Bifurcations in the interplay of messenger RNA, protein and nonprotein coding RNA

2008 J. Phys. A: Math. Theor. 41 285101

(http://iopscience.iop.org/1751-8121/41/28/285101)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 285101 (8pp) doi:10.1088/1751-8113/41/28/285101

Bifurcations in the interplay of messenger RNA, protein and nonprotein coding RNA

Vladimir P Zhdanov

Department of Applied Physics, Chalmers University of Technology, S-412 96 Göteborg, Sweden
and
Boreskov Institute of Catalysis, Russian Academy of Sciences, Novosibirsk 630090, Russia

E-mail: zhdanov@catalysis.ru

Received 23 March 2008, in final form 22 May 2008
Published 17 June 2008
Online at stacks.iop.org/JPhysA/41/285101

Abstract
The interplay of messenger RNA (mRNA), protein, produced via translation of this RNA, and nonprotein coding RNA (ncRNA) may include regulation of the ncRNA production by protein and (i) ncRNA–protein association resulting in suppression of the protein regulatory activity or (ii) ncRNA–mRNA association resulting in degradation of the miRNA–mRNA complex. The kinetic models describing these two scenarios are found to predict bistability provided that protein suppresses the ncRNA formation.

PACS numbers: 87.16.−b, 05.10.−a

Introduction

The conventional scheme of gene expression includes mRNA formation via gene transcription, performed by RNA polymerase (RNAP), and protein synthesis via mRNA translation by ribosomes [1]. The gene transcription is often controlled by master regulatory proteins. Due to this feedback between mRNA formation and protein synthesis, the kinetics of mRNA and protein formation may be complex even in the simplest genetic networks. The corresponding models focused on bistability and oscillations in simple genetic networks and on complex networks are numerous (see recent reviews [2–5], articles [6–17], and references therein).

The conventional view outlined above is fully applicable to prokaryotes whose genomes consist of tightly packed protein-coding sequences. In eukaryotic cells, the protein-coding genes constitute however only a part of the genome [18, 19]. The rest of the genome is nevertheless transcribed as well. In particular, the transcription of many genes results in the formation of nonprotein coding RNA (ncRNA). Only recently, it has become clear that such RNAs form the cornerstone of a regulatory network of signalling that operates in concert with the protein network [20–28]. The numerous biological functions of ncRNAs are based on their

abilities (i) to bind to and modulate the activity of proteins, or (ii) to pair with target mRNAs [22] (in addition, ncRNA may catalyse biochemical reactions).

In analogy with mRNA, the ncRNA formation can be controlled by transcription factors (proteins) involved in the regulation of conventional genes (see e.g. reports [29–31] and a recent review by Kulshreshtha *et al* [26]).

The first kinetic models taking into account ncRNAs have recently been proposed in [32–36]. In the treatments presented, there is no regulation of the ncRNA production by protein (P). To clarify what may happen with this regulation, we outline in this communication two mean-field (MF) kinetic models describing the interplay of mRNA, P and ncRNA. Specifically, we analyse scenarios (i) and (ii) outlined above.

Scenario (i). In our first model, the reaction scheme consists of mRNA, ncRNA and P formation and degradation,

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}, \tag{1}$$

$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{ncRNA}, \tag{2}$$

$$\mathrm{mRNA} \to \mathrm{mRNA} + P, \tag{3}$$

$$\mathrm{mRNA} \to \emptyset, \tag{4}$$

$$\mathrm{ncRNA} \to \emptyset, \tag{5}$$

$$P \to \emptyset, \tag{6}$$

and ncRNA and P association and degradation,

$$\mathrm{ncRNA} + P \to \mathrm{ncRNA} * P \to \emptyset. \tag{7}$$

In this scheme, steps (1)–(6) are conventional. Step (7) is typical for ncRNA. Four real examples of the latter step are described in detail by Goodrich and Kugel [22] (in their review and in the review by Yazgan and Krebs [28], one can also find references to other examples).

P produced via mRNA translation is considered to regulate the ncRNA synthesis. This regulation may be positive or negative. Our attention is focused on a negative feedback, because this case is found to be more intriguing. Specifically, the ncRNA formation is assumed to run provided that all the regulatory sites are free of P. The P association with and dissociation from the gene are considered to be rapid so that these steps are at equilibrium. These approximations are used in conventional kinetic models focused on the interplay of proteins and mRNAs [5]. Alternatively, one can employ the standard Hill expression in order to describe the effect of P on the gene-expression rate (it does not change our conclusions). In addition, as usual in models of genetic networks, we assume that the association of RNAP with DNA does not limit gene transcription.

With the specification above, the MF kinetic equations for the mRNA, ncRNA and P numbers in a cell are as follows:

$$\frac{\mathrm{d}N_m}{\mathrm{d}t} = \kappa_m - k_m N_m, \tag{8}$$

$$\frac{\mathrm{d}N_{nc}}{\mathrm{d}t} = \kappa_{nc}\left(\frac{K_P}{K_P + N_P}\right)^n - rN_P N_{nc} - k_{nc}N_{nc}, \tag{9}$$

$$\frac{\mathrm{d}N_P}{\mathrm{d}t} = k_s N_m - k_P N_P - rN_P N_{nc}, \tag{10}$$

where κ_m and κ_{nc} are the rate constants of mRNA and ncRNA formation, $[K_P/(K_P + N_P)]^n$ is the probability that all the regulatory sites of gene 2 are free of P (this probability corresponds

to the P association–dissociation equilibrium), n is the number of regulatory sites, K_P is the constant describing the P association–dissociation equilibrium, r is the rate constant of ncRNA and P association, k_s is the rate constant of P synthesis, and k_m, k_{nc} and k_P are the rate constants of mRNA, ncRNA and P degradation.

In combination, steps (1)–(7) and equations (8)–(9) form a minimal basis for illustrating the likely interplay of mRNA, ncRNA and P. In reality, the interaction between these species may of course be more complex. For example, ncRNA may participate in a few steps of conversion between the formation and step (7) (in our model, these steps are described by a lumped step (2)). The kinetics may include conformational changes of P, etc. Due to additional steps, the interplay of mRNA, ncRNA and P may be more subtle compared to that shown below. Our results are nevertheless generic. This means that the effects we discuss may often hold even if one introduces additional steps.

Assuming the growth of the cell to be slow and the system to be close to steady-state conditions and using equations (8) and (9), we have

$$N_m = \kappa_m / k_m, \tag{11}$$

and

$$N_{nc} = \frac{\kappa_m k_s}{r k_m N_P} - \frac{k_P}{r}. \tag{12}$$

Substituting the latter expression into equation (10) yields

$$\kappa_{nc} \left(\frac{K_P}{K_P + N_P} \right)^n - \left(N_P + \frac{k_{nc}}{r} \right) \left(\frac{\kappa_m k_s}{k_m N_P} - k_P \right) = 0. \tag{13}$$

Equation (13) always has at least one solution, because its left-hand part is obviously negative and positive at low and high N_P, respectively. If $n \geqslant 2$, the left-hand part of equation (13) may have a local maximum and accordingly equation (13) may have three solutions. In such cases, as usual, the lower and upper solutions are stable and the intermediary solution is unstable. If the rate constant r (for step (7)) is used as a governing parameter, the transition from a unique steady state to bistability occurs provided that r is sufficiently high. Mathematically, this transition represents a saddle-node bifurcation.

To illustrate explicitly the model predictions it is instructive to choose biologically reasonable values of the model parameters. For P and mRNA, typical kinetic parameters are presented, e.g. in the review by Kaern *et al* [2]. Detailed studies of the kinetics of steps occurring with participation of ncRNA are now unfortunately lacking. Nevertheless, we can select reasonable parameters for ncRNA. In particular, taking into account that the mechanisms of formation and degradation of ncRNA are similar to those of mRNA, we can use for ncRNA the same range of parameters as for mRNA. The values of r were chosen to illustrate bistability. *A posteriori*, the values found for r were validated by using the theory of diffusion-limited reactions as described below.

In general, the range of the values of the rate constants under consideration is rather wide. Often, the scale of the rate constants of mRNA and protein degradation is about 0.1 min^{-1} [2]. For this reason, we use $k_P = k_m = k_{nc} = 0.1\ \text{min}^{-1}$. The value of the rate constant of the mRNA synthesis, $\kappa_m = 10\ \text{min}^{-1}$, was chosen to have $N_m = 100$ (see equation (11)). The value of the rate constant of the ncRNA synthesis (in the absence of suppression by P), $\kappa_{nc} = 10^3\ \text{min}^{-1}$, was selected to have $N_{nc} = 10^4$ at $N_P \to 0$ (see equation (9)). The value of the rate constant of the P synthesis, $k_s = 1\ \text{min}^{-1}$, was chosen to have $N_P = 10^3$ at $N_{nc} \to 0$ (see equation (10)). All these values are obviously physically reasonable. In particular, the values of the rate constants are in line with [2].

The P association–dissociation constant, K_P, may be in the range from 10 (for strong association) to ∞ (for no association). The number of regulatory sites, n, is in the range from

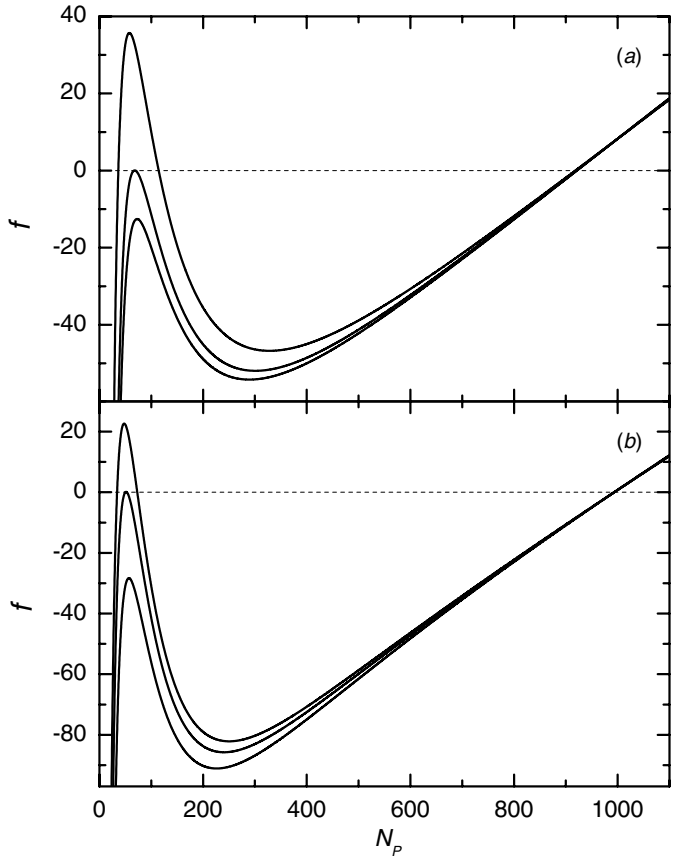

Figure 1. Left-hand part of equation (13) as a function of N_P for (*a*) $n = 2$, $K_P = 100$, $r = 5 \times 10^{-4}$ (lower curve), 5.25×10^{-4} (medium curve) and 6×10^{-4} min^{-1} (upper curve), and (*b*) $n = 4$, $K_P = 200$, $r = 5.5 \times 10^{-4}$ (lower curve), 6.04×10^{-4} (medium curve) and 6.5×10^{-4} min^{-1} (upper curve). For the other parameters, see the text.

0 to 8. To be specific, we use $n = 2$ and $K_P = 100$. In this case, the bistability occurs at $r > 5.25 \times 10^{-4}$ min^{-1} (see figures 1(*a*) and 2). If for example $r = 6 \times 10^{-4}$ min^{-1} (the upper curve in figure 1(*a*)), the reactant populations for the regime with low rate of the P synthesis are predicted to be $N_P = 36$, $N_m = 100$ and $N_{nc} = 4412$, while for the regime with high rate of the P synthesis one has $N_P = 918$, $N_m = 100$ and $N_{nc} = 15$.

For $n > 2$, the bistability can of course be observed as well for slightly different parameters. If for example $n = 4$ and $K_P = 200$, the bistability occurs at $r > 6.04 \times 10^{-4}$ min^{-1} (figure 1(*b*)).

Figure 2 shows that in the bistable case the P population is either high or low. Physically, it is clear that in order to maintain the regime with low P population, the rate of the ncRNA synthesis should be sufficiently high. This makes it possible to derive a simple necessary criterion for observation of bistability. In particular, equation (9) indicates that the maximum possible ncRNA population is $N_{nc} = \kappa_{nc}/k_{nc}$. This number should be sufficient in order to suppress the P population. This is possible provided that $rN_{nc} \gg k_P$. Using for N_{nc} the expression above, we can rewrite this condition as $r\kappa_{nc} \gg k_{nc}k_P$. In addition to this

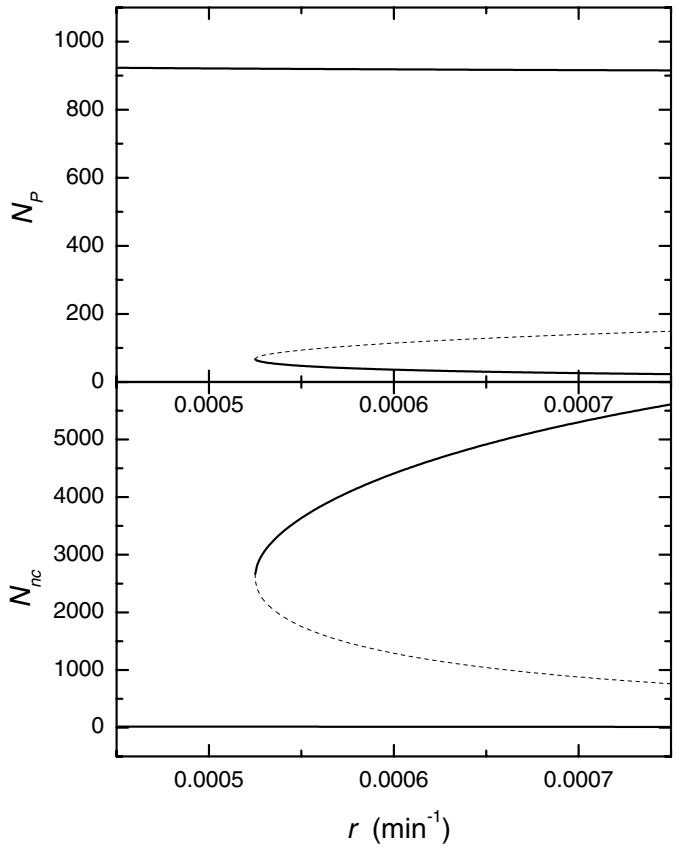

Figure 2. P and ncRNA numbers as a function of r according to equations (8)–(10) under steady-state conditions for $n = 2$ and $K_P = 100$ (for the other parameters, see the text). The thick solid and dashed lines correspond respectively to the stable and unstable steady states. For comparison, note that in this case the mRNA number is equal to 100 irrespective of r (see equation (11)).

criterion, there is another desirable condition. If the P population is suppressed, we have $rN_PN_{nc} \simeq k_s\kappa_m/k_m$ (see equations (10) and (11)). The high ncRNA population can be maintained if this rate is small compared to the rate of the ncRNA synthesis in the absence of its suppression by P, i.e., $\kappa_{nc} \gg k_s\kappa_m/k_m$. In our calculations, these conditions are fulfilled with a wide margin, and accordingly the bistability can be observed even if one somewhat reduces κ_{nc} (e.g., down to 500 min^{-1}).

In relation with the values of r used in our calculations presented above, it is appropriate to note that the biochemical reactions of association (step (7) belongs to this class) are often relatively rapid so that the corresponding rate constants are comparable to those predicted for the diffusion-limited case. Thus, the upper value of r is $4\pi D\rho/V$, where D is the protein diffusion coefficient, ρ is the length comparable to the protein size, and V the cell volume. According to hydrodynamics, for example, the coefficient of diffusion of spherically shaped particles in water is given by $D = k_BT/(6\pi\eta\varrho)$, where ϱ is the particle radius, and η is the viscosity. For folded proteins, ϱ is about 2×10^{-7} cm. Using this value and $\eta = 0.01$ g cm s^{-1}, one obtains $D \simeq 10^{-6}$ cm^2 s^{-1}. Inside cells, the diffusion coefficient is however lower by about one order of magnitude due to macromolecular crowding resulting in

steric constraints on diffusion and influencing diffusion via weak intermolecular interactions, i.e., $D \simeq 10^{-7}$ cm^2 s^{-1} [37]. This value can also be used for small RNAs. For long RNAs, the diffusion coefficient is smaller [37]. For ρ, one can use $\rho = \varrho = 2 \times 10^{-7}$. It would imply that any contact results in reaction. In reality, however, only a fraction of contacts is effective. Thus, ρ is expected to be at least a few times lower than ϱ. Employing $\rho = 5 \times 10^{-8}$ cm, $D = 10^{-7}$ cm^2 s^{-1} and $V = 10^{-9}$ cm^3, we get $4\pi D\rho/V \simeq 5 \times 10^{-5}$ s^{-1} or $\simeq 3 \times 10^{-3}$ min^{-1}. The latter value is only slightly larger than the values of r used in our calculations. Thus, our values of r are physically reasonable.

Our treatment above is focused on the situation when the feedback between the ncRNA and P production is negative. The case of positive feedback was analysed as well by modifying the first term in the right-hand part of equation (9) as

$$\frac{\mathrm{d}N_{nc}}{\mathrm{d}t} = \kappa_{nc}\left(\frac{N_P}{K_P + N_P}\right)^n - rN_mN_{nc} - k_{nc}N_{nc}, \tag{14}$$

where $[N_P/(K_P + N_P)]^n$ is the probability that all the regulatory sites of gene 2 are occupied by P. With this modification, the solution of the corresponding equations was found to be unique. If one adds the delay between the ncRNA formation (2) and reaction (7) (e.g., due to the ncRNA conversion), the model was proved [38] to predict oscillations provided the parameters are suitable.

Scenario (ii). In our second model, the reaction scheme includes mRNA, ncRNA and P formation and degradation (steps (1)–(6)) and ncRNA and mRNA association and degradation,

$$\text{mRNA} + \text{ncRNA} \rightarrow \text{mRNA} * \text{ncRNA} \rightarrow \emptyset. \tag{15}$$

In reality, step (15) often occurs with participation of microRNAs (miRNAs) which are 20–22 nucleotides long [18–27]. The latter RNAs are transcribed as long ncRNA and then generated via a two-step processing pathway including first the formation of a few different ∼65-nt pre-miRNAs and then conversion of each of them into the corresponding miRNA. In our generic scheme, as already noted in the previous section, the formation of a long ncRNA and its conversion to miRNA is represented as a single lumped step (2). If necessary, the formation of miRNA can easily be described more explicitly.

Assuming P suppresses the ncRNA formation and using for steps (1)–(6) and (15) the kinetic equations similar to equations (8)–(10), one can easily get the following equation for N_P:

$$\kappa_{nc}\left(\frac{K_P}{K_P + N_P}\right)^n - \left(\frac{k_PN_P}{k_s} + \frac{k_{nc}}{r}\right)\left(\frac{\kappa_mk_s}{k_PN_P} - k_m\right) = 0, \tag{16}$$

where r is the rate constant of mRNA and ncRNA association (the other symbols are as in equations (8)–(10)). The structure of this equation for N_P is the same as that of equation (13). Thus, all our conclusions, drawn above for scenario (i) with the negative feedback between the ncRNA and P production, are applicable to scenario (ii) as well. For the positive feedback, the situation is similar (cf equation (14)).

Conclusion

We have proposed two generic kinetic models describing the mRNA, ncRNA and P interplay. Specifically, our analysis implies that ncRNA either binds P or pairs with mRNA and then results in degradation of the corresponding complex. In both cases, the steady state is found to be unique if the feedback between the ncRNA and P production is positive. For negative feedback, the models predict either unique steady state or bistable kinetics.

In the bistable case, the protein population is predicted to be either high or low. In relation with the latter regime, we may note that the formal condition for applicability of equation (9) for N_P is $N_P \gg n$. Although this condition is fulfilled in our calculations, we may note that if N_P is low (e.g., comparable to n), equation (9) predicts that the average number of P on the gene regulatory sites is close to zero and that the factor $[K_P/(K_P + N_P)]^n$ is close to 1. These predictions are physically correct. This means that the corrections to equation (9) are irrelevant, and accordingly equation (9) can often be used even if N_P is low.

In bistable systems, the limits of the applicability of the MF kinetic equations are related to fluctuations. If the number of reactants controlling the stability of steady states is relatively low, the fluctuations of this number may result from time to time in transitions between steady states, i.e., one can observe stochastic transcriptional 'bursts'. For the conventional models describing the interplay of protein and mRNA, this effect was studied in detail in many works [2, 5, 11, 15]. In our present calculations, the stability of the regime with low P population is guaranteed by high value of N_{nc}, and the stability of the regime with high P population is guaranteed by high value of N_P. Thus, the stochastic effects are not significant. For other sets of model parameters, in the situations when the maximum values of N_P and N_{nc} are not high, the models proposed were proved (not shown) to predict stochastic bursts. Qualitatively, these burst are similar to those observed earlier [2, 5, 11, 15].

Finally, it is appropriate to articulate that our two models describe important novel aspects of the kinetics of gene expression. The results presented help to understand the functions of ncRNA and may guide the experiments.

Acknowledgment

This work was partially funded by the Swedish Science Council.

References

[1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K and Walter P 2002 *Molecular Biology of the Cell* (New York: Garland)
[2] Kaern M, Elston T C, Blake W J and Collins J J 2005 *Nature Rev. Genet.* **6** 451
[3] Bornholdt S 2005 *Science* **310** 449
[4] Kaufmann B B and van Oudenaarden A 2007 *Curr. Opin. Gen. Dev.* **17** 107
[5] Mariani L, Löhning M, Radbruch A and Höfer T 2004 *Prog. Bipohys. Mol. Biol.* **86** 45
[6] Bratsun D, Volfson D, Tsimring L S and Hasty J 2005 *Proc. Natl Acad. Sci. USA* **102** 14593
[7] Raser J M and O'Shea E K 2005 *Science* **309** 2010
[8] Chang H H, Oh P Y, Ingber D E and Huang S 2006 *BMC Cell Biol.* **7** art. no. 11
[9] Isaacson S A and Peskin C S 2006 *SIAM J. Sci. Comput.* **28** 47
[10] Guido N J, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor C R, Elston T C and Collins J J 2006 *Nature* **439** 856
[11] Zhdanov V P 2006 *Chem. Phys. Lett.* **424** 394
[12] Huang S, Guo Y P, May G and Enver T 2007 *Dev. Biol.* **305** 695
[13] Rateitschak K and Wolkenhauer O 2007 *Math. Biosci.* **205** 163
[14] Yoda M, Ushikubo T, Inoue W and Sasai M 2007 *J. Chem. Phys.* **126** 115101
[15] Zhdanov V P 2007 *JETP Lett.* **85** 302
[16] Zhdanov V P 2007 *JETP Lett.* **86** 492
[17] Zhdanov V P 2008 *Chem. Phys. Lett.* **458** 247
[18] Bertone P *et al* 2004 *Science* **306** 2242
[19] RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) and FANTOM Concortium 2005 *Science* **309** 1564
[20] Bartel D P 2004 *Cell* **116** 281
[21] Costa F F 2005 *Gene* **357** 83
[22] Goodrich J A and Kugel J F 2006 *Nat. Rev. Mol. Cell. Biol.* **7** 612

[23] Michalak P 2006 *J. Evol. Biol.* **19** 1768
[24] Costa F F 2007 *Gene* **386** 1
[25] Gammell P 2007 *Cytotechnology* **53** 55
[26] Kulshreshtha R, Ferracin M, Negrini M, Calin G A, Davuluri R V and Ivan M 2007 *Cell Cycle* **6** 1426
[27] Wilfred B R, Wang W X and Nelson P T 2007 *Mol. Gen. Metabol.* **91** 209
[28] Yazgan O and Krebs J E 2007 *Biochem. Cell Biol.* **8** 484
[29] Rao P K, Kumar R M, Farkhondeh M, Baskerville S and Lodish H F 2005 *Proc. Natl Acad. Sci. USA* **103** 8721
[30] Zhao Y, Samal E and Srivastava D 2005 *Nature* **436** 214
[31] Duan Z, Person R E, Lee H H, Huang S, Donadieu J, Badolato R, Grimes H L, Papayannopoulou T and Horwitz M S 2007 *Mol. Cell. Biol.* **27** 6889
[32] Levine E, Zhang Z, Kuhlman T and Hwa T 2007 *PLOS Biol.* **5** 1998
[33] Levine E, McHale P and Levine H 2007 *PLOS Comput. Biol.* **3** 2356
[34] Levine E, Ben Jacob E and Levine H 2007 *Biophys. J.* **93** L52
[35] Mitarai N, Andersson A M C, Krishna S, Semesey S and Sneppen K 2007 *Phys. Biol.* **4** 164
[36] Shimoni Y, Friedlander G, Hetzroni G, Altuvia S, Biham O and Margalit H 2007 *Mol. Syst. Biol.* **3** 138
[37] Misteli T 2008 *Histochem. Cell Biol.* **129** 5
[38] Zhdanov V P 2008 *Chem. Phys. Lett.* **458** 359