Cooperative RNA Polymerase Molecules Behavior on a Stochastic Sequence-Dependent Model for Transcription Elongation

Pedro Rafael Costa, Marcio Luis Acencio, Ney Lemke

Departamento de Física e Biofísica, Instituto de Biociências de Botucatu, UNESP - Univ Estadual Paulista, Botucatu, São Paulo, Brazil

Abstract

The transcription process is crucial to life and the enzyme RNA polymerase (RNAP) is the major component of the transcription machinery. The development of single-molecule techniques, such as magnetic and optical tweezers, atomic-force microscopy and single-molecule fluorescence, increased our understanding of the transcription process and complements traditional biochemical studies. Based on these studies, theoretical models have been proposed to explain and predict the kinetics of the RNAP during the polymerization, highlighting the results achieved by models based on the thermodynamic stability of the transcription elongation complex. However, experiments showed that if more than one RNAP initiates from the same promoter, the transcription behavior slightly changes and new phenomena are observed. We proposed and implemented a theoretical model that considers collisions between RNAPs and predicts their cooperative behavior during multi-round transcription generalizing the Bai et al. stochastic sequence-dependent model. In our approach, collisions between elongating enzymes modify their transcription rate values. We performed the simulations in Mathematica® and compared the results of the single and the multiple-molecule transcription with experimental results and other theoretical models. Our multi-round approach can recover several expected behaviors, showing that the transcription process for the studied sequences can be accelerated up to 48% when collisions are allowed: the dwell times on pause sites are reduced as well as the distance that the RNAPs backtracked from backtracking sites.

Introduction

The first step of the Central Dogma of Molecular Biology concerns the transport of information from the DNA to an RNA molecule. This process, known as transcription, must be exquisitely controlled during the development and maintenance of living beings. It can be divided into three phases—initiation, elongation and termination—and is carried out by the RNA polymerase enzyme (RNAP).

The RNAP scans the duplex DNA to find the sites for transcription initiation, known as promoters, and bind to them, exposing the DNA template. Once the active site of RNAP is in the correct position, the Transcriptional Elongation Complex (TEC) starts the elongation phase. During this phase, the RNAP polymerizes RNA chains, incorporating a ribonucleoside complementary to the nucleotide present in its active site. After incorporation, the RNAP moves along one nucleotide in the template strand and then restarts the process. When it recognizes the site for termination, the TEC is disassembled and the RNAP releases the transcript and disengages from DNA. Throughout this process, RNAP is able to recruit accessory proteins for several of these activities.

The development of single-molecule techniques, such as magnetic and optical tweezers, atomic-force microscopy and single-molecule fluorescence, increased our understanding of the transcription process, complementing traditional biochemical studies (for a review, see Herbert et al. [1]). Kinetics studies showed the occurrence of transcriptional pauses. There are two different possible explanations for this phenomenon: interaction between nascent RNA and RNAP or sequence-dependent interactions among RNA, RNAP and DNA. These pauses play an important role in the mechanism of transcriptional regulation. Several works showed that pauses allow the recruitment of regulatory factors and are important for transcriptional termination, among other functions (see the introduction of Herbert et al. [2] work for a brief review).

Different approaches have been proposed to model the elongation kinetics. Models based on thermodynamic analysis presented in the pioneering paper of von Hippel and colleagues [3] made possible the prediction of DNA sequence dependent pause sites [4,5]. Although these models consider the presence of only one transcribing RNAP on the DNA strand, experimental evidence has been gathered to suggest that multiple RNAPs can simultaneously participate in the process. Ribosomal genes, for instance, are highly transcribed and electron micrographs of the process are remarkable for their Christmas-tree-like molecular organization, where the “trunk” of the tree are the DNA strand, and the “branches” are the nascent RNAs in increasing length. It
is possible to observe even hundreds of branches, indicating the high RNAP concentration in these genes [6,7]. There are two different views on the impact of this multi-round transcription. The first view is that the leading RNAP blocks the advancement of the trailing ones at pause sites, inducing traffic jams [8]. The other one is that the trailing RNAP prevents backtracking and pushes the leading RNAP out of pause sites [9]. The later view corroborates Epshtein and Nudler [10] work, that compared the single and the multi-round transcriptions, showing that in the presence of multiple active enzymes a larger number of RNAP can leave an arrest site and continue the elongation. In another work, Epshtein et al. [11] showed that efficient transcription through proteins roadblocks depends on multiple rounds of initiation. They also conclude that the trailing RNAP molecules can rescue roadblocked complexes in vivo by “pushing” them forward.

Klumpp [12] presented a model to study the pauses and backtracking under dense traffic conditions. His approach distributes the pauses and the backtracking sites along the DNA template according to a density parameter, neglecting the dependence of the local DNA sequence on the transcription rates. He drew the conclusion that the pauses are suppressed due to the fact that the leading polymerase backtracking is restricted in the presence of a trailing RNAP. Rajala and colleagues [13] concluded, using a delayed stochastic model, that the rate of occurrence and duration of the pauses affect the “microbursting” of RNA production. According to them, these bursts occur when two or more RNAPs complete the transcription of the same gene within a shorter interval than the expected minimum interval between consecutive initiations. A long-pause site can drastically alter the distribution of bursts without changing mean expression levels.

Tripathi and Chowdhury [14] proposed a sequence-independent model that also considers collisions and traffic jams and investigated the impact of RNAP interactions on the fluctuations in the synthesis of RNA.

Table 1. Parameters Values.

|          | ATP   | UTP   | GTP   | CTP   |
|----------|-------|-------|-------|-------|
| $k_{NTP}^{max}$ (s$^{-1}$) | 50 ± 6 | 18 ± 1 | 36 ± 5 | 33 ± 6 |
| $K_{NTP}^*$ (µM)         | 38 ± 7 | 24 ± 4 | 62 ± 18 | 7 ± 4  |

Parameters values for Eq. 3. These NTP-dependent parameters were experimentally determined by Bai et al. [15].

doi:10.1371/journal.pone.0057328.t001

Figure 1. Representation of transcription elongation complex (TEC): The TEC is given by the RNA length, $n$, and by the relative position of the Active Site, AS, to the 3’ end of the RNA. a) Post-translocated state: active site free. The TEC structure is given by an RNA-DNA hybrid 8 bp long and by a DNA bubble with 12 bp. b) Post-translocated state, incorporation phase: active site just occupied, NTP being incorporated to RNA. c) Pre-translocated state: active site occupied and the NTP incorporated in RNA. Here, the RNA-DNA hybrid is 9 bp long and the TEC will move one nucleotide forward, returning to a).

doi:10.1371/journal.pone.0057328.g001

Figure 2. Representation of time evolution of RNAP: The red lines represent the last simulated step. a) The time evolution for a single molecule. We consider that the movement is uniform between each step. The velocity is assigned by the Gillespie algorithm. b) The method used to solve collisions. In blue we show the leader RNAP and in black the trailing RNAP. $t_i$ represents the required time for the collision. $t_{min}$ is the time for the first event among the following possibilities: nucleotide inclusion by trailing RNAP, nucleotide inclusion by leading RNAP or collision.

doi:10.1371/journal.pone.0057328.g002
Figure 3. Diagram representing the overall algorithm: The subroutine "Solving collision" is described in Figure 4.

doi:10.1371/journal.pone.0057328.g003
In this current work, we propose a multiple-round elongation approach (MRA) that identifies and solves collisions between the RNAP molecules. This approach is based on the thermodynamic sequence-dependent model developed by Bai et al. [4], hereafter referred as Bai model. A brief recapitulation of this model is presented on the Methods section. The Bai model was implemented using updated parameters and referred as single-round transcription approach, SRA.

Methods
A brief recapitulation of Bai’s thermodynamic sequence-dependent model
The Bai model has been chosen due to its simplicity and consistency with experimental results. This model is based on the thermal ratchet model for molecular motors and considers the elongation as a three-step reaction:

Figure 4. Diagram representing the subroutine to solve collisions: TN mean trailing RNAP in normal elongation. LN means leading RNAP in normal elongation. LBb means leading RNAP in backtracking backwards movement. LBf means leading RNAP in backtracking forwards movement. TBb means trailing RNAP in backtracking backwards movement. TBf means trailing RNAP in backtracking forwards movement.

doi:10.1371/journal.pone.0057328.g004
anion is abbreviated as PPi. The \textit{pre} and \textit{post} represent the TEC configuration and \( n \) is the transcript length. Figure 1 depicts the TEC structure. \( K_{NTP}^{d} \) and \( K_{NTP}^{\text{max}} \) were experimentally established and depend on the nucleotide to be incorporated. Their values are shown on Table 1. \( K_i(n) \) is given by

\[ K_i(n) = \exp\left(\frac{\Delta G(n,\text{post}) - \Delta G(n,\text{pre}) - Fd/k_B T}{k_B T}\right) \]

where \( \Delta G(n,m) \) is the standard Gibbs free energy for the TEC conformation and \( F \) is an external force applied to RNAP, with \( d \approx 0.34 \) nm being the distance between adjacent nucleotides in the DNA strand [15]. \( m \) gives us the position of the active site of RNAP on the template. It is equivalent to use \( m = 0 \) or \( m = \text{pre} \) and \( m = 1 \) or \( m = \text{post} \).

Using the Michaelis-Menten kinetics, we can obtain the overall rate for each ribonucleoside incorporation [16]:

\[ k_{\text{main}}(n) = \frac{k_{\text{NTP}}^{\text{max}} [\text{NTP}]}{K_{\text{NTP}}^d [\text{NTP}] + [\text{NTP}]} \]

Finally, there is also a backtracking/forwardtracking rate given by

\[ k_{n,m\rightarrow n\pm 1} = k_0 \exp\left(\frac{(\Delta G^d - \Delta G(n,m)) + Fd(F)}{k_B T}\right) \]

where \( k_0 = 1.0 \times 10^9 \) s\(^{-1}\) is a pre-factor constant and \( \Delta G^d = 41.2 k_B T \) is the energy barrier for this phenomenon.

The standard Gibbs free energy, (\( \Delta G \)), for the TEC. The \( \Delta G(n,m) \) for the TEC, first described by Yager and von Hippel [3], is a measure of its stability, and is given by the sum of the isolated components:

\[ \Delta G(n,m) = \Delta G(n,m,\text{DNA bubble}) + \Delta G(n,m,\text{DNA hybrid}) + \Delta G(n,m,\text{RNAP binding}) \]
The first term is the free energy due to the break of the hydrogen bonds between complementary nucleotides on the double-strand DNA and the second term is the free energy due to the formation of the RNA-DNA hybrid duplex. They are clearly sequence-dependent and we use the nearest-neighbor model for both [17]. We used SantaLucia et al. [18] values for the DNA-DNA energy and Sugimoto et al. [19] values for the RNA-DNA hybrid energy. The third term represents the interactions between the RNAP and the nucleic acids, considered sequence-independent and, for simplification, as zero. For further details and considerations on these calculation, see Bai et al. [4].

The multiple-round approach

The multiple-round approach extends the Gillespie algorithm [20] to consider collisions between RNAP molecules. Essentially, the algorithm is summarized below:

1. We generate a simulation for each RNAP on the DNA strand. If the RNAP is in the normal elongation, we calculate the reaction rates using Eq. 3 and Eq. 4. If the RNAP is backtracking, we use Eq. 4 to determine the rates for the backtracking/forwardtracking reaction. For both cases, we set \( F = 0 \). The simulations will return the required time for each RNAP complete its movement.

2. A linear approximation was used to determine the time-dependent position equation for each enzyme \( i, S_i(t) \). Figure 2a shows a diagram to illustrate the method.

3. The algorithm determines the required time for collision between each subsequent pair of enzymes by solving \( S_i(t) - d_2 = S_{i+1}(t) + d_1 \), where \( d_1 \) and \( d_2 \) represent the distances between the active site of the RNAP to its ends. A sketch of a collision case between two RNAPs is presented in Figure 2b. The algorithm determines which event occurs first: if an RNAP completes its movement or if we have a collision.

4. All the RNAPs are repositioned: using the time of the event determined in item 3 in their respective \( S_i(t) \), we determine where each molecule is, and a new position equation for all of them. If the event was a collision, we solve the collision according to its type, what will change the position equation for both molecules (see below) and go back to item 3. Otherwise, we go back to item 1.

During the simulation, both the trailing RNAP molecule (T) and the leader one (L) can be in normal elongation, or in backtracking, moving backwards or forwards. Due to this, the collision between the RNAP molecules can be categorized in six different types. Basically, if \( T \) is in normal elongation, it applies \( F_0 \) on \( L \). As the result, \( L \) applies \( F_0 \) on \( T \) in the opposite direction. If \( L \) is backtracking and after the collision \( T \) moves forwards, \( T \) pushes \( L \). If they both are backtracking, we treat the collision as being perfectly elastic. In Figures 3 and 4 we present diagrams depicting the algorithm. Figure 4 details the subroutine to solve collisions.

![Figure 7. Distribution of the distances covered by backtracked RNAPs from the backtracking site](doi:10.1371/journal.pone.0057328.g007)

![Figure 8. Schematic view of the roadblock effect](doi:10.1371/journal.pone.0057328.g008)
Simulations

We used Monte Carlo simulation with the Gillespie algorithm (first reaction method) [20], using the SSA (Stochastic Simulation Algorithm) package [21] for Mathematica®. A complementary package was developed with functions to implement the single and multiple-round approaches. We performed 4800 independent simulations of the kinetics up to 10 RNAP molecules on the same DNA strand and we set the required time for initiation to 0.5 s. We also simulate different cases varying the force parameter, \( F_0 \), from 6.25 pN to 100 pN. The sequences considered correspond to the deletions D104, D111, D112, D123, D167 and D387 of the early genetic region of the bacteriophage T7 [22] ([NTP]: 10 \( \mu M \) for pause sites and 25 \( \mu M \) for the simulated gel) and seq10 to seq13 from Tadigota et al. [5] ([NTP]: 40 \( \mu M \) for seq10, 30 \( \mu M \) for the others). All the reactions were simulated at 24°C.

Pause Criteria

Here, we used the same pause criteria proposed by Bai and Wang in [23]. A pause site is defined by

\[
\tau(n) > (1/\eta) \text{Min}\{\tau(n)\}
\]  

(6)

where \( \tau(n) \) is the dwell time for each the template position \( n \) and \( \text{Min}\{\tau(n)\} \) is the shortest \( \tau(n) \) on the template at a given NTP.
concentration. The parameter \( g \) is tuned to optimize the pause predictions for all the studied sequences. As Tadigotla et al. in [5], we choose to minimize the proportion of incorrect to correct predictions (false positives plus false negatives divided by true positives). Here, we use the third quartile values for \( \eta(t) \). Although the \( \eta \) value for Bai model was 0.05, here we set \( g \sim 0.64 \) for SRA, and \( \eta \sim 0.38 \) for MRA. Expressing them as a dwell time threshold, we have about 0.8 s for SRA and 1.4 s for MRA.

Results and Discussion

We developed an extension of the Bai model for RNAP transcription kinetics that considers interactions among multiple RNAPs. We compared the model with existing experimental data and discussed the possible biological implications emerging from these interactions. Initially, we set \( F_o \sim 25 \) pN. This value was used because it is the average stalling force of RNAP [24].

Figure 5 shows a result obtained during a simulation run of the MRA, i.e., the kinetic of the transcribing RNAPs in the DNA strand. As we can observe the dynamics is intermittent: random walk-like movements coexist with strongly directed progressive motion. The regions where the random walks prevail are considered strong pause sites. The random walk movement is usually interrupted after a collision between two RNAPs. From a qualitative point of view, we can see that collisions can increase the

Figure 10. Relative transcription efficiency, \( E \): Each color represents a sequence, as indicated at the end of the respective curve. The \( E \) values are shown as a function of the maximum number of enzymes \( M \) allowed during the simulation. Note that the asymptotic values of curves are sequence dependent.

doi:10.1371/journal.pone.0057328.g010

Figure 11. Comparison among the simulated transcription gels in different times: a) Template D167. b) Template D387. Each column represents the reaction time for that simulation. The intensity of the bands is proportional to the amount of produced transcripts of that length in the allowed time. The experimental results shown here are graphical reproductions of the gels presented in the experimental work of Bai and Wang [23]. The experimentally observed bands are indicated with their respective transcript size. Some false positives present in the SRA appear less intense in the MRA. For the single-round and the multi-round elongation: some bands appear earlier in MRA than in SRA and the predicted intensities are different for most bands.

doi:10.1371/journal.pone.0057328.g011
pace of transcription. Moreover, we can easily see that there are regions on templates more prone to be pause sites indicating that this behavior is sequence-dependent. Experimentally, these regions can be detected on transcription gels. In the remaining article, we will explore this qualitative picture in further details.

We compared the biological consequences of collisions shown in Figure 6 for sequences D167 and D387. Figure 6a shows the third quartile of dwell time (TQDT) distribution for nucleotide inclusion in function of RNA length. In Figure 6b we present the ranked TQDT obtained on single round approach. We compared these values with the respective TQDT obtained in the same RNA length for multiple round approach. We restricted ourselves to the 40 largest dwell time values because these dwell times are more susceptible to collision effects. In general longer TQDTs are more affected by collisions, but the pause position on the sequence also plays a role. For example, in D387 we can clearly see an strong pause in the beginning of the sequence that remains unaltered even considering collisions. This can be explained since the active enzyme does not allow the attachment of another RNAP to the DNA strand. Another example of the effect of pause position can be seen at the positions 60 and 70. At position 70, the inclusion time is not affected while the time for position 60 is almost halved. This behavior can be understood by observing that at the position 70, the TQDT is shorter than the typical time that a RNAP takes to move from position 60 to 70. Nevertheless, the TQDTs are always reduced when we consider multiple RNAPs. Despite the strong sequence dependence of this effect, the collisions significantly accelerate the passage of RNAPs by the pause sites. The gain in performance occurs mainly on the top 5% dwell times: without collisions they cost 44% of the required time for elongation. When we consider collisions this value is reduced to 27%.

The suppression of backtracking, noticed by Epshtein et al. [11] and reproduced by Klumpp’s model [12], is also observed in our model. The distribution of the distances covered by backtracked RNAPs from the backtracking site during all the simulations can be seen in Figure 7. These distributions have approximately the same median, but it is possible to notice a 2.6-fold reduction in the third quartile of backtracking distance for multiple RNAPs in comparison to single RNAP. Trailing RNAPs are physical barriers that prevent backtracking and the collisions between trailing and leader RNAPs induce a forward movement (see Figure 8). Among the possible biological consequences, we can highlight the reduction of transcriptional proofreading efficiency as suggested by the theoretical study of Sahoo and Klumpp [25].

To evaluate the overall effects of multi-round transcription, we determined the distribution of required times for a complete elongation for each enzyme in each DNA template. These results are presented in Figure 9. The distribution of transcription times weakly depends on the RNAP order except for the last enzyme. The last RNAP is not accelerated by collisions and is restrained to move forward by other enzymes. This explains why this molecule takes longer time to transcribe than the RNAP in the SRA simulation. This behavior persists for all studied sequences.

To characterize the efficiency of the multi-round transcription, we defined the average transcription time by:

**Figure 12. Relation between the incorrect and correct prediction, Q:** The error bars represent the standard deviation obtained by a bootstrap technique. In this case, we evaluated the impact of the finite number of studied sequences on Q.

doi:10.1371/journal.pone.0057328.g012

**Figure 13. Force influence on multiple-round approach:** a) Force influence on the required time for complete elongation. Red bars represent the first RNAP, blue bars represent the fifth RNAP and the green bar represents the tenth RNAP. b) Force influence on the dwell times for two representative nucleotide inclusions.

doi:10.1371/journal.pone.0057328.g013
\[ A_M = \frac{1}{M} \sum_{i} A_i \]

where \( M \) is the maximum number of molecules on the template, \( \tau_i \) is the median of the distribution of required times for the completion of transcription for \( i \)th enzyme and the relative transcription efficiency by:

\[ E(M) = \frac{A_1}{A_M} \]

The values for \( E \) are shown in Figure 10. We can see that \( E \) is an asymptotic increasing function of \( M \) where the asymptote is strongly sequence dependent, clearly indicating the cooperative nature of RNAP interactions. For example, the asymptotic \( E \) for sequence D112 is only 1.03, while the asymptotic \( E \) for Seq13 is 1.48. The finite length of the template imposes a definite limit for \( E \), since there is a finite number of simultaneously transcribing RNAPs. Furthermore, the number, intensity and position of the pauses are also limiting factors: for example, intense pauses at the beginning of the sequence, such as in the D387, restrict the presence of several active RNAPs on the template.

Figure 11 shows the comparison of the predicted transcription gels by MRA and SRA in particular times with the experimental results [23]. It is important to stress that the multiplex round approach does not represent these experimental conditions, since in this case we have a single transcribing RNAP. To quantitatively measure the relative predictive performance of the different models, we used a quality metric originally proposed by Tadigotla et al. [5]: first, pause sites are grouped in clusters in which pauses within 3 bp are considered as a single cluster, and then the number of detected (True Positives, TP) and undetected (False Positives, FP) experimental clusters are counted. Finally, these factors are incorporated in a quality metric defined as:

\[ Q_{model} = \frac{TP_{model}}{FP_{model}} \]

Figure 12 shows \( Q \) for the different models: Bai model, SRA, and MRA. The use of updated parameters improved pause detection compared to Bai model. The RNAP collisions do not impact significantly on the pauses position, but the \( q \) parameter, that is strongly related to transcription efficiency, varies from \( q = 0.64 \) (SRA) to \( q = 0.38 \) (MRA). The large error bars are due to the small number of experimental data.

On Figure 13a we compared the required time for complete elongation for different values of the force parameter for sequence D167. For the first and the fifth RNAP we observe a decrease on the total required elongation time with increasing force. For the last RNAP the effect is negligible. We also present the result for a single RNAP for comparison purposes. The effect of the pushing force is small when compared to the roadblock effect (see Figure 8). On Figure 13b we show that the force influence is strongly sequence dependent by comparing dwell times for two different transcripts lengths. The pause site located at the position 99 with a longer dwell time than the one located at position 78 suffers a stronger influence of the cooperative RNAP behavior.

In summary, our results are in accordance with the literature, showing that the multi-round transcription increases the overall transcription rate by reducing pauses duration and by suppressing RNAP backtracking. The low \( Q \) parameter indicates that some key ingredients to describe this scenario are still missing. An interesting possibility is to propose a more detailed model for the nascent transcript considering different possible RNA structures and their influence on the pace of transcription. Although we can speculate that pauses can be used as regulation factors by the cell, the cooperative behavior of RNAPs reduces the variability of the required time for complete elongation. Our model could be extended to deal with other experimental conditions such as the cooperative action of RNAPs in overcoming the nucleosomal barrier [26].

Acknowledgments

We would like to thank Lucy Bai for his willingness to answer our questions during the implementation of the model and the anonymous referees for useful suggestions.

Author Contributions

Conceived and designed the experiments: PRC NL. Performed the experiments: PRC. Analyzed the data: PRC NL MLA. Wrote the paper: PRC NL MLA.

References

1. Herbert KM, Greenleaf WJ, Block SM (2008) Single-molecule studies of RNA polymerase: Motoring along. Annu Rev Biochem 77: 149–176.
2. Herbert KM, La Porta A, Wong BJ, Mooney RA, Neuman KC, et al. (2006) Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell 125: 1083–1094.
3. Yager TD, vonHippel PH (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. Biochemistry 30: 1097–1118.
4. Bai L, Shundrovsky A, Wang MD (2004) Sequence-dependent kinetic model for transcription elongation by RNA polymerase. J Mol Biol 344: 335–349.
5. Tadigotla VR, O’Malley Be, Sengupta AM, Epstein V, Ehrlich RH, et al. (2006) Thermodynamic and kinetic modeling of transcriptional pausing. Proc Natl Acad Sci U S A 103: 4439–4444.
6. J. 1977. Modulation of ribosomal RNA synthesis in Oenothera fasciculata. An electron microscopic study of the relationship between changes in chromatin structure and transcriptional activity. Cold Spring Harbor Symp Quant Biol 42: 723–740.
7. F. 1981. Electron microscopic analysis of transcription of a ribosomal RNA operon of E. coli. Nucleic Acids Res 9: 1339–1350.
8. Klumpp S, Hwa T (2000) Stochasticity and traffic jams in transcription of ribosomal RNA: Intriguing role of termination and antitermination. Proc Natl Acad Sci U S A 103: 18159–18164.
9. Herbert KM, Greenleaf WJ, Block SM, Neuman KC, et al. (2006) Single-molecule studies of RNA polymerase: Motoring along. Annu Rev Biochem 77: 149–176.
10. Herbert KM, La Porta A, Wong BJ, Mooney RA, Neuman KC, et al. (2006) Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell 125: 1083–1094.
11. Yager TD, vonHippel PH (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. Biochemistry 30: 1097–1118.
12. Bai L, Shundrovsky A, Wang MD (2004) Sequence-dependent kinetic model for transcription elongation by RNA polymerase. J Mol Biol 344: 335–349.
13. Tadigotla VR, O’Malley Be, Sengupta AM, Epstein V, Ehrlich RH, et al. (2006) Thermodynamic and kinetic modeling of transcriptional pausing. Proc Natl Acad Sci U S A 103: 4439–4444.
14. J. 1977. Modulation of ribosomal RNA synthesis in Oenothera fasciculata. An electron microscopic study of the relationship between changes in chromatin structure and transcriptional activity. Cold Spring Harbor Symp Quant Biol 42: 723–740.
15. F. 1981. Electron microscopic analysis of transcription of a ribosomal RNA operon of E. coli. Nucleic Acids Res 9: 1339–1350.
16. Klumpp S, Hwa T (2000) Stochasticity and traffic jams in transcription of ribosomal RNA: Intriguing role of termination and antitermination. Proc Natl Acad Sci U S A 103: 18159–18164.
19. Sugimoto N, Nakano S, Katoh M, Matsumra A, Nakamura H, et al. (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry 34: 11211–11216.

20. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81: 2340–2361.

21. Shapiro BE, Levchenko A, Meyerowitz EM, Wold BJ, Mjolsness ED (2003) Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. Bioinformatics 19: 677–678.

22. Levin JR, Chamberlin MJ (1987) Mapping and characterization of transcriptional pause sites in the early genetic region of bacteriophage T7. J Mol Biol 196: 61–84.

23. Bai L, Wang MD (2010) Comparison of pause predictions of two sequence-dependent transcription models. J Stat Mech-Theory Exp.

24. Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, et al. (1998) Force and velocity measured for single molecules of RNA polymerase. Science 282: 902–907.

25. Sahoo M, Klumpp S (2011) Transcriptional proofreading in dense rna polymerase traffic. EPL (Europhysics Letters) 96: 60004.

26. Jin J, Bai L, Johnson DS, Fulbright RM, Kireeva ML, et al. (2010) Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. NATURE STRUCTURAL & MOLECULAR BIOLOGY 17: 745–U122.