradation linéaire [$f(p) = p$, $h(m) = m$], on trouve $u = s = 1$. Des valeurs petites (ou même négatives) des pentes s et u dénotent généralement des mécanismes de dégradation fortement non linéaires [2,11], par exemple des saturations.

Pour déterminer si les équations (3a)–(3c) possèdent un régime d'oscillation spontanée, nous avons cherché les valeurs des paramètres pour lesquelles le point fixe perd sa stabilité via une bifurcation de Hopf. Dans un souci de simplicité, nous nous plaçons dans le cas d'une répression parfaite ($\mu = 0$) et d'un seuil de commutation du gène P_0 élevé ($\alpha \sim 0$). Dans ces conditions, le critère de Routh-Hürwitz nous indique que le point fixe devient instable lorsque la quantité

$$\mathcal{H} = \sigma + g_* (-\delta \lambda g_*^2 + \sigma^2) \tau + \pi g_*^2 \sigma \tau^2 \quad (6)$$

devient négative. Dans l'équation (6), la variable $\tau = 1/\theta$ est le temps de réponse du gène (à strictement parler, le temps moyen de résidence de la protéine sur le gène) ; la somme $\sigma = \delta s + u$ et le produit $\pi = \delta s u$ sont des fonctions symétriques des taux de dégradation δs et u .

Le critère (6) permet de retrouver nombre de résultats connus relatifs à la dynamique du gène auto-réprimé. Ainsi, on remarque par exemple qu'une rétroaction négative importante (λ élevé) déstabilise le système alors que des taux de dégradation importants (σ et π élevés) tendent à le stabiliser. Une approximation couramment utilisée est de considérer que le gène réagit instantanément à la concentration de protéine. Dans ce cas, $\tau = 0$ et $\mathcal{H} = \sigma$: l'existence d'oscillations implique alors $\sigma < 0$. Il est effectivement bien connu que des taux de dégradation négatifs peuvent induire des oscillations [2,11]. On aussi peut montrer que \mathcal{H} est toujours positif quand les protéines et les ARN ont des dégradations linéaires : il ne peut apparaître d'oscillations dans ce cas. De même, quand les dégradations des protéines et des ARN sont complètement saturées, ($s = u = \sigma = 0$), $\mathcal{H} = -\tau \delta \lambda g_*^3$ est *constamment négatif* : des oscillations apparaîtront systématiquement. Le comportement intermédiaire entre les régimes de dégradation linéaires et saturés dépend du temps de réponse τ .

Par la suite, nous considérons que $u, s > 0$ et démontrons que des oscillations peuvent apparaître pour des valeurs de τ finies. Plus précisément, nous souhaitons comprendre comment les oscillations peuvent apparaître en l'absence d'une dégradation saturée. Dans ce but, nous utilisons un premier indice $\nu = \sqrt{us} = \sqrt{\pi/\delta}$ correspondant à la moyenne géométrique des pentes ($\nu = 1$ pour des dégradations linéaires et $\nu = 0$ pour une dégradation saturée). Nous nous proposons de déterminer la valeur de τ donnant des oscillations pour des valeurs de ν les plus grandes possibles.

Remarquons d'abord que dans (6), les quantités σ et π jouent des rôles complémentaires. Le taux global de dégradation σ contrôle l'instabilité pour les valeurs de τ faibles et modérés, alors qu'un produit π élevé bloque les oscillations pour les grandes valeurs de τ , lorsque la dynamique du gène est lente. Pour une valeur donnée du taux de dégradation global σ , π varie entre $\sigma^2/4$ et 0 selon que les taux de dégradation des protéines et des ARN sont égaux ou déséquilibrés. Nous introduisons un deuxième indice $\epsilon = 2\sqrt{\pi}/\sigma \in [0, 1]$ caractérisant l'équilibre des taux de dégradations ($\epsilon = 1$ indique le cas symétrique et $\epsilon = 0$ le cas déséquilibré).

Il apparaît clairement dans (6) que

$$\sigma < \sigma_c = g_* \sqrt{\delta \lambda} = \frac{g_* \sqrt{2}}{t_{sw}} \quad (7)$$

est une condition nécessaire d'oscillation, où la quantité t_{sw} est le temps durant lequel un gène non réprimé ($G = 1$) synthétise à partir de zéro la quantité de protéines correspondant au seuil de demi-répression. En particulier, il suffit lorsque l'une des dégradations est complètement saturée ($\pi = 0$) que la condition (7) soit vérifiée pour que les oscillations apparaissent systématiquement pour des valeurs de τ suffisamment grandes. Cette condition suggère la mise à l'échelle suivante

$$\sigma = \sigma_c \Sigma, \quad \pi = \left(\frac{\epsilon \sigma_c}{2} \right)^2 \Sigma^2, \quad \tau = \frac{T}{g_* \sigma_c} \quad (8)$$

qui permet de réécrire la condition d'oscillation sans dépendance explicite par rapport aux paramètres

$$\mathcal{H}_\epsilon(\Sigma, T) = \Sigma \times \left[\frac{\epsilon^2 \Sigma^2}{4} T^2 + \left(\Sigma - \frac{1}{\Sigma} \right) T + 1 \right] < 0. \quad (9)$$

La condition (9) définit une série de courbes $\Sigma_\epsilon(T)$ spécifiant le taux de dégradation Σ au seuil d'oscillation comme une fonction du temps de réponse T du gène et de l'indice d'équilibre ϵ . Pour un indice ϵ donné, les oscillations apparaissent pour $\Sigma \leq \Sigma_\epsilon(T)$. Parallèlement, lorsque l'on maintient T fixe $\Sigma_\epsilon(T)$ décroît de manière monotone avec ϵ . La figure 2 représente les courbes limites $\Sigma_1(T)$ et $\Sigma_0(T)$ qui sont particulièrement importantes pour comprendre le diagramme de bifurcation : quel que soit l'indice d'équilibre ϵ , le circuit oscille toujours quand $\Sigma < \Sigma_1(T)$ et jamais quand $\Sigma > \Sigma_0(T)$.

Pour vérifier la pertinence de notre analyse, nous avons exploré numériquement l'espace des paramètres des équations (3a–3c) à la recherche d'oscillations. Pour cela, nous avons considéré le cas d'une dégradation de la protéine via une enzyme allostérique et d'une dégradation de Michaelis-Menten pour l'ARN :

$$f(p) = \frac{p \times (a + p/\kappa)}{a + 2a(p/\kappa) + (p/\kappa)^2}, \quad h(m) = \frac{\chi m}{\chi + m}. \quad (10)$$

Les points obtenus sont représentés par des points noirs sur la figure 2. Bien que les simulations numériques aient été menées avec $\mu, \alpha \neq 0$, l'accord est excellent : tous les points sont en dessous de la courbe $\Sigma_0(T)$ et les quelques points significativement au-dessus de $\Sigma_1(T)$ ont généralement un des taux de dégradation petit. Notre analyse semble donc pouvoir établir des résultats généraux quant aux comportements dynamiques des équations (1a–1c).

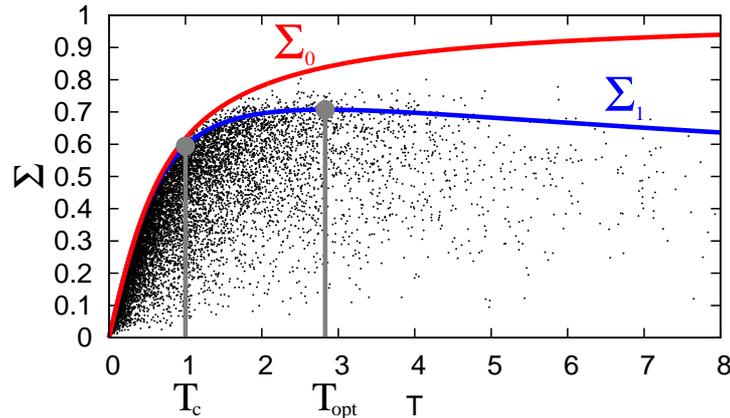


Fig.2. Diagramme de bifurcation de (1a–1c) dans le plan (Σ, T) obtenu d'après (9). Un système avec un indice d'équilibre ϵ oscille pour $\Sigma < \Sigma_\epsilon(T)$. Les points noirs indiquent les paramètres d'oscillation de (1a–1c), avec $\theta, \delta \in [10^{-1}, 10]$, $\theta/\alpha \in [10, 1000]$, en supposant les mécanismes de dégradation (10) avec $a \in [10^{-6}, 1]$. $T_c = 1$ est l'échelle de temps en deçà de laquelle la dynamique transcriptionnelle ne peut être négligée. $T_{\text{opt}} = 2\sqrt{2}$ est la position du maximum de $\Sigma_1(T)$.

Sur la figure 2, deux régions peuvent être distinguées. Pour $T < 1$, le seuil d'instabilité $\Sigma_\epsilon(T)$ augmente rapidement avec T et sa dépendance en ϵ est négligeable. Pour les petits T , nous trouvons que $\Sigma_\epsilon(T) \sim T$, soit en retranscrivant à l'aide de (8) :

$$\sigma < \lambda g_*^3 \delta \tau \quad (11)$$

Dans la région où $T > 1$, $\Sigma_\epsilon(T)$ atteint une valeur maximale $\Sigma_m(\epsilon) = 1/\sqrt{\epsilon+1}$ pour $T = T_m(\epsilon) = 2\sqrt{\epsilon+1}/\epsilon$, et décroît comme T^{-1} pour $T \rightarrow \infty$, si $\epsilon \neq 0$. $\Sigma_0(T)$ croît de manière monotone jusque $\Sigma = 1$. Pour une valeur de Σ fixée, les oscillations apparaissent dans un domaine de T restreint ; ce domaine devient infini lorsque $\epsilon = 0$

Notons que des valeurs plus importantes de Σ au seuil d'oscillation peuvent être obtenues pour des valeurs de ϵ plus petites (Fig. 2). Cependant, celles-ci correspondent aussi à des plus petites valeurs de l'indice $\nu \sim \epsilon \Sigma$, c'est-à-dire à des configurations plus proches d'une dégradation saturée ($s = 0$ ou $u = 0$). Il est facile de montrer qu'au seuil d'oscillation, la valeur maximale de l'indice ν est atteinte pour $\epsilon = 1$, au maximum de la courbe $\Sigma_1(T)$ situé en $T_{\text{opt}} = 2\sqrt{2}$ avec

$$\nu_{\text{opt}} = \frac{g_* \sqrt{\lambda}}{2\sqrt{2}} \quad (12)$$

Ainsi, notre analyse suggère l'existence d'un phénomène de résonance dans la dynamique d'un gène auto-régulé tenant compte de la dynamique de la transcription : le circuit bifurque plus facilement vers un comportement périodique, ou plus généralement devient moins stable, pour une valeur finie du temps de relaxation du gène donné par

$$\tau_{\text{opt}} = 2\sqrt{2} \tau_c, \quad \tau_c = \frac{1}{g_*^2 \sqrt{\delta \lambda}} = \delta_M \times \frac{1}{g_*^2} \sqrt{\frac{P_0}{\lambda_0 \beta_0}}. \quad (13)$$

Le comportement dynamique du circuit diffère significativement du cas de la régulation instantanée du gène pour $\tau > \tau_c$ (notons que τ_c est proportionnel à t_{sw} défini par (7)). Cela suggère que la dynamique de transcription d'un gène peut être négligée seulement quand τ_c est démesurément grand. Comme g_* dépend de λ , μ , et des fonctions de dégradation à travers (5), la détermination de la borne inférieure de τ_c nécessite la connaissance des fonctions de dégradation. Néanmoins, il est remarquable que pour des valeurs réalistes des paramètres $\lambda_0 = \beta_0 = 10 \text{ min}^{-1}$, $P_0 = 100$, et $g_*^2 = 0,5$, on obtienne $t_c = \tau_c/\delta_M = 2$ minutes, ce qui correspond à une valeur très réaliste.

3 Dynamique stochastique

Nous allons voir maintenant que même lorsque G est une variable stochastique binaire (le gène est soit inactif soit actif), notre résultat principal reste valable : il existe une échelle de temps voisine de τ_{opt} favorisant les oscillations. Pour cela, nous avons effectué des simulations stochastiques du circuit génétique de la figure 1 pour différents temps de résidence τ en maintenant P_0 fixé. Dès lors, les oscillations régulières semblables à celle issue du modèle déterministe ne peuvent bien entendu plus être observées. L'évolution du nombre de protéines en fonction du temps présente des séquences de pics irrégulièrement espacés. Il est naturel de se demander si ces pics apparaissent plus régulièrement pour les valeurs des paramètres donnant une oscillation dans le modèle déterministe et en particulier quand $\tau = \tau_{\text{opt}}$.

Les intervalles de temps Δt entre les pics sont définis entre deux passages par $P = 1, 2P_{\text{avg}}$ séparés par au moins un passage par $P = 0, 8P_{\text{avg}}$ (c'est-à-dire en imposant une amplitude minimum de 40%), P_{avg} étant le niveau moyen de protéine. La distribution des intervalles de temps est alors caractérisée par son coefficient de variation $\eta = \frac{\sigma_{\Delta t}}{\langle \Delta t \rangle}$. Sans boucle de rétroaction, Δt est la somme de deux variables suivant chacune une distribution de Poisson et donc $\eta \in [1/\sqrt{2}, 1]$. La figure 3 présente une variation typique de η avec τ dans notre système. Cela suggère clairement que notre analyse déterministe reste pertinente dans le régime stochastique. En effet, il y a clairement une échelle de temps minimisant la fluctuation de l'intervalle temporel entre les pics, et celle-ci est voisine de τ_{opt} . De plus, il y a un domaine étendu autour du domaine d'oscillation du modèle déterministe pour lequel $\eta < 1/\sqrt{2}$, c'est-à-dire où l'intervalle de temps entre les pics est plus régulier que pour une cascade de deux processus de Poisson.

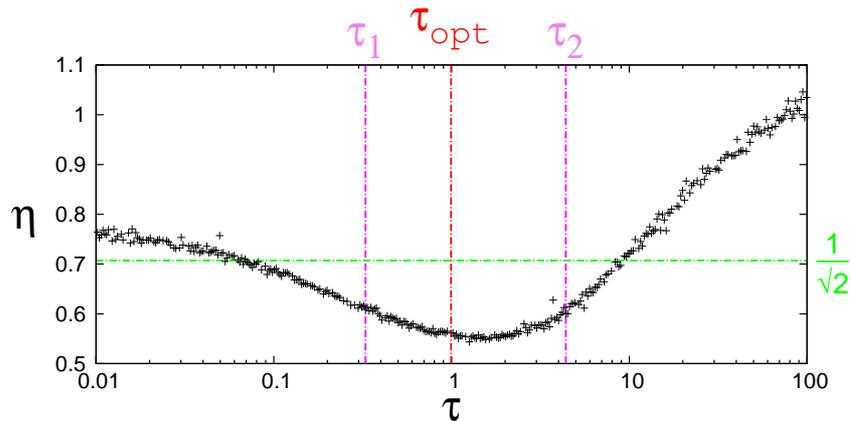


Fig.3. Coefficient de variation de l'intervalle de temps séparant deux pics en fonction du temps moyen de résidence τ pour une simulation stochastique du gène auto-réprimé.

4 Conclusion

En conclusion, nous avons montré qu'une dynamique de transcription non triviale peut déstabiliser un gène auto-réprimé. Bien qu'il soit déjà connu qu'un mécanisme de dégradation non linéaire puisse induire des oscillations dans ce système, nous observons un phénomène de résonance telle qu'une moindre non linéarité des mécanismes de dégradation est nécessaire pour induire des oscillations lorsque le temps de réponse du gène avoisine un certain temps caractéristique. L'expression de ce temps peut être déterminée analytiquement, ce qui nous permet d'identifier la région de l'espace des paramètres où la dynamique de la transcription ne peut être négligée. Les stimulations stochastiques effectuées confirment l'existence de ce temps caractéristique. La dynamique de la transcription doit ainsi être considérée comme une source potentielle d'oscillations à l'instar d'autres effets déterministes [5,2,4,7,6,10,8,11] ou stochastiques [9].

Références

1. X. DARZACQ, Y. SHAV-TAL, V. DE TURRIS, Y. BRODY, S. M SHENOY, R. D. PHAIR, & R. H. SINGER, In vivo dynamics of RNA polymerase II transcription, *Nat. Struct. Mol. Biol.* **14**, 796–806 (2007).
2. C. P. FALL, E. S. MARLAND, J. M. WAGNER, & J. J. TYSON, *Computational Cell Biology* (Springer, New York, 2002).
3. D. B. FORGER & C. S. PESKIN, Stochastic simulation of the mammalian circadian clock, *Proc. Nat. Acad. Sci.* **102**, 321–324 (2005).
4. A. GOLDBETER, A model for circadian oscillations in the *Drosophila* period protein (Per), *Proc. R. Soc. Lond. B* **261**, 319–324 (1995).
5. A. GOLDBETER. *Biochemical Oscillations and Cellular Rhythms : The molecular bases of periodic and chaotic behaviour* (Cambridge University Press, Cambridge, 1996).
6. M. H. JENSEN, K. SNEPPEN, AND G. TIANA. Sustained oscillations and time delays in gene expression of protein hes1, *FEBS Lett.* **541**, 176–177 (2003).
7. 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**, 433–448 (1999).
8. J. LEWIS, Autoinhibition with transcriptional delay : a simple mechanism for the zebrafish somitogenesis oscillator, *Curr. Biol.* **13**, 1398–1408 (2003).
9. A. LOINGER & O. BIHAM, Stochastic simulations of the repressilator circuit, *Phys. Rev. E* **76**, 051917 (2007).
10. N. A. M. MONK, Oscillatory expression of Hes1, p53 and NF- κ B driven by transcriptional time delays, *Curr. Biol.* **13**, 1409–1413 (2003).
11. 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**, 2411–2417 (1999).