PCCP

ROYAL SOCIETY OF CHEMISTRY

PERSPECTIVE

Cite this: *Phys. Chem. Chem. Phys.*, 2021, **23**, 21399

Understanding the molecular mechanisms of transcriptional bursting

Alena Klindziuk[a] and Anatoly B. Kolomeisky *[b]

Received 9th August 2021,
Accepted 13th September 2021

DOI: 10.1039/d1cp03665c

rsc.li/pccp

In recent years, it has been experimentally established that transcription, a fundamental biological process that involves the synthesis of messenger RNA molecules from DNA templates, does not proceed continuously as was expected. Rather, it exhibits a distinct dynamic behavior of alternating between productive phases when RNA molecules are actively synthesized and inactive phases when there is no RNA production at all. The bimodal transcriptional dynamics is now confirmed to be present in most living systems. This phenomenon is known as transcriptional bursting and it attracts significant amounts of attention from researchers in different fields. However, despite multiple experimental and theoretical investigations, the microscopic origin and biological functions of the transcriptional bursting remain unclear. Here we discuss the recent developments in uncovering the underlying molecular mechanisms of transcriptional bursting and our current understanding of them. Our analysis presents a physicochemical view of the processes that govern transcriptional bursting in living cells.

1 Introduction

It is well known that biological cells store their genetic material in DNA molecules and that the information contained in the DNA is transferred *via* a process called transcription.[1–3] From the chemical point of view, this process involves special enzymatic molecules, known as RNA polymerases (RNAPs), that catalyze the synthesis of messenger RNA molecules.[4] An RNAP sequentially translocates along the specific DNA segment, known as a gene, producing RNA copies complementary to this DNA segment. Several other types of protein molecules also participate in this process to ensure that the RNAPs start on time from the correct position on the DNA strand (beginning of the to-be-transcribed gene) and that the transcription process terminates at the right position (end of the transcribed gene).[1,2,4–6] Because of the fundamental importance of transcription in the functioning of all living systems, it has been intensively investigated for many years using a wide spectrum of experimental and theoretical methods.[5–10] However, there are still many unanswered questions about the molecular mechanisms that govern transcription and related processes.[8]

Messenger RNA molecules produced during transcription are utilized in the synthesis of corresponding proteins that are needed to maintain the operation of the living organisms. However, there are cellular processes that actively degrade RNA, requiring a constant supply of newly synthesized species.[11] Based on these arguments, it was generally assumed that transcription is more or less a continuous process of RNA synthesis. Surprisingly, recent experimental studies that measured with high temporal and spatial resolutions single-cell transcription dynamics revealed that this process is very noisy and discontinuous: periods with active RNA synthesis alternate with the periods of no RNA production.[7,12–14] Importantly, in these experiments, the transcription dynamics was monitored at the single-cell level, avoiding problems with ensemble-averaging measurements.[15] This observed phenomenon of discontinuous transcription dynamics is called the transcriptional bursting. It attracted the attention of researchers from different fields and prompted them to try to uncover the microscopic origin and biological relevance of these observations.[6,9,10,15,16] Some progress has been achieved in recent years, allowing for a better understanding of molecular mechanisms and regulation of the transcription process.[9,10] However, many questions remain unanswered.

In this article, we discuss the underlying processes that lead to bursting dynamics in transcription. We do not aim to present a comprehensive description of transcriptional bursting accounting for all results and observations in this field since there are already several recent excellent reviews that cover most biological, biochemical and biophysical aspects of this fascinating phenomenon.[9,10,15] Rather, we would like to present a slightly different, physicochemical, view of the microscopic mechanisms behind transcriptional bursting focusing on several of its most important features and discussing the physics underlying the observations. Our goal is to stimulate more discussions and studies on the microscopic origin of transcriptional bursting that would better clarify its biological role.

[a] *Department of Chemistry, Center for Theoretical Biological Physics and Applied Physics Graduate Program, Rice University, Houston, TX 77005-1892, USA. E-mail: ak62@rice.edu*

[b] *Department of Chemistry, Department of Physics and Astronomy, Department of Chemical and Biomolecular Engineering and Center for Theoretical Biological Physics, Rice University, Houston, TX 77005-1892, USA. E-mail: tolya@rice.edu*

2 Experimental observations

A significant advance in our understanding of the transcriptional process occurred after the developments in the single-molecule fluorescent microscopy methods enabled researchers to monitor and count the produced RNA molecules with unprecedented temporal and spatial resolutions.[7,12,14,17–21] In these experiments, the newly synthesized messenger RNA molecules were labeled by chemically associating them to fluorescent proteins. Then, measuring the fluorescence signal, which is proportional to the number of messenger RNAs in the system, allowed researchers to directly observe the appearance and disappearance of the RNA molecules as a function of time. An example of such investigation on mammalian cells is presented in Fig. 1.[19] One can see that, although all these cells are genetically identical, they exhibit a very heterogeneous behavior in the production of RNA molecules. This leads to a broad distribution in the number of RNA transcripts per cell. The analysis of these experimental results suggested that this is a result of the transcriptional bursting, *i.e.*, the production of RNA was not a continuous process. Similar observations have also been reported for various bacterial and yeast systems, suggesting that transcriptional bursting is a universal common phenomenon in living cells.[6,7,12–14,21,22]

Furthermore, the fluorescent measurements of transcription dynamics provided a direct way to visualize its bursting behavior.[23] This is because the fluorescence signal from the produced RNA is proportional to its amount in the system and as soon as the RNA is degraded its signal disappears. Fig. 2 shows the very noisy real-time dynamics of the production of messenger RNA molecules in a *in vitro* single-molecule assay.[23] Interestingly, the times when the transcription is ON ($\sim$100 seconds) are significantly smaller than the times when the transcription is OFF ($\sim$1000 seconds). Although these measurements were made *in vitro*, similar results have been also found *in vivo*.[7] These observations raised questions on the efficiency of the transcriptional process and the purpose of such stochastic behavior in biological systems. One would expect that to maintain the cellular processes the production of RNA molecules should always be active. But these experimental results do not agree with such arguments.

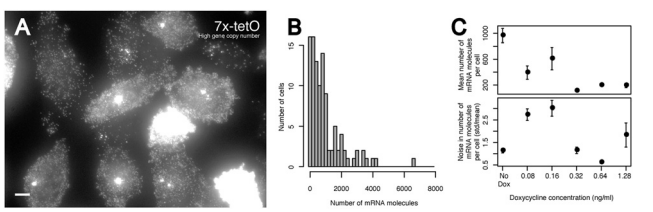

Fig. 1 Observation of variations in the produced messenger RNA number in genetically identical mammalian cells from the fluorescence signal of labeled RNA molecules. (A) A view of cells with different levels of fluorescence due to different numbers of synthesized RNA. (B) Distribution of produced RNA molecules per cell. (C) Mean (top) and normalized standard deviation (bottom) of the distribution at different experimental conditions. The figure is reproduced with permission from ref. 19.

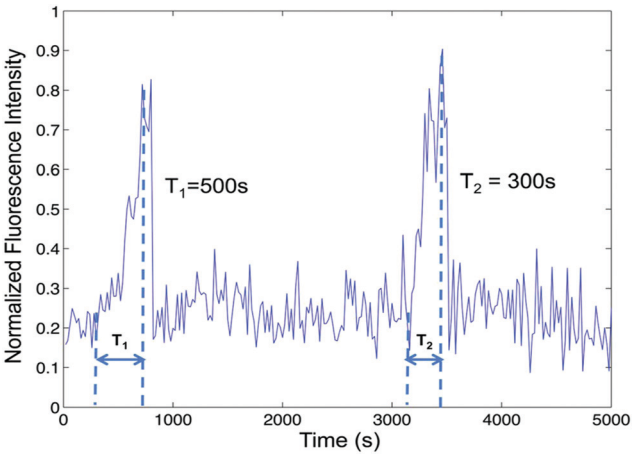

Fig. 2 Observation of fluctuations in the fluorescence signal that measures the amount of the produced RNA transcripts in single-molecule experiments *in vitro*. The figure is adopted with permission from ref. 23.

3 Theoretical analysis

Experimental observations of transcriptional bursting at the single-cell level stimulated extensive theoretical efforts to clarify its underlying mechanisms.[9,15,16,23–33] Many aspects of transcriptional dynamics have been discussed and are much better understood now.[7,9,10] But at the same time, some features of the transcriptional bursting remain unexplained. In this article, we would like to concentrate on two topics that, as we subjectively believe, should help better understand the molecular mechanisms and biological relevance of transcriptional bursting.

The first topic that we discuss is related to the question of how many states are needed to properly describe transcriptional bursting.[9] It is unclear how to define a state in a system where multiple biochemical reactions are taking place. In addition, we address the problem of how to determine the optimal set of states to describe this complex process. The second discussed topic focuses on the molecular origins of transcriptional bursting. In other words, we try to answer the questions of what microscopic processes lead to the appearance of these states, what specifically determines the number of states, and what is the biological relevance of these occurrences.

3.1 Multi-state theoretical models

The production of messenger RNA molecules involves multiple chemical reactions such as the binding and dissociation of various proteins to and from DNA chains, the addition of nucleotides to growing RNA molecules, RNA degradation and many others.[1,2] Because of the stochastic nature of the underlying chemical processes, it is expected that the overall synthesis of RNA should be quite a noisy process. But why do the experimental observations of transcriptional bursting imply that a "multi-state" description is needed? And how many states of transcription are there?

To begin answering these questions, let us start with the simplest minimal theoretical model of RNA production presented in Fig. 3a.[34] In this model, the RNA is continuously synthesized with a rate α and it

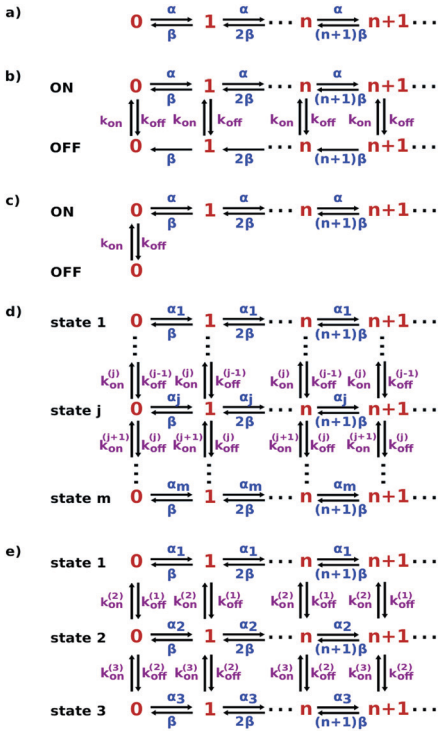

Fig. 3 Various multi-state kinetic models used to describe transcriptional bursting. (a) one-state model, (b) two-state model, (c) "Poisson with zero spike" model,[23] d) general multi-state model, and (e) three-state model. Reprinted with permission from the *J. Phys. Chem. B*, 2018, **122**, 11969–11977. Copyright 2018 American Chemical Society.

is also degraded with a first-order rate constant β. Each chemical state $n = 0, 1, 2,\ldots$ corresponds to having exactly nRNA molecules in the system. We are interested in the stationary state dynamics that can be reached at large times ($t \to \infty$). It can be shown that in this limit most dynamical properties of the system depend only on the parameter $x = \alpha/\beta$, which can be viewed as an equilibrium constant for RNA synthesis/degradation.[24,34] Defining the stationary probability to find the system in the state n as P_n, the calculations yield

$$P_n = \frac{x^n e^{-x}}{n!}, \tag{1}$$

which is a well-known Poisson distribution.[12,24] One can now evaluate the average number of RNA transcripts, $\langle n \rangle$,[24]

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_n = x. \tag{2}$$

In a similar way, higher moments of the distribution can be easily calculated. For example, the second moment is given by[24]

$$\langle n^2 \rangle = \sum_{n=0}^{\infty} n^2 P_n = x^2 + x. \tag{3}$$

A dimensionless parameter F, known as the Fano factor, has been widely utilized for clarifying the mechanisms of transcriptional bursting. It is defined as the normalized variance in the number of produced RNA transcripts, and it is a convenient measure of noise and stochastic fluctuations in the system. It can be shown that for the one-state model in Fig. 3a

$$F = \frac{\langle n^2 \rangle - \langle n \rangle^2}{\langle n \rangle} = 1, \tag{4}$$

which is a signature of the Poisson process. This is an important result because the Fano factors obtained from experimental measurements of transcriptional processes in various living organisms deviate significantly from unity ($F > 1$),[12,14,22] suggesting that the one-state kinetic model from Fig. 3a cannot describe the appearance of transcriptional bursting. This was the main reason for researchers to explore more complex multi-state kinetic models illustrated in Fig. 3b–e.[24,29,35]

At this point, it is important to explain the confusing terminology that exists in this field. The model presented in Fig. 3a is called a "one-state" model even though the system can be found in one of an infinite number of chemical states n (n= 0,1,2,...). All these states are chemically different because they contain a different number of RNA transcripts. The label "one-state" is associated then with a single mode of production of the RNA molecules and with a single specific set of synthesis and degradation rates $\{\alpha, \beta\}$. In other words, the stationary distribution of the produced RNA transcripts has only a single peak: see eqn (1). Then the multi-state models, shown in Fig. 3b–e, would correspond to systems with multiple modes or channels of RNA transcript production, *i.e.*, multiple sets of synthesis/degradation rates $\{\alpha_i, \beta_i\}$ for $i = 1, 2,\ldots,m$: see Fig. 3d. To distinguish these macro-states from real individual chemical states of the system, we sometimes call them "biochemical states", emphasizing that the biochemical products (RNA molecules) are made in different ways for each set of the synthesis/degradation rates.[24] Thus, the multi-state models of transcription reflect the multiple modes of RNA production and degradation, and this is exhibited by the multiple peaks in the stationary distributions of RNA transcripts.

The deviations of experimentally measured Fano factors from unity not only demanded the development of multi-state kinetic models, but it also raised a question of what is the appropriate model to fully describe the transcriptional bursting phenomenon. The majority of theoretical and experimental studies utilized for their studies the two-state model presented in Fig. 3b.[16,34,36] In this model, one state (ON) involves both the production and degradation of RNA molecules while the other state (OFF) only involves the degradation of RNA transcripts. However, no reasons have been given for using the two-state model in transcriptional bursting analysis beyond it being the simplest non-trivial extension of the unsuccessful one-state model. At the same time, several experimental studies found that more than two biochemical states might be involved in transcription.[9,37–41] These observations stimulated the study of more complex multi-state kinetic models; some examples of them are presented in Fig. 3.[24,29,35]

While the multi-state kinetic models have existed for a while,[9,29,35] a comprehensive theoretical framework to analyze transcription processes has been developed only recently.[24] In this framework, it was assumed that the system can follow one

of m possible biochemical pathways of RNA production and degradation, as presented in Fig. 3d. The synthesis of RNA was assumed to be state-dependent with a rate α_j (j = 1, 2,...,m), while the degradation rate constant β was assumed to be state-independent. The transitions between different biochemical states *via* state-dependent transition rates $k_{\mathrm{on}}^{(j)}$ ($j \to j-1$) and $k_{\mathrm{off}}^{(j)}$ ($j \to j+1$) are also allowed: see Fig. 3d. Using a master equations approach together with a generating functions method,[42] it was shown explicitly how to evaluate all stationary dynamic properties of the system, including the most relevant mean number of produced RNA transcripts and the Fano factor for a given system.[24] This general theoretical framework became a convenient tool for understanding the mechanisms of transcriptional bursting.

Several interesting results have been obtained by analyzing the general theoretical framework for multi-state kinetic models of transcription.[24] First of all, it was shown that in the system with originally m biochemical states, the actual number of observed states, which corresponds to the number of peaks in the stationary distribution of produced RNA molecules, surprisingly might be less than m. This can be clearly seen in Fig. 4 for the system with m = 3 states (the kinetic scheme for this three-state system is shown in Fig. 3e). Depending on the choice of kinetic parameters, the stationary distributions of the produced RNA molecules exhibit one, two or three peaks that correspond to the one, two or three channels for RNA synthesis and degradation. The key factor that determines the specific outcome is the choice of values of the state-switching transition rates $k_{\mathrm{on}}^{(j)}$ and $k_{\mathrm{off}}^{(j)}$ in comparison with the synthesis and degradation rates.

To understand these observations better, let us consider two limiting cases. If the switching rates are much smaller than synthesis and degradation rates α_j and β, then the system has enough time to fully explore each biochemical state (each pathway in Fig. 3e) individually and the number of observed states (as well as the number of peaks in the stationary distribution) will be equal to m. In the opposite limit, if the state transition rates are much faster than the synthesis/degradation rates, the system will not have time to explore each state individually, so it will exhibit a single equilibrated biochemical state with synthesis and degradation rates averaged over all the states. This would lead to a single-peak stationary distribution of produced RNA transcripts. For all other ranges of parameters, the number of observed states will vary from 2 to m − 1 depending on how fast are the local switching rates in comparison with the local synthesis and degradation rates. In this range, some local equilibria might be reached lowering the number of observed biochemical states. These arguments suggest that the question of what kinetic models should be used to describe the transcriptional bursting is quite complex since the number of experimentally realized biochemical states might differ significantly from the number of actual states that participate in transcription.[9]

Theoretical calculations[24] also found an interesting correlation. Increasing the number of observed biochemical states increases the degree of stochastic fluctuations as measured by the Fano factor (Fig. 4f). In other words, the system exhibits more noise if it is capable of exploring all different modes of the RNA production and degradation during transcription.

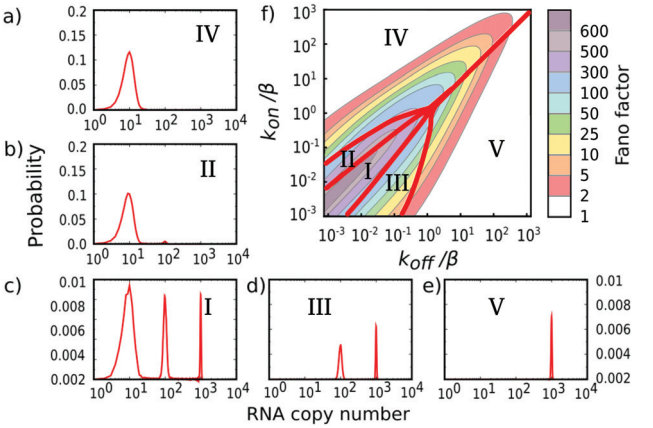

Fig. 4 Transcription dynamics of the three-state kinetic model. (a)–(e) Examples of five different stationary distributions of produced RNA molecules. (f) A dynamic phase diagram that using a contour plot shows how the Fano factor varies as a function of the normalized switching rates. Solid lines give qualitative boundaries between different dynamic regimes of RNA production presented in parts (a)–(e). Reprinted with permission from the *J.Phys. Chem. B*, 2018, **122**, 11969–11977. Copyright 2018 American Chemical Society.

Another important result from the analysis of the multi-state kinetic models is the proposed procedure to estimate the minimal number of biochemical states. It can explain the experimental observations of various multi-state transcription dynamics.[24] Since the number of experimentally observed states might deviate from the actual number of biochemical states, it seems reasonable to study the optimal multi-state models with a minimal number of parameters. It was shown that the parameter m (minimal number of biochemical states) can be found from the simultaneous knowledge of both the mean $\langle n \rangle$ and the Fano factor F of the produced RNA transcripts. For realistic cellular conditions, the following approximate relation was proposed,

$$m \simeq 1 + \frac{F-1}{\langle n \rangle}. \tag{5}$$

This result was used to analyze the minimal number of states in many genes of *E. coli* bacteria, illustrated in Fig. 5.[14] The two-state kinetic model (m = 2) provides an adequate description for the majority of genes and more complex multi-state kinetic models typically are not needed.[24] However, some genes require as many as $m \simeq 10$ states to properly describe the transcription process. These results suggest that the two-state kinetic models are not the universal descriptions of the transcriptional bursting, but that the proper minimal model can be selected using experimentally measured mean and Fano factor of the stationary RNA number distribution.

At the same time, we have to emphasize that the theoretical predictions of the dynamic phase diagrams (Fig. 4f) and of the number of minimal biochemical states (Fig. 5) should be

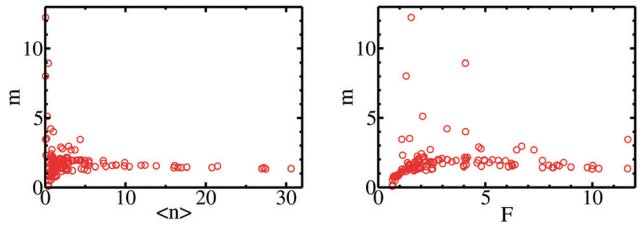

Fig. 5 Estimates of the minimal number of biochemical states m from experimental measurements of transcription dynamics (mean and Fano factor) in different genes of *E. coli* bacteria. Data are taken from ref. 14 Reprinted with permission from the *J. Phys. Chem. B*, 2018, **122**, 11969–11977. Copyright 2018 American Chemical Society.

considered very cautiously. Such characterizations of transcriptional system's dynamics might change depending on the existence of additional genetic regulation mechanisms like feedback loops.[30,31] In terms of the multi-state kinetic model shown in Fig. 3, this would correspond to a situation where transition rates would also depend on the number of the produced RNA molecules n. Furthermore, there are experimental observations indicating that the transcriptional burst frequency and the burst size may be time-dependent, at least for some genes.[43,44] For the multi-state kinetic scheme in Fig. 3 this would mean that the transition rates would also depend on time.

3.2 Origins of transcriptional bursting

Clearly, one of the most important questions is what causes the transcriptional bursting at the microscopic level.[9,10] Since major biochemical events taking place during transcription are reasonably well known,[1,2] it has been suggested that the rate-limiting step of the binding of transcriptional factors to the promoter region on DNA is the main event involved in starting the transcriptional burst.[9,10] There are multiple experimental observations from biochemical and single-molecule studies of various organisms that support this claim.[10]

However, a more complex question is what ends the transcriptional burst. It was proposed that the duration of the burst is determined by the lifetime of the RNAP bound to the DNA chain.[10] But the fact that there are typically several simultaneously transcribing RNAP enzymes puts this argument in doubt. Adding only a few more polymerases can significantly increase the overall time of active transcription as shown by the following simple arguments. If we assume that the probability of one RNAP to dissociate during the time of transcription is p then for n independently operating polymerases, the probability to dissociate simultaneously from the DNA decreases exponentially to p^n. An alternative suggestion was that chromatin remodeling is somehow responsible for regulating the dynamics of transcription. However, a detailed mechanism of how this might lead to the transcriptional bursting remains unclear.[9,10]

Several bulk[45–47] as well as single-molecule[23,48] studies pointed out to the importance of topological and mechanical properties of DNA during transcription. When the RNAP enzyme is engaged in transcription, it rotationally moves along the DNA double helix. It is known that DNA molecules are frequently topologically constrained in living cells (possibly *via* binding to other components of the cell's nucleus), and this leads to the buildup of positive supercoiling in front and negative supercoiling behind the RNAP. The supercoiling slows transcription down, but this buildup of mechanical stress on the DNA can be released by the action of several classes of topoisomerase enzymes. These observations led to the development of multiple theoretical models which propose that transcriptional bursting is a result of coupling between mechanical and chemical processes that are taking place at the microscopic level in the system.[23,25–28,49–51]

To better understand the interplay between chemical and mechanical forces involved in transcription, let us consider a discrete-state stochastic mechanochemical model illustrated in Fig. 6.[25] This model takes into account the most relevant physicochemical processes using a thermodynamically consistent approach. It considers transcription by a single RNAP molecule in the presence or absence of the topoisomerase gyrase that can relieve the supercoiling stress. When the gyrase is bound to DNA, the stress buildup is not taking place due to its enzyme action. This is labeled as the ON state in Fig. 6. In

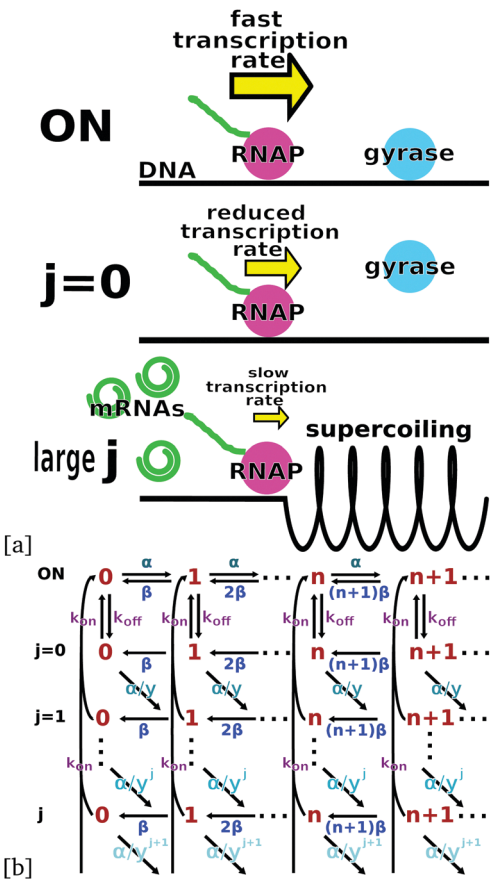

Fig. 6 A schematic overview of the discrete-state mechanochemical model of transcriptional bursting. (a) A pictorial view of different states; (b) a corresponding chemical-kinetic scheme for the model. Reprinted from A. Klindziuk, B. Meadowcroft and A. B. Kolomeisky, A Mechanochemical Model of Transcriptional Bursting, *Biophys. J.*, **118**, 1213–1220, Copyright 2020, with permission from Elsevier.

this case, RNA transcripts are produced with a rate α and destroyed with a rate constant β. However, after the gyrase dissociates from the DNA chain with a rate k_{off}, the supercoiling starts to increase with every new produced RNA molecule and this slows down the transcription: see Fig. 6. The synthesis rate is now given by the rate α/y^{j+1}, where j is the number of produced RNA transcripts after the last dissociation of the gyrase molecule. The parameter $y = \exp(\varepsilon/k_BT)$ is called the mechanochemical parameter and it plays a crucial role in the dynamics of the system. It quantifies how the mechanical forces created by supercoiling are resisting the chemical reaction of RNA synthesis.

Note that the parameter j can be viewed as a quantitative measure of the degree of supercoiling on the DNA strand. The synthesis of every new RNA molecule (in the absence of the gyrase) increases the mechanical stress on DNA. Therefore, j can also be thought of as the number of transcripts produced after gyrase detached. Additionally, j along with the parameter y, help describe the mechanochemical coupling in the system in a thermodynamically consistent way. This is because the parameter y is associated with an additional energy ε that is needed for RNAP to transcribe the supercoiled DNA. The stronger is the degree of supercoiling, the slower is the rate of the RNA synthesis. Then, following Kramer's description of chemical rates, the decrease in the RNA production rates will be exponential, α/y^{j+1}.[25]

The discrete-state mechanochemical model can be solved explicitly by two complementary theoretical methods: the forward and backward master equations.[25] This provides a full dynamic description of the system that allows to clarify the molecular mechanisms of transcription. Fig. 7a presents a stationary distribution of produced RNA molecules for different mechanochemical coupling strengths. If supercoiling does not affect the RNA production ($y = 1$ and $\varepsilon = 0$), a single-peak distribution is obtained (orange symbols) because, in this case, the system has a single biochemical state with one set of synthesis/degradation rates. However, the situation changes if the mechanochemical coupling is strong (large y). Then, the system exhibits two peaks in the distribution (blue symbols) and, therefore, has two biochemical states. This bimodal distribution has one peak at $n = 0$, where the gyrase is unbound so the synthesis rate approaches zero, and a peak at $n > 0$, where the gyrase is bound and normal synthesis and degradation processes are taking place. In agreement with these arguments, increasing the strength of the mechanochemical coupling lowers the average number of produced RNA molecules, as indicated in Fig. 7b. Importantly, the bimodal distribution of mechanochemical model explains the appearance of multiple states of transcriptional bursting from a more microscopic point of view.[25]

Another advantage of the mechanochemical model is its ability to quantify the degree of supercoiling and explain how that influences transcription dynamics, which is illustrated in Fig. 8. The bimodal distribution of the mechanical stress in the system is predicted for all ranges of parameters. This reflects the fact that there are two types of states in the system at all

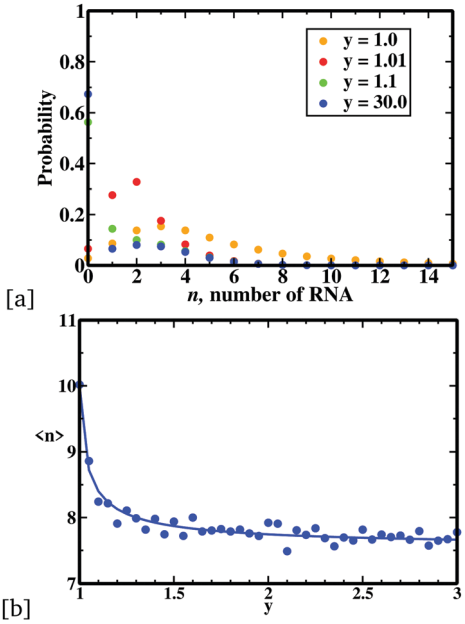

Fig. 7 (a) Stationary-state distributions for the production of RNA molecules for different mechanochemical coupling streangths. (b) The mean number of synthesized RNA molecules as a function of the energetic cost of supercoiling. The symbols are from computer simulations and solid lines are analytical results. Reprinted from A. Klindziuk, B. Meadowcroft and A. B. Kolomeisky, A Mechanochemical Model of Transcriptional Bursting, *Biophys. J.*, **118**, 1213–1220, Copyright 2020, with permission from Elsevier.

times. The first type corresponds to the states where gyrase is bound to DNA and no stress is present (ON), while another type describes the states where the gyrase is unbound and the supercoiling is taking place ($j > 0$). Increasing the rate of gyrase association k_{on} lowers the overall degree of the mechanical stress and shifts the equilibrium between two peaks in favor of the stress-free states: see Fig. 8a. The degree of mechanochemical coupling also influences the distribution of stress in the system (Fig. 8b). For a small energy expenditure in supercoiling ($y \sim 1$ and $\varepsilon \sim 0k_BT$), a broad distribution is

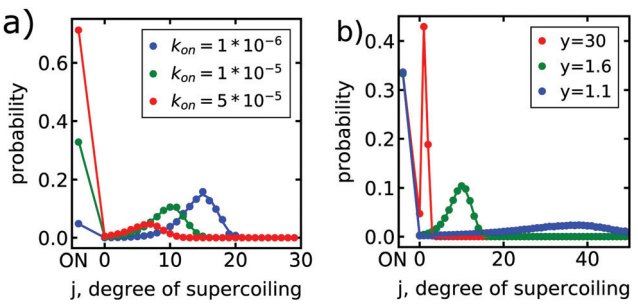

Fig. 8 Stationary-state distributions of mechanical stress on DNA. The symbols are from computer simulations and solid lines between numerical values of j are theoretical predictions. (a) The effect of varying the association rates of gyrase. (b) The effect of varying the energetic cost of supercoiling. Reprinted from A. Klindziuk, B. Meadowcroft and A. B. Kolomeisky, A Mechanochemical Model of Transcriptional Bursting, *Biophys. J.*, **118**, 1213–1220, Copyright 2020, with permission from Elsevier.

observed because the state with any level of supercoiling j can be reached (blue symbols). Making the mechanochemical coupling stronger ($y \gg 1$ and $\varepsilon \gg 0k_BT$) narrows the distribution and lowers the overall level of the mechanical stress (red symbols). In this case, the states with a large j cannot be reached due to the fast decrease in the synthesis rates after the gyrase detaches.

The mechanochemical model has been applied for analyzing *in vitro* measurements of transcriptional bursting in T7 and *E. coli* bacteria.[23] Using the experimental data of the decrease in elongation speeds in the presence of supercoiling, the values of mechanochemical coupling were quantitatively estimated for both bacterial systems.[25] It was found that for T7, $y = 1.61$ and the energetic cost of supercoiling is $\varepsilon = 0.48k_BT$, while for *E. coli* bacteria the analysis predicts $y = 1.89$ and $\varepsilon = 0.64k_BT$. These results suggest the supercoiling has a relatively modest slowing effect on RNA production. But at the same time, it was argued that this leads to the most optimal regulation of transcription.[25] If the supercoiling would have a very strong effect, this would lead to very little RNA production, which is not good for transcription. In the opposite limit of very weak mechanochemical coupling between RNA production and supercoiling, the ability to tune the number of produced RNAs would be lost, which is apparently also not beneficiary for the cell.

The mechanochemical model proposes the following microscopic picture of the appearance of the transcriptional bursting. After the gyrase molecule binds to DNA, the RNAP begins to actively transcribe the gene with a constant RNA production rate. This is the beginning of the transcriptional burst. After some time, the gyrase will dissociate from the DNA strand and the mechanical stress will start to build up quickly slowing the RNA production. Soon after the dissociation event, the RNA synthesis rate becomes negligibly small and the burst ends. Thus, transcriptional bursting results from the balance between the chemical properties of gyrase and RNA synthesis, and the mechanical properties of DNA double helix that resist the RNA production.

Although we explained the transcriptional bursting mainly by multi-state model and the origin of the bursting phenomenon using the mechanochemical model, it is important to emphasize that these models should be still viewed as hypothetical. Available experimental observations can be reasonably well explained using these theoretical approaches, but this does not exclude the possibility that other microscopic pictures might be the origin of transcriptional bursting. More experimental studies are needed to test the validity of these theoretical models.

4 Future directions and open questions

In recent years, big experimental and theoretical advances helped researchers to better clarify the mechanisms of transcription. While many aspects of this process are now understood, it is important to note that our knowledge of the underlying microscopic picture is still very limited. Multiple questions need to be addressed in future studies of transcriptional bursting. Let us briefly discuss several of them.

First of all, experiments show that the frequency and the size of transcriptional bursts are not constant and change with time.[43,44,52] Moreover, it looks like each gene has its own transcriptional signature. These observations suggest that new theoretical tools for the analysis of transcriptional busting need to be developed since the current multi-state models can only have constant transitions rates. Another important question is what causes the modulation of transcriptional burst frequency and size. Is it a mechanochemical coupling issue due to sequence-dependent supercoiling or is it another biochemical regulation pathway that we still do not know?

Another interesting future direction is to understand the collective behavior of RNA polymerases in transcription. The relevant questions are: Why does transcription sometimes involve multiple RNAPs transcribing simultaneously? How does this influence the transcriptional bursting? And what are the benefits of these collective dynamics for the cellular systems? There are some indications that multiple polymerases simultaneously engaged in transcription might cooperate to achieve faster elongation rates, decrease the transcriptional noise and avoid the buildup of the mechanical stress.[33,49,53–56] However, the molecular mechanisms of such cooperativity largely remain unexplained.

Finally, a fundamentally important problem is to understand how the transcriptional bursting relates to subsequent cellular development. More specifically, it will be interesting to clarify how the spatial modulation of transcriptional dynamics leads to complex gene expression patterns in living embryos.[57–59] The recent advances in single-molecule experimental methods already allowed researchers to collect a significant amount of quantitative information on gene pattern formation and its relation to transcriptional bursting.[59] This should stimulate the development of new theoretical methods to investigate these important questions.

Transcriptional bursting is a fascinating biological phenomenon that has been at the center of recent research activities. Many microscopic features of this phenomenon have been uncovered due to the outstanding work of scientists in different fields. But, in our opinion, this is only the beginning of the road to a full understanding of transcriptional bursting. The future looks promising as the close collaboration between experimental and theoretical studies should help us uncover the mysteries of the fundamental biological process of transcription.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The work was supported by the Welch Foundation (C-1559), by the NSF (CHE-1953453 and MCB-1941106), and by the Center

for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

Notes and references

1 B. Alberts, *Molecular Biology of the Cell*, Garland Science, 2008.
2 H. F. Lodish, *Molecular Cell Biology*, W. H. Freeman and Company, 6th edn, 2008.
3 R. Phillips, J. Kondev, J. Theriot and H. Garcia, *Physical Biology of the Cell*, Garland Science, 2012.
4 M. Dangkulwanich, T. Ishibashi, L. Bintu and C. Bustamante, *Chem. Rev.*, 2014, **114**, 3203–3223.
5 S. Sainsbury, C. Bernecky and P. Cramer, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 129–143.
6 T. L. Lenstra, J. Rodriguez, H. Chen and D. R. Larson, *Annu. Rev. Biophys.*, 2016, **45**, 25–47.
7 H. Chen and D. R. Larson, *Genes Dev.*, 2016, **30**, 1796–1810.
8 P. Cramer, *Nature*, 2019, **573**, 45–54.
9 E. Tunnacliffe and J. R. Chubb, *Trends Genet.*, 2020, **36**, 288–297.
10 J. Rodriguez and D. R. Larson, *Annu. Rev. Biochem.*, 2020, **89**, 189–212.
11 J. Houseley and D. Tollervey, *Cell*, 2009, **136**, 763–776.
12 I. Golding, J. Paulsson, S. M. Zawilski and E. C. Cox, *Cell*, 2005, **123**, 1025–1036.
13 D. Zenklusen, D. R. Larson and R. H. Singer, *Nat. Struct. Mol. Biol.*, 2008, **15**, 1263–1271.
14 L. So, A. Ghosh, C. Zong, L. Sepulveda, R. Segev and I. Golding, *Nat. Genet.*, 2013, **43**, 554–560.
15 A. Sanchez, S. Choubey and J. Kondev, *Annu. Rev. Biophys.*, 2013, **42**, 469–491.
16 B. Munsky, G. Neuert and A. van Oudenaarden, *Science*, 2012, **336**, 183–187.
17 M. H. Larson, R. A. Mooney, J. M. Peters, T. Windgassen, D. Nayak, C. A. Gross, S. M. Block, W. J. Greenleaf, R. Landick and J. S. Weissman, *Science*, 2014, **344**, 1042–1047.
18 J. Ma, L. Bai and M. D. Wang, *Science*, 2013, **340**, 1580–1583.
19 A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas and S. Tyagi, *PLoS Biol.*, 2006, **4**, e309.
20 M. Wang, J. Zhang, H. Xu and I. Golding, *Nat. Microbiol.*, 2019, **4**, 2118–2127.
21 A. Sanchez and I. Golding, *Science*, 2013, **342**, 1188–1193.
22 D. R. Larson, D. Zenklusen, B. Wu, J. Chao and R. Singer, *Science*, 2011, **332**, 475–478.
23 S. Chong, C. Chen, H. Ge and X. Xie, *Cell*, 2014, **158**, 314–326.
24 A. Klindziuk and A. B. Kolomeisky, *J. Phys. Chem. B*, 2018, **122**, 11969–11977.
25 A. Klindziuk, B. Meadowcroft and A. B. Kolomeisky, *Biophys. J.*, 2020, **118**, 1213–1220.
26 S. A. Sevier, D. A. Kessler and H. Levine, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 13983–13988.
27 S. A. Sevier and H. Levine, *Phys. Rev. Lett.*, 2017, **118**, 268101.
28 S. A. Sevier and H. Levine, *Nucleic Acids Res.*, 2018, **46**, 5924–5934.
29 A. Mugler, A. M. Walczak and C. H. Wiggins, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2009, **80**, 041921.
30 C. Jia, P. Xie, M. Chen and M. Q. Zhang, *Sci. Rep.*, 2017, **7**, 16037.
31 N. Kumar, T. Platini and R. V. Kulkarni, *Phys. Rev. Lett.*, 2014, **113**, 268105.
32 N. Kumar, A. Singh and R. V. Kulkarni, *PLoS Comput. Biol.*, 2015, **11**, e1004292.
33 A. Klindziuk and A. B. Kolomeisky, *J. Phys. Chem. B*, 2021, **125**, 4692–4700.
34 A. Raj and A. Oudenaarden, *Cell*, 2008, **135**, 216–226.
35 T. Zhou and J. Zhang, *SIAM J. Appl. Math.*, 2012, **72**, 789–818.
36 J. Paulsson, *Phys. Life Rev.*, 2005, **2**, 157–175.
37 A. M. Corrigan, E. Tunnacliffe, D. Cannon and J. R. Chubb, *eLife*, 2016, **5**, e13051.
38 K. Featherstone, K. Hey, H. Momiji, A. V. McNamara, A. L. Patist, J. Woodburn, D. G. Spiller, H. C. Christian, A. S. McNeilly, J. J. Mullins, B. F. Finkenstädt, D. A. Rand, M. R. White and J. R. Davis, *eLife*, 2016, **5**, e08494.
39 G. Rieckh and G. Tkacik, *Biophys. J.*, 2014, **106**, 1194–1204.
40 K. L. Hey, H. Momiji, K. Featherstone, J. R. Davis, M. R. White, D. A. Rand and B. Finkenstädt, *Biostatistics*, 2015, **16**, 655–669.
41 D. J. Jenkins, B. Finkenstädt and D. A. Rand, *Bioinformatics*, 2013, **29**, 1158–1165.
42 J. Peccoud and B. Ycart, *Theor. Popul. Biol.*, 1995, **48**, 222–234.
43 R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson and L. S. Weinberger, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 17454–17459.
44 T. Muramoto, D. Cannon, M. Gierliński, A. Corrigan, G. J. Barton and J. R. Chubb, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7350–7355.
45 S. J. Brill, S. DiNardo, K. Voelkel-Meiman and R. Sternglanz, *Nature*, 1987, **326**, 414–416.
46 M. Yamagishi and M. Nomura, *Curr. Genet.*, 1988, **13**, 305–314.
47 R. S. Joshi, B. Piña and J. Roca, *EMBO J.*, 2010, **29**, 740–748.
48 K. Tantale, F. Mueller, A. Kozulic-Pirher, A. Lesne, J.-M. Victor, M.-C. Robert, S. Capozi, R. Chouaib, V. Bäcker and J. Mateos-Langerak, *et al.*, *Nat. Commun.*, 2016, **7**, 1–14.
49 X. Jing, P. Loskot and J. Yu, *Phys. Biol.*, 2018, **15**, 066007.
50 N. Mitarai, I. B. Dodd, M. T. Crooks and K. Sneppen, *PLoS Comput. Biol.*, 2008, **4**, e1000109.
51 M. Ancona, A. Bentivoglio, C. A. Brackley, G. Gonnella and D. Marenduzzo, *Biophys. J.*, 2019, **117**, 369–376.
52 C. Li, F. Cesbron, M. Oehler, M. Brunner and T. Höfer, *Cell Syst.*, 2018, **6**, 409–423.
53 S. Kim, *BioEssays*, 2020, **42**, 1900215.
54 S. Kim, B. Beltran, I. Irnov and C. Jacobs-Wagner, *Cell*, 2019, **179**, 106–119.
55 P. Guptasarma, *BioEssays*, 1996, **18**, 325–332.
56 T. Heberling, L. Davis, J. Gedeon, C. Morgan and T. Gedeon, *PLoS Comput. Biol.*, 2016, **12**, e1005069.
57 B. J. Vincent, J. Estrada and A. H. DePace, *Integr. Biol.*, 2016, **8**, 475–484.
58 B. Zoller, S. C. Little and T. Gregor, *Cell*, 2018, **175**, 835–847.
59 N. C. Lammers, V. Galstyan, A. Reimer, S. A. Medin, C. H. Wiggins and H. G. Garcia, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 836–847.

