Physics Reports 500 (2011) 1–42

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Kinetic models of gene expression including non-coding RNAs

Vladimir P. Zhdanov *

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

ARTICLE INFO

Article history:
Accepted 4 December 2010
Available online 14 December 2010
editor: M. Klein

ABSTRACT

In cells, genes are transcribed into mRNAs, and the latter are translated into proteins. Due to the feedbacks between these processes, the kinetics of gene expression may be complex even in the simplest genetic networks. The corresponding models have already been reviewed in the literature. A new avenue in this field is related to the recognition that the conventional scenario of gene expression is fully applicable only to prokaryotes whose genomes consist of tightly packed protein-coding sequences. In eukaryotic cells, in contrast, such sequences are relatively rare, and the rest of the genome includes numerous transcript units representing non-coding RNAs (ncRNAs). During the past decade, it has become clear that such RNAs play a crucial role in gene expression and accordingly influence a multitude of cellular processes both in the normal state and during diseases. The numerous biological functions of ncRNAs are based primarily on their abilities to silence genes via pairing with a target mRNA and subsequently preventing its translation or facilitating degradation of the mRNA–ncRNA complex. Many other abilities of ncRNAs have been discovered as well. Our review is focused on the available kinetic models describing the mRNA, ncRNA and protein interplay. In particular, we systematically present the simplest models without kinetic feedbacks, models containing feedbacks and predicting bistability and oscillations in simple genetic networks, and models describing the effect of ncRNAs on complex genetic networks. Mathematically, the presentation is based primarily on temporal mean-field kinetic equations. The stochastic and spatio-temporal effects are also briefly discussed.

Contents

Abbreviations: **C**, ncRNA–protein complex; **DNA**, deoxyribonucleic acid; **mRNA**, messenger RNA; **miRNA**, microRNA; **ncRNA**, non-coding RNA; **P**, protein; **Pol II**, RNAP II; **rRNA**, ribosomal RNA; **RNAP**, RNA polymerase; **RNA**, ribonucleic acid; **snRNA**, small nuclear RNA; **tRNA**, transfer RNA; **3D**, three-dimensional.

* Corresponding address: Department of Applied Physics, Chalmers University of Technology, S-412 96 Göteborg, Sweden.
E-mail address: zhdanov@catalysis.ru.

0370-1573/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physrep.2010.12.002

1. Introduction

In living cells, the information containing the biological heritage is stored in and transferred from generation to generation via deoxyribonucleic acid (DNA) or, more specifically, genes representing segments of DNA [1]. The understanding of the mechanism and kinetics of the expression of this information is clearly of high interest from very different perspectives of natural sciences, medicine, agriculture, and philosophy. In particular, our responses to such basic questions as "how did we get here" and "where do we go from here" are inherently related to DNA or, specifically, to gene expression [2].

Historically, the modern era in the study of gene expression started in 1944 with the discovery by Avery et al. [3] that DNA contains the genetic information. The next big landmark was the suggestion of the DNA structure by Watson and Crick in 1953 [4]. Since then, the *central dogma of molecular biology* (for the origin of this term, see Ref. [5]) is that the information encoded in genes is expressed via their transcription (by polymerase), resulting in the synthesis of ribonucleic acids (RNAs) or, more specifically, messenger RNAs (mRNAs) and subsequent translation of these RNAs to proteins by ribosomes. Detailed description of the species and processes involved into gene expression can now be found in common textbooks [1]. The original experimental studies scrutinizing the mechanisms of the mRNA and protein synthesis and degradation are countless (see, e.g., recent reviews of gene transcription [6–9], mRNA translation [10–13], mRNA and protein degradation [14–16], or reviews [17–19] focused on the spatio-temporal effects in gene expression).

The important point is that the whole process of gene expression is regulated at every step. In particular, the gene transcription is often controlled by the master regulatory proteins (transcription factors). Such proteins associate with DNA and either facilitate or suppress the RNA synthesis. Due to these feedbacks, the kinetics of gene expression may be fairly complex. In addition, most genes exist in cells at single or low copy numbers, the numbers of mRNA and protein copies are not large (from a few up to about 10^4 copies [14]), and the mRNAs and proteins expressed in small numbers of copies are integrally much more abundant than those expressed in relatively large numbers of copies (quantitatively, the latter is described by Zipf's law [20]). For these reasons, the kinetics of gene expression often exhibit stochastic features.

To scrutinize the gene-expression kinetics and/or to explain the concepts used in this field, one can employ kinetic models. The available models can be divided into two complementary categories. The models of the first category are aimed at detailed description of the basic steps of gene expression, including, first of all, gene transcription [21–31] and mRNA translation [32–35]. The models of the second category are focused on the kinetics of the mRNA–protein networks. The latter models are more coarse-grained. In particular, the synthesis or degradation of a mRNA or protein is usually described as a single step and the emphasis is on the interplay of such steps. At present, there are plenty of models of this type (see reviews focused on stochastic bursts and bistability in simple genetic networks [36–40], kinetic oscillations [41,42], and complex genetic networks [43–47]).

The paradigm outlined above is fully applicable to prokaryotes (cells without nuclei) whose genomes consist of tightly packed sequences transcribed into protein-coding RNAs (e.g., to mRNAs). In contrast, the genomes of the eukaryotes (cells with nuclei) contain relatively rare protein-coding sequences (in the human genome, for example, there are only $\sim 2.3 \times 10^4$ protein-coding genes, representing about 2% of the total genomic sequence [48]). During the past fifteen years, it has been established that an appreciable part (up to 90%) of the rest of the genome of the latter cells is transcribed as well into ncRNAs. Most of such RNAs are transcribed at low levels. Many of ncRNAs are, however, abundant. Altogether, they form the cornerstone of a regulatory network that operates in concert with the protein network (see recent reviews focused on long ncRNAs [49–53] and small ncRNAs generated by processing long ncRNAs [54–58]). In prokaryotes, ncRNAs are less abundant and nevertheless should also be taken into account [59].

The numerous biological functions of ncRNAs in general and especially of small ncRNAs are based primarily on their ability to pair with target mRNAs and then either to prevent translation or to result in rapid degradation of the mRNA–ncRNA complex [54–56] (the relative role of these channels is discussed in Ref. [60]). This mechanism of regulation of genetic networks is different compared to conventional protein-mediated regulation. The important point is that miRNAs often have many (up to a few hundreds) mRNA targets [54]. For this reason, such RNAs are expected to play a role of global regulators. Long ncRNAs and, to some extent, small ncRNAs possess many other abilities as well [49–51].

The important role of ncRNAs has been mapped out in a wide variety of cellular processes. One can safely say that ncRNAs influence in one or another way almost every intracellular process. For example, thousands of mammalian mRNAs are highly expressed at developmental stages before small ncRNA expression and their levels tend to fall as the small ncRNAs that target them begin to accumulate [61] (for more specific aspects, see Ref. [62]). Stem cell proliferation and differentiation depend on small ncRNAs [63–65]. Glucose and lipid metabolism depend on small ncRNAs as well [66]. Small ncRNAs are crucial for brain development [67] and are expressed at a high level in the brain in adulthood and participate in normal and abnormal brain functions [68]. Abnormal levels of ncRNA expression were observed in many types of human cancer [69,70]. Misexpression of ncRNAs occurs also in many other diseases including diabetes, obesity, leukemia, heart disease and inflammation [66,71]. Viral ncRNAs may manipulate host-cell gene expression by suppressing the population of some of host ncRNAs [72]. All these observations are indicative of various specific and global functions of ncRNAs.

Taking the diversity of ncRNAs and their abilities into account, one can conjecture that their formation is one of the reasons of the diversity of eukaryotes and cells in different tissues of eukaryotes [53]. The protein-coding genes in such cells are often similar. For example, the worm *Caenorhabditis elegans* has roughly the same amount of protein-coding genes as humans. Such a small difference might not be large enough to result in the difference between these species, and the role of ncRNAs may be critical here.

The observations than many diseases are accompanied by misexpression of ncRNAs is potentially important for applications. The manipulation of ncRNA levels can control disease phenotypes. The latter opens up new ways in treatments of various diseases. The race to bring the corresponding therapeutic to the market has already begun [71].

The field of ncRNAs is now rapidly expanding. The reviews describing various aspects of the ncRNA biogenesis and regulatory functions are already numerous (a few of them have already been mentioned above). This area is truly interdisciplinary. Although the central role is played here by biochemists, the physicists can also find niches for the use of their skills. One of the challenges for theorists is, for example, to devise a genome-wide computational search that captures most of the ncRNA targets without also bringing in too many false predictions [54,73]. Another challenge is to construct sound kinetic models describing various aspects of the mRNA, ncRNA and protein interplay. Activities in these areas have been seen already for several years. The goal of our review is to describe the results obtained in the latter area.

Aiming at a broad readership, we first outline the relevant biological, physical and mathematical backgrounds (Sections 2 and 3). The main part of the presentation (Sections 4–9) is focused on the kinetic models including ncRNAs. Such models can formally be divided into three groups. The simplest models describe the mRNA–ncRNA interplay in the absence of the protein-mediated feedbacks and accordingly predict a unique steady state (Section 4). To clarify the conditions of the appearance of bistability and/or oscillations, one should take the protein-mediated feedbacks between the mRNA and ncRNA synthesis into account. Here, there are two complementary strategies. First, one can introduce the terms describing the mRNA–ncRNA association and degradations into the conventional models, containing the feedbacks in the mRNA–protein interplay, and consider that there is no regulation of the ncRNA synthesis by proteins. Alternatively, one can ignore the effect of proteins on the mRNA synthesis and focus on the protein-regulated ncRNA synthesis. The results obtained following these two lines are reviewed, respectively, in Sections 5 and 6 with emphasis on simple genetic networks. More complex networks are discussed afterwards (Section 7). The results presented in Sections 4–7 are based on the temporary mean-field

kinetic equations ignoring fluctuations and concentration gradients. The stochastic and spatio-temporal effects are outlined in Sections 8 and 9, respectively.

The models presented in Sections 4–9 can be used to describe transient and steady-state kinetics of gene expression including ncRNAs. All the examples shown in the figures there correspond to steady-state conditions. This restriction is motivated by four points: (i) for eukaryotic cells, the steady-state approximation often holds (see Section 3.7); (ii) the steady-state kinetics should be classified first of all; (iii) the available original studies are primarily focused on such kinetics; and (iv) the corresponding transient kinetics are too diverse and at present their classification is hardly possible.

In general, the goals of and driving forces behind the construction of kinetic models may be different. One of the common standpoints is that ideally (i) all the ingredients of a model should be based on the fundamental physical laws and/or validated experimentally, (ii) the parameters should be obtained from theory and/or independent experiments, and (iii) the predictions should be compared with the results of experiments. Another standpoint, shared by the author of this review, is that with correct validation a model should (i) predict what may happen in reality, (ii) contribute to the formation of a conceptual basis in its field, and (iii) guide experiments. Allegiance to one of these standpoints depends partly on the state of the art in a field and partly on personal taste. In this context, we may notice that the models described in this review include mechanistic details which were firmly identified in numerous experiments by using various schemes of measurements. On the other hand, as a rule, the available experiments focused on ncRNAs did not directly track temporal kinetics of gene expression or the corresponding steady-state kinetics as a function of governing parameters. The rate constants for various steps were not measured either. For these reasons, direct comparison of the results of calculations with the available experiment is now hardly possible. In the field of the ncRNA biochemistry, this situation is typical because the quantitative information on kinetic processes occurring with participation of ncRNAs is still very limited despite high current interest in ncRNAs. The lack of reliable parameters does not, however, mean that the kinetics should not be analyzed theoretically. For example, the first models [74,75] describing stochastic effects in the mRNA–protein interplay were proposed in the 1970s, 20 years before their actual experimental observation, and now those models are considered to be seminal [37]. The current experimental facilities are much better (ncRNAs can, for example, be detected on single-molecule level [76]), and there is no doubt that the kinetics of gene expression including ncRNAs will soon be studied experimentally in detail. Our humble goal is to contribute to rapid progress in this field.

2. Biological background

In this section written primarily for the readers who do not work in the biosciences, we briefly describe the biological background of the kinetic models discussed in the main part of the review.

2.1. DNA, mRNAs, proteins, ncRNAs

DNA consists of two polynucleotide chains (strands) composed of deoxyribonucleotide subunits of four types [1]. Each nucleotide is composed of a five-carbon sugar (deoxyribose) to which are attached a phosphate group and a nitrogen-containing base. The base may be either adenine (A), cytosine (C), guanine (G), or thymine (T). The nucleotides are covalently linked through the sugars and phosphates, while the hydrogen bonds between the complementary bases (two or three bonds per a base-base contact) belonging to different chains maintain them together in the spiral shape. The number of nucleotide pairs in genomes is in the range from a few thousands (in bacteria) to $\sim 3.2 \times 10^9$ (in humans). In eukaryotes, DNA is located in the nucleus and divided between a set of different chromosomes (e.g., the human genome is distributed over 24 chromosomes). Each chromosome consists of a DNA chain associated with proteins (histones) that fold DNA (this complex is called *chromatin*). In addition, it carries proteins required for the DNA transcription, replication, and repair.

RNA represents a chain of subunits of four types similar to those forming DNA [a sugar is slightly different (ribose instead of deoxyribose), and one of the four bases is slightly different as well (uracil (U) in place of T)]. The RNA synthesis is performed by RNA polymerases (RNAPs; chemically, RNAP is an enzyme). This process can be divided into five phases including (i) the RNAP recruitment on DNA near the promoter sequence, (ii) RNAP and DNA isomerization resulting in the formation of a locally open DNA chain (in other words, this is isomerization from a closed to an open complex), (iii) promoter escape, (iv) steps of the RNA initiation and elongation by adding nucleotides one by one (the incoming nucleotides are in the form of ribonucleoside triphosphates), and (v) RNA and RNAP release after reading the final base of DNA [6–8].

With proper specification, the scheme described above is applicable both to prokaryotes and eukaryotes. For example, bacteria contain only one form of RNAP including a core enzyme consisting of five subunits associated with a σ factor which is used to recognize promoter DNA sequences (the promoter escape is typically accompanied by loss of this factor). In contrast, eukaryotic cells contain three nuclear RNAPs, with RNAP II (Pol II) responsible for transcribing all mRNAs and numerous ncRNAs. Pol II, a 12-subunit enzyme, does not recognize promoter DNA by itself, but rather as part of the basal Pol II machinery including general transcription factors (proteins). Like σ factors, these general transcription factors usually dissociate from Pol II during the transition between initiation and elongation. Plants have distinct RNAP complexes, Pol IV and Pol V [77].

The kinetic models of the RNA synthesis are presented in Refs. [21–31].

Cells produce several types of RNA including mRNAs, ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and small nuclear RNAs (snRNAs) [1]. snRNAs direct the splicing of pre-mRNAs to form mRNAs. mRNAs encode proteins and ultimately are

translated into proteins by ribosomes. rRNAs form (together with proteins) the core of ribosomes. tRNAs associate with specific amino acids and participate in their incorporation into growing proteins on ribosomes.

mRNA can be viewed as a set of codons. Each codon contains three nucleotides. With four different nucleotides, there are 64 (4^3) different codons. Each tRNA contains an anticodon which is complementary to one of the mRNA codons. tRNA functions as an adaptor. It recognizes and binds both to a codon and, at another side, to an amino acid. There are 20 different amino acids used to form proteins step by step on ribosomes (some amino acids are specified by more than one codon). This process is described in detail, e.g., in recent reviews [10–13]. For the corresponding kinetic models, see Refs. [32–35].

Structurally, a proteins represents a linear chain of amino acid residues joined via the peptide bonds between the carboxyl and amino groups. The protein native state is usually globular. Often, it contains well-ordered domains consisting of α-helices and/or β-sheets. Such low-dimensional fragments of proteins are stabilized primarily by hydrogen bonds.

In the kinetics models focused on the mRNA–protein interplay, a protein-coding gene is usually considered to be transcribed into a specific mRNA which in turn is translated into a specific protein. In reality, the situation may be more complex (especially in eukaryotes). In particular, the initiation and termination of transcription may sometimes occur on different regions of a gene or two adjacent genes, and accordingly a gene may be transcribed into distinct mRNAs (see Ref. [78] and references therein). In addition, mRNAs are often first formed as pre-mRNAs and then converted into mRNAs. The latter process may sometimes occur in two or more ways resulting in different mRNAs [80].

Above, we have mentioned rRNAs, tRNAs, and snRNAs. Although these RNAs are formally non-coding, they are directly involved in the synthesis of proteins. In addition to these RNAs, many sequences of DNA of eukaryotic cells are transcribed into other non-coding RNAs which are not directly related to the synthesis of proteins. Our review is focused on the latter RNAs, and the abbreviation "ncRNAs" is reserved for the RNAs of this category.

The simplest classification of ncRNAs is based on their size. Usually, ncRNAs are divided into two groups, including (i) long ncRNAs (from ∼200–300 to thousands of nucleotides) obtained directly after gene transcription and (ii) small ncRNAs (from 20 to ∼200–300 nucleotides) obtained by cleavage of long ncRNAs. In turn, small ncRNAs can be divided at least into five subgroups [81].

One of the most important, abundant and interesting subgroups of small ncRNAs includes microRNAs (miRNAs) that are 20–22 nucleotides long. These abundant RNAs are transcribed as long ncRNA and then generated via a two-step processing pathway including the formation of a few different ∼65-nt pre-miRNAs followed by conversion of each of them into the corresponding miRNA [82]. Typically, miRNAs are associated and operate with protein (AGO). Interestingly, some miRNAs can be formed via viral transcription [72,83].

The RNA and protein degradation is performed by various enzymes [14–16]. Chemically, enzymes are proteins incorporating usually a metal ion within their structure to stabilize the protein folded conformation and to form a catalytically active site.

2.2. Regulation of gene expression by proteins

As already noted in the introduction, every step on the pathway to gene expression and especially the transcription initiation can be regulated. In the conventional regulation schemes, activator or repressor proteins bind to specific DNA sequences near or overlapping with the RNAP binding sites [84,85]. Unconventional regulators found more recently bind directly to RNAP without binding to DNA [86].

The first kinetic models of the regulation of gene transcription by proteins were proposed in the beginning of the eighties [87,88]. Since then, many aspects of this regulation have been rationalized in detail (for numerous models, see recent articles [89–94], review [95] and references therein). In general, the statistics of occupation of regulatory sites depends on the lateral (along DNA) interactions between bound proteins. In the corresponding treatments, the protein association to and dissociation from the regulatory sites is usually considered to be rapid compared to the transcription steps or, in other words, the system is assumed to be close to an association–dissociation equilibrium. In this case, the probabilities of various arrangements of proteins on the regulatory sites can be calculated by employing the grand canonical distribution or the steady-state kinetic equations describing protein association and dissociation. In addition, one should specify the transcription rate for each arrangement of regulatory proteins. Then, these rates and the protein-arrangement probabilities are used to calculate the average transcription rate.

In the generic models focused on the RNA–protein interplay, binding of proteins to the regulatory sites is usually considered to be independent and fast (close to the association–dissociation equilibrium) and to occur at $n \gg m$, where n is the protein population in a cell, and m is the number of regulatory sites for a given gene. In addition, the sites are considered to be equivalent and the lateral protein–protein interactions are assumed to be negligible. In this case, the regulation of gene transcription by protein is described by simple power-law equations.

For example, let us consider that the transcription rate is high only provided that m equivalent regulatory sites of the gene are occupied by proteins. The probability of this arrangement is equal to p_1^m, where p_1 is the probability that a site is occupied. The latter probability is given by $p_1 = k_a c_P/(k_d + k_a c_P)$, where k_a and k_d are the association and dissociation rate constants, and c_P is the protein concentration. Taking into account that the protein concentration is related with the number of protein copies and cell volume as $c_P = n/V$, the expression for the probability that a regulatory site is occupied by protein can be rewritten as $n/(K + n)$, where $K \equiv k_d V/k_a$ is the constant characterizing the protein–gene association–dissociation

equilibrium. The probability that all the regulatory sites are occupied is accordingly given by $[n/(K+n)]^m$. With this specification, the transcription rate can be represented as

$$w = w_b + w_\circ \left(\frac{n}{K+n}\right)^m, \tag{1}$$

where w_b and $w_\circ$ ($w_b \ll w_\circ$) are the rates of the basal and protein-regulated gene transcription.

Eq. (1) corresponds to positive regulation of gene transcription. In the simplest case of negative regulation, the transcription occurs provided that all the regulatory sites are vacant. The corresponding expression for the transcription rate can be derived in analogy with Eq. (1),

$$w = w_\circ \left(\frac{K}{K+n}\right)^m. \tag{2}$$

Expressions (1) and (2) will be widely used in our presentation below. Alternatively, one can employ the conventional Hill expressions, e.g.,

$$w = w_b + \frac{w_\circ n^m}{K^m + n^m} \quad \text{and} \quad w = \frac{w_\circ K^m}{K^m + n^m}, \tag{3}$$

or more specific expressions taking the correlations in the arrangement of proteins into account.

2.3. Regulation of gene expression by small ncRNAs

According to reviews [54–56], the beginning of studies of small ncRNAs is associated with the discovery of *lin*-4 in the worm *C. elegans* in 1993 [96,97]. Since then, it has been demonstrated that such ncRNA can direct mRNA cleavage and/or, more often, mRNA translational repression and/or destabilization by binding to target mRNAs [54–56]. Other more specific types of regulation include ncRNA-mediated protein recruitment for translational activation [98] or inactivation [99].

2.4. Regulation of gene expression by long ncRNAs

According to the review by Mercer et al. [49], long ncRNAs were first identified as a a transcriptional class during the large-scale sequencing of the mouse DNA [100]. In analogy with small ncRNAs, long ncRNAs often pair with mRNAs and/or proteins. In general, the latter RNAs are, however, associated with many other processes (see recent reviews [49–51]). In brief, the emerging paradigms in this area are as follows:

(i) *Chromatin modification.* In eukaryotic cells, chromosomes are packaged by histones (positevely charged proteins) into a condensed structure called chromatin. ncRNAs can silence gene expression by recruitment of chromatin modifying complexes, exclusion of the transcription machinery from the chromosome, modification of histones, and subsequent changes in the chromatin structure [50]. There are also other scenarios of the ncRNA-induced chromatin modification resulting in transcriptional activation [51].

(ii) *Transcriptional interference.* The transcription of a ncRNA across the promoter region of a protein-coding gene can interfere with transcription factor binding and prevent this gene from the mRNA synthesis [51].

(iii) *Direct regulation of transcription.* ncRNAs can directly regulate transcription through a range of mechanisms [49]: ncRNA can associate with a gene and recruit a ncRNA-binding protein with subsequent regulation of the gene activity by this protein or the ncRNA–protein complex; ncRNA can associate with protein, and the ncRNA–protein complex can then associate with a gene and regulate transcription; ncRNA and regulatory protein can competitively adsorb on a regulatory site of a gene and the transcription rate can depend on the state of this site.

(iv) *Indirect regulation of transcription.* ncRNAs can associate with mRNAs, other ncRNAs or proteins and reduce the population of these species and/or affect their processing [51]. In addition, long ncRNAs can be processed to yield small RNAs [51]. Such diverse processes can indirectly regulate the transcription of many genes.

(v) *Spatial aspects.* ncRNAs can serve as structural RNAs. Within the nucleus, for example, a number of RNA-binding proteins localize to paraspeckles. These irregularly shaped compartments seem to be partly formed of ncRNAs [51]. More globally, ncRNAs have been found to participate in the organization and maintenance of the cellular cytoskeleton [51].

3. Mathematical and physical background

The kinetics of gene expression are described by using the standard methods of chemical kinetics and statistical physics. Recently, this area was reviewed in detail by Kulasiri et al. [38] and Wilkinson [101]. For general readership, we may notice here that the bulk of the kinetic models of gene expression are based on the assumption that the RNA and protein diffusion in a cell is rapid on the time scale of other processes and that accordingly these species are distributed at random in a cell or in the nucleus and cytoplasm. This assumption is usually valid. In this case, the kinetics of gene expression are temporal and can be described in terms of the activity of genes (Section 3.1) or RNA and protein populations (Sections 3.2–3.5).

3.1. Boolean networks

Since Kauffman's seminal works [102], the Boolean networks are widely employed to describe complex genetic networks [38,46]. In this approach, each gene is represented as a node which can be active or inactive. Each node is connected with one or a few other nodes. The time is changed discretely. During each step, the states of all the nodes are changed (usually simultaneously) depending on the states of other nodes.

For the networks including ncRNAs, the effect of ncRNAs on other species is gradual and can hardly be described by the Boolean rules. For this reason, to our knowledge, the Boolean networks have not been used in this subfield of the theory of genetic networks.

3.2. Mean-field kinetic equations

The temporal mean-field kinetic equations operating with the intracellular RNA and protein concentrations or populations are very widely employed to describe genetic networks. This approach is simple and flexible and allows one to take various details into account. For example, the simplest equation for the mRNA and protein populations, N and n, are as follows

$$\mathrm{d}N/\mathrm{d}t = w - kN, \tag{4}$$

$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n, \tag{5}$$

where w is the transcription rate, υ is the mRNA translation rate constant, and k and κ are the mRNA and protein degradation rate constants.

The use of Eqs. (4) and (5) or other similar equations implies that the cell machinery functions in a time-independent manner. For example, the RNA and protein degradation is performed by special enzymes, and the RNA and protein degradation rates should be proportional to the enzyme concentrations. In Eqs. (4) and (5), these concentrations are included in the degradation rate constants. This is possible provided that the enzymes concentrations are constant. The latter is often valid provided that the cell growth is negligible and may also be valid during growth, because in the latter case the concentrations of many proteins and enzymes remain nearly constant (see Section 3.7 below).

To describe the spatio-temporal kinetics, one can employ the mean-field reaction–diffusion kinetic equations.

3.3. Chemical master equations

The mean-field kinetic equations ignore fluctuations. To take fluctuations into account, one can use the chemical master equations for the probabilities of realization of the states with given discrete RNA and protein populations [37,38]. These equations can be solved in simple cases. The models including ncRNAs and containing the corresponding bimolecular reaction steps do not belong to this category. For this reason, to our knowledge, the chemical master equations have not been employed so far to analyze the gene-expression kinetics occurring with with participation of ncRNAs.

3.4. Chemical Langevin equations

To take fluctuations into account, one can also use the chemical Langevin equations [103]. In this approach, the mean-field kinetic equations for the average populations are complemented by the terms describing noise.

3.5. Monte Carlo simulations

The Monte Carlo technique is simple, flexible and accurate. For these reason, it is widely used in various branches of natural science in general and in simulations of the kinetics of gene expression in particular. The latter simulations are usually focused on temporal kinetics. In this case, the standard Gillespie algorithm [104] includes the calculation of the total rate of all the possible steps, $W_t = \sum_i W_i$, realization of one of the steps chosen with probability W_i/W_t, and the corresponding increment of time by $|\ln(\rho)|/W_t$, where ρ ($0 < \rho \leq 1$) is a random number. This approach is equally applicable to simulations of transient kinetics and kinetics occurring under steady-state conditions.

Concerning the Monte Carlo simulations of the kinetics of gene expression, we may notice that in general the total rate, W_t, should include not only the rates of synthesis, association and degradation of RNAs and proteins but also the rates of protein association to and dissociation from the regulatory sites. In practice, the latter rates are usually neglected. This approximation is valid provided that the association and dissociation are rapid and $n \gg m$.

In general, there are plenty of Monte Carlo algorithms [105,106]. In the case of spatio-temporal kinetics, many algorithms are more convenient compared to the Gillespie algorithm.

The examples of 3D spatio-temporal Monte Carlo simulations of the kinetics of gene expression are still rare [107,108].

3.6. Typical values of the rate constants

To describe genetic networks by using kinetic equations or Monte Carlo simulations, one usually needs a lot of kinetic parameters. In specific cases, as a rule, the quantitative information about these parameters is scarce and/or lacking. This is especially the case for the networks including ncRNAs. The typical ranges of various parameters are, however, well known.

According to experiments, the transcription rate calculated in base per min is usually considered to be 0.4–6 kb/min [9,109]. The corresponding rate calculated in mRNA per min is about 1–10 min^{-1}.

The time scale of the mRNA translation per codon is often about 0.1–1 s [13]. The translation rate constant is accordingly in the range 0.1–1 min^{-1}.

The mRNA and protein degradation usually occurs on the time scale between a few minutes and one hour or longer [110,111], and accordingly the corresponding rate constants are often in the range 0.01–0.1 min^{-1} or lower.

The parameter ranges above are presented for orientation. In reality, the transcription, translation and degradation rates or rate constants depend on various factors, e.g., on the availability of substrates, and in specific cases can often be lower or higher than those indicated. This is clear, for example, from the graphs showing the protein populations *versus* mRNA populations (e.g., for bacteria, yeast and human [14]). For human, these populations are in the ranges $10^{1.5}$–10^{5} and 10^{2}–$10^{4.5}$, respectively. This means that roughly the ratio of the transcription rate and the mRNA degradation rate is 10^{2}–$10^{4.5}$, and that the ratio of the translation rate constant and the protein degradation rate constant is on average somewhat larger than unity but in specific cases can be both higher and lower than unity.

Direct experimental data about the rates of ncRNA formation and the rate constants of ncRNA degradation are still lacking. The mechanisms of gene transcription into mRNAs and ncRNAs are, however, similar (for example, the transcription of mRNA genes and most miRNA genes is mediated by RNAP II [7,56]). The mechanisms of conventional degradation of these species are similar as well. For these reasons, the kinetic parameters for ncRNAs are expected to be in the same ranges as those for mRNAs. This assumption is widely used in order to validate the choice of parameters in theoretical studies of the kinetics of gene expression including ncRNAs.

In addition, we should should briefly discuss association of ncRNA and mRNA (or protein). The rate of this process (per cell) can be represented as

$$W_{\mathrm{as}} = rNN_{*}, \tag{6}$$

where r is the corresponding rate constant, and N and N_{*} are the mRNA and ncRNA populations. In relation with the value of r, it is appropriate to notice that the biochemical reactions of association are often relatively rapid so that the corresponding rate constants are comparable to those predicted for the diffusion-limited case [112], i.e., the upper value of r is $4\pi D\rho/V$, where D the RNA diffusion coefficient, ρ is the length comparable to the RNA size, and V the cell volume. According to hydrodynamics, the coefficient of diffusion of spherically shaped particles in water is given by $D = k_B T/(6\pi\eta\varrho)$, where ϱ is the particle radius, and η is the viscosity. For mRNAs and proteins, this equation typically yields $D \simeq 5 \times 10^{-7}$ cm^2/s. Inside cells, the diffusion coefficient is however usually 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 [113]. Using for estimates this value in combination with $\rho = 5 \times 10^{-8}$ cm 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}. Thus, we should have $r \leq 3 \times 10^{-3}\mathrm{min}^{-1}$.

With the parameters above, the generic kinetic models can be used to illustrate general concepts in the field under consideration.

3.7. Steady-state conditions

We have already noticed in the introduction that although the models presented in this review can be used to describe transient kinetics of the RNA–protein interplay the main attention is paid to the steady-state kinetics. The applicability of the steady-state approximation is validated by two factors.

(i) The duration of the cycle of eukaryotic cells is well known to be in a very wide range from about 10 min to a few years. For example, the cycle duration in the fission yeast *Schizosaccharomyces pombe* is about 3 h [114]. For fast-dividing mammalian cells (e.g., for rat neural stem stem cells [115]), the cycle duration is typically about one day. On the other hand, the mRNA and protein degradation occurs on the time scale from a few minutes to one hour (Section 3.6). Thus, the mRNA and protein degradation is often (but not always) fast on the time scale of the cell cycle, and accordingly the steady-state approximation is usually valid during each phase of the cell cycle.

(ii) The results obtained in the steady-state approximation are instructive provided that the intracellular biochemistry does not change appreciably during the cell cycle. This condition does not hold in the case of the cycle-controlling and cycle-related RNAs and proteins. The contribution of such RNAs and proteins to the global RNA and protein populations is, however, relatively small. The most proteins are expressed in a cell size dependent manner so that the protein concentrations in cells are nearly constant irrespective of the cell size (see, e.g., experiments [116,117]). This is indicative that for the bulk of RNAs and proteins the steady-state kinetics are representative for the whole cell cycle.

The kinetic models focused on the cycle-controlling mRNAs and proteins are available in the literature (see, e.g., Refs. [118–121]). Some of ncRNAs are expected to participate in the cell-cycle control as well. The corresponding models are, however, lacking, and this subject is not discussed in our review.

4. Models without the protein-mediated feedbacks

The numerous biological functions of ncRNAs in general and miRNAs in particular are often based on their ability to associate with target mRNAs and then either to prevent translation or to result in rapid degradation of the mRNA–ncRNA complex. The simplest kinetic models of gene expression, focused on the mRNA–ncRNA association, do not contain the protein-mediated feedbacks and operate only with the intracellular mRNA and ncRNA populations.

4.1. mRNA and ncRNA with irreversible association

In general, the mRNA–ncRNA association is reversible. If the subsequent degradation of the mRNA–ncRNA complex is rapid, one can neglect its dissociation and consider that the mRNA–ncRNA association is irreversible. The generic kinetic model describing this case includes transcription of two genes, respectively, into mRNA and ncRNA,

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}, \tag{7}$$

$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{ncRNA}, \tag{8}$$

conventional enzyme-mediated degradation of these species,

$$\mathrm{mRNA} \to \varnothing, \qquad \mathrm{ncRNA} \to \varnothing, \tag{9}$$

and irreversible mRNA–ncRNA association and degradation,

$$\mathrm{mRNA} + \mathrm{ncRNA} \to \mathrm{mRNA} * \mathrm{ncRNA} \to \varnothing. \tag{10}$$

The corresponding mean-field kinetic equations for the mRNA and ncRNA populations in a cell are as follows

$$\mathrm{d}N/\mathrm{d}t = w - kN - rNN_*, \tag{11}$$

$$\mathrm{d}N_*/\mathrm{d}t = u - k_*N_* - rNN_*, \tag{12}$$

where w and u are the transcription rates, k and k_* are the degradation rate constants, and r is the association rate constant.

Under steady-state conditions, Eqs. (11) and (12) can be elementarily solved. In particular, the mRNA population is represented as

$$N = \frac{rw - ru - kk_*}{2rk} + \left[\left(\frac{rw - ru - kk_*}{2rk}\right)^2 + \frac{k_*w}{rk}\right]^{1/2}. \tag{13}$$

The expression for the ncRNA population is similar.

If the mRNA–ncRNA association is slow, one can neglect the last term in Eqs. (11) and (12) and get

$$N \simeq w/k \quad \text{and} \quad N_* \simeq u/k_*. \tag{14}$$

If the mRNA–ncRNA association is fast and $w > u$, Eqs. (11) and (12) (or (13)) yield

$$N \simeq (w - u)/k \quad \text{and} \quad N_* \simeq 0. \tag{15}$$

For $u > w$, one has

$$N \simeq 0 \quad \text{and} \quad N_* \simeq (u - w)/k_*. \tag{16}$$

Physically, expressions (15) and (16) correspond to the situation when the specie produced with lower rate degrades rapidly and almost completely via step (10) and accordingly its synthesis rate simply reduces the synthesis rate of the specie produced with a higher rate. In other words, this means that the gene with a lower transcription rate is effectively silenced. If for example the rate of the ncRNA synthesis is fixed (Fig. 1), the gene transcribed into mRNA is silenced as long as its transcription rate is low, $w < u$ (Eq. (16)). With increasing the transcription rate above the threshold, $w > u$, the mRNA population linearly increases (Eq. (15)).

The model including steps (7)–(10) and Eqs. (11) and (12) was widely used in the literature. In particular, Lenz et al. [122] have suggested that high sensitivity to external signals near the threshold (Eqs. (15) and (16)) can be behind the directional diffusion of quorum-sensing bacteria.

Levine et al. [123] (see also review [124]) have observed and scrutinized the threshold in the mRNA–ncRNA interplay in *Escherichia coli* using a synthetic reporter target.

Semsey et al. [125] have analyzed temporal kinetics in *E. coli* in the case of sudden decrease or increase in the extracellular iron level. The complementary results have been presented for this case by Mitarai et al. [126].

Shimoni et al. [127] and Legewie et al. [128] have also analyzed temporal kinetics including the manifestation of the ncRNA-related threshold in such kinetics. In particular, both groups have noticed that the mRNA response to the change of the rate of the ncRNA formation may be delayed (reviewed in Ref. [124]). If for example the cell is under steady-state conditions, the ncRNA population is large, the mRNA population is low due to the interaction with ncRNA, and then the gene transcription into ncRNA is switched off, the ncRNA pool present originally in the cell should first be depleted down

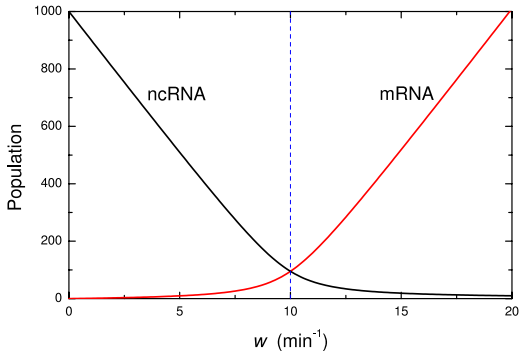


Fig. 1. mRNA and ncRNA populations as a function of the mRNA synthesis rate according to Eqs. (11) and (12) (or (13)) with $u = 10\ \text{min}^{-1}$, $k = k_* = 0.01\ \text{min}^{-1}$, and $r_a = 10^{-3}\ \text{min}^{-1}$.

to the threshold level and only then, i.e., after the delay the target mRNA and corresponding protein will accumulate. The duration of the delay period can easily be regulated by changing the initial ncRNA population. During this period, the small-amplitude signals perturbing the mRNA level will be washed out. These features are complementary to those inherent to the protein-mediated regulation and accordingly open up new ways of temporal regulation of intracellular processes.

Mehta et al. [129] have compared the regulations of protein expression by conventional transcription factors and ncRNAs with emphasis on the threshold, transient kinetics and noise. To large extent, their conclusions are similar to those drawn in Refs. [127,128]. Globally, their analysis suggests that although the transcription factors appear to be better suited for quantitative adjustment of the mRNA and protein levels, the ncRNAs likely fill a niche in allowing cells to transition quickly yet reliably between distinct states.

Mitarai et al. [130] have scrutinized the steady-state and dynamic potential of two of the best-characterized bacterial ncRNAs, *Spot42* and *RyhB*, with respect to binding to target mRNAs.

Jia et al. [131] have analyzed fluctuations of the mRNA and ncRNA populations by using the chemical Langevin equations.

Elgart et al. [132] have suggested experimental protocols for determining parameters controlling the efficiency of regulation by small RNAs and for analyzing factors and processes regulating changes in mRNA transcription and decay. Elgart et al. [133] have also analyzed stochastic effects in detail.

4.2. mRNA and ncRNA with reversible association

If the dissociation of the mRNA–ncRNA complex is not negligible, the mRNA–ncRNA interplay may occur as

$$\text{Gene}_1 \to \text{Gene}_1 + \text{mRNA}, \tag{17}$$
$$\text{Gene}_2 \to \text{Gene}_2 + \text{ncRNA}, \tag{18}$$
$$\text{mRNA} \to \varnothing, \qquad \text{ncRNA} \to \varnothing, \tag{19}$$
$$\text{mRNA} + \text{ncRNA} \rightleftharpoons \text{mRNA} * \text{ncRNA}, \tag{20}$$
$$\text{mRNA} * \text{ncRNA} \to \varnothing, \tag{21}$$
$$\text{mRNA} * \text{ncRNA} \to \text{ncRNA}, \tag{22}$$
$$\text{mRNA} * \text{ncRNA} \to \text{mRNA}. \tag{23}$$

Steps (17)–(19) are here the same as steps (7)–(9). Steps (20)–(23) represent reversible association of mRNA and ncRNA and three channels of degradation of the mRNA–ncRNA complex.

The kinetic equations describing steps (17)–(23) are defined as

$$\mathrm{d}N/\mathrm{d}t = w - kN - r_a N N_* + r_d N_C + k_{d3} N_C, \tag{24}$$
$$\mathrm{d}N_*/\mathrm{d}t = u - k_* N_* - r_a N N_* + r_d N_C + k_{d2} N_C, \tag{25}$$
$$\mathrm{d}N_C/\mathrm{d}t = r_a N N_* - (r_d + k_{d1} + k_{d2} + k_{d3}) N_C, \tag{26}$$

where N_C is the mRNA–ncRNA complex population, r_a and r_d are the association and dissociation rate constants, and k_{d1}, and k_{d2}, k_{d3} are the degradation rate constants corresponding to channels (21)–(23) (the other designations are as in Eqs. (11) and (12)).

If steps (20) are rapid or the system is at or close to the steady state, Eq. (26) can be solved in the steady-state approximation, i.e., the mRNA–ncRNA complex population can be represented as

$$N_C = r_a N N_* / (r_d + k_{d1} + k_{d2} + k_{d3}). \tag{27}$$

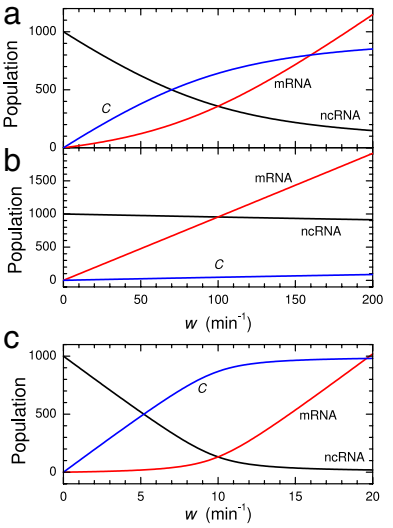


Fig. 2. mRNA and ncRNA mRNA–ncRNA complex (C) populations as a function of the mRNA synthesis rate according to Eqs. (24)–(26) in the case of negligible degradation via channels (22) and (23): (a) $u = 100\ \text{min}^{-1}$, $k = k_* = r_d = k_{d1} = 0.1\ \text{min}^{-1}$, and $r_a = 10^{-3}\ \text{min}^{-1}$; (b) as (a) for $r = 10^{-5}\ \text{min}^{-1}$; (c) $u = 10\ \text{min}^{-1}$, $k = k_* = r_d = k_{d1} = 0.01\ \text{min}^{-1}$, and $r_a = 10^{-3}\ \text{min}^{-1}$. Note that in case (a) the degradation rate constants are relatively high and the mRNA threshold is not manifested. In case (b), the mRNA–ncRNA association is slow and the effect of ncRNA on mRNA is nearly negligible. In case (c), the degradation rate constants are relatively low and the mRNA threshold is well manifested.

Using this expression, one can rewrite Eqs. (24) and (25) as

$$\text{d}N/\text{d}t = w - kN - rNN_*, \tag{28}$$

$$\text{d}N_*/\text{d}t = u - k_*N_* - r_*NN_*, \tag{29}$$

where r and r_* are the effective rate constants given by

$$r = r_a(k_{d1} + k_{d2})/(r_d + k_{d1} + k_{d2} + k_{d3}), \tag{30}$$

$$r_* = r_a(k_{d1} + k_{d3})/(r_d + k_{d1} + k_{d2} + k_{d3}). \tag{31}$$

If the degradation channels (22) and (23) are negligible, i.e., $k_{d2} = k_{d3} = 0$, expressions (30) and (31) are reduced to

$$r = r_* = r_a k_{d1}/(r_d + k_{d1}). \tag{32}$$

In this case, Eqs. (28) and (29) are formally identical to Eqs. (11) and (12).

Typical steady-state kinetics predicted by Eqs. (24)–(26) are shown in Fig. 2.

The scheme described in this section implies that the mRNA–ncRNA complex degradation occurs via irreversible steps (21)–(23). A slightly more complex scheme including reversible substeps of the degradation has been analyzed by Levine et al. [134].

4.3. ncRNA and distinct mRNAs

Many ncRNAs and especially miRNAs have many (up to a few hundreds) mRNA targets [54]. The direct experimental identification and validation and reliable theoretical predictions of the targets still represent challenges in the ncRNA/miRNA biochemistry (see reviews [135] and [136], respectively). The kinetic models including association of ncRNA with two mRNAs and many mRNAs were analyzed in Refs. [127,131] and [126,137], respectively. This subject has also been reviewed in Ref. [124]. All these studies are indicative that ncRNAs can play a role of global regulators.

For M distinct mRNAs, the kinetic scheme and equations presented in Section 4.1 are naturally generalized as

$$\text{Gene}_i \to \text{Gene}_i + \text{mRNA}_i, \tag{33}$$

$$\text{Gene}_* \to \text{Gene}_* + \text{ncRNA}, \tag{34}$$

$$\text{mRNA}_i \to \varnothing, \qquad \text{ncRNA} \to \varnothing, \tag{35}$$

$$\text{mRNA}_i + \text{ncRNA} \to \text{mRNA}_i * \text{ncRNA} \to \varnothing, \tag{36}$$

and

$$\text{d}N_i/\text{d}t = w_i - k_iN - r_iN_iN_*, \tag{37}$$

$$\text{d}N_*/\text{d}t = u - k_*N_* - \sum_i^M r_iN_iN_*, \tag{38}$$

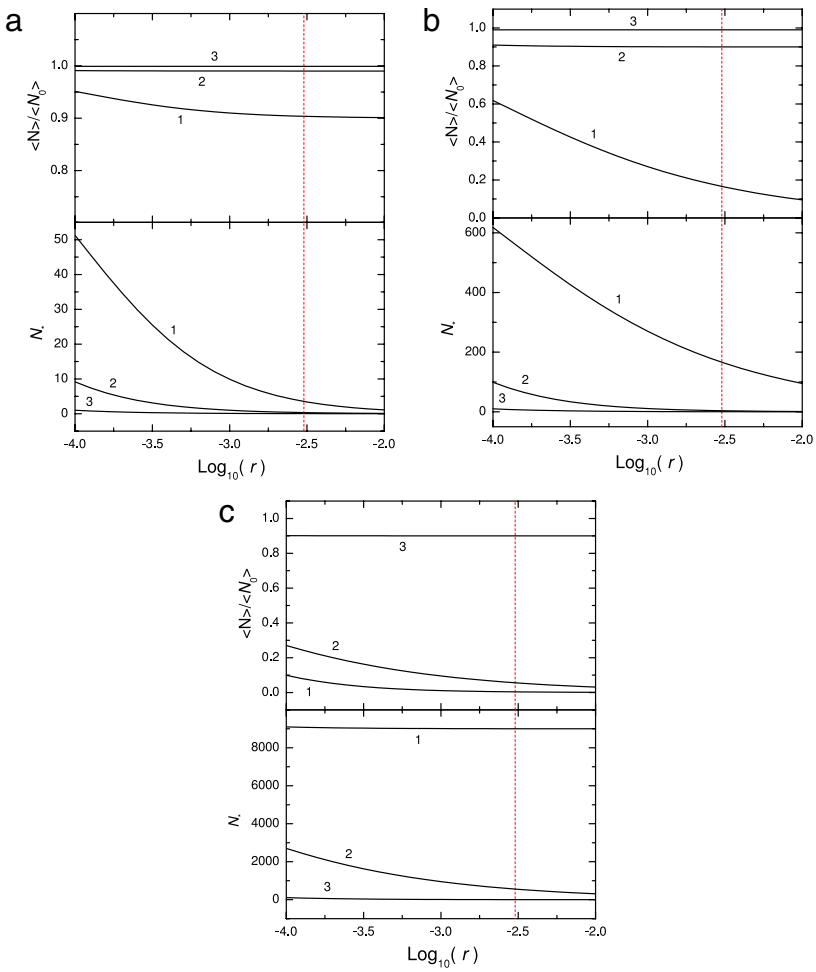


Fig. 3. Average mRNA population and miRNA population as a function of r for $w_* = 10$ (a), 10^2 (b), and 10^3 min^{-1} (c). Curves 1, 2 and 3 correspond, respectively, to the cases when in the absence of the interaction with miRNA the average mRNA population is $\langle N_0\rangle = 10$, 10^2, and 10^3. The dashed line marks $r = 3\times 10^{-3}$ min^{-1} (this value corresponds to the diffusion-limited regime of the miRNA–mRNA interaction). (According to Eqs. (37) and (38) with $r_i = r$; for the parameters, see Ref. [137].)

where N_i and N_* are the mRNA and ncRNA populations, w_i and u are the transcription rates, k_i and k_* are the degradation rate constants, and r_i is the association rate constant.

Typical results of calculations [137] performed for $M = 100$ by using Eqs. (37) and (38) under steady-state conditions are shown in Fig. 3. For such a large number of targets, the decrease in the average population of mRNAs due to interaction with miRNA is found to be appreciable (about 1.5–2-fold) only if the rate of the miRNA synthesis is high. In the absence of the miRNA–mRNA interaction, it should be sufficient to maintain the miRNA population of the order of 10^4 per cell. In addition, the average mRNA population should not be too high (lower than or comparable to 100 for each kind of mRNA). For lower miRNA synthesis rates, the significant influence of miRNA on mRNAs is only possible provided that the average mRNA population is very low (of the order of 10). These conclusions are in line with the results of experiments [138,139] indicating that although a single miRNA can repress the production of hundreds of proteins due to interaction with mRNAs, this repression is typically relatively mild. For smaller M and/or with the presence of protein-mediated regulation of the gene transcription into mRNAs, the effect of ncRNAs on the mRNA population may be larger.

5. Models with regulation of the mRNA synthesis

The models described in the preceding section do not contain the protein-mediated feedbacks. In such situations, the mRNA and ncRNA synthesis rate can be considered to be constant, and one can often operate only with the mRNA and ncRNA populations. With the feedbacks, the dependence of these rates on the protein populations should be described explicitly, i.e., one should operate with the mRNA, ncRNA and protein populations. Following this line, we show in this section the effect of ncRNA on the simplest bistable and oscillatory mRNA–protein networks including feedbacks between the mRNA and protein synthesis. The attention is focused on the role of the mRNA–ncRNA association and degradation. The regulation of

the ncRNA synthesis by proteins is neglected. In this case, the situation is especially simple if the rates of the ncRNA synthesis, u, and degradation, k_*N_*, are rapid compared to the rate of the mRNA–ncRNA association and degradation. Under such circumstances, the ncRNA population is nearly constant, $N_* \simeq u/k_*$, and the effect of ncRNA on the mRNA–protein network is reduced just to the change of the mRNA degradation rate constant, i.e., one can use the effective degradation rate constant, $k_{\rm eff} = k + ru/k_*$, instead of k. This means that basically the bistable and oscillatory mRNA–protein kinetics remain the same as in the absence of ncRNA (the only difference is that they are shifted if k is employed as a governing parameter). In reality, the situation is often more complex and interesting, because the rates the ncRNA synthesis, mRNA–ncRNA association and degradation can be comparable.

5.1. Scenario 1 of bistability

The simplest scheme of gene expression including mRNA, ncRNA and protein (P) is as follows

$$\text{Gene}_1 \to \text{Gene}_1 + \text{mRNA}, \tag{39}$$
$$\text{Gene}_2 \to \text{Gene}_2 + \text{ncRNA}, \tag{40}$$
$$\text{mRNA} \to \text{mRNA} + \text{P}, \tag{41}$$
$$\text{mRNA} \to \varnothing, \tag{42}$$
$$\text{ncRNA} \to \varnothing, \tag{43}$$
$$\text{P} \to \varnothing, \tag{44}$$
$$\text{mRNA} + \text{ncRNA} \to \text{mRNA} * \text{ncRNA} \to \varnothing, \tag{45}$$

where steps (39) and (40), (41), (42)–(44), and (45) represent transcription, translation, degradation, and association, respectively.

In the absence of ncRNA, steps (39), (42) and (44) are well known to predict bistability provided that the mRNA formation is positively regulated by protein associated with two or more regulatory sites [140,141]. In fact, this is positive selfregulation, because protein is formed via mRNA translation. In the literature, this generic scenario of bistability was widely discussed in different contexts including, for example, helper T cell differentiation (for the corresponding model, see Ref. [140]).

With ncRNA, steps (39)–(45) can be described as [142]

$$dN/dt = w_b + w_\circ \left(\frac{n}{K+n}\right)^m - kN - rNN_*, \tag{46}$$
$$dN_*/dt = u - k_*N_* - rNN_*, \tag{47}$$
$$dn/dt = \upsilon N - \kappa n, \tag{48}$$

where N, N_* and n are the mRNA, ncRNA and protein populations, w_b, $w_\circ$, K and m are the parameters describing the protein-regulated mRNA formation (cf. Eq. (1)), and u, υ, k, k_*, κ, and r are the rate constants of the other steps.

Under steady-state conditions (i.e., at $dN/dt = dN_*/dt = dn/dt = 0$), Eqs. (47) and (48) yield $n = \upsilon N/\kappa n$ and $N_* = u/(k_* + rN)$. Substituting these expressions into Eq. (46) results in

$$w_b + w_\circ \left(\frac{N}{\kappa K/\upsilon + N}\right)^m = kN + \frac{ruN}{k_* + rN}. \tag{49}$$

Solving this equation graphically (Fig. 4(a)), one can easily find that at $m \geq 2$ it has either one solution or three solutions. In the former case, the solution is stable. In the latter case, as usual, the lower and upper solutions are stable while the intermediate solution is unstable.

To clarify the physics behind bistability, let us use the mRNA degradation rate constant, k, as a governing parameter. If k is high, the mRNA and protein populations are small. In the absence of ncRNA, these populations increase with decreasing k due to a decrease of the mRNA degradation rate and also due to positive regulation of Gene_1 by protein. These two processes facilitate each other, and if $m \geq 2$ the model predicts a stepwise jump (saddle–node bifurcation) to the state with large mRNA and protein populations. With increasing k, one can observe a stepwise jump back to the state with small mRNA and protein populations. The corresponding kinetics (two stable solutions) exhibit a hysteresis loop inherent to bistability. With ncRNA, the bistability can be observed in a wide range of parameters (see, e.g., Fig. 4(b) and (c)).

Without and with ncRNA, the biochemistry behind the stable states may be appreciably different. Without ncRNA, for example, the mRNA degradation always occurs via step (42). With ncRNA, it occurs via steps (42) and (45). Step (42) may dominate in the case of the steady state with a large mRNA population, while step (45) may dominate in the case of the steady state with a small mRNA population. In the former case, the ncRNA population is relatively small ($N_* \ll u/k_*$), while in the latter case the ncRNA population is large ($N_* \simeq u/k_*$). Thus, basically, the model with ncRNA can describe a ncRNA switch.

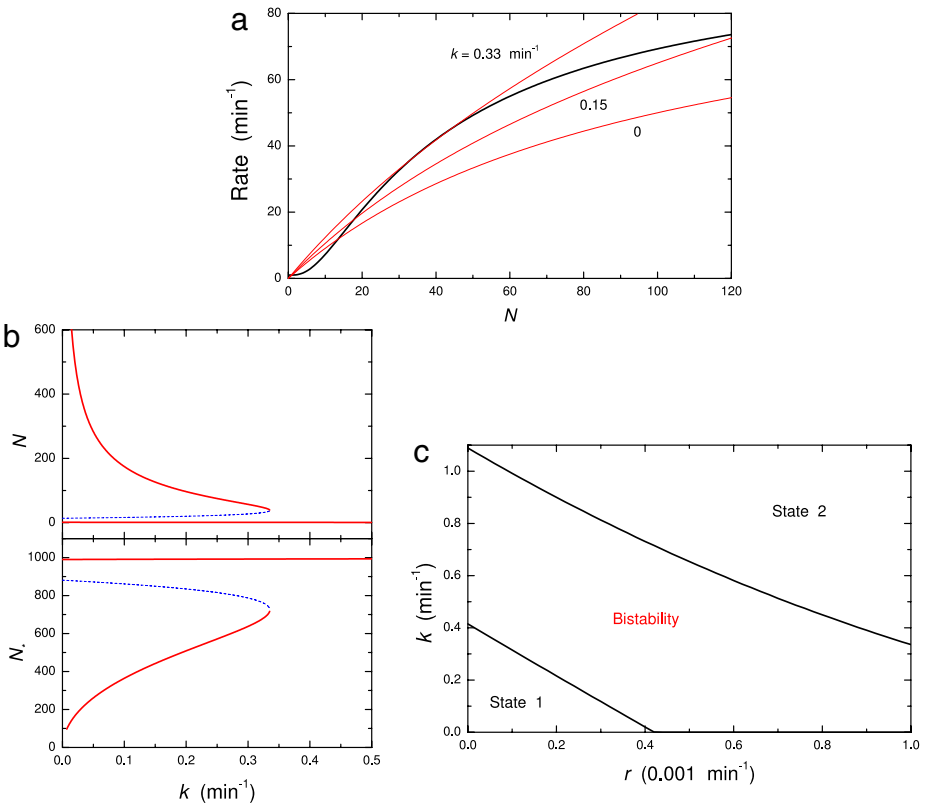


Fig. 4. Analysis of scheme (39)–(45): (a) Left- and right-hand parts of Eq. (49) as a function of N for $m = 4$, $r = 10^{-3}$ min^{-1} and $k = 0, 0.15$ and 0.33 min^{-1}; (b) mRNA and ncRNA populations as a function of k for $r = 10^{-3}$ min^{-1} (the solid and dashed lines correspond respectively to the stable and unstable steady states); (c) kinetic phase diagram in the plane of r and k for the scheme (states 1 and 2 exhibit high and low mRNA populations, respectively). For the other parameters, see Ref. [142].

In the absence of ncRNA, as already noted, steps (39), (42) and (44) predict bistability provided that the mRNA formation is positively regulated by protein associated with two or more regulatory sites. With ncRNA, bistability is possible with one regulatory site [143].

The model described above has been extended [144] to the case when there are mRNAs of M types with one of them positively regulated by protein produced via its translation. For large M (e.g., $M = 100$), the bistability is found to be possible provided that the formation rate of the latter mRNA is sufficiently high. The ncRNA formation rate should be high as well.

5.2. Scenario 2 of bistability

The scheme presented in the preceding subsection includes ncRNA association with mRNA and regulation of the mRNA synthesis by protein. Alternatively, the gene transcription can sometimes be regulated by the ncRNA–protein complex [49]. In the case of positive self-regulation of the mRNA synthesis by this complex, this scheme may result in bistability. Physically, the mechanism of the bistability under consideration is similar to that described in the preceding section. Mathematically, this case is equivalent to the complementary case with self-regulation of the ncRNA synthesis by the ncRNA–protein complex (see Section 6.3 for details).

5.3. Scenario 3 of bistability

One of the best and most important examples of bistability in the mRNA–protein interplay is the so-called toggle switch [145,146]. It includes two protein-coding genes,

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}_1, \tag{50}$$

$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{mRNA}_2, \tag{51}$$

$$\mathrm{mRNA}_1 \to \mathrm{mRNA}_1 + \mathrm{P}_1, \tag{52}$$

$$\mathrm{mRNA}_2 \to \mathrm{mRNA}_2 + \mathrm{P}_2, \tag{53}$$

$$\mathrm{mRNA}_1 \to \varnothing, \qquad \mathrm{mRNA}_2 \to \varnothing, \tag{54}$$

$$P_1 \to \varnothing, \qquad P_2 \to \varnothing, \tag{55}$$

with mutual protein-mediated suppression of their transcription. This scheme is bistable provided that in the absence of suppression the $mRNA_1$ and $mRNA_2$ populations are sufficiently large. With regulation of gene transcription (by proteins associated with two or more regulatory sites), both these populations cannot be large simultaneously, because this is not compatible with suppression of the transcription rates. For this reason, the model predicts two stable states so that either one or another gene is active under steady-state conditions.

The toggle switch was constructed in *E. coli* [145] and is now widely used to describe various intracellular phenomena. For example, this model was employed to interpret induced neutrophil differentiation [147] or as an ingredient of a model of spontaneous cell differentiation [148].

In the presence of ncRNA, we have additional steps

$$\mathrm{Gene}_* \to \mathrm{Gene}_* + \mathrm{ncRNA}, \tag{56}$$

$$\mathrm{ncRNA} \to \varnothing, \tag{57}$$

$$\mathrm{mRNA}_1 + \mathrm{ncRNA} \to \mathrm{mRNA}_1 * \mathrm{ncRNA} \to \varnothing, \tag{58}$$

$$\mathrm{mRNA}_2 + \mathrm{ncRNA} \to \mathrm{mRNA}_2 * \mathrm{ncRNA} \to \varnothing. \tag{59}$$

With these steps, the model has been proved [142] to remain bistable even if the rate constants of the mRNA–ncRNA association are high and the ncRNA population is relatively large. The type of the corresponding solutions of the kinetics equations depends on the relative rates of steps (58) and (59).

For example, let us consider the simplest situation when the genes transcribed into mRNAs are kinetically equivalent and steps (58) and (59) are kinetically equivalent as well. This case is described as

$$\mathrm{d}N_1/\mathrm{d}t = w_\circ \left(\frac{K}{K+n_2}\right)^m - kN_1 - rN_1N_*, \tag{60}$$

$$\mathrm{d}N_2/\mathrm{d}t = w_\circ \left(\frac{K}{K+n_1}\right)^m - kN_2 - rN_2N_*, \tag{61}$$

$$\mathrm{d}N_*/\mathrm{d}t = u - k_*N_{\mathcal{R}} - r(N_1+N_2)N_*, \tag{62}$$

$$\mathrm{d}n_1/\mathrm{d}t = \upsilon N_1 - \kappa n_1, \tag{63}$$

$$\mathrm{d}n_1/\mathrm{d}t = \upsilon N_2 - \kappa n_2, \tag{64}$$

where N_1, N_2, N_*, n_1 and n_2 are the mRNA, ncRNA and protein populations, $w_\circ$, K and m are the parameters describing the protein-regulated mRNA formation (cf. Eq. (2)), and u, υ, k, k_*, κ, and r are the rate constants of the other steps.

Figs. 5 and 6 constructed by using Eqs. (60)–(64) show that the bistability can be observed provided that the mRNA degradation rate constant is relatively low even if the mRNA–ncRNA association rate constant is close to that corresponding to the diffusion-limited regime.

Ref. [142] contains also other examples showing what happens when steps (58) and (59) are kinetically not equivalent.

5.4. Kinetic oscillations

The generic model predicting oscillations in the mRNA–protein interplay includes a gene with negative regulation of the mRNA production by protein formed via mRNA translation and a few steps of conversion [149–153] (for additional relevant references, see Refs. [41,42]). In combination with the ncRNA formation and degradation and mRNA–ncRNA association and degradation, the simplest scheme of this category is as follows [142] (for slightly different schemes, see Refs. [154,155])

$$\mathrm{Gene} \to \mathrm{Gene} + \mathrm{mRMA}, \tag{65}$$

$$\mathrm{mRNA} \to \mathrm{mRRNA} + P_1, \tag{66}$$

$$P_1 \to P_2, \tag{67}$$

$$P_2 \to P_3, \tag{68}$$

$$\mathrm{mRNA} \to \varnothing, \tag{69}$$

$$P_3 \to \varnothing, \tag{70}$$

$$\mathrm{Gene}_* \to \mathrm{Gene}_* + \mathrm{ncRNA}, \tag{71}$$

$$\mathrm{ncRNA} \to \varnothing, \tag{72}$$

$$\mathrm{mRNA} + \mathrm{ncRNA} \to \mathrm{mRNA} * \mathrm{ncRNA} \to \varnothing. \tag{73}$$

These steps lead directly to the kinetic equations for the mRNA, P_1, P_3, P_3 and ncRNA populations,

$$\mathrm{d}N/\mathrm{d}t = w_\circ \left(\frac{K}{K+n_3}\right)^m - kN - rNN_*, \tag{74}$$

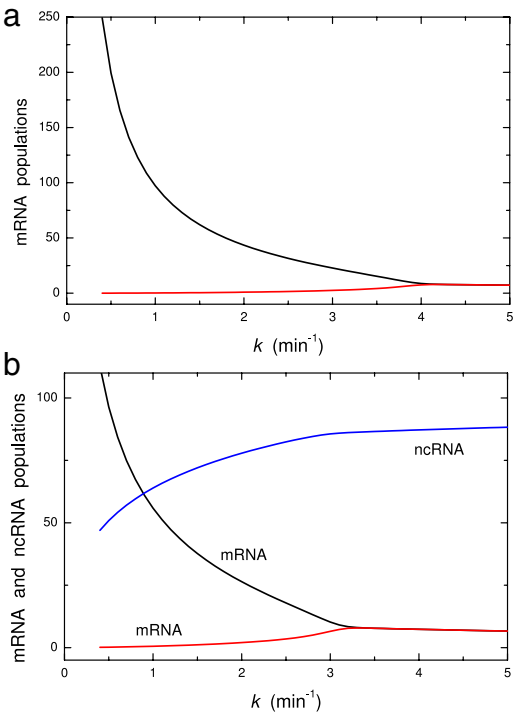


Fig. 5. mRNA and ncRNA populations as a function of k according to Eqs. (60)–(64) with $r = 0$ (a) and 10^{-3} min^{-1} (b) (for the other parameters, see Ref. [142]).

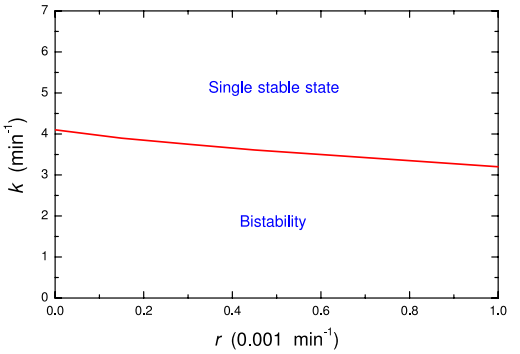


Fig. 6. Kinetic phase diagram predicted by Eqs. (60)–(64). The parameters are the same as in the case of Fig. 5(b).

$$dn_1/dt = \upsilon N - \kappa_{12} n_1, \tag{75}$$

$$dn_2/dt = \kappa_{12} n_1 - \kappa_{23} n_2, \tag{76}$$

$$dn_3/dt = \kappa_{23} n_2 - \kappa n_3, \tag{77}$$

$$dN_*/dt = u - k_* N_* - rNN_*, \tag{78}$$

where κ_{12} and κ_{23} are the rate constants of protein conversion (67) and (68), and the other parameters are as in the preceding subsections.

Oscillations in the mRNA–protein interplay (steps (65)–(70)) are related to the delay (steps (67) and (68)) between the protein synthesis (via steps (65) and (66)) and the ability of protein to regulate these steps. In particular, the phase with relatively large mRNA population is accompanied by growth of the P_1 population. With delay, the P_3 population grows as well. With negative regulation of the gene transcription, the latter suppresses the rate of the mRNA formation, and the mRNA population starts to fall. The decrease of the mRNA population results in a decrease of the P_1 population. With delay, the P_3 population decreases as well. Then, the mRNA population starts to grow and so on.

The extensive synthesis of ncRNA may obviously suppress these oscillations. On the other hand, the analysis [142] indicates that the role of ncRNA in kinetic oscillations may also be positive as, for example, shown in Figs. 7 and 8. In particular, the oscillatory window may become appreciably wider (Fig. 8).

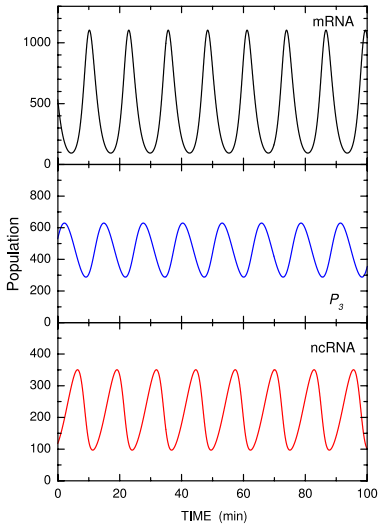


Fig. 7. mRNA, ncRNA and protein populations as a function of time according to Eqs. (74)–(78) (for the parameters, see Ref. [142]).

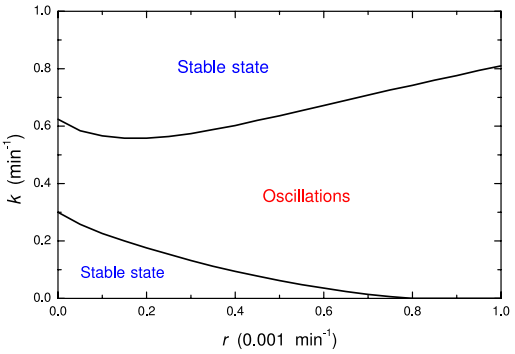


Fig. 8. Kinetic phase diagram predicted by Eqs. (74)–(78) (for the parameters, see Ref. [142]).

A more complex oscillatory model, including two genes transcribed into mRNAs, one gene transcribed into ncRNA, and mRNA–ncRNA association and degradation, has been analyzed by Nandi et al. [156].

6. Models with regulation of the ncRNA synthesis

In this section, we present the simplest models describing the mRNA–ncRNA–protein interplay with regulation of the ncRNA synthesis by protein. As one could expect, the gene expression kinetics with this regulation can be bistable or oscillatory.

6.1. Scenario 1 of bistability

As already noticed in Section 5.1, the simplest scheme of gene expression including mRNA, ncRNA and protein is as follows

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}, \tag{79}$$

$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{ncRNA}, \tag{80}$$

$$\mathrm{mRNA} \to \mathrm{mRNA} + \mathrm{P}, \tag{81}$$

$$\mathrm{mRNA} \to \varnothing, \qquad \mathrm{ncRNA} \to \varnothing, \qquad \mathrm{P} \to \varnothing, \tag{82}$$

$$\mathrm{mRNA} + \mathrm{ncRNA} \to \mathrm{mRNA} * \mathrm{ncRNA} \to \varnothing. \tag{83}$$

With regulation of step (80) by protein, this scheme is bistable provided that the regulation is negative (with $m \geq 2$). The corresponding kinetic equations can be represented as [157,158]

$$\mathrm{d}N/\mathrm{d}t = w - kN - rNN_*, \tag{84}$$

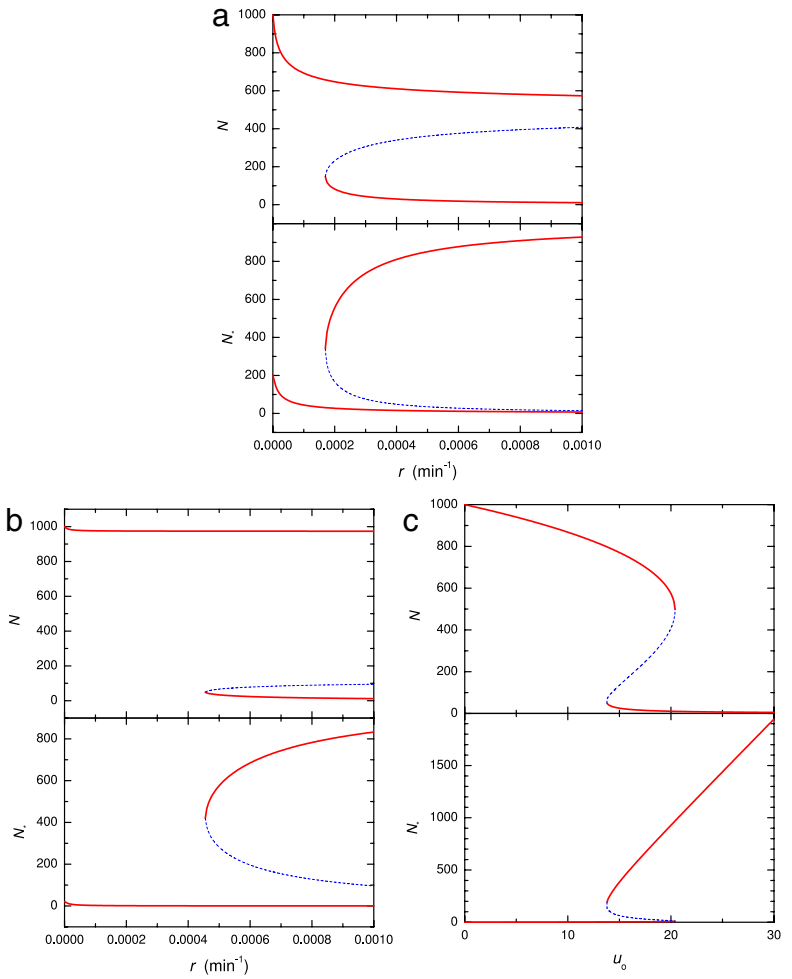


Fig. 9. [(a) and (b)] mRNA and ncRNA populations as a function of r according to Eqs. (84)–(86) with $w = 10\ \mathrm{min}^{-1}$, $u_\circ = 20\ \mathrm{min}^{-1}$, $\upsilon = k = k_* = \kappa = 0.01\ \mathrm{min}^{-1}$, $K = 500$, and $m = 2$ (a) and 4 (b). (c) The same populations as a function $u_\circ$ for $m = 2$ and $r = 10^{-3}\ \mathrm{min}^{-1}$. The solid and dashed lines correspond respectively to the stable and unstable steady states. (Note that with these parameters the protein population is equal to the mRNA population.)

$$\mathrm{d}N_*/\mathrm{d}t = u_\circ \left(\frac{K}{K+n}\right)^m - k_* N_* - rNN_*, \tag{85}$$

$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n, \tag{86}$$

where N, N_* and n are the mRNA, ncRNA and protein populations, $u_\circ$, K and m are the parameters describing the protein-regulated ncRNA formation (cf. Eq. (1)), and w, υ, k, k_*, κ, and r are the rate constants of the other steps.

A typical example of bistable steady-state kinetics is shown in Fig. 9(a) and (b) where the mRNA–ncRNA association rate constant, r, is chosen as a governing parameter. If r is low, the mRNA and protein populations are large, the transcription of Gene_2 to ncRNA is suppressed by protein, and accordingly the ncRNA population is small. This state is also stable at high r. In the latter case, there is, however, another stable steady state with large ncRNA population and small mRNA and protein populations. In this state, the mRNA population is suppressed by ncRNA via step (83).

Switches similar to those exhibited in Fig. 9(a) and (b) can be observed by using other governing parameters parameters. For example, the rate of the ncRNA formation can be regulated by signals (e.g., by other proteins). This regulation can be described by changing $u_\circ$. With increasing $u_\circ$, the model predicts (Fig. 9(c)) a switch from the state with low ncRNA and high mRNA populations to the state with high ncRNA and low mRNA populations.

The model presented above has been extended [159] to the case when there are mRNAs of M types with one of them translated into protein regulating the ncRNA synthesis. With physically reasonable model parameters, the model predicts bistability or, in other words, switch in the expression of hundreds of genes. The model can also be easily extended to the situation when P regulates many genes.

The two key ingredients of this model, ncRNA–mRNA association and negative regulation of the ncRNA synthesis by protein formed via mRNA translation, are common. For this reason, the model is expected to be applicable in various

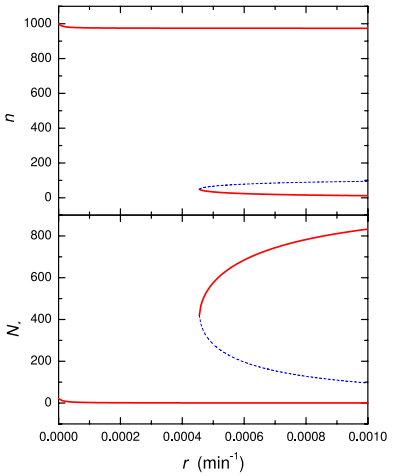


Fig. 10. Protein and ncRNA populations as a function of r according to Eqs. (88)–(90) with $w = 10\,\mathrm{min}^{-1}$, $u_\circ = 20\,\mathrm{min}^{-1}$, $\upsilon = k = k_* = \kappa = 0.01\,\mathrm{min}^{-1}$, $K = 500$, and $m = 4$. The solid and dashed lines correspond respectively to the stable and unstable steady states. (Note that in this case the mRNA population is constant, $N = w/k = 1000$.)

situations. For example, the best studied miRNA in neuronal development, *miR-124*, negatively downregulates REST, a transcriptional repressor of numerous neuronal genes including *miR-124* itself (reviewed in Ref. [67]). The interplay of these processes may result in switch in the expression of many genes. Interestingly, *miR-124* also represses the activity of BTBP and Sox9. All these proteins are known to antagonize the formation of neuronal cells during development. Thus, *miR-124* acts as a neuronal differentiation-inducing factor, and the model described shows (see, e.g., Fig. 9(c)) how this miRNA can operate. Recent experiments indicate that miRNAs may play an important role also in differentiation of other cells (see, e.g., reviews [61–65]).

The other available kinetic models of deterministic (signal-induced) cell differentiation are usually based on bistable or multistable mRNA–protein modules. The corresponding studies are focused on embryonic stem cells [160,161], bone marrow-derived progenitor cells [162–164], blood cell precursors (Refs. [140,165–169]), neural progenitor cells [170], and induced pluripotency [171] (for general analysis of multistable switches, see Ref. [172]; for spontaneous differentiation, see Ref. [148]).

6.2. Scenario 2 of bistability

ncRNAs may associate with proteins. The simplest scheme of this category is obtained by replacing step (83) in scheme (79)–(83) by

$$\mathrm{P} + \mathrm{ncRNA} \to \mathrm{P} * \mathrm{ncRNA} \to \varnothing. \tag{87}$$

With this modification, Eqs. (84)–(86) can be rewritten as

$$\mathrm{d}N/\mathrm{d}t = w - kN, \tag{88}$$

$$\mathrm{d}N_*/\mathrm{d}t = u_\circ \left(\frac{K}{K+n}\right)^m - k_* N_* - rnN_*, \tag{89}$$

$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n - rnN_*. \tag{90}$$

With suitable kinetic parameters, this model predicts bistability (see, e.g., Fig. 10) provided that $m \geq 2$ [157]. One of the stable regimes exhibits high protein population. This population suppresses the ncRNA formation due to negative regulation of this step, and accordingly the ncRNA population is low. Another stable regime exhibits high ncRNA population and low protein population. The latter population is suppressed via step (87).

6.3. Scenario 3 of bistability

The scheme presented in the preceding subsection includes ncRNA association with protein and regulation of the ncRNA synthesis by protein. Alternatively, the gene transcription can sometimes be regulated by the ncRNA–protein complex. In particular, Feng et al. [173] identified a generic mechanism whereby the transcription is controlled by the cooperative action of an ncRNA and a protein. They have shown in detail that in mouse brain cells the ncRNA *Evf2*, transcribed from the region between the Dlx-5 and Dlx-6 protein-coding genes, pairs with the Dlx-2 protein, and then this complex binds the same

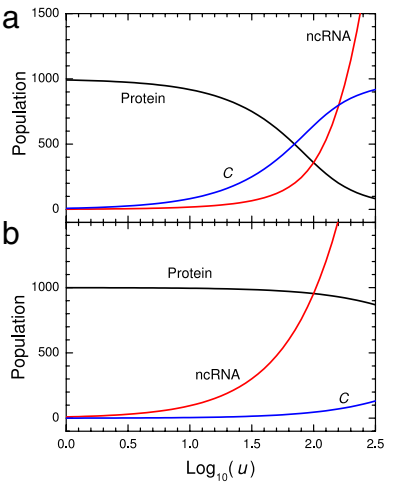


Fig. 11. Protein, ncRNA and ncRNA–protein-complex populations as a function of the ncRNA synthesis rate according to Eqs. (101)–(103) for $w_P = 100\ \mathrm{min}^{-1}$, $k_* = \kappa = k_C = r_d = 0.1\ \mathrm{min}^{-1}$, and $r_a = 10^{-3}$ (a) and $10^{-5}\ \mathrm{min}^{-1}$ (b). The mRNA population (not shown) is constant in this case (see Eq. (101)).

region and enhances expression of the *Dlx-5* and *Dlx-6* genes. Now, this is believed to be a general strategy for regulating the expression of key developmental genes [49].

The simplest scheme with the formation of the ncRNA–protein complex (C) is as follows

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}, \tag{91}$$
$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{ncRNA}, \tag{92}$$
$$\mathrm{mRNA} \to \mathrm{mRNA} + \mathrm{P}, \tag{93}$$
$$\mathrm{mRNA} \to \varnothing, \qquad \mathrm{ncRNA} \to \varnothing, \qquad \mathrm{P} \to \varnothing, \tag{94}$$
$$\mathrm{mRNA} + \mathrm{P} \rightleftharpoons \mathrm{C}, \tag{95}$$
$$\mathrm{C} \to \varnothing. \tag{96}$$

This formulation leads to the equations

$$\mathrm{d}N/\mathrm{d}t = w - kN, \tag{97}$$
$$\mathrm{d}N_*/\mathrm{d}t = u - k_*N_* - r_a n N_* + r_d N_C, \tag{98}$$
$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n - r_a n N_* + r_d N_C, \tag{99}$$
$$\mathrm{d}N_C/\mathrm{d}t = r_a n N_* - r_d N_C - k_C N_C, \tag{100}$$

where N, N_*, n and N_C are the mRNA, ncRNA, protein and C populations, w and u are the transcription rates, υ is the translation rate constant, r_a and r_d are the association and dissociation rate constants, and k, k_*, κ and k_C are the degradation rate constants.

Under steady-state conditions, Eqs. (97) and (100) yield

$$N = w/k \quad \text{and} \quad N_C = r_a n N_*/(r_d + k_C). \tag{101}$$

Substituting these expressions into Eqs. (98) and (99), we obtain

$$u = k_* N_* + r_{\mathrm{eff}} n N_*, \tag{102}$$
$$w_P = \kappa n + r_{\mathrm{eff}} n N_*, \tag{103}$$

where $w_P \equiv \upsilon w/k$ is the protein-synthesis rate, and $r_{\mathrm{eff}} \equiv r_a k_C/(r_d + k_C)$ is the effective rate constant of the protein degradation via steps (95) and (96).

Mathematically, Eqs. (102) and (103) are identical to Eqs. (11) and (12) and can be elementarily solved (cf. Eq. (13)), i.e., one can calculate N_* and n. Then, N and N_C can be calculated by using expressions (101). For example, we show (Fig. 11) a typical dependence of N_*, n and N_C on the ncRNA synthesis rate.

In our analysis above, the rates of the mRNA and ncRNA synthesis were considered to be independent of the ncRNA–protein-complex concentration. This complex may, however, attach to DNA and regulate one of these rates. The likely manifestations of these two feedbacks are similar, because the protein synthesis rate is proportional to the rate of the mRNA synthesis and the structures of Eqs. (102) and (103) for the ncRNA and protein populations are similar. For this reason, it is sufficient to scrutinize the manifestation of only one of them. To be specific, we assume that the rate of the ncRNA

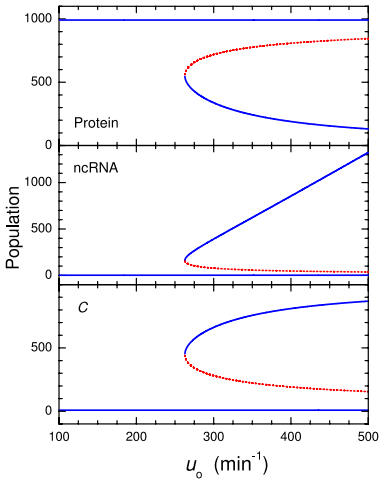


Fig. 12. Protein, ncRNA and ncRNA–protein-complex populations as a function of $u_\circ$ according to Eqs. (101)–(104) with $u_b = 1$ min^{-1}, $K = 200$ and $m = 4$ (the other parameters are as in Fig. 11(a)). Stable and unstable steady-state kinetics are indicated by solid and dashed lines, respectively.

synthesis depends on the ncRNA–protein-complex population. If this regulation is positive, the kinetics may be bistable. To illustrate bistability, the ncRNA-synthesis rate can be represented as

$$u = u_b + u_\circ \left(\frac{N_C}{K + N_C} \right)^m , \tag{104}$$

where u_b, $u_\circ$, K and m are the corresponding parameters (cf. Eq. (1)). With this expression for u and expression (101) for N_C, Eqs. (102) and (103) can be solved numerically. A typical dependence of N_*, n and N_C on $u_\circ$ is shown in Fig. 12.

6.4. Scenario 4 of bistability

In Section 5.3, we have described the conventional toggle switch including two protein-coding genes with mutual protein-mediated suppression of transcription. The experiments (Ref. [174] and review [175]) indicate that a similar switch may function with the regulation of the ncRNA synthesis by proteins. Specifically, such a switch seems to operate in nematode chemosensory neuronal cells to assign their fate. In this system, the binary outcome is specified by the interplay of the transcription factors DIE-1 and COG-1 and two miRNAs, *lsy-6* and one of the *miR-273* family [174,175]. Specifically, the miRNA *lsy-6* is activated downstream of DIE-1, and it represses *cog-1*. In turn, miRNAs of the *miR-273* family are activated downstream of COG-1, and they repress *die-1*. Although the types of the mutual regulations of these proteins and miRNAs have been firmly identified by using various experiments, the mechanistic details of the regulation and bistability are open for discussion (see Ref. [174]). The simplest interpretation of the results can be based on two biologically reasonable conjectures. (i) The fact that the miRNAs *lsy-6* and *miR-273* are activated, respectively, by the transcription factors DIE-1 and COG-1 seems to indicate that these proteins directly positively regulate the synthesis of these miRNAs. (ii) The suppression of the synthesis of the transcription factors COG-1 and DIE-1, respectively, by the miRNAs *lsy-6* and *miR-273* can be explained by association of these miRNAs with the mRNAs responsible for the synthesis of DIE-1 and COG-1.

The scheme corresponding to the scenario outlined above includes transcription of two genes into mRNAs,

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}_1, \tag{105}$$

$$\mathrm{Gene}_2 \to \mathrm{Gene}_2 + \mathrm{mRNA}_2, \tag{106}$$

translation of these mRNAs to proteins (P_1 and P_2),

$$\mathrm{mRNA}_1 \to \mathrm{mRNA}_1 + \mathrm{P}_1, \tag{107}$$

$$\mathrm{mRNA}_2 \to \mathrm{mRNA}_2 + \mathrm{P}_2, \tag{108}$$

transcription of two other genes into ncRNAs,

$$\mathrm{Gene}_1^* \to \mathrm{Gene}_1^* + \mathrm{ncRNA}_1, \tag{109}$$

$$\mathrm{Gene}_2^* \to \mathrm{Gene}_2^* + \mathrm{ncRNA}_2, \tag{110}$$

conventional degradation of these species,

$$\mathrm{mRNA}_1 \to \varnothing, \qquad \mathrm{mRNA}_2 \to \varnothing, \tag{111}$$

$$\mathrm{P}_1 \to \varnothing, \qquad \mathrm{P}_2 \to \varnothing, \tag{112}$$

$$\mathrm{ncRNA}_1 \to \varnothing, \qquad \mathrm{ncRNA}_2 \to \varnothing, \tag{113}$$

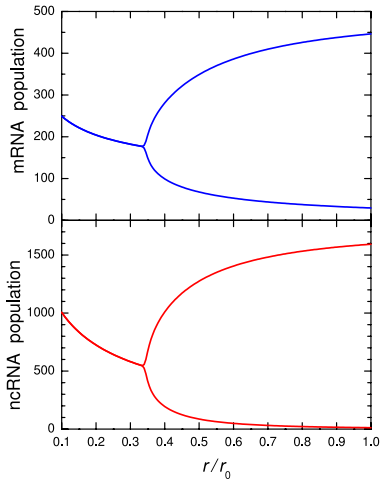


Fig. 13. mRNA and ncRNA populations as a function of r/r_0 ($r_0 \equiv 10^{-3}$ min^{-1}) according to Eqs. (116)–(121) with $w = 50$ min^{-1}, $\upsilon = 0.2$ min^{-1}, $u_b = 1$ min^{-1}, $u_\circ = 500$ min^{-1}, $K = 500$, $m = 2$ and $k = k_* = \kappa = 0.1$ min^{-1}.

and association and subsequent degradation of mRNA and ncRNA,

$$\mathrm{mRNA}_1 + \mathrm{ncRNA}_2 \to \mathrm{mRNA}_1 * \mathrm{ncRNA}_2 \to \varnothing. \tag{114}$$

$$\mathrm{mRNA}_2 + \mathrm{ncRNA}_1 \to \mathrm{mRNA}_2 * \mathrm{ncRNA}_1 \to \varnothing. \tag{115}$$

This scheme can be bistable provided steps (109) and (110) are positively regulated by P_1 and P_2, respectively.

To reduce the number of model parameters, let us consider that Gene_1 and Gene_2 are kinetically equivalent, and Gene_1^* and Gene_2^* are kinetically equivalent as well. In this case, the scheme above is described as

$$\mathrm{d}N_1/\mathrm{d}t = w - kN_1 - rN_1N_2^*, \tag{116}$$

$$\mathrm{d}N_2/\mathrm{d}t = w - kN_2 - rN_2N_1^*, \tag{117}$$

$$\mathrm{d}n_1/\mathrm{d}t = \upsilon N_1 - \kappa n_1, \tag{118}$$

$$\mathrm{d}n_2/\mathrm{d}t = \upsilon N_2 - \kappa n_2, \tag{119}$$

$$\mathrm{d}N_1^*/\mathrm{d}t = u_b + u_\circ \left(\frac{n_1}{K + n_1}\right)^m - k_*N_1^* - rN_2N_1^*, \tag{120}$$

$$\mathrm{d}N_2^*/\mathrm{d}t = u_b + u_\circ \left(\frac{n_2}{K + n_2}\right)^m - k_*N_2^* - rN_1N_2^*, \tag{121}$$

where N_1, N_2, N_1^*, N_2^*, n_1, and n_2 are the mRNA, ncRNA and protein populations in a cell, w is the rate of the mRNA synthesis, u_b, $u_\circ$, K and m are the parameters for the protein-regulated rates of the ncRNA synthesis (cf. Eq. (1)), υ is the rate constant of the protein formation, k, k_* and κ are the degradation rate constants, and r is the rate constant of steps (114) and (115).

According to Eqs. (116)–(121), the bistability is predicted provided that $m \geq 1$ and r is sufficiently high (Figs. 13 and 14). If for example N_1 is large under steady-state conditions, N_2^* will be small (see Eq. (116)) while n_1 and N_1^* will be large (Eqs. (118) and (120)). Large N_1^* will in turn result in small N_2 (Eq. (117)), and the latter will result in small n_2 (Eq. (119)). The other solution yields, alternatively, small N_1, n_1 and N_1^* and large N_2, n_2 and N_2^*.

6.5. Scenario 1 of kinetic oscillations

Oscillations in the mRNA–protein interplay can be related to a few steps of protein conversion resulting in a delay between the protein synthesis and the protein ability to regulate the mRNA formation (Section 5.3). In the case of regulation of the ncRNA formation, oscillations may appear if the reaction scheme includes a few steps of ncRNA conversion. In reality, such steps are common. For example, miRNAs are transcribed as long ncRNA and then generated via a processing pathway including first the formation of pre-miRNAs and then conversion of them into the corresponding miRNA. The simplest scheme of this category is as follows [176]

$$\mathrm{Gene} \to \mathrm{Gene} + \mathrm{mRNA}, \tag{122}$$

$$\mathrm{mRNA} \to \mathrm{mRNA} + \mathrm{P}, \tag{123}$$

$$\mathrm{Gene}_* \to \mathrm{Gene}_* + \mathrm{ncRNA}_1, \tag{124}$$

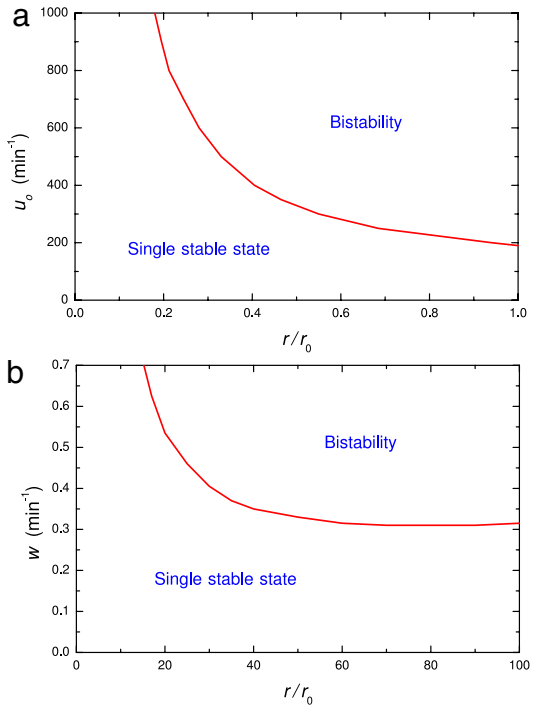


Fig. 14. Kinetic phase diagrams in the plane of (a) r and $u_\circ$ and (b) r and w for the model defined by Eqs. (116)–(121). The parameters are as in Fig. 13.

$$\mathrm{ncRNA}_1 \to \mathrm{ncRNA}_2, \tag{125}$$

$$\mathrm{ncRNA}_2 \to \mathrm{ncRNA}_3, \tag{126}$$

$$\mathrm{mRNA} \to \varnothing, \qquad \mathrm{P} \to \varnothing, \qquad \mathrm{ncRNA}_3 \to \varnothing, \tag{127}$$

$$\mathrm{mRNA} + \mathrm{ncRNA}_3 \to \mathrm{mRNA} * \mathrm{ncRNA}_3 \to \varnothing. \tag{128}$$

This scheme predicts oscillations (Fig. 15) provided that the regulation of the ncRNA synthesis by protein (P) is positive. The corresponding kinetic equations for the mRNA, ncRNA and protein populations, N, N_1^*, N_2^*, N_3^* and n, are read as

$$\mathrm{d}N/\mathrm{d}t = w - kN - rNN_3^*, \tag{129}$$

$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n, \tag{130}$$

$$\mathrm{d}N_1^*/\mathrm{d}t = u_\circ \left(\frac{n}{K+n}\right)^m - k_{12}N_1^*, \tag{131}$$

$$\mathrm{d}N_2^*/\mathrm{d}t = k_{12}N_1^* - k_{23}N_2^*, \tag{132}$$

$$\mathrm{d}N_3^*/\mathrm{d}t = k_{23}N_2^* - k_*N_3^* - rNN_3^*, \tag{133}$$

where k_{12} and k_{23} are the rate constants of the ncRNA conversion, and the other parameters are as in the preceding subsections.

To elucidate the mechanism of oscillations, let us start from the situation when the protein population is high. In this case, the protein-regulated rate of the ncRNA_1 formation and the ncRNA_1 population are high as well. The latter results in the growth of the ncRNA_2 and ncRNA_3 populations. After a while, the ncRNA_3 population becomes relatively high and accordingly suppresses the mRNA population via step (128). In turn, this results in the decrease of the protein population and accordingly the rate of the ncRNA_1 formation and the ncRNA_1 population decrease as well. After a while, the drop in the ncRNA_1 population causes a decrease of the ncRNA_2 and ncRNA_3 populations. With decreasing ncRNA_3 population, step (128) becomes less important, and accordingly the mRNA population starts to grow and reaches a maximum. The protein population also start to grow and again reaches a maximum.

6.6. Scenario 2 of kinetic oscillations

The model described in the preceding subsection includes the ncRNA association with mRNA (step (128)). ncRNA may, however, associate with protein (P) as well, i.e., step (128) can be replaced by

$$\mathrm{P} + \mathrm{ncRNA}_3 \to \mathrm{P} * \mathrm{ncRNA}_3 \to \varnothing. \tag{134}$$

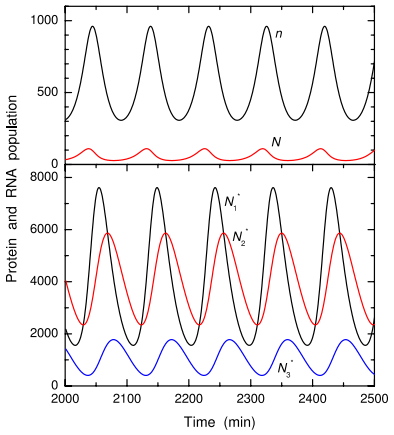


Fig. 15. Kinetic oscillations predicted by the model defined by Eqs. (129)–(133) (for the parameters, see Ref. [176]).

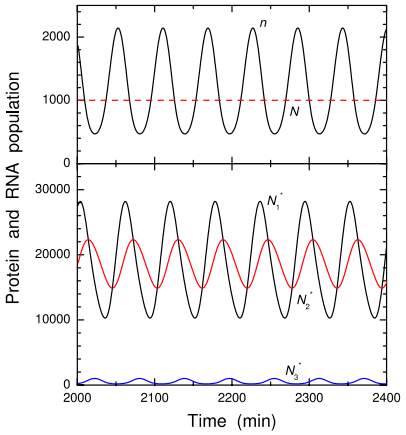


Fig. 16. Kinetic oscillations predicted by the model defined by Eqs. (135)–(139) (for the parameters, see Ref. [176]).

With this step, Eqs. (129)–(133) are modified as

$$dN/dt = w - kN, \tag{135}$$

$$dn/dt = \upsilon N - \kappa n - rnN_3^*, \tag{136}$$

$$dN_1^*/dt = u_\circ \left(\frac{n}{K+n}\right)^m - k_{12}N_1^*, \tag{137}$$

$$dN_2^*/dt = k_{12}N_1^* - k_{23}N_2^*, \tag{138}$$

$$dN_3^*/dt = k_{23}N_2^* - k_*N_3^* - rnN_3^*. \tag{139}$$

This model also predicts oscillations as, e.g., shown in Fig. 16. The mechanism of these oscillations can easily be elucidated in analogy with the reasoning presented in the preceding section.

7. Complex genetic networks including ncRNAs

7.1. General aspects

In Sections 4–6, we have discussed simple genetic networks including only a few genes or many genes but with simple regulation. In reality, the genetic networks often include many genes with complex regulation. The understanding of such networks is now limited. This state of the art is typical for many areas of contemporary cell biology, where the gaps in the understanding are often at the level of systems. The available experimental and theoretical studies of complex genetic networks are focused primarily on the mRNA–protein interplay (see, e.g., reviews [43–46]). Via association with mRNAs and proteins, ncRNAs may play a role of a global regulator in complex mRNA–protein networks. This function of ncRNAs was experimentally tracked in many systems both in the normal state and during diseases (see the introduction; Refs. [49–56]).

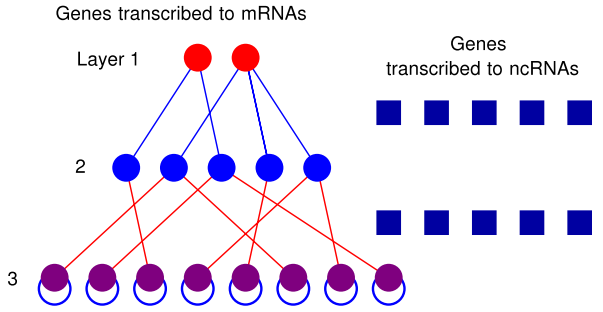


Fig. 17. Network including 15 genes (filled circles) transcribed into mRNAs and 10 genes (squares) transcribed into ncRNAs. The former genes form three layers. Via the mRNA–protein links, the upper layer governs the medium layer which in turn controls the bottom layer. The negative self-regulation of the genes belonging to the bottom layer is shown by loops. The genes transcribed into mRNAs interact (not shown) with the genes transcribed into ncRNAs via the mRNA–ncRNA association and degradation.

The detailed studies of the corresponding kinetics are, however, still rare. Theoretical studies of the kinetics of complex genetic networks including ncRNAs are also just beginning [177–182].

Although complex mRNA–protein networks have long attracted attention, their detailed classification is still lacking. Qualitatively, one can distinguish here two globally different types of regulation and many more specific situations somewhere between. On one side, there are networks with hierarchical (or, in other words, "autocratic") regulation of downstream genes by upstream genes. Such networks are typically organized in the shape of a pyramid, with each row of genes linked to genes directly beneath it. On the other side, there are distributed (or, in other words, "democratic") networks relying on mutual regulation of genes. The former networks have been studied experimentally in more detail because their structure and function are often more evident. The latter networks gained empirical support only recently, partly because their characterization involves studies of genome-wide dynamical processes [45].

As already noted, one of the key functions of ncRNAs is global regulation of complex mRNA–protein networks. Below, we first show the effect of such regulation on the mRNA–protein networks with the hierarchical and distributed architectures (Sections 7.2 and 7.3; Refs. [180–182]). Then, we outline two more specific models focused on regulation of a cancer network (Section 7.4; Ref. [177]) and post-transcriptional regulation activity of small ncRNAs in *E. coli* (Section 7.5; Ref. [179]). In addition, we briefly discuss the likely role of ncRNAs in the global regulation of the mRNA and protein populations during the growth of eukaryotic cells (Section 7.6).

7.2. Hierarchical genetic networks

Hierarchical mRNA–protein networks often have a layered architecture or include a few subgroups of genes with this architecture [43]. The layered structure is typical, for example, for the transcriptional regulatory networks both in prokaryotes and eukaryotes (e.g., in *E. coli* [183,184] and *C. elegans* [185,186], respectively). In the corresponding networks, the subgroups of genes forming three- or two-layer clusters are numerous while the four- or five-layers clusters are less frequent. As a rule, the upper layers contain a smaller number of genes compared to the lower layers.

The effect of ncRNA on a two-layer mRNA–protein network has recently been analyzed in Ref. [180]. As a more general example, let us consider the interaction of ncRNAs with a three-layer mRNA–protein network shown in Fig. 17 [181]. In this case, the upper layer containing two genes governs the medium layer containing five genes, and the latter layer controls in turn the bottom layer containing eight genes. The genes forming the bottom layer are negatively self-regulated (this feature is often observed in real networks). In addition, there are ten genes transcribed into ncRNAs. ncRNAs are considered to associate with mRNAs and inhibit their translation and/or facilitate degradation. The corresponding mean-field kinetic equations for the mRNA, ncRNA and protein populations, N_i, N_j^*, and n_i, in a cell are as follows

$$\mathrm{d}N_i/\mathrm{d}t = w_i - k_i N_i - \sum_j r_{ij} N_i N_j^*, \tag{140}$$

$$\mathrm{d}N_j^*/\mathrm{d}t = u_j - k_j^* N_j^* - \sum_i r_{ij} N_i N_j^*, \tag{141}$$

$$\mathrm{d}n_i/\mathrm{d}t = \upsilon_i N_i - \kappa_i n_i, \tag{142}$$

where w_i, u_j, υ_i, k_i, k_j^* and κ_i are the rates or rate constants of the reactant synthesis and degradation, and r_{ij} are the ncRNA–mRNA association rate constants.

To mimic hierarchical networks, we should specify the regulation of the mRNA synthesis by proteins or, in other words, the dependence of the rates of the mRNA synthesis, w_i, on the protein populations. In the model under consideration (Fig. 17), the two genes forming the upper layer are not regulated by proteins, and the corresponding transcription rates, w_1 and w_2,

are constant. Each of the five genes (with $3 \le i \le 7$) forming the medium layer is considered to be governed by a randomly chosen gene of the upper layer, and the corresponding transcription rates are represented as

$$w_i = w_i^b + w_i^\circ F(n_l), \tag{143}$$

where w_i^b is the basal synthesis rate, w_i° is the maximum value of the regulated rate, and $F(n_l)$ is the dimensionless function dependent on the protein populations n_l ($l = 1$ or 2). Each of the eight genes (with $8 \le i \le 15$) forming the bottom layer is assumed to be negatively self-regulated and to be governed by a randomly chosen gene of the medium layer, and in this case we use

$$w_i = w_i^b + w_i^\circ F(n_l)G(n_i), \tag{144}$$

where $F(n_l)$ and $G(n_i)$ are the dimensionless functions dependent on the protein populations n_l ($3 \le l \le 7$) and n_i ($8 \le i \le 15$). The positive regulation is described as (cf. Eq. (1))

$$F(n_l) = [n_l/(K_l + n_l)]^m, \tag{145}$$

while the negative regulation as (cf. Eq. (2))

$$F(n_l) = [K_l/(K_l + n_l)]^m, \tag{146}$$

$$G(n_i) = [K_i/(K_i + n_i)]^m, \tag{147}$$

where K_l (or K_i) is the protein association constant, and m is the number of regulatory sites for a given protein.

The model above contains a lot of parameters which can be chosen at random in biologically reasonable ranges. Our detailed analysis of the corresponding steady-state kinetics with $m = 2$ for self-regulation and $m = 2$ or 4 for hierarchical regulation indicates that the possible kinetic regimes depend in a simple manner on the types of the regulation of the layers. If the regulation of layer 2 by layer 1 is positive, the regulation of layer 3 by layer 2 is positive as well, and, for example, the average value of u_j is used as a governing parameter, the model often predicts bistability (see, e.g., Fig. 18(a)–(c)) with a fairly narrow hysteresis loop (this feature may be beneficial for the function of cells because it makes it possible to switch without appreciable retardation from one regime of the gene expression to another regime) or a unique steady state (Fig. 18(d) and (e)). If these regulations are negative or mixed (e.g., the regulation of layer 2 by layer 1 is negative while the regulation of layer 3 by layer 2 is positive), the steady state is unique. For the two-layer architecture, the situation is similar [180].

7.3. Distributed genetic networks

A generic model of complex mRNA–protein networks with distributed regulation of gene transcription implies that the protein-mediated contacts and the type of contacts are chosen at random. Such "random" networks were widely mimicked by using the Boolean networks (see, e.g., Refs. [187–189] and references therein). In this section, employing the mean-field kinetics equations, we show the behaviour of a "random" mRNA–protein network regulated globally by ncRNA [182]. The corresponding reaction scheme includes synthesis and conventional degradation of ncRNA, M different mRNAs and proteins (P_i; $1 \le i \le M$), and association of ncRNA and mRNA,

$$\mathrm{Gene}_* \to \mathrm{Gene}_* + \mathrm{ncRNA}, \tag{148}$$

$$\mathrm{Gene}_i \to \mathrm{Gene}_i + \mathrm{mRNA}_i, \tag{149}$$

$$\mathrm{mRNA}_i \to \mathrm{mRNA}_i + \mathrm{P}_i, \tag{150}$$

$$\mathrm{ncRNA} \to \varnothing, \qquad \mathrm{mRNA}_i \to \varnothing, \qquad \mathrm{P}_i \to \varnothing, \tag{151}$$

$$\mathrm{ncRNA} + \mathrm{mRNA}_i \to \mathrm{ncRNA} * \mathrm{mRNA}_i \to \varnothing. \tag{152}$$

This formulation leads to the equations for the ncRNA, mRNA and protein populations,

$$\mathrm{d}N_*/\mathrm{d}t = u - k_* N_* - \sum_i r_i N_* N_i, \tag{153}$$

$$\mathrm{d}N_i/\mathrm{d}t = w_i - k_i N_i - r_i N_* N_i, \tag{154}$$

$$\mathrm{d}n_i/\mathrm{d}t = \upsilon_i N_i - \kappa_i n_i, \tag{155}$$

where u, w_i, υ_i, k_*, k_i and κ_i are the rates or rate constants of the reactant synthesis and degradation, and r_i are the ncRNA–mRNA association rate constants.

To mimic a distributed network, each gene transcribed into mRNA is assumed to be governed by two other randomly chosen genes (including self-regulation) via association of the corresponding proteins to the regulatory sites. For simplicity, the regulatory sites are considered to operate independently. Each of two proteins governing the transcription of a given gene is assumed to be able to associate with one or two sites. If there are two sites for each protein, the sites are considered to be equivalent. In particular, the mRNA synthesis rate is represented as

$$w_i = w_i^b + w_i^\circ F_1(n_1)F_2(n_2), \tag{156}$$

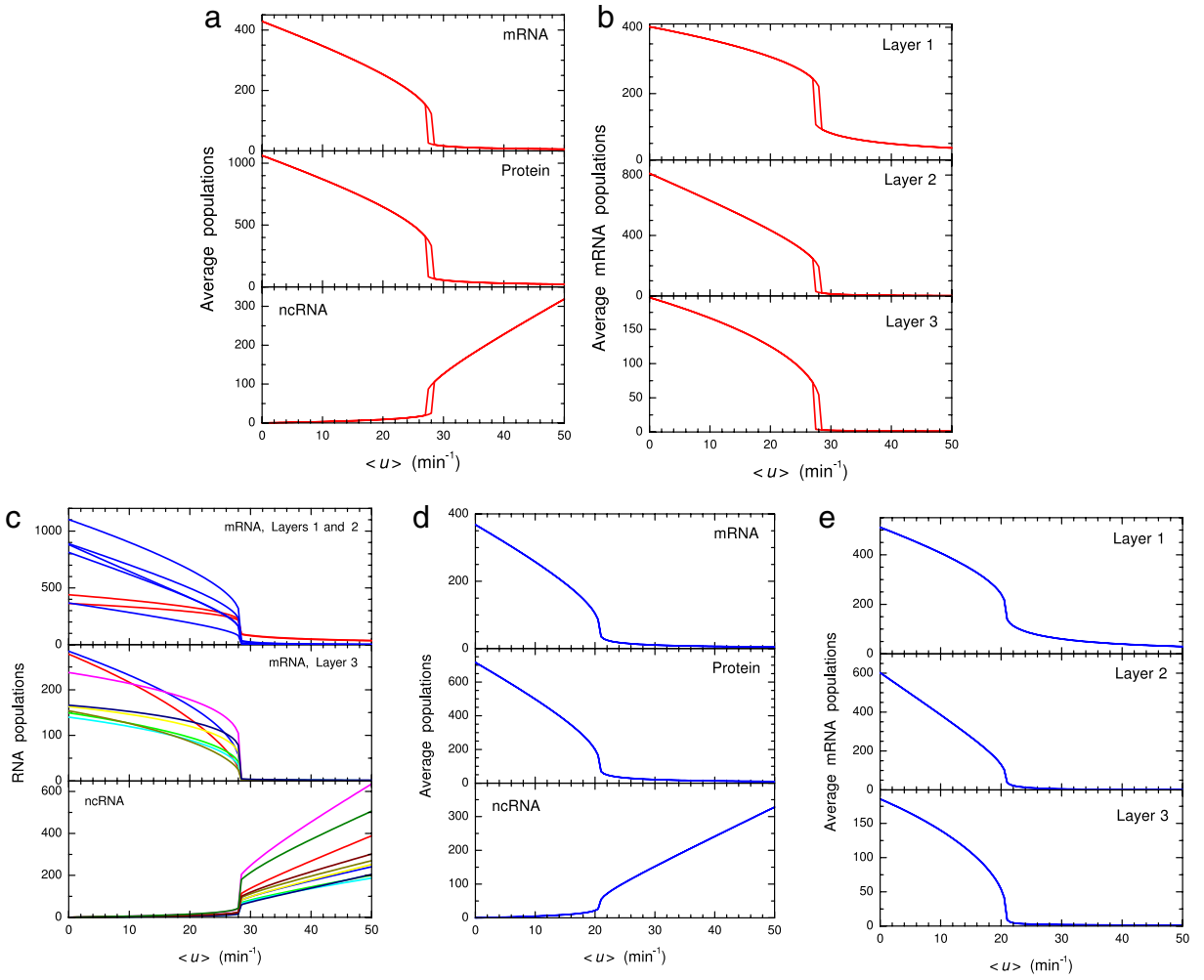


Fig. 18. Average RNA and protein populations as a function of the average value of u_j according to Eqs. (140)–(147) in the case of positive regulation of layer 2 by layer 1 and positive regulation of layer 3 by layer 2. Panels (a)–(c) show an example of bistable kinetics: (a) average values for the whole network; (b) average mRNA populations corresponding to different layers; (c) populations of specific RNAs (note that in order to avoid overlapping of curves this panel exhibits only one branch of the hysteresis). Panels (d) and (e) show an example of the kinetics with a unique steady state: (d) average values for the whole network; (e) average mRNA populations corresponding to different layers. The calculations were performed with $m = 4$ for hierarchial regulation. The kinetic parameters are chosen at random in the ranges from 0.05 to 0.15 min^{-1} for k_i, k_j^* and κ_i, from 1 to 2 min^{-1} for w_i^b, from 0.1 to 0.3 min^{-1} for υ_i, from 0 to 10^{-3} min^{-1} for r_i, and from 200 to 500 for K_l. For genes 1 and 2, $w_1^\circ = w_2^\circ = 50$ min^{-1}. For the other genes, w_i° are selected at random in the range between 100 and 300 min^{-1}. The ncRNA synthesis rates are represented as $u_j = \langle u\rangle\eta_i$, where η_i is a dimensionless factor chosen at random in the range between 0.5 and 1.5, and $\langle u\rangle$ is the average rate.

where w_i^b is the basal synthesis rate, w_i° is the maximum value of the regulated rate, and $F_1(n_1)$ and $F_2(n_2)$ are the dimensionless functions (≤ 1) dependent on the protein populations n_1 and n_2 (the locally used subscripts 1 and 2 correspond to two proteins regulating gene i). The former function is defined (cf. Eqs. (1) and (2)) for positive regulation as

$$F_1(n_1) = [n_1/(K_1 + n_1)]^m \tag{157}$$

and for negative regulations as

$$F_1(n_1) = [K_1/(K_1 + n_1)]^m, \tag{158}$$

where K_1 is the protein association constant, and m (1 or 2) is the number of regulatory sites for a given protein. $F_2(n_2)$ is defined in analogy.

For example, let us consider that a gene has two regulatory sites for each regulatory protein, and the regulation by one protein is positive while the regulation by another protein is negative. In this case, the transcription rate is given by

$$w_i = w_i^b + \frac{w_i^\circ n_1^2 K_2^2}{(K_1 + n_1)^2 (K_2 + n_2)^2}. \tag{159}$$

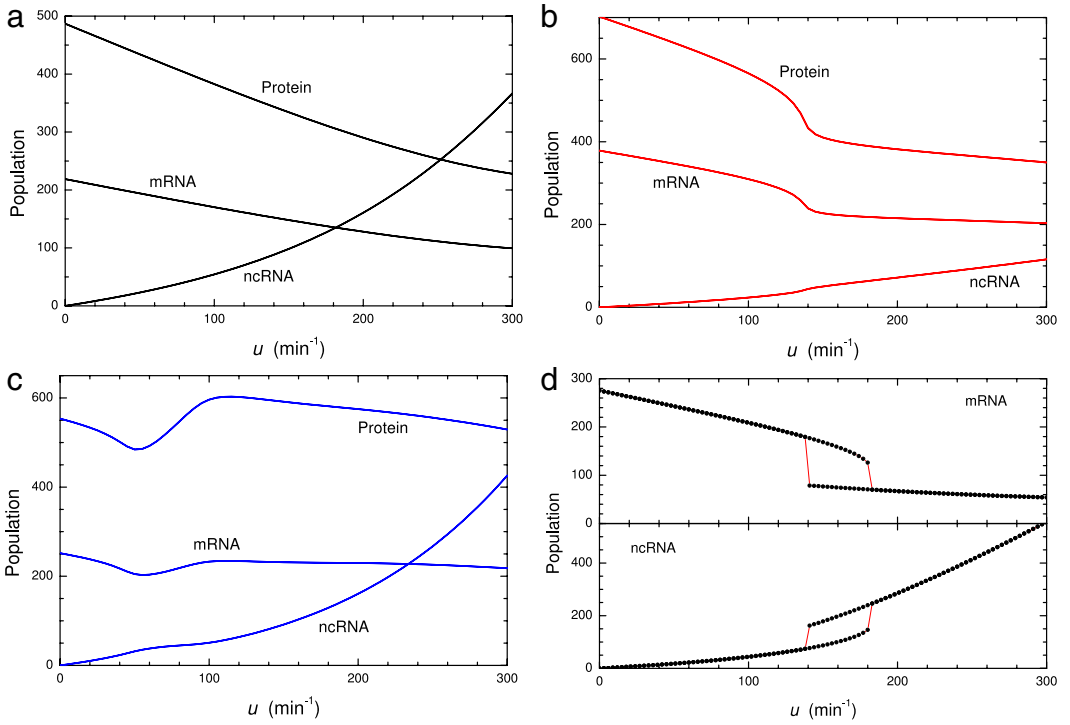


Fig. 19. ((a)–(i)) Average mRNA and ncRAN populations as a function of u according to Eqs. (153)–(158) in the case when each of two proteins governing the transcription of a gene is able to associate with two sites ($m = 2$) and each feedback is either positive or negative with probability 0.5. The average protein population is not shown (except panels (a)–(c)), because its dependence on u is similar to that of mRNA. In case (g), one can notice tiny irregular features at $u \simeq 100\ \mathrm{min}^{-1}$. These features correspond to stable oscillations with a very small amplitude. In addition, it is of interest to notice that in case (h) the oscillations in the mRNA populations are appreciable while the oscillations in the ncRNA populations are nearly negligible. This means that ncRNA serves in this case as a trigger for oscillations. (The kinetic parameters, except $k_* = 0.1\ \mathrm{min}^{-1}$, are chosen at random in the ranges from 0.05 to 0.15 min^{-1} for k_i and κ_i, from 1 to 2 min^{-1} for w_i^b, from 100 and 200 min^{-1} for w_i°, from 0.1 to 0.3 min^{-1} for υ_i, from 0 to 10^{-3} min^{-1} for r_i, and from 100 to 200 for K_j.)

In the model described above, ncRNA plays a role of a global regulator, and accordingly it is instructive to use u as a governing parameter. The analysis of the model indicates that in the case of entirely positive regulation the kinetics are either bistable or exhibit a single stable steady state. If the regulation is entirely negative, there is a single stable steady state with monotonous dependence of the reactant populations on u. If $m = 1$ and each feedback is chosen to be either positive or negative at random, the negative feedbacks dominate, and there is a single stable steady state. If $m = 2$ and each feedback is positive or negative at random, the predictions of the model are much richer as shown, for example, in Fig. 19. In particular, one can observe a single steady state with various types of the dependence of the reactant populations on u (Fig. 19[(a)–(c)]), various bistable regimes (Fig. 19[(d)–(g)]), and sometimes oscillations (Fig. 19(h)) or oscillations and bistability simultaneously (Fig. 19(i)). For $m = 4$, the results are qualitatively similar.

Referring to the random Boolean networks [187], one could expect to find kinetic chaos in the model under consideration. This feature is, however, not observed. In the absence of ncRNA, the model does not predict chaos either [190].

7.4. miRNA regulation of a cancer network

The proliferative potential and some other features of cancer cells are similar to those of stem cells, and in both these cases miRNAs seem to play a key role in regulation of the corresponding genetic networks (see the introduction). In particular, miRNAs are found to be involved in the pathophysiology of all types of human cancers mainly by aberrant gene expression with abnormal levels of mature and/or precursor miRNA transcripts [69,70]. The model proposed by Aguda et al. [177] is focused on three transcription factors E2F (E2F1, E2F2, and E2F3), transcription factor Myc, and a cluster of miRNAs called *miR-17-92*. These transcription factors participate in the control of cell proliferation and apoptosis and can act as oncogenes (i.e., genes that are mutated or expressed at high levels) or tumor suppressors depending on their levels of expression. The *miR-17-92* cluster includes 7 mature miRNAs. These miRNAs are over-expressed during diffuse large lymhoma, B cell lymphoma, lung cancer and some other cancers [69,177]. The model includes the genes responsible for the formation of these transcription factors and miRNAs and a few other genes.

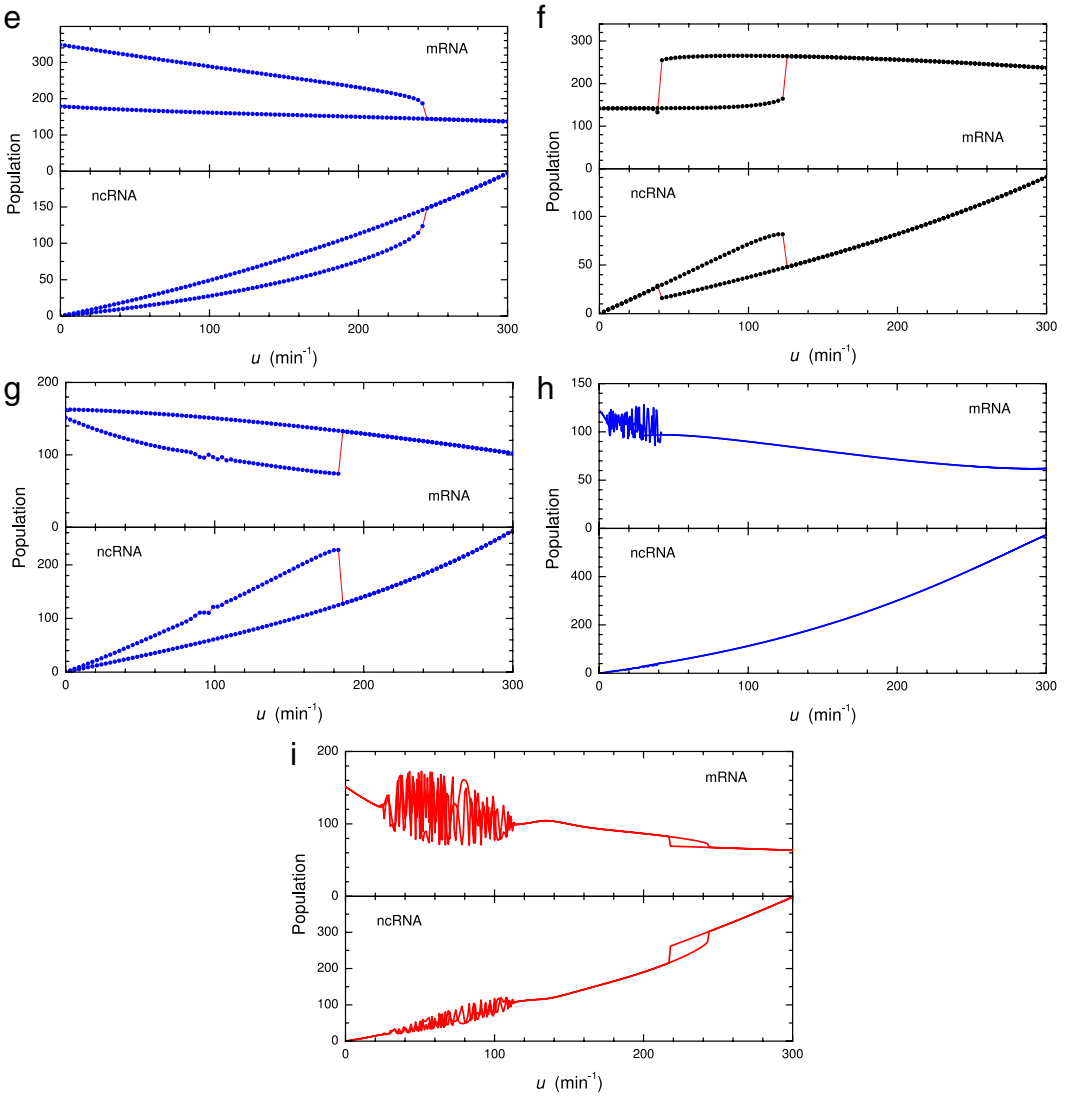


Fig. 19. (*continued*)

Step by step, Aguda et al. [177] have simplified their model, and its final coarse-grained version includes only two components representing the protein (Myc and E2Fs) module and the *miR-17-92* cluster. The equations for the corresponding protein and miRNA populations, p and m, are as follows (we keep the original designations)

$$\frac{dp}{dt} = \alpha + \frac{k_1 p^2}{\Gamma_1 + p^2 + \Gamma_2 m} - \delta p, \tag{160}$$

$$\frac{dm}{dt} = \beta + k_2 p - \gamma m. \tag{161}$$

The three terms on the right-hand parts of these equations represent, respectively, the protein and miRNA basal synthesis rates, regulated synthesis rates, and degradation rates (α, β, k_1, k_2, Γ_1, Γ_2, δ and γ are the corresponding parameters).

The model defined by Eqs. (160) and (161) is seen to include positive self-regulation of the protein formation, positive regulation of the miRNA formation by protein, and negative regulation of the protein formation by miRNA. With these ingredients, the model predicts bistability (cf. Section 5.1). The stable regime corresponding to quiescence exhibits low protein and miRNA populations. With changing governing parameters (e.g., with increasing α), there is another stable regime with high protein and miRNA populations. The latter regime is postulated to correspond to healthy cell cycle, cancer, or apoptosis depending on the protein and miRNA populations. It is of interest that under steady-state conditions these

populations change in the same direction, although slow non-steady-state or transient dynamics are possible where the changes could be in opposite directions.

Aguda et al. [177] have also extended the model described above by introducing delay into the protein formation (to the second term in the right-hand parts of Eq. (160)). With this feature, as expected (cf. Sections 5.4 or 6.5), the model predicts kinetic oscillations. Some other modifications of this model can be found in Ref. [178].

The model presented by Aguda et al. [177] is especially interesting as one of the first attempts to clarify the role of miRNAs in specific complex genetic networks. The crucial ingredient of such models is validation. The detailed discussion of the validation of Aguda's model is beyond our goals. We may only notice that in addition to *miR-17-92* the cancer is usually accompanied by aberrant expression of many other miRNAs. During the lung cancer, for example, the expression of *miR-19a, 21, 92, 155, 191, 205, 210* is also increased, the expression of the *let-7* family is decreased, and the unfavorable prognosis is related to the changes of expression of *miR-155* and *let-7* [69]. The role of all these changes in gene expression is open for debate.

7.5. Small ncRNAs in Escherichia coli

Although ncRNAs are inherent to eukaryotes (see the introduction), their role in prokaryotes is also important. The latter is illustrated by the analysis performed by Wang et al. [179]. Their goal was to clarify the changes in gene expression in *E. coli* during the carbon source transition from glucose to acetate. If *E. coli* grows on excess glucose, acetate is produced and excreted as excess carbon (the excreted acetate can be consumed as a carbon source when the glucose level drops) [191]. The metabolic switch from glucose to acetate is accompanied by activation of genes involved in acetate uptake. To describe this transition, Wang et al. [179] have used a model including 38 regulators (22 transcription factors and 16 small ncRNAs) and 150 target genes. The regulation by transcription factors was described by the simplest power-law equations (like in the mass-action law). ncRNAs were considered to associate with mRNAs with subsequent degradation. As usual in such models, the reconstructed gene activities were dependent on the initial setting of regulation matrices. Despite this problem, Wang et al. [179] were able to describe reasonably the time evolution of various regulators. Their results show that including ncRNAs into a model can shed light on the hidden effects in the regulation activity of transcription factors and thus can uncover the biological mechanisms in gene regulation in a more accurate manner.

7.6. ncRNAs and global protein population

In relation with complex genetic networks including ncRNAs, it is of interest to discuss the interplay of mRNA, protein and ncRNA during the growth of eukaryotic cells.

In prokaryotes, the growth of cells is well known to be exponential and the average mRNA and protein concentrations are nearly constant during the growth [192]. The interpretation of these features appears to be straightforward if one takes into account that in this case the cell cycle is relatively short, the DNA replication occurs during the whole cycle, the dependence of the DNA amount on time is nearly exponential, and the rate of the mRNA synthesis is proportional to the DNA amount (for the corresponding models, see Ref. [192] and references therein).

Although in eukaryotes the cell growth is often considered to be exponential as well, the linear and bilinear growth models have also been proposed [114]. In yeasts, for example, the deviation from the exponential growth are well visible in the end of the cycle [114]. On the global scale, the experiments [116,117] indicate that the concentration of the majority of proteins during the growth of eukaryotic cells is nearly constant like in prokaryotes. Weak dependence of concentration of most proteins on the cell volume seems to be beneficial for the cell function, because it may facilitate the control of various intracellular processes. The interpretation of this feature in eukaryotic cells is, however, not straightforward. In such cells, the cycle is traditionally divided into four sequential phases: G_1 (gap phase), S (DNA replication), G_2 (gap phase), and M (division) [1]. The cycle duration is typically about one day (it may be shorter in simple organisms, e.g., in yeasts, or longer, e.g., in mammals). The duration of the S phase is often relatively short, and one could expect that the growth would be different before and after this phase. In reality, this does not seem to be the case. The understanding why the growth is apparently insensitive to the S phase is still limited [193]. One of the likely related factors is the effect of ncRNAs on the mRNA–protein network [194].

The simplest scheme allowing us to analyze the interplay of mRNA, protein and ncRNA during the growth of cells includes steps (39)–(45). The corresponding equations for the mRNA, ncRNA and protein populations are represented as

$$\mathrm{d}N/\mathrm{d}t = w - kN - rNN_*, \tag{162}$$

$$\mathrm{d}N_*/\mathrm{d}t = u - k_*N_* - rNN_*, \tag{163}$$

$$\mathrm{d}n/\mathrm{d}t = \upsilon N - \kappa n, \tag{164}$$

where w, u, υ, k_*, k and κ are the rates or rate constants of the reactant synthesis and degradation, and r is the ncRNA–mRNA association rate constant.

During the cell cycle, all the parameters in Eqs. (162)–(164) may depend on time, and we should take the key factors behind this dependence into account. One of the factors might be the regulation of the mRNA, ncRNA and protein synthesis by proteins. Concerning this aspect, we may notice that in cells the number of cycle-related mRNAs and proteins is large but

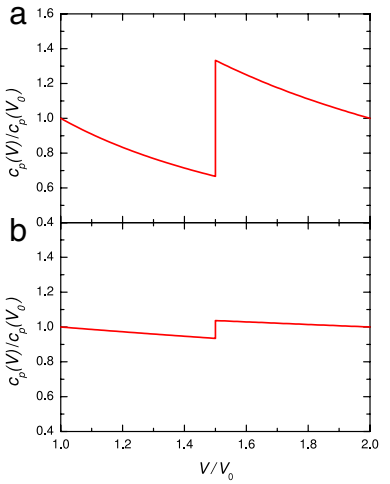


Fig. 20. Normalized protein (or mRNA) concentration as a function of the cell volume during the cell cycle (a) in the absence of the mRNA–ncRNA association ($r_0 = 0$) and (b) with association ($r_0 = 10^{-3}$ min^{-1}; for the other parameters, see Ref. [194]). The DNA replication is assumed to occur at $V = 1.5V_0$.

their relative abundance in the global mRNA and protein pool is modest [195,196] and they can hardly control the population of the majority of proteins. The latter population, as already noticed in the beginning of this subsection, is approximately proportional to the cell volume and their concentration is nearly constant [116,117]. For these reasons, the protein-mediated feedbacks can be ignored, i.e., υ can be considered to be constant.

On the other hand, the DNA replication occurring during the S phase of the cell cycle results in the twofold increase of the number of genes and the corresponding increase of the gene-transcription rate resulting in the mRNA and ncRNA synthesis. Taking into account that the time scale of the S phase is often relatively short compared to the duration of the cell cycle, the DNA replication can be mimicked by a stepwise increase in w,

$$w(t) = \begin{cases} w_\circ & \text{for } V_0 \le V(t) < V_r, \\ 2w_\circ & \text{for } V_r < V(t) \le 2V_0, \end{cases} \tag{165}$$

where $V_0 \equiv V(0)$ is the initial cell volume, and V_r is the cell volume corresponding to the DNA replication. The rate of the ncRNA synthesis can be described in analogy with Eq. (165).

The rate constants of conventional degradation of mRNAs, ncRNAs and proteins are proportional to the concentrations of the corresponding enzymes. Basically, enzymes are proteins, and accordingly in our course-grained model these rate constants are expected to be proportional to the protein concentration. In analogy with the majority of proteins [116,117], the concentration of the enzymes under consideration can be considered to be constant or, at least, to be weakly dependent on time so that this dependence can be neglected, and accordingly k, k_* and κ can be assumed to be constant. (The validity of this approximation can be verified *a posteriori*, or the model can be extended as described in Ref. [193].)

The ncRNA–mRNA association rate per unit volume can be represented as $W = \rho c_m c_n$, where ρ is the volume- and time-independent rate constant, and c_m and c_n are the mRNA and ncRNA concentrations. Taking into account that $c_m = N/V(t)$ and $c_n = N_*/V(t)$, we have $W = \kappa NN_*/V^2(t)$. The total rate of the mRNA and ncRNA association is $W_{tot} \equiv WV(t) = \rho NN_*/V(t)$. On the other hand, the total rate is defined by Eqs. (162) and (163) as $W_{tot} = rNN_*$. Comparing these expressions, we conclude that the dependence of r on volume (or time) can be represented as

$$r(t) = r_\circ V_0/V(t), \tag{166}$$

where $r_\circ = \rho/V_0$.

To complete the description of the model, we may notice that the number of distinct potential ncRNA-targets is often high, and the number of distinct ncRNAs is high as well. Under such conditions, the full set of equations describing the interplay between distinct mRNAs, ncRNAs and protein is large. In such situations, Eqs. (162)–(164) can nevertheless be used by assuming that N, N_* and n represent the average numbers of large groups of mRNAs, ncRNAs and proteins.

To solve Eqs. (162)–(164), one can use a steady-state approximation, because the growth of eukaryotic cells is relatively slow (Section 3.7). The corresponding calculations indicate (Fig. 20) that ncRNA can help to maintain nearly constant concentrations of mRNA and protein provided that the ncRNA population is higher than the mRNA population and the mRNA–ncRNA association and degradation are rapid. This effect is related to the increase of the rate of the mRNA degradation via association with ncRNA due to the increase of the ncRNA concentration just after the DNA replication. The increase of the mRNA degradation rate compensates for the increase of the mRNA synthesis rate after the DNA replication.

The conclusion above has been drawn by using a model describing three species, ncRNA, mRNA and protein. In this context, it is of interest to speculate whether ncRNAs or, more specifically, miRNAs can contribute to the global control of the

the mRNA and protein population. Concerning this point, one can notice that at present the number of confidently identified miRNA genes has surpassed 400 in humans [54]. This number of protein-coding genes, $\simeq 3 \times 10^4$, is much larger, and the first impression might be that the global control by miRNAs is unlikely. There are, however, at least three factors indicating that the reality is more subtle. (i) The number of miRNAs will undoubtedly increase as high-throughput sequencing continues to be applied [54]. (ii) The transcriptional activity of many protein-coding genes is very low, while the miRNA studies, using e.g. the ncRNA microarray technique [197], tend to be focused on miRNAs with relatively high populations. For this reason, on average, the identified miRNA genes are more active that the protein-coding genes. (iii) miRNAs are transcribed as long ncRNA and then generated via a two-step processing pathway including the formation of a few different $\sim$65-nt pre-miRNAs followed by conversion of each of them into the corresponding miRNA [82]. This is an additional reason why the efficiency of the genes generating miRNAs may be a few times higher than that of the protein-coding genes. Taking all these points into account, one cannot exclude that miRNAs can contribute to the global control of the the mRNA and protein population. To have this situation, many miRNAs must be appreciably expressed in different tissues. Interestingly, this is the case in normal human tissues [198].

8. Stochastic effects

8.1. General aspects

In cells, many genes exist in a single copy and the populations of mRNAs and proteins are often relatively low. For these reasons, the kinetics of gene expression frequently exhibit stochastic features or, more specifically, bursts representing sequential periods of high and low transcriptional and/or translational activity. Stochasticity may be related, for example, to low population of a transcription factor regulating the gene transcription, while the mRNA or protein population transcribed or translated during bursts may actually be appreciable. Such features of the kinetics of gene expression have long attracted appreciable attention because this phenomenon is of interest from very different perspectives.

The biological functions of stochastic bursts in the kinetics of gene expression are still not fully clear. Already now, there are, however, indications (reviewed in Ref. [199]), that this mode of gene expression may play an important role in very diverse intracellular phenomena including, for example, signals processing [200], embryonic and adult stem cell biology [201] (see also the discussion in Ref. [148]), tumor maintenance [202], transient increased resistance to chemotherapy [79], and survival and responsiveness of microbes in dynamic or diverse environments (e.g., resistance of bacteria to antibiotics [203]). These and other examples show that, depending on circumstances, stochastic bursts may be advantageous, disadvantageous, or, e.g., advantageous for bacteria and disadvantageous for a host.

The experimental and theoretical studies of stochastic features in the kinetics of gene expression are focused primarily on the interplay of mRNAs and proteins (see, e.g., reviews [36–39] and [199]). In eukaryotic cells, such features can also be related to ncRNAs. Despite the current boom in the studies of ncRNAs, the experimental data on the corresponding kinetics are still scarce. In particular, to our knowledge, there are no experimental reports on observation of ncRNA-related stochastic bursts. To tackle this problem theoretically, we recall below the main types of stochastic bursts in the mRNAs–protein interplay, outline the key abilities of ncRNAs with emphasis on associated stochastic effects, and consider in more detail a few possible scenarios of the corresponding stochastic bursts.

8.2. Bursts in the mRNA–protein interplay

In genetic networks including mRNAs and proteins, there are three general scenarios of stochastic bursts (the experiments are reviewed in Ref. [199]; the corresponding models are numerous; see reviews [36–39], one of the formative articles [204], more recent articles [205–221], and references therein):

(i) *Translational bursts*. The simplest scheme of gene expression consists of gene transcription into mRNA, mRNA translation to protein (P), and mRNA and protein degradation,

$$\text{Gene} \to \text{Gene} + \text{mRNA}, \tag{167}$$

$$\text{mRNA} \to \text{mRNA} + \text{P}, \tag{168}$$

$$\text{mRNA} \to \varnothing, \tag{169}$$

$$\text{P} \to \varnothing. \tag{170}$$

Stochastic bursts can be observed even if a network contains only these four steps and there is no any regulation [36,37]. This is possible provided that the rate constant of step (168) is high compared to that of step (170). In this case, the synthesis of each mRNA results in a stochastic burst in the P population. These bursts are well-manifested provided that the mRNA population is low and step (167) is slow compared to steps (168) and (170). The former is the case provided that the rate constant of step (169) is comparable to or higher than the rate of step (167). In general, step (169) can be regulated by proteins, and this regulation can influence the bursts under consideration [220].

(ii) *Bursts related to transcription*. If step (167) or, more specifically, its initiation is controlled by slow attachment of a transcription factor [protein ($\mathcal{P}$)] to and detachment from a regulatory site, one can observe stochastic bursts in the mRNA

and P population (Refs. [36,37,204–214]). In particular, these populations will be relatively small or large depending on the state of a regulatory site. Alternatively, the transitions between the active and inactive gene states can be induced by chromatin remodeling. Stochastic bursts may also be caused by delays in the RNAP traffic along DNA during mRNA elongation [26,199,221] or presumably by the return of RNAP to the beginning of a transcriptional cycle [31].

(iii) *Bursts related to bistability* (Refs. [204,215–219]). In the presence of feedbacks between the mRNA and protein synthesis, the kinetics of gene expression are often bistable (in the mean-field approximation). For example, steps (167)–(170) may result in bistability provided that step (167) is positively regulated by P and the number of the corresponding regulatory sites is equal or larger than 2 (Section 5.1). If, in addition, the mRNA and protein populations are relatively small, one can observe stochastic bursts representing sequential transitions between the states which are close to the stable mean-field steady states. These bursts are more collective compared to those described in item (ii). In particular, the periods of high and low activity cannot be expressed in terms of the rate of one of the steps.

8.3. ncRNA abilities and stochastic effects

In the preceding section, we have outlined three general scenarios of stochastic bursts in the mRNA–protein interplay. If the corresponding kinetic schemes are complemented by ncRNA(s) or, more specifically, by association of mRNA(s) and ncRNA(s) with subsequent degradation (in analogy with the models described in Section 5), the ncRNA population will be influenced by bursts in the mRNA and protein populations and accordingly may also exhibit stochastic bursts and/or in its turn influence the mRNA and protein bursts. In such situation, basically, ncRNAs provide an additional channel of mRNA degradation (see the discussion in the beginning of Section 5), and accordingly the ncRNA function is not too specific, because often the same role can be played by enzymes (to some extent, it does not matter whether the mRNA population is reduced by ncRNA or by enzyme). Compared to proteins or, more specifically, transcription factors, the function of ncRNAs may, however, be qualitatively different. Let us consider the simplest situation (Ref. [123]; reviewed in Ref. [124]) when the mRNA degradation rate is low, the mRNA population is large, and the goal is to reduce this population. This goal can be reached by using a transcription factor suppressing the gene transcription into mRMA. In this case, the mRNA synthesis and degradation rates will be low and the mRNA population may exhibit bursts [scenario (ii) in the preceding subsection]. Alternatively, the goal can be reached by employing ncRNA. In the latter case, the mRNA synthesis and degradation rate will be high and the mRNA population will exhibit conventional fluctuations without distinct bursts.

In this subsection and subsections below, we will discuss and analyze more complex situations when stochastic bursts are inherently related to ncRNAs. In such situations, scenario (i) is obviously irrelevant (because ncRNAs are not translated into proteins) while scenarios (ii) and (iii) can be realized. In particular, scenario (ii) is equally applicable to ncRNAs. Due to the interplay between ncRNA and mRNA (see, e.g., the simplest scheme including steps (39)–(45)), the corresponding bursts in the ncRNA population may result in the appearance of anti-phase bursts in the mRNA and protein populations. Many other more specific mechanisms of bursts related to the abilities of ncRNAs are also possible. The ncRNA abilities have already been described in Section 2.4. Here, we outline the corresponding paradigms again (for the references, see Section 2.4) in combination with the commentaries focused on stochastic bursts.

(i) *Indirect regulation of transcription.* ncRNAs can associate with mRNAs, other ncRNAs or proteins and reduce the population of these species and/or affect their processing. In addition, long ncRNAs can be processed to yield small RNAs. Such diverse processes can indirectly regulate the transcription of many genes, the kinetics of gene expression can be bistable (Sections 5–7), and it may result in stochastic bursts (see, e.g., Section 8.4 below) similar to those described in item (iii) of Section 8.2.

(ii) *Direct regulation of transcription.* ncRNAs can directly regulate transcription through a range of mechanisms: ncRNA can associate with a gene and recruit a ncRNA-binding protein with subsequent regulation of the gene activity by this protein or the ncRNa–protein complex; ncRNA can associate with protein, and the ncRNa–protein complex can then associate with a gene and regulate transcription; ncRNA and regulatory protein can competitively adsorb on a regulatory site of a gene and the transcription rate can depend on the state of this site. Under suitable conditions, all these scenarios may result in stochastic bursts (see, e.g., Section 8.5 below) similar to those described in item (ii) of Section 8.2. In the presence of feedbacks, stochastic bursts similar to those described in item (iii) of Section 8.2 are possible as well.

(iii) *Chromatin modification.* In eukaryotic cells, chromosomes are packaged by histones (positively charged proteins) into a condensed structure called chromatin. ncRNAs can silence gene expression by recruitment of chromatin modifying complexes, exclusion of the transcription machinery from the chromosome, modification of histones, and subsequent changes in the chromatin structure. There are also other scenarios of the ncRNA-induced chromatin modification resulting in transcriptional activation. In general, chromatin modification seems to occur via a few steps. The simplest coarse-grained description of this process can be based on the two-state (‘‘unmodified’’ and ‘‘modified’’) approximation. Following this line, one can conclude that the ncRNA-induced chromatin modification may result in stochastic bursts (see Section 8.6 below) as in item (ii) of Section 8.2.

(iv) *Transcriptional interference.* The transcription of a ncRNA across the promoter region of a protein-coding gene can interfere with transcription factor binding and prevent this gene from the mRNA synthesis. In this case, the gene-expression kinetics can exhibit stochastic bursts (see Section 8.7 below) similar to those described in item (ii) of Section 8.2.

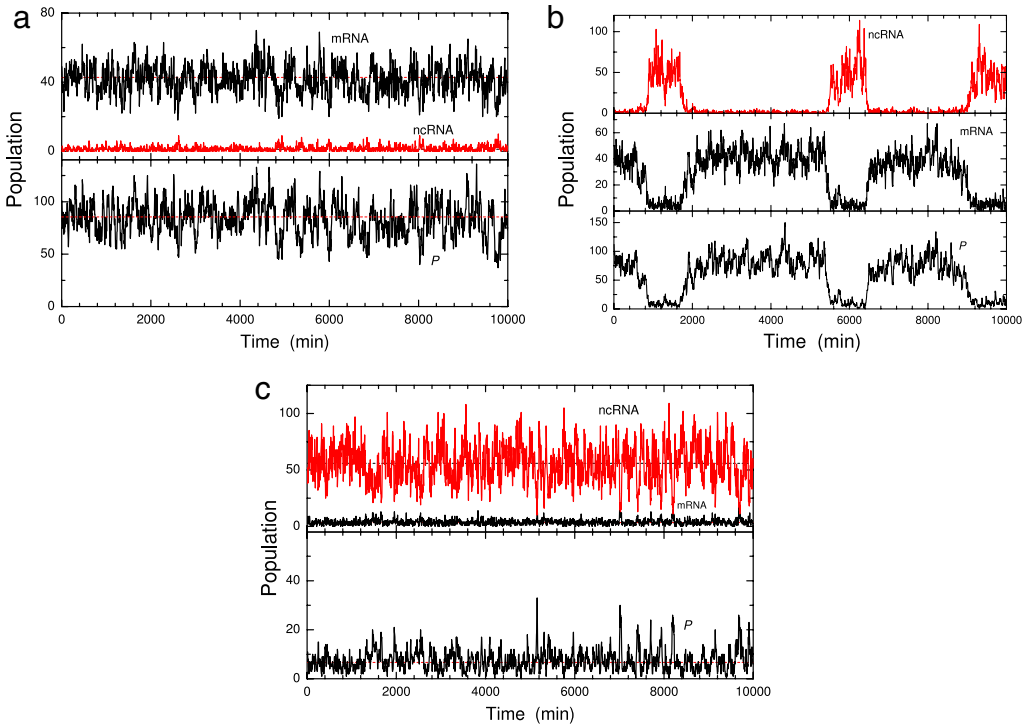


Fig. 21. Protein, mRNA, and ncRNA populations as a function of time according to Monte Carlo simulations performed in the framework of the model described in Section 6.1: (a) stable kinetics (outside the bistability window) with low ncRNA population; (b) kinetic bursts inside the bistability window; (c) stable kinetics with relatively high ncRNA population. The mean-field values of the populations are indicated in panels (a) and (c) by dashed lines. (For the details of the simulations, see Ref. [158].)

8.4. Bursts related to bistability

The kinetics of the mRNA–ncRNA–protein interplay including the regulation of gene transcription into ncRNAs can often be bistable (Sections 5–7). If, in addition, the ncRNA, mRNA and/or protein populations are relatively small, one can observe stochastic bursts representing sequential transitions between the states which are close to the stable mean-field steady states. Such transitions are inherently related to fluctuations. To induce a transition from one state to another, the fluctuations should be appreciable. If the populations of all the species participating in the game are large, the probability of realization of such fluctuations is very low, and the bursts cannot be observed.

The simplest bistable scheme including ncRNA has been described in detail in Section 6.1. The corresponding stochastic bursts obtained by using Monte Carlo simulations [158] are shown in Fig. 21.

8.5. Mixed regulation of transcription

As already noted in item (ii) of Section 8.3, ncRNA and regulatory protein can competitively associate to a regulatory site of a gene and the transcription rate can depend on the state of this site. For example, a ncRNA transcribed from the *DHFR* minor promoter in humans can associate with the major promoter, occlude the binding of the general transcription factor TFIID, and thereby silence *DHFR* gene expression (see experiments [222] and review [49]). A generic kinetic model of gene expression with this type of the regulation includes synthesis of mRNA, ncRNA and protein (P),

$$\mathrm{Gene}_1 \to \mathrm{Gene}_1 + \mathrm{mRNA}, \tag{171}$$

$$\mathrm{Gene}_* \to \mathrm{Gene}_* + \mathrm{ncRNA}, \tag{172}$$

$$\mathrm{mRNA} \to \mathrm{mRNA} + \mathrm{P}, \tag{173}$$

reversible association of ncRNA and protein with the regulatory site of the gene transcribed into RNA (the type of this RNA does not matter),

$$\mathrm{Gene}_2 + \mathrm{ncRNA} \rightleftharpoons \mathrm{Gene}_2 * \mathrm{ncRNA}, \tag{174}$$

$$\mathrm{Gene}_2 + \mathrm{P} \rightleftharpoons \mathrm{Gene}_2 * \mathrm{P}, \tag{175}$$

$$\mathrm{Gene}_2 * \mathrm{P} \to \mathrm{Gene}_2 * \mathrm{P} + \mathrm{RNA}, \tag{176}$$

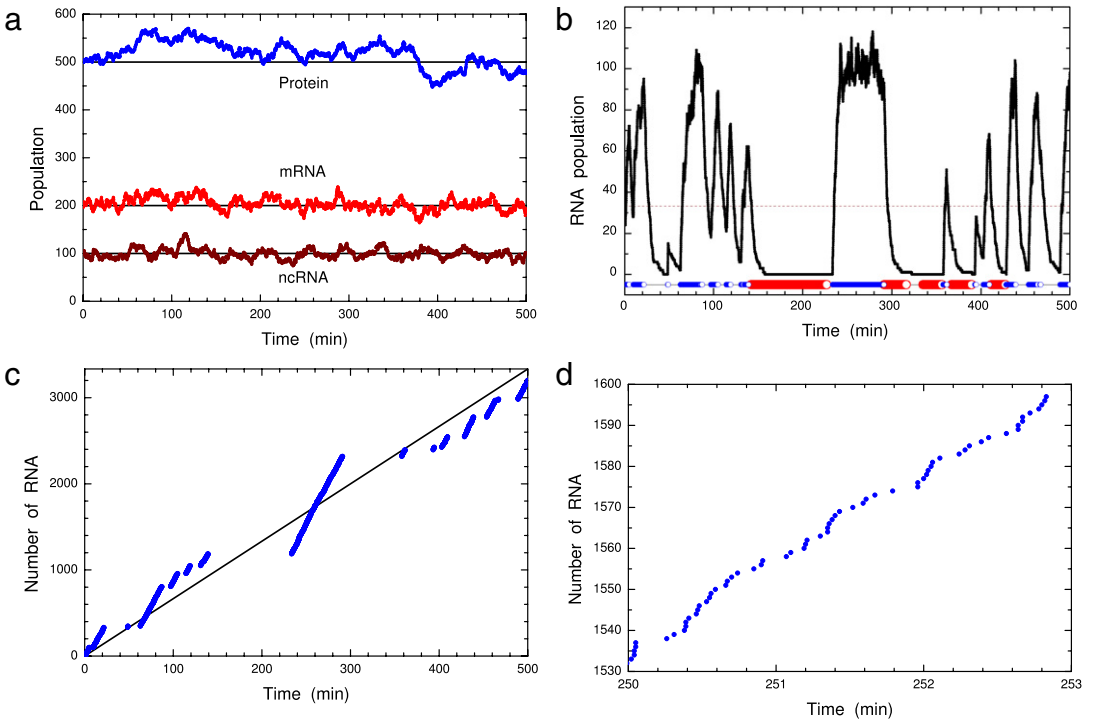


Fig. 22. [(a) and (b)] Populations of mRNA, ncRNA, protein and RNA and [(c) and (d)] the total number of synthesized RNA as a function of time according to Monte Carlo simulations performed in the framework of the model defined by steps (171)–(177). The state of the regulatory site of $Gene_2$ is shown on the bottom line in panel (b) (the large and small circles correspond to occupation of this site by ncRNA and protein, respectively). The predictions of the mean-field kinetic equations are indicated by straight solid or dashed lines. Panel (d) exhibits a fragment of panel (c) in order to show the fine structure of the kinetics. (For the details of the simulations, see Ref. [223].)

and degradation of all these species,

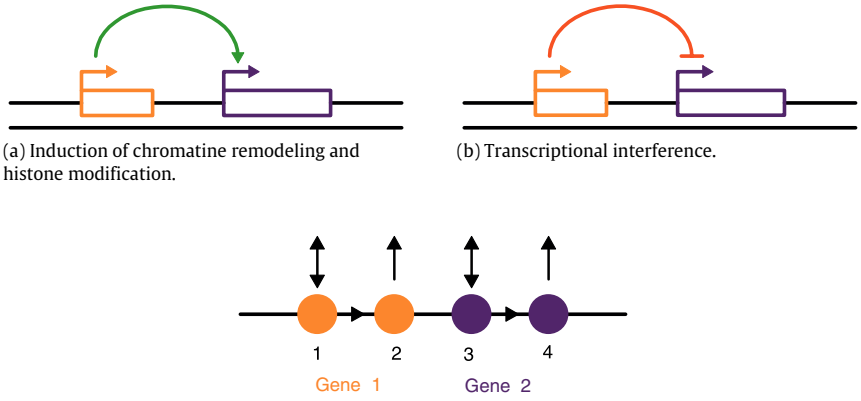
$$\text{mRNA} \to \varnothing, \qquad \text{ncRNA} \to \varnothing, \qquad \text{P} \to \varnothing, \qquad \text{RNA} \to \varnothing. \tag{177}$$

The synthesis of RNA (step (176)) is assumed to occur only in the presence of P on the regulatory site. Association of ncRNA and P with the regulatory site is considered to be competitive, i.e., the site can be occupied either by ncRNA or by P (as in the case of the *DHFR* gene [222]).

Steps (171)–(177) were recently analyzed by using the mean-field equations and Monte Carlo simulations [223]. The mean-field kinetics are simple. In particular, the steady state is unique and the populations of all the species monotonously depend on the model's parameters. With fluctuations, the model can, however, predict distinct stochastic bursts of the synthesis of RNA by $Gene_2$. Although the bursts are basically of the type described in item (ii) of Section 8.2, some of their features are novel. In particular, the RNA population as a function of time can exhibit relatively long bursts composed of short bursts (Fig. 22). This feature can be observed provided that the rates of the ncRNA attachment to and detachment from the regulatory site of $Gene_2$ are much lower than the corresponding rates for P.

8.6. Chromatin modification

According to Hirota et al. [224] (see also review [51]), the RNAP II transcription of ncRNAs upstream of the *S. pombe* $fbp1^+$ results in chromatin remodeling that is critical for transcriptional activation of the downstream protein-coding gene (Fig. 23(a)). One of the indicators of this mechanism was that the insertion of a transcription terminator into this upstream region abolishes both the cascade of transcription of ncRNAs and the chromatin alteration. To describe this mode of the expression of two genes, one can use the two-site approximation [31] for each gene. In this model, the gene transcribed into ncRNA is represented by Sites 1 and 2 (S_1 and S_2). The ncRNA formation is considered to start at Site 1 and to end at Site 2. The gene transcribed into mRNA is represented by Sites 3 and 4 (S_3 and S_4). The mRNA formation is assumed to start at Site 3 and to end at Site 4. The two genes are considered to be nearest-neighbour. The latter gene is assumed to be regulated by the former gene. Specifically, the RNAP association with Site 2 results in chromatin modification, and after RNAP departure from this site the chromatin rapidly returns to the initial state. The initiation of the transcription of the downstream gene into mRNA is assumed to be possible only provided that the chromatin is modified, i.e., provided that Site 2 is occupied by RNAP.



(a) Induction of chromatine remodeling and histone modification.

(b) Transcriptional interference.

(c) Sites representing genes.

Fig. 23. Schemes of (a) positive and (b) negative regulation of the downstream gene (on the right) transcribed into mRNA by an upstream gene (on the left) transcribed into ncRNA (adapted from Ref. [51]). The straight arrows indicate the direction of the RNAP move. Panel (c) exhibits schematically the model used to describe the interference of the gene transcription shown in panels (a) and (b). Gene 1 transcribed into ncRNA is represented by Sites (circles) 1 and 2. Gene 2 transcribed into mRNA is represented by Sites (circles) 3 and 4. The pathways of RNAP are shown by arrows. The initiation of the transcription on Site 3 depends on the state of Site 2.

With the specification above, the gene transcription into ncRNA schematically includes reversible RNAP recruitment (the superscripts $\star$ and $*$ corresponds below to the situations when the DNA chain is locally closed and open, respectively),

$$\mathrm{RNAP} + \mathrm{S}_1 \rightleftharpoons \mathrm{RNAP}^{\star} \cdot \mathrm{S}_1, \tag{178}$$

transcription initiation via formation of an open DNA chain,

$$\mathrm{RNAP}^{\star} \cdot \mathrm{S}_1 \to \mathrm{RNAP}^{*} \cdot \mathrm{S}_1, \tag{179}$$

RNAP transition along DNA (this step results in the ncRNA elongation),

$$\mathrm{RNAP}^{*} \cdot \mathrm{S}_1 + \mathrm{S}_2 \to \mathrm{S}_1 + \mathrm{RNAP}^{*} \cdot \mathrm{S}_2, \tag{180}$$

and RNAP and ncRNA release,

$$\mathrm{RNAP}^{*} \cdot \mathrm{S}_2 \to \mathrm{RNAP} + \mathrm{ncRNA} + \mathrm{S}_2. \tag{181}$$

In this scheme the initial RNAP association with DNA (step (178)) is reversible, while the open-complex formation (step (179)) and the following steps of the RNA synthesis are irreversible, because these steps operate much like motor vehicles.

The gene transcription into mRNA occurs in analogy with the formation of ncRNA, i.e.,

$$\mathrm{RNAP} + \mathrm{S}_3 \rightleftharpoons \mathrm{RNAP}^{\star} \cdot \mathrm{S}_3, \tag{182}$$

$$\mathrm{RNAP}^{\star} \cdot \mathrm{S}_3 \to \mathrm{RNAP}^{*} \cdot \mathrm{S}_3, \tag{183}$$

$$\mathrm{RNAP}^{*} \cdot \mathrm{S}_3 + \mathrm{S}_4 \to \mathrm{S}_3 + \mathrm{RNAP}^{*} \cdot \mathrm{S}_4, \tag{184}$$

$$\mathrm{RNAP}^{*} \cdot \mathrm{S}_4 \to \mathrm{RNAP} + \mathrm{mRNA} + \mathrm{S}_4. \tag{185}$$

Step (183) is realized here provided that Site 2 is occupied by RNAP (this condition corresponds to the RNAP-induced chromatin modification near site 2).

Steps (178)–(185) can be described by using the mean-field approximation or Monte Carlo simulations. If steps (178)–(181) are slow, the Monte Carlo simulations predict stochastic bursts in the synthesis of mRNA (Fig. 24). The explanation of this effect is simple: the mRNA bursts are observed during long periods when Site 2 is occupied by RNAP.

8.7. *Transcriptional interference*

According to Martens et al. [225] (see also review [51]), the RNAP II transcription of the *Saccharomyces cerevisiae SER3* gene is tightly repressed during growth in rich medium. The regulatory region of this gene is highly transcribed under these conditions and produces ncRNA (*SRG*1). Expression of this ncRNA is required for repression of *SER3*. Additional experiments have demonstrated that the repression occurs by a transcription-interference mechanism. Specifically, the *SRG*1 transcription interferes with the binding of activators of the *SER*3 transcription and results in repression of *SER*3. This mode of the expression of two genes (Fig. 23(b)) can be described by using a model similar to that presented in the preceding section. In particular, steps (178)–(185) can be complemented by reversible association of regulatory protein (P) with Site 2,

$$\mathrm{P} + \mathrm{S}_2 \rightleftharpoons \mathrm{P} \cdot \mathrm{S}_2. \tag{186}$$

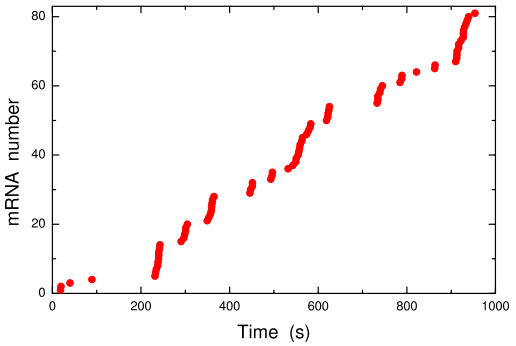


Fig. 24. Typical stochastic bursts in the the number of transcribed mRNAs according to scheme (178)–(185) in the case of slow steps (178)–(181).

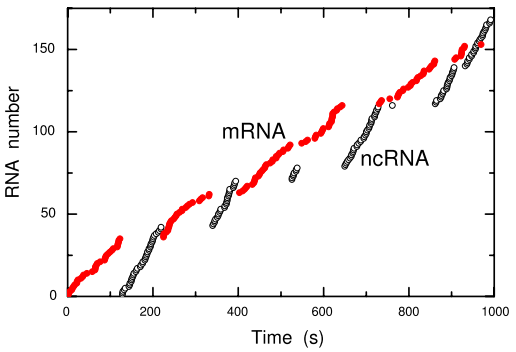


Fig. 25. Typical stochastic bursts in the the numbers of transcribed mRNAs and ncRNAs according to scheme (178)–(186) in the case of slow step (186).

This means that Site 2 can be vacant or occupied either by RNAP or by P, while the other sites (S_1, S_3 and S_4) can be either vacant or occupied by RNAP. P is considered to be an activator. In particular, the initiation of gene transcription into mRNA (step (183)) is here assumed to occur provided that Site 2 is occupied by P. In this case, the competition between RNAP and P for occupation of Site 2 results in the interference of gene transcription into ncRNA and mRNA.

Concerning stochastic effects in the model under consideration, we may notice that if the transcription into ncRNA is slow, each act of transcription may result in a stochastic burst in the mRNA synthesis. This effect (not shown) is similar to that discussed in the preceding section. A more interesting and novel effect can be observed if the protein association to and dissociation from a regulatory site is slow. In this case, acts of protein association and dissociation can be accompanied by sequential anti-phase bursts in the ncRNA and mRNA synthesis as shown in Fig. 25. The explanation of the bursts is simple: the mRNA bursts are observed if the regulatory site is occupied by a protein, while the ncRNA bursts occur if the regulatory site is vacant.

9. Spatial effects

The kinetic models described in Sections 4–8 are temporal. In this section, we will briefly discuss spatio-temporal effects in the formation and degradation of ncRNAs. Concerning this subject, it is appropriate to repeat that ncRNAs can participate in the spatial organization of eukaryotic cells (Section 2.4). The latter aspects of the ncRNA biochemistry are, however, beyond our scope.

9.1. Inside a single cell

Eukaryotic cells are highly heterogeneous, and every kinetic process occurring in such cells is spatio-temporal to some extent. The RNA and protein diffusion in cells is however rather rapid and often the gradients in the distribution of these species can be neglected. Although this approximation is very widely used in kinetic models of gene expression [36–39], its applicability is limited. For example, the penetration of RNAs and/or proteins via the membrane separating the nucleus and cytoplasm may be slow and the concentrations of these species in these regions may be different. If the RNA and protein diffusion inside the nucleus and cytoplasm is rapid, the temporal models remain nevertheless valid on the level of compartments (see, e.g., Ref. [226] and references therein). The rate of the RNA diffusion is however not always sufficient in order to maintain the well-mixed conditions. In particular, the experiments [227] indicate that the gradients in the mRNA

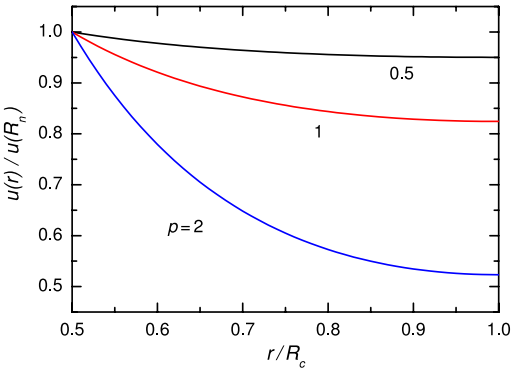


Fig. 26. Normalized ncRNA concentration in the cytoplasm as a function of r for $R_n/R_c = 0.5$ and $p \equiv (\eta c/D_*)^{1/2}R_c = 0.5$, 1, and 2. R_n is the nucleus radius. (According to Ref. [228].)

distribution can sometimes be significant even in prokaryotic cells. Monte Carlo simulations [107,108] of the mRNA–protein interplay with realistic parameters are also indicative that the gradients in the mRNA distribution can be significant.

To scrutinize the role of the ncRNA diffusion in the context of the formation and degradation of ncRNAs, let us consider the simplest situation when the cell is spherical, the ncRNA formation occurs in the center, the ncRNA degradation is related with association with the second reactant, R (mRNA or protein), and the gradients in the distribution of the latter specie are negligible [228]. In this case, the reaction–diffusion equation for the ncRNA concentration, c_*, is represented as

$$\frac{\partial c_*}{\partial t} = D_* \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial c_*}{\partial r} - \eta c_* c, \tag{187}$$

where D_* is the ncRNA diffusion coefficient, c is the R concentration, and η is the association rate constant.

The first boundary condition for Eq. (187) is obtained taking into account that the ncRNA diffusion flux calculated at $r \to 0$ should be equal to the rate of the ncRNA formation, u, i.e.,

$$-4\pi r^2 D_* \left.\frac{\partial c_*}{\partial r}\right|_{r\to 0} = u. \tag{188}$$

The second boundary condition takes into account that ncRNA does not penetrate the cell membrane, and accordingly the ncRNA diffusion flux should be equal to zero at $r = R_c$, where R_c is the cell radius, i.e.,

$$\left.\frac{\partial c_*}{\partial r}\right|_{r=R_c} = 0. \tag{189}$$

With the boundary conditions (188) and (189) under steady-state conditions, Eq. (187) can be elementarily solved [228], and one can obtain that the ncRNA gradients are appreciable (see, e.g., Fig. 26) provided that

$$\eta c R_c^2/D_* > 1. \tag{190}$$

In this condition, the R concentration can be replaced via the number of R copies in the cell as $c = N/V \equiv 3N/(4\pi R_c^3)$, where $V = 4\pi R_c^3/3$ is the cell volume, i.e.,

$$3\eta N/(4\pi D_* R_c) > 1. \tag{191}$$

The upper value of η is $4\pi\rho D$ (see Section 3.6), where ρ is the length comparable to the ncRNA or R size, and D is the R diffusion coefficient (R is considered to diffuse faster than ncRNA). Substituting this estimate for k_a into condition (191), we obtain

$$3\rho D N/(D_* R_c) > 1. \tag{192}$$

For estimates, one can consider that ρ is comparable with the protein radius, $\simeq 2\times 10^{-7}$ cm, and employ $R_c = 10^{-3}$ cm and $D/D_* = 5$. With these parameters, condition (192) holds provided that N exceeds 300. The latter condition can obviously be met.

9.2. Array of cells

As already noted in the introduction, ncRNAs play an important role already at the earliest stages of development of eukaryotes. The cell proliferation and differentiation during these states is governed by morphogens (transcription factors or ligands that bind cell-surface receptors) diffusing between cells. The gradients in the morphogen concentration induce

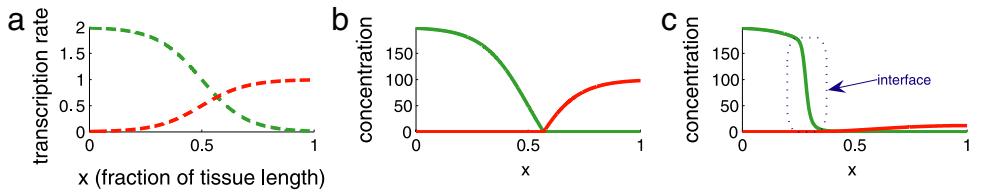


Fig. 27. (a) mRNA and miRNA transcription profiles. [(b) and (c)] mRNA and miRNA concentrations as a function of the normalized coordinate without and with miRNA diffusion, respectively. The synthesis of mRNA and miRNA occurs primarily on the left- and right-hand parts of the domain, respectively. *Source:* Adapted from Ref. [230].

the corresponding gradients in the target RNAs, i.e., the RNA populations may change along an array of cells. Recent experiments (see Ref. [229] and references therein) indicate than miRNAs may play a role of a morphogen. In particular, miRNAS may diffuse between cells (via active export or transport through gap junctions between neighboring cells) and induce or influence the gradients in the targeted mRNAs.

A generic 1D model describing the effect of the intercellular miRNA diffusion on the gradients in the mRNA concentration has been proposed by Levine et al. [230]. In their analysis, miRNA and mRNA are considered to be transcribed in a space-dependent manner. Association of miRNA and mRNA is considered to result in degradation of the miRNA–mRNA complex. The miRNA diffusion is described on the scale larger than the cell size by using the conventional reaction–diffusion equations. In particular, the equations for the miRNA and mRNA concentrations, c_* and c, are represented as

$$\frac{\partial c_*}{\partial t} = D_* \frac{\partial^2 c_*}{\partial x^2} + u(x) - k_* c_* - \eta c_* c, \tag{193}$$

$$\frac{\partial c}{\partial t} = w(x) - kc - \eta c_* c, \tag{194}$$

where D_* is the miRNA effective diffusion coefficient, $u(x)$ and $w(x)$ are the prescribed rates of the miRNA and mRNA formation, k_* and k are the rate constant of conventional degradation, and η is the association rate constant.

Eqs. (193) and (194) were integrated [230] under steady-state conditions inside a domain with the no-flux boundary conditions for miRNA (cf. Eq. (189)). The functions $u(x)$ and $w(x)$ (Fig. 27) were chosen so that the synthesis of miRNA and mRNA occurs primarily on the right- and left-hand parts of the domain, respectively. For rapid miRNA–ncRNA association in the absence of the intercellular miRNA diffusion, this model predicts (cf. Section 4.1) that there exist critical distance, x_c, defined by the condition $u(x_c) = w(x_c)$, so that the mRNA and miRNA concentrations are negligible at $x < x_c$ and $x > x_c$, respectively. With miRNA diffusion (this process is relatively slow on the scale larger than the cell size), this specie is able to migrate to the left for a while and to react there with mRNA, and accordingly the mRNA expression pattern becomes sharper at $x \simeq x_c$.

10. Conclusion

The discovery of numerous ncRNAs in eukaryotic cells is one of the main recent breakthroughs in cellular biology [49–58]. The progress in this field was explosive. Only a few years ago, one of the common lines of thinking in this area was that ncRNAs produced by pervasive transcription are non-functional and just transcriptional noise or spurious transcription generated by the RNA polymerase [53]. Now, however, the evidence of the important role of ncRNAs in various intracellular processes is fairly solid [49–58]. Already available experimental data indicate that ncRNAs influence in one or another way almost every cellular function. The corresponding epigenetic mechanisms often include association of ncRNAs with mRNAs or proteins. Many other pathways are also possible [49–51]. The important specifics of ncRNAs is that they can play a role of global regulators.

The experimentally obtained information about mechanistic details of the influence of ncRNAs on the eukaryotic cell machinery has made it possible to construct kinetic models describing various aspects of the ncRNA–mRNA–protein interplay. In our review, we have systematically presented such models. Almost all the mechanisms implied in the models are experimentally validated. On the other hand, detailed experimental kinetic measurements of the interaction of ncRNAs with mRNAs and/or proteins, including variation of governing parameters, are still rare. For this reason, the values of the parameters, employed in the models, usually do not correspond to specific systems. For the same reason, direct comparison of the model predictions and experimental observations has been hardly possible so far. There is no doubt, however, that detailed kinetics will soon be measured.

The models presented are based primarily on the temporal mean-field kinetic equations. The main attention has been paid to the corresponding kinetics occurring under steady-state conditions. The studies of transient, stochastic and/or spatio-temporal kinetics including ncRNAs are just beginning. These aspects will obviously attract much more attention already in the nearest future.

Depending on the type of the protein-mediated regulation of gene transcription, the available models predict under steady-state conditions either a unique steady state (with, e.g., a threshold-like behaviour) or bistability, oscillations, and

stochastic bursts. Although all these features are obviously to be expected in nonlinear feedback systems, the models described are instructive because they clarify the likely mechanistic details behind the kinetic complexity, contribute to the formation of a conceptual basis in the field under consideration, and may guide experiments.

Referring to the random Boolean networks, one could expect to find kinetic chaos as well. This feature was however not observed in the available models including ncRNAs.

Outlined together, the models presented allow us to classify the scenarios of kinetics of gene expression including ncRNAs. This is perhaps the most important outcome of this review.

Finally, we may articulate that many aspects of the kinetics of gene expression including ncRNAs are still open for further experimental and theoretical studies. In fact, we are here in the beginning of a long way with many bright and unexpected findings.

References

[1] B. Alberts, A. Johnson, J. Lewis, et al., Molecular Biology of the Cell, Garland, New York, 2002.
[2] R.E. Green, J. Krause, A.W. Briggs, et al., Science 328 (2010) 710.
[3] O. Avery, C. Macleod, M. Mccarty, J. Exp. Med. 79 (1944) 137.
[4] J.D. Watson, F.H.C. Crick, Nature 171 (1953) 737.
[5] F. Crick, Nature 227 (1970) 561.
[6] K.M. Herbert, W.J. Greenleaf, S.M. Block, Annu. Rev. Biochem. 77 (2008) 149.
[7] J.T. Wade, K. Struhl, Curr. Opin. Genet. Developm. 18 (2008) 130.
[8] N.J. Fuda, M.B. Ardehali, J.T. Lis, Nature 461 (2009) 186.
[9] G.L. Hager, J.G. Mcnally, T. Misteli, Molec. Cell 35 (2009) 741.
[10] J. Frank, C.M.T. Spahn, Rep. Progr. Phys. 69 (2006) 1383.
[11] A. Korostelev, D.N. Ermolenko, H.F. Noller, Curr. Opin. Chem. Biol. 12 (2008) 674.
[12] R.A. Marshall, C.E. Aitken, M. Dorywalska, J.D. Puglisi, Ann. Rev. Biochem. 77 (2008) 177.
[13] J.B. Munro, K.Y. Sanbonmatsu, C.M.T. Spahn, S.C. Blanchard, Trends Biochem. Sci. 34 (2009) 390.
[14] R. de Sousa Abreu, L.O. Penalva, E.M. Marcotte, C. Vogel, Mol. Biosyst. 5 (2009) 1512.
[15] J. Houseley, D. Tollervey, Cell 136 (2009) 763.
[16] N.L. Garneau, J. Wilusz, C.J. Wilusz, Nature Rev. Mol. Cell Biol. 8 (2007) 113.
[17] C.E. Holt, S.L. Bullock, Science 326 (2009) 1212.
[18] H.D. Lipshitz, Nature Rev. Mol. Cell Biol. 10 (2009) 509.
[19] L. Shapiro, H.H. Mcadams, R. Losick, Science 326 (2009) 1225.
[20] C. Furusawa, K. Kaneko, Phys. Rev. Lett. 90 (2003) 088102.
[21] F. Julicher, R. Bruinsma, Biohphys. J. 74 (1998) 1169.
[22] M.A. Gibson, J. Bruck, in: J.M. Bower, H. Bolouri (Eds.), Computational Modeling of Genetic and Biochemical Networks, MIT Press, Cambridge, 2001, pp. 49–71.
[23] H.J. Woo, Phys. Rev. E 74 (2006) 011907.
[24] L. Bai, R.M. Fulbright, M.D. Wang, Phys. Rev. Lett. 98 (2007) 068103.
[25] M. Djordjevic, R. Bundschuh, Biophys. J. 94 (2008) 4233.
[26] M. Voliotis, N. Cohen, C. Molina-Paris, T.B. Liverpool, Biophys. J. 94 (2008) 334.
[27] T. Tripathi, D. Chowdhury, Phys. Rev. E 77 (2008) 011921.
[28] Y.R. Yamada, C.S. Peskin, Biophys. J. 96 (2009) 3015.
[29] A.S. Ribeiro, O.-P. Smolander, T. Rajala, et al., J. Comput. Biol. 16 (2009) 539.
[30] A.S. Ribeiro, A. Häkkinen, H. Mannerström, et al., Phys. Rev. E 81 (2010) 011912.
[31] V.P. Zhdanov, Phys. Rev. E 80 (2009) 051925.
[32] T. Chou, G. Lakatos, Phys. Rev. Lett. 93 (2004) 198101.
[33] H. Zouridis, V. Hatzimanikatis, Biophys. J. 92 (2007) 717.
[34] A. Garai, D. Chowdhury, D. Chowdhury, T.V. Ramakrishnan, Phys. Rev. E 80 (2009) 011908.
[35] L. Ciandrini, I. Stansfield, M.C. Romano, Phys. Rev. E. 81 (2010) 051904.
[36] M. Kaern, T.C. Elston, W.J. Blake, J.J. Collins, Nature Rev. Genet. 6 (2005) 451.
[37] J. Paulsson, Phys. Life Rev. 2 (2005) 157.
[38] D. Kulasiri, L.K. Nguyen, S. Samarasinghe, Z. Xie, Curr. Bioinf. 3 (2008) 197.
[39] J. Rausenberger, C. Fleck, J. Timmer, M. Kollmann, Progr. Biophys. Molec. Biol. 100 (2009) 57.
[40] A.S. Ribeiro, Math. Biosci. 223 (2010) 1.
[41] G. Tiana, S. Krishna, S. Pigolotti, et al., Phys. Biol. 4 (2007) R1.
[42] D. Lepzelter, H. Feng, J. Wang, Chem. Phys. Lett. 490 (2010) 216.
[43] X. Zhu, M. Gerstein, M. Snyder, Genes. Dev. 21 (2007) 1010.
[44] G. Karlebach, R. Shamir, Nature Rev. Molec. Cell Biol. 9 (2008) 771.
[45] Y. Bar-Yam, D. Harmon, B.de. Bivort, Science 323 (2009) 1016.
[46] M.C. Lagomarsino, B. Bassetti, G. Castellani, D. Remondini, Molec. Biosyst. 5 (2009) 335.
[47] D.F.T. Veiga, B. Dutta, G. Balazsi, Molec. Biosyst. 6 (2010) 469.
[48] International human genome sequencing consortium, Nature 431 (2004) 931.
[49] T.R. Mercer, M.E. Dinger, J.S. Mattick, Nature Rev. Genet. 10 (2009) 155.
[50] J. Whitehead, G.K. Pandey, C. Kanduri, Biochim. Biophys. Acta 1790 (2009) 936.
[51] J.E. Wilusz, H. Sunwoo, D.L. Spector, Gen. Devel. 23 (2009) 1494.
[52] C.P. Ponting, P.L. Oliver, W. Reik, Cell 136 (2009) 629.
[53] F.F. Costa, Bioess. 32 (2010) 599.
[54] D.P. Bartel, Cell 136 (2009) 215.
[55] M. Ghildiyal, P.D. Zamore, Nature Rev. Genet. 10 (2009) 94.
[56] V.N. Kim, J. Han, M.C. Siomi, Nature Rev. Molec. Cell. Biol. 10 (2009) 126.
[57] J. Krol, I. Loedige, W. Filipowicz, Nature Rev. Genet. 11 (2010) 597.
[58] K. Breving, A. Esquela-Kerscher, Intern. J. Biochem. Cell Biol. 42 (2010) 1316.
[59] C.L. Beisel, G. Storz, FEMS Microbiol. Rev. 34 (2010) 866.
[60] H. Guo, N.T. Ingolia, J.S. Weissman, D.P. Bartel, Nature 466 (2010) 835.
[61] K.K.H. Farh, A. Grimson, C. Jan, et al., Science 310 (2005) 1817.
[62] K.M. Foshay, G.I. Gallicano, Develop. Biol. 326 (2009) 431.
[63] J.B. Johnnidis, M.H. Harris, R.T. Wheeler, et al., Nature 451 (2008) 1125.

[64] B.M. Stadler, H. Ruohola-Baker, Cell 132 (2008) 563.
[65] U. Lakshmipathy, J. Davila, P.R. Hart, Regen. Medic. 5 (2010) 545.
[66] F.C. Lynn, Trends Endocrin. Metabol. 20 (2009) 452.
[67] E. Vreugdenhil, E. Berezikov, Front. Neuroendocrin. 31 (2010) 128.
[68] J. Gao, W.-Y. Wang, Y.W. Mao, et al., Nature 466 (2010) 1105.
[69] E. Barbarotto, T.G. Schmittgen, G.A. Calin, Internat. J. Cancer 122 (2008) 969.
[70] R. Baffa, M. Fassan, S. Volinia, et al., J. Pathol. 219 (2009) 214.
[71] A.G. Seto, Intern. J. Biochem. Cell Biol. 42 (2010) 1298.
[72] D. Cazalla, T. Yario, J. Steitz, Science 328 (2010) 1563.
[73] Y. Dai, X.-F. Zhou, Open Acc. Bioinform. 2 (2010) 29.
[74] D.R. Rigney, W.C. Schieve, J. Theor. Biol. 69 (1977) 761.
[75] O.G. Berg, J. Theor. Biol. 73 (1978) 587.
[76] H.-M. Chan, L.-S. Chan, R.N.-S. Wong, H.-W. Li, Anal. Chem. 82 (2010) 6911.
[77] K.F. Erhard, J.L. Stonaker, S.E. Parkinson, et al., Science 323 (2009) 1201.
[78] S.A. Shabalina, A.Y. Ogurtsov, 1.A.N. Spiridonov, et al., Mol. Biol. Evol. 27 (2010) 1745.
[79] S.V. Sharma, D.Y. Lee, B.H. Li, et al., Cell 141 2010 (2010) 69.
[80] E. Kim, A. Goren, G. Ast, BioEssays 30 (2007) 38.
[81] O. Yazgan, J.E. Krebs, Biochem. Cell Biol. 85 (2007) 484.
[82] P. Gammell, Cytotechn. 53 (2007) 55.
[83] S.L. Lamersa, G.B. Fogel, M.S. Mcgrath, Biosyst. 101 (2010) 88.
[84] D.F. Browning, S.J. Busby, Nature Rev. Microbiol. 2 (2004) 57.
[85] W. Ross, R.L. Gourse, Methods 47 (2009) 13.
[86] S.P. Haugen, W. Ross, R.L. Gourse, Nature Rev. Microbiol. 6 (2008) 507.
[87] G.K. Ackers, A.D. Johnson, M.A. Shea, Proc. Natl. Acad. Sci. USA 79 (1982) 1129.
[88] G.K. Ackers, M.A. Shea, F.R. Smith, J. Mol. Biol. 170 (1983) 223.
[89] L. Bintu, N.E. Buchler, H.G. Garcia, et al., Curr. Opin. Genet. Devel. 15 (2005) 116.
[90] T. Kuhlman, Z. Zhang, M.H. Saier, T. Hwa, Proc. Natl. Acad. Sci. USA 104 (2007) 6043.
[91] T. Gedeon, K. Mischaikow, K. Patterson, E. Traldid, Bull. Math. Biol. 70 (2008) 1660.
[92] D. Müller, J. Stelling, PLoS Comput. Biol. 5 (2009) e1000279.
[93] A. Coulon, O. Gandrillon, G. Beslon, BMC Syst. Biol. 4 (2010) 2.
[94] R. Murugan, J. Phys. A 43 (2010) 195003.
[95] D. Michel, Prog. Biophys. Mol. Biol. 102 (2010) 16.
[96] R.C. Lee, R.L. Feinbaum, V. Ambros, Cell 75 (1993) 843.
[97] B. Wightman, I. Ha, G. Ruvkun, Cell 75 (1993) 855.
[98] S. Vasudevan, Y.C. Tong, J.A. Steitz, Science 318 (2007) 1931.
[99] D.H. Kim, P. Saetrom, O. Snove, J.J. Rossi, Proc. Natl. Acad. Sci. USA 105 (2008) 16230.
[100] Y. Okazaki, M. Furuno, T. Kasukawa, et al., Nature 420 (2002) 563.
[101] D.J. Wilkinson, Nature Rev. Genet. 10 (2009) 122.
[102] S.A. Kauffman, J. Theoret. Biol. 22 (1969) 437.
[103] D.T. Gillespie, J. Phys. Chem. 113 (2000) 297.
[104] D.T. Gillespie, J. Phys. Chem. 81 (1977) 2340.
[105] K. Binder, in: K. Binder (Ed.), Monte Carlo Methods in Statistical Physics, Springer, Berlin, 1979, p. 1.
[106] D.P. Landau, K. Binder, A Guide to Monte Carlo Simulations in Statistical Physics, Cambridge University Press, Cambridge, 2000.
[107] V.P. Zhdanov, JETP Lett. 85 (2007) 302.
[108] V.P. Zhdanov, Chem. Phys. Lett. 456 (2008) 247.
[109] X. Darzacq, Y. Shav-Tal, V. de Turris, et al., Nature Str. Mol. Biol. 14 (2007) 796.
[110] E. Yang, E. Van nimwegen, M. Zavolan, et al., Genome Res. 13 (2003) 1863.
[111] A. Beyer, J. Hollunder, H.-P. Nasheuer, T. Wilhelm, Mol. Cell. Proteom. 3 (2004) 1083.
[112] S.A. Rice, Diffusion-Limited Reactions, Elsevier, Amsterdam, 1985.
[113] T. Misteli, Histochem. Cell Biol. 129 (2008) 5.
[114] S. Baumgartner, I.M. Tolic-Norrelykke, Biophys. J. 96 (2009) 4336.
[115] V.P. Zhdanov, D. Steel, B. Kasemo, J. Gold, Phys. Chem. Chem. Phys. 7 (2005) 3496.
[116] A. Sigal, R. Milo, A. Cohen, et al., Nature 444 (2006) 643.
[117] E. Lundberg, M. Gry, P. Oksvold, et al., J. Proteomics 71 (2008) 448.
[118] A. Csikász-Nagy, D. Battogtokh, K.C. Chen, et al., Biophys. J. 90 (2006) 4361.
[119] B. Pfeuty, K. Kaneko, Phys. Biol. 4 (2007) 194.
[120] K. Sriram, G. Bernot, F. Képés, IET Syst. Biol. 1 (2007) 326.
[121] I. Mura, A. Csikász-Nagy, J. Theor. Biol. 254 (2008) 850.
[122] D.H. Lenz, K.C. Mok, B.N. Lilley, et al., Cell 118 (2004) 69.
[123] E. Levine, Z. Zhang, T. Kuhlman, T. Hwa, PLoS Biol. 5 (2007) e229.
[124] E. Levine, T. Hwa, Curr. Opin. Microbiol. 11 (2008) 574.
[125] S. Semsey, A.M.C. Andersson, S. Krishna, et al., Nucl. Acid. Res. 34 (2006) 4960.
[126] N. Mitarai, A.M.C. Andersson, S. Krishna, et al., Phys. Biol. 4 (2007) 164.
[127] Y. Shimoni, G. Friedlander, G. Hetzroni, et al., Molec. Syst. Biol. 3 (2007) 138.
[128] S. Legewie, D. Dienst, A. Wilde, et al., Biophys. J. 95 (2008) 3232.
[129] P. Mehta, S. Goyal, N.S. Wingreen, Molec. Syst. Biol. 4 (2008) 221.
[130] N. Mitarai, J.-A.M. Benjamin, S. Krishna, et al., Proc. Natl Acad. Sci. USA 106 (2009) 10655.
[131] Y. Jia, W.H. Liu, A.B. Li, et al., Biophys. Chem. 143 (2009) 60.
[132] V. Elgart, T. Jia, R. Kulkarni, Biophys. J. 98 (2010) 2780.
[133] V. Elgart, T. Jia, R.V. Kulkarni, Phys. Rev. E 82 (2010) 021901.
[134] E. Levine, E.B. Jacob, H. Levine, Biophys. J. 93 (2007) L52.
[135] D.E. Kuhn, M.M. Martin, D.S. Feldman, et al., Methods 44 (2008) 47.
[136] L. Li, J.Z. Xu, D. Ying, et al., Mamm. Genome 21 (2010) 1.
[137] V.P. Zhdanov, Molec. Biosyst. 5 (2009) 638.
[138] M. Selbach, B. Schwanhausser, N. Thierfelder, et al., Nature 455 (2008) 58.
[139] D. Baek, J. Villen, C. Shin, et al., Nature 455 (2008) 64.
[140] L. Mariani, M. Löhning, A. Radbruch, T. Höfer, Prog. Biophys. Mol. Biol. 86 (2004) 45.
[141] V.P. Zhdanov, Chem. Phys. Lett. 424 (2006) 394.
[142] V.P. Zhdanov, Biophys. Rev. Lett. 5 (2010) 89.
[143] L.P. Xiong, Y.-Q. Ma, L.H. Tang, Chin. Phys. Lett. 27 (2010) 098701.
[144] V.P. Zhdanov, Physica A 389 (2010) 887.

[145] T.S. Gardner, C.R. Cantor, J.J. Collins, Nature 403 (2000) 339.
[146] J.L. Cherry, F.R. Adler, J. Theor. Biol. 203 (2000) 117.
[147] H.H. Chang, P.Y. Oh, D.E. Ingber, S. Huang, BMC Cell Biol. 7 (2006) Art. No. 11.
[148] V.P. Zhdanov, J. Phys. A 42 (2009) 065102.
[149] D. Bratsun, D. Volfson, L.S. Tsimring, J. Hasty, Proc. Natl. Acad. Sci. USA 102 (2005) 14593.
[150] K. Rateitschak, O. Wolkenhauer, Math. Biosci. 205 (2007) 163.
[151] V.P. Zhdanov, JETP Lett. 86 (2007) 492.
[152] S. Krishna, S. Semsey, M.H. Jensen, Phys. Biol. 6 (2009) 036009.
[153] V.P. Zhdanov, J. Exp. Theor. Phys. 108 (2009) 1050.
[154] Z.-R. Xie, H.-T. Yang, W.-C. Liu, M.-J. Hwang, Biochem. Biophys. Res. Comm. 358 (2007) 722.
[155] J.W. Shen, Z.R. Liu, W.X. Zheng, et al., Physica A 388 (2009) 2995.
[156] A. Nandi, C. Vaz, A. Bhattacharya, R. Ramaswamy, BMC Syst. Biol. 3 (2009) 45.
[157] V.P. Zhdanov, J. Phys. A 41 (2008) 285101.
[158] V.P. Zhdanov, Biosyst. 95 (2009) 75.
[159] V.P. Zhdanov, JETP Lett. 88 (2008) 466.
[160] V. Chickarmane, C. Troein, U.A. Nuber, et al., PLOS Comput. Biol. 2 (2006) 1080.
[161] T. Kalmar, C. Lim, P. Hayward, et al., PLOS Biol. 7 (2009) e1000149.
[162] S. Huang, Y.P. Guo, G. May, T. Enver, Develop. Biol. 305 (2007) 695.
[163] L. Wang, B.L. Walker, S. Iannaccone, et al., Proc. Natl Acad. Sci. USA 106 (2009) 6638.
[164] D. Schittler, J. Hasenauer, F. Allgower, S. Waldherr, Chaos 20 (2010) 045121.
[165] R.E. Callard, Immun. Cell. Biol. 85 (2007) 300.
[166] H.-J. Van den ham, R.J.de. Boer, Intern. Immun. 20 (2008) 1269.
[167] V. Chickarmane, T. Enver, C. Peterson, PLoS Comput. Biol. 5 (2009) e1000268.
[168] S. Bhattacharya, R.B. Conolly, N.E. Kaminski, et al., Toxicol. Sci. 115 (2010) 51.
[169] Q. Zhang, S. Bhattacharya, D.E. Kline, et al., BMC Syst. Biol. 4 (2010) 40.
[170] K. Lai, M.J. Robertson, D.V. Schaffer, Biophys. J. 86 (2004) 2748.
[171] B.D. Macarthur, C.P. Please, R.O.C. Oreffo, PLoS One 3 (2008) e3086.
[172] R. Guantes, J.F. Poyatos, PLoS Comput. Biol. 4 (2008) e1000235.
[173] J. Feng, C. Bi, B.S. Clark, et al., Gen. Develop. 20 (2006) 1470.
[174] R.J. Johnston, S. Chang, J.F. Etchberger, et al., Proc. Natl Acad. Sci. USA 102 (2005) 12449.
[175] A.S. Flynt, E.C. Lai, Nature Rev. Genet. 9 (2008) 831.
[176] V.P. Zhdanov, Chem. Phys. Lett. 458 (2008) 359.
[177] B.D. Aguda, Y. Kim, M.G. Piper-Hunter, et al., Proc. Nat. Acad. Sci. USA 105 (2008) 19678.
[178] S. Nikolov, J. Vera, U. Schmitz, O. Wolkenhauer, Theory Biosci. (2011), in press (doi:10.1007/s12064-010-0109-5).
[179] R.S. Wang, G. Jin, X.S. Zhang, L.N. Chen, BMC Bioinformatics 10 (Suppl. 4) (2009) S6.
[180] V.P. Zhdanov, Cent. Eur. J. Phys. 8 (2010) 864.
[181] V.P. Zhdanov, Chaos 20 (2010) 045112.
[182] V.P. Zhdanov, Cent. Eur. J. Phys. (2011), in press (doi:10.2478/s11534-010-0104-y).
[183] H.-W. Ma, J. Buer, A.-P. Zeng, BMC Bioinform. 5 (2004) 199.
[184] A. Martynez-Antonio, S.C. Janga, D. Thieffry, J. Mol. Biol. 381 (2008) 238.
[185] N.J. Martinez, M.C. Ow, M. Inmaculada Barrasa, et al., Gen. Devel. 22 (2008) 2535.
[186] V. Vermeirssen, A. Joshi, T. Michoel, et al., Molec. Biosyst. 5 (2009) 1817.
[187] B. Derrida, Y. Pomeau, Europhys. Lett. 1 (1986) 45.
[188] M. Andrecut, S.A. Kauffman, Phys. Lett. A 372 (2008) 4757.
[189] D.M. Wittmann, C. Marr, F.J. Theis, J. Theor. Biol. 266 (2010) 436.
[190] V.P. Zhdanov, JETP Lett. 2011 (in press).
[191] K.C. Kao, L.M. Tran, J.C. Liao, J. Biol. Chem. 280 (2005) 36079.
[192] S. Tsuru, J. Ichinose, A. Kashiwagi, et al., Phys. Biol. 6 (2009) 036015.
[193] V.P. Zhdanov, Biophys. Rev. Lett. (2011) (in press).
[194] V.P. Zhdanov, J. Exp. Theor. Phys. 111 (2010) 699.
[195] R.J. Cho, M. Huang, M.J. Campbell, et al., Nature Genet. 27 (2001) 48.
[196] M.L. Whitfield, G. Sherlock, A.J. Saldanha, et al., Molec. Biol. Cell 13 (2002) 1977.
[197] J.Q. Yin, R.C. Zhao, K.V. Morris, Trends Biotechn. 26 (2008) 70.
[198] Y. Liang, D. Ridzon, L. Wong, C. Chen, BMC Genomics 8 (2007) 166.
[199] J.R. Chubb, T.B. Liverpool, Curr. Opin. Genet. Develop. 20 (2010) 478.
[200] J.E. Lisman, Trends Neurosci. 20 (1997) 38.
[201] R. Losick, C. Desplan, Science 320 (2008) 65.
[202] A. Roesch, M. Fukunaga-Kalabis, E.C. Schmidt, et al., Cell 141 (2010) 583.
[203] N.Q. Balaban, J. Merrin, R. Chait, et al., Science 305 (2004) 1622.
[204] T.B. Kepler, T.C. Elston, Biophys. J. 81 (2001) 3116.
[205] R. Karmakar, I. Bose, Phys. Biol. 1 (2004) 197.
[206] J.M. Raser, E.K. O'Shea, Science 304 (2004) 1811.
[207] J.E.M. Hornos, D. Schultz, G.C.P. Innocentini, et al., Phys. Rev. E 72 (2005) 051907.
[208] Y. Morishita, T.J. Kobayashi, K. Aihara, J. Theor. Biol. 235 (2005) 241.
[209] A.M. Walczak, J.N. Onuchic, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 102 (2005) 18926.
[210] A. Lipshtat, A. Loinger, N.Q. Balaban, O. Biham, Phys. Rev. Lett. 96 (2006) 188101.
[211] A. Loinger, A. Lipshtat, N.Q. Balaban, O. Biham, Phys. Rev. E 75 (2007) 021904.
[212] Y. Takenaka, H. Nagahara, H. Kitahata, K. Yoshikawa, Phys. Rev. E 77 (2008) 031905.
[213] S. Iyer-Biswas, F. Hayot, C. Jayaprakash, Phys. Rev. E 79 (2009) 031911.
[214] J.Z. Lei, J. Theor. Biol. 256 (2009) 485.
[215] Q. Liu, Y. Jia, Phys. Rev. E 70 (2004) 041907.
[216] Y. Okabe, Y. Yagi, M. Sasai, J. Chem. Phys. 127 (2007) 105107.
[217] V.P. Zhdanov, JETP Lett. 84 (2007) 632.
[218] P. Mehta, R. Mukhopadhyay, N.S. Wingreen, Phys. Biol. 5 (2008) 026005.
[219] B. Pfeuty, K. Kaneko, Phys. Biol. 6 (2009) 046013.
[220] T. Jia, R.V. Kulkarni, Phys. Rev. Lett. 105 (2010) 018101.
[221] M. Dobrzynski, F.J. Bruggeman, Proc. Natl. Acad. Sci. USA 106 (2009) 2583.
[222] I. Martianov, A. Ramadass, A.S. Barros, et al., Nature 445 (2007) 666.
[223] V.P. Zhdanov, JETP Lett. 92 (2010) 410.
[224] K. Hirota, T. Miyoshi, K. Kugou, et al., Nature 456 (2008) 130.
[225] J.A. Martens, L. Laprade, F. Winston, Nature 429 (2004) 571.
[226] V.P. Zhdanov, J. Exp. Theor. Phys. 108 (2009) 1050.
[227] P.M. Llopis, A.F. Jackson, O. Sliusarenko, et al., Nature 466 (2010) 77–82.
[228] V.P. Zhdanov, Biophys. Rev. Lett. 4 (2009) 267.
[229] D.H. Chitwood, F.T.S. Nogueira, M.D. Howell, et al., Gen. Devel. 23 (2009) 549.
[230] E. Levine, P. Mchale, H. Levine, PLoS Comput. Biol. 3 (2007) e233.