Single-site transcription rates through fitting of ensemble-averaged data from fluorescence recovery after photobleaching: A fat-tailed distribution

Liat Rosenfeld,[1] Eldad Kepten,[1] Sharon Yunger,[2] Yaron Shav-Tal,[2] and Yuval Garini[1]
[1]*The Department of Physics and Institute of Nanotechnology, Bar Ilan University, Ramat Gan 52900, Israel*
[2]*The Mina and Everard Goodman Faculty of Life Sciences and Institute of Nanotechnology, Bar Ilan University, Ramat Gan 52900, Israel*
(Received 11 May 2015; published 17 September 2015)

The stochastic process of gene expression is commonly controlled at the level of RNA transcription. The synthesis of messenger RNA (mRNA) is a multistep process, performed by RNA polymerase II and controlled by many transcription factors. Although mRNA transcription is intensively studied, real-time *in vivo* dynamic rates of a single transcribing polymerase are still not available. A popular method for examining transcription kinetics is the fluorescence recovery after photobleaching (FRAP) approach followed by kinetic modeling. Such analysis has yielded a surprisingly broad range of transcription rates. As transcription depends on many variables such as the chromatin state, binding and unbinding of transcription factors, and cell phase, transcription rates are stochastic variables. Thus, the distribution of rates is expected to follow Poissonian statistics, which does not coincide with the wide range of transcription rate results. Here we present an approach for analyzing FRAP data for single-gene transcription. We find that the transcription dynamics of a single gene can be described with a constant rate for all transcribing polymerases, while cell population transcription rates follow a fat-tailed distribution. This distribution suggests a larger probability for extreme rates than would be implied by normal distribution. Our analysis supports experimental results of transcription from two different promoters, and it explains the puzzling observation of extreme average rate values of transcription.

DOI: 10.1103/PhysRevE.92.032715 PACS number(s): 82.39.−k, 87.16.−b, 87.80.−y, 82.20.Pm

I. INTRODUCTION

Regulating gene expression in space and time is crucial for controlling cells development and function. The gene expression process is commonly controlled at the level of transcription where the first step is the synthesis of messenger RNA (mRNA) by RNA polymerase II (Pol II). Pol II can initiate transcription after binding to a specific DNA sequence, named the promoter, located upstream of the start site of the transcribed gene. The promoter sequence together with different transcription factors determine the affinity of Pol II and thereby the level of expression of the gene. After initiation, the Pol II slides along the gene template and incorporates adequate nucleotides (elongation) to generate the mRNA molecule. Transcription ends when the polymerase reaches a termination sequence and detaches from the DNA together with the nascent mRNA transcript. When Pol II progresses along the gene, a new polymerase can bind with a certain probability and begin transcribing, so that often there are several transcribing polymerases along the gene.

A simple gene transcription model is based on the assumption that a gene transcribes mRNAs at a constant average rate. This model predicts a Poisson distribution for the mRNA numbers in individual cells [1]. Individual transcription events consist of small numbers of molecules in which individual reaction events dominate the behavior. To properly describe this system, stochastic kinetics methods are commonly used [2], which are usually based on chemical master equations [3,4]. This stochastic nature can generate considerable cell-to-cell variability across isogenic cell populations [5]. Analysis of this stochastic variation leads to the statistical distributions of random variables of interest. By comparing the mathematical models of the stochastic kinetics with experimental observations, it is possible to draw a better understanding of the mechanisms involved in transcription.

A common experimental approach for studying transcription *in vivo* is by imaging mRNA using the MS2 system [6–8]. The system contains two elements, an MS2 coat protein fused to a fluorescent protein such as GFP (MS2-CP-GFP), and an insertion of a repeated MS2 DNA sequence into a gene of interest. The MS2 repeats form stem-loop structures in the mRNA that are bound efficiently and stably [9] by MS2-CP-GFP. When the gene is active and transcribes mRNA with MS2 repeats, several MS2-CP-GFP proteins bind to the mRNA and, thus, mRNA transcription can be followed in real time. In order to measure transcription rates, the recovery of the fluorescent signal from photobleached [fluorescence recovery after photobleaching (FRAP)] GFP-MS2-CP on the site of transcription is monitored over time [10,11]. The recovery kinetics are determined by the effective elongation rate of the new stem loops and therefore contain kinetic information of the actual transcription rates.

Initial FRAP experiment estimates of Pol II elongation rates have used tandem gene arrays [10–12], where multicopy gene arrays expressing MS2 stem loops were integrated. However, under these conditions, the kinetics of a single gene remains hidden within the averaged population of active genes in the array, and hence the precise dynamics cannot be measured. Recent studies have reported the kinetics of transcription of single-copy genes [13–16] and have led to a better understanding of the transcription rates of Pol II, but even in these studies, the analysis was performed over the ensemble average.

Previous experiments have provided a very wide range of transcription rates varying from 0.31 to 100 kb/min [17]. Even when comparing data from a single gene to a tandem array of the same gene in the same cell type, the rates are not identical [18]. These results emphasize that transcription dynamics on the single-gene level are obscure and we still lack

full understanding of the process. Is there a single elongation rate for all genes? Do all Pol II's elongate at the same rate? If not, what is the nature of the rates distribution?

In this study we present an analysis method for FRAP data measured on a single copy of a transcribing gene. Moreover, using this method we show that the transcription rates are drawn from a distinct distribution that we find analytically. The results reveal that transcription rates do not follow the central limit theorem and predict large fluctuations in rate constants leading to interesting and innovative insights. This behavior can explain the inhomogeneity and variance in previously published data of transcription rates.

II. MATERIALS AND METHODS

A. Sample preparation

HEK-293 Flp-In cells with an integrated cyclin D1 were previously described [13,19]. Transient transfections were performed by calcium phosphate precipitation (for transient MS2-CP-GFP expression).

B. Microscopy

Time-lapse transcription fluorescent imaging experiments were performed on a laser scanning confocal microscope (Olympus FV1000, Japan) with a $60\times/1.35$ NA PlanApo objective. eGFP fluorescence was detected using an Argon laser (488 nm, 3 mV output) with 1%–2% laser power. Cells were maintained at 37 °C in a 5% CO_2 humidified atmosphere using a stage adapted incubator in conjunction with an objective heater (Tokai, Japan).

C. FRAP

Four-dimensional FRAP experiments were performed measuring three-dimensional images. FRAP experiments consisted of 40 slices with Z intervals of 0.3 μm. Bleaching was achieved with a 488-nm laser illuminated for 250 ms at full power (3 mW) in a circular Region of Interest (ROI). After bleaching, image sequences were acquired at time intervals of 30 s for a total time of 40 min. Image sequences were then analyzed using IMAGEJ SPOT TRACKER plugin [20].

Bleach correction was applied to time-lapse images using [13]

$$\frac{[I_s(t)-I_n(t)]/[I_s(t_0)-I_n(t_0)]}{I_n(t)/I_n(t_0)}, \tag{1}$$

where t_0 and t are the measurement times before and after bleaching, respectively. I_n is the intensity of an arbitrary area in the nucleus. I_s is the intensity at the transcription site. Exponential FRAP curves were analyzed using the "reaction dominant" FRAP model [21]. The dissociation rate γ was extracted by fitting the experimental curve according to Eq. (2) where C is the intensity at the transcription site area at the first image after bleach.

III. RESULTS AND DISCUSSION

We measured transcription in real time on a single integrated *cyclin D1* gene containing 24 MS2 stem-loop sequence repeats integrated in human embryonic kidney cells (HEK-293). For more information see the Methods section. The gene was controlled by an efficient Cytomegalovirus (CMV) promoter that ensures an active transcription. We combined the measurements from the present study with previously measured data [13]. The kinetic behavior of the transcribing gene was explored using FRAP [Fig. 1(a)].

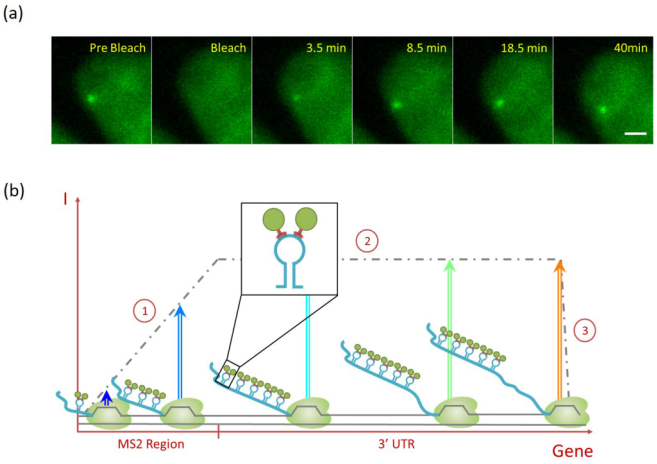

FIG. 1. (Color online) Single-gene FRAP experimental results. Top: Images from a FRAP experiment. Scale bar 1.5 μm. Bottom: A scheme of the fluorescent intensity as the Pol II (oval lobes) moves along the gene. While a polymerase transcribes the MS2 region, the nascent transcript folds into a structure of stem and loop (enlarged inset). Each stem and loop is bound by up to two MS2-CP-GFP molecules (circles). This region can contribute to an increase in the intensity signal of a transcribed mRNA. From this region on, there are no stem-loop structures that can bind fluorescent molecules and therefore there is no further contribution to the intensity signal; the intensity remains at a constant level until the Pol II and the transcript detach from the gene and the intensity drops abruptly.

For the case of a single binding reaction, ensemble-averaged FRAP data can be fitted with an exponential equation of the form [21]

$$f(t) = 1 - e^{-\gamma t} + C, \tag{2}$$

where C is the steady state at equilibrium, and γ is the dissociation rate. This model is adequate also for the transcription site where occasionally a multiexponential fit has to be used depending on the number of binding states in the process [11].

FRAP analysis of single-gene data can be accurately fitted with a single exponent, even after the averaging of ten single-site measurements [13].

The FRAP recovery curve of a single transcription site results from polymerases moving along the gene. Under the approximation that the Pol II has a constant speed, the fluorescence time trace can be described with a naive dynamic model [Fig. 1(b)]. Fluorescence intensity increases only where the polymerase transcribes the region of MS2 repeats, which can bind MS2-CP-GFP. Accordingly, the fluorescent time-dependent signal is expected to be a trapezoidlike curve with three distinct segments: (1) Linear growth during the transcription of the stem loops; (2) steady signal until the end of the gene; and finally (3) a sharp drop due to detachment

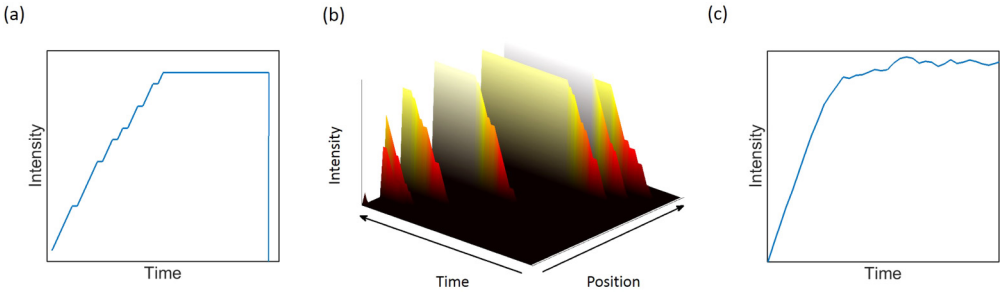

FIG. 2. (Color online) Idealized fluorescent transcription curve. (a) A single polymerase transcription fluorescent curve. Transcription of the MS2 region is shown as discrete steps followed by a steady signal while transcribing the 3′ UTR. The release of the nascent transcript and the Pol II leads to a fast decrease in the intensity. (b) Representation of intensity distribution on a gene as a function of time and Pol II's position along the gene. Each intensity profile represents a single Pol II and reflects on its elongation progress. (c) FRAP intensity profile of a transcribing gene. The curve represents the intensity sum of several transcripts scattered along the gene at different polymerization states.

and diffusion of the polymerase and the transcript [Fig. 1(b)] [15,22].

As the binding of the MS2-CP-GFP proteins to the stem loops is stochastic, a more realistic model is shown in Fig. 2(a) [7], with a stepwise slope which represents the statistics of MS2-CP-GFP binding to the RNA. As there are several Pol II's that can transcribe along the gene at the same time [Fig. 2(b)], the sum of several trapezoid time traces represents the recovery curve arising from a single transcribing gene. However, when we consider the stochastic binding of each Pol II and calculate the sum, it does not demonstrate an exponential growth, as expected from experimental data, but rather a linear recovery at the short time range and a roundoff to a plateau at longer times [Fig. 2(c)]. This suggests that studying transcription through a simplified model using a single transcription rate for elongating polymerases demonstrates discrepancies between experimental FRAP results and the model.

To solve the discrepancy between our calculations and experimental data, we examined the FRAP curves of single cells (where only a single gene is being transcribed). As shown in Figs. 3(a)–3(e), the recovery curves mostly demonstrate a trapezoidlike behavior as predicted. However, the transcription rates are not identical between cells, but rather distributed. Nevertheless, the average of these measurements shows up as an exponential recovery curve [Fig. 3(f)]. Attempting to fit the averaged recovery curve to a linear form gave correlated residuals, and was thus erroneous (Fig. 4).

Realizing that an exponential recovery curve can result from an average of multiple trapezoid signals (Fig. 3), each with a different slope, allows us to extract significant information on the transcription rates. Each slope is analogous to the single-gene transcription data where every experiment has a slightly different effective polymerization rate. Assuming that the transcription rates are randomly drawn from a well-defined distribution, it is possible to extract the whole distribution of rate constants from the resulting ensemble-averaged recovery curve, $f(t)$.

The slope of $f(t)$ at time t, $K_f(t)$, is the mean of all the single-gene transcription slopes that did not reach the steady state:

$$K_f(t) = \langle \{k\}_t \rangle = \frac{\partial f}{\partial t} = \gamma e^{-\gamma t}, \tag{3}$$

where $\{k\}_t \equiv \{k_i; k_i < 1/t\}$ are the rate constants of cells i which have not reached the plateau by time t, and γ is the exponent power as described in Eq. (2).

This average can be calculated by

$$\langle \{k\}_t \rangle = \int_0^{1/t} P(k) k dk, \tag{4}$$

where $P(k)$ is the distribution probability for a rate k. The upper limit of the integration is $k_{\max} = 1/t$ as faster transcription rates have already finished transcribing at time t and therefore they do not contribute to the slope of the recovery curve.

By solving Eq. (4), one finds that

$$P(k)k = \frac{\partial [K_f(t)]}{\partial \left(\frac{1}{t}\right)} = \gamma^2 t^2 e^{-\gamma t}, \quad P(k) = \gamma^2 \frac{1}{k^3} e^{-\frac{\gamma}{k}},$$

where we have used $k = 1/t$.

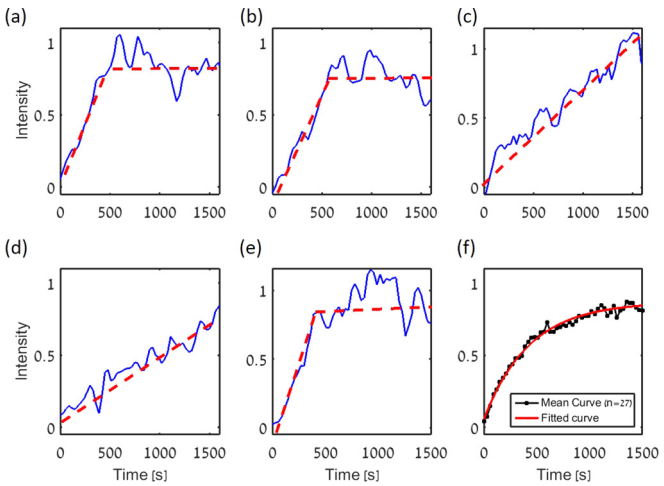

FIG. 3. (Color online) Single-gene FRAP experimental results. (a–e) Single-cell FRAP experiments are shown. Slopes that do not reach steady state represent slow elongation rates that did not yet reach the UTR region in the measurement time (for this presentation data were minimally smoothed to decrease noise fluctuations). Dashed curves are a guide to the eye to emphasize the trapezoid shape of the data. (f) Average of $n = 27$ single-gene FRAP curves with an exponential fit according to Eq. (2),$\gamma = (2.3 \pm 0.2) \times 10^{-3}$ s^{-1}.

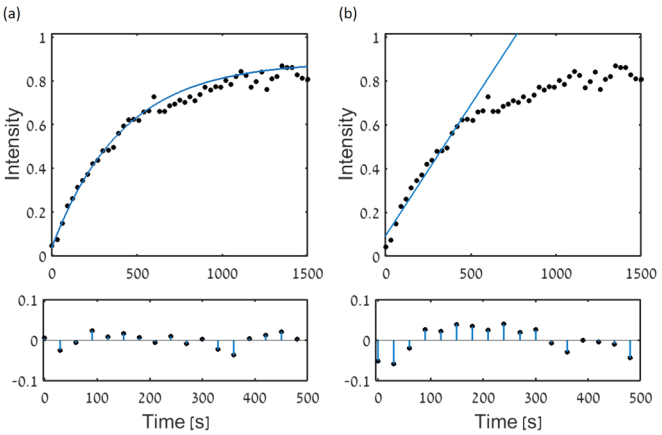

FIG. 4. (Color online) Fit to FRAP experimental data. (a) Fitting FRAP data with Eq. (2). (b) Fitting FRAP data to a linear equation. Both curves are the average of $n = 27$ experimental results. Data are marked with dots and the curve fit with a continuous line. Goodness of fit is evaluated by viewing the residuals distribution which is plotted at the bottom of each panel. The fitting was done only to the first 15 data points.

Therefore, the distribution of transcription rates for an exponential recovery following Eq. (2) of the ensemble-averaged single-site FRAP is

$$P(k) = \gamma^2 \frac{1}{k^3} e^{-\frac{\gamma}{k}}. \tag{5}$$

It is important to note that the distribution $P(k)$ that we found has an asymptotic power-law tail (Fig. 5). It has a finite mean but the second moment does not converge; i.e., the variance is infinite. It therefore falls under the category of "fat-tailed distributions" and does not follow the central limit theorem [23], which requires data to stem from a distribution with finite variance. This divergence increases the probability

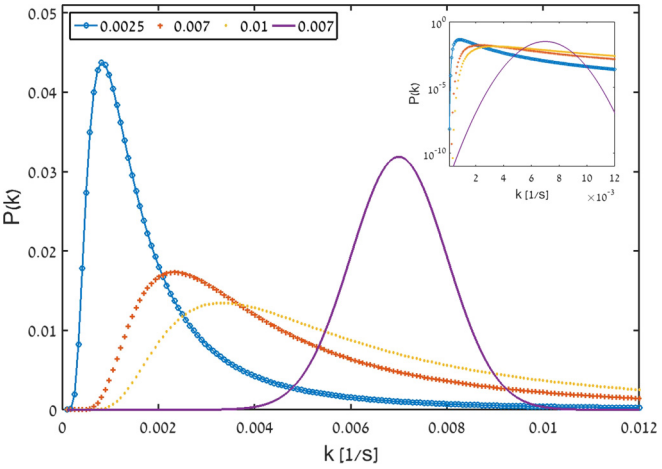

FIG. 5. (Color online) The distribution of Eq. (5) for different γ values (circles, plus sign, and dots) and a normal distribution with $\gamma = 0.007$ (solid line).The inset shows the same distribution on a semilogarithmic scale to emphasize the difference in the tails of the distributions according to Eq. (5) and the normal distribution.

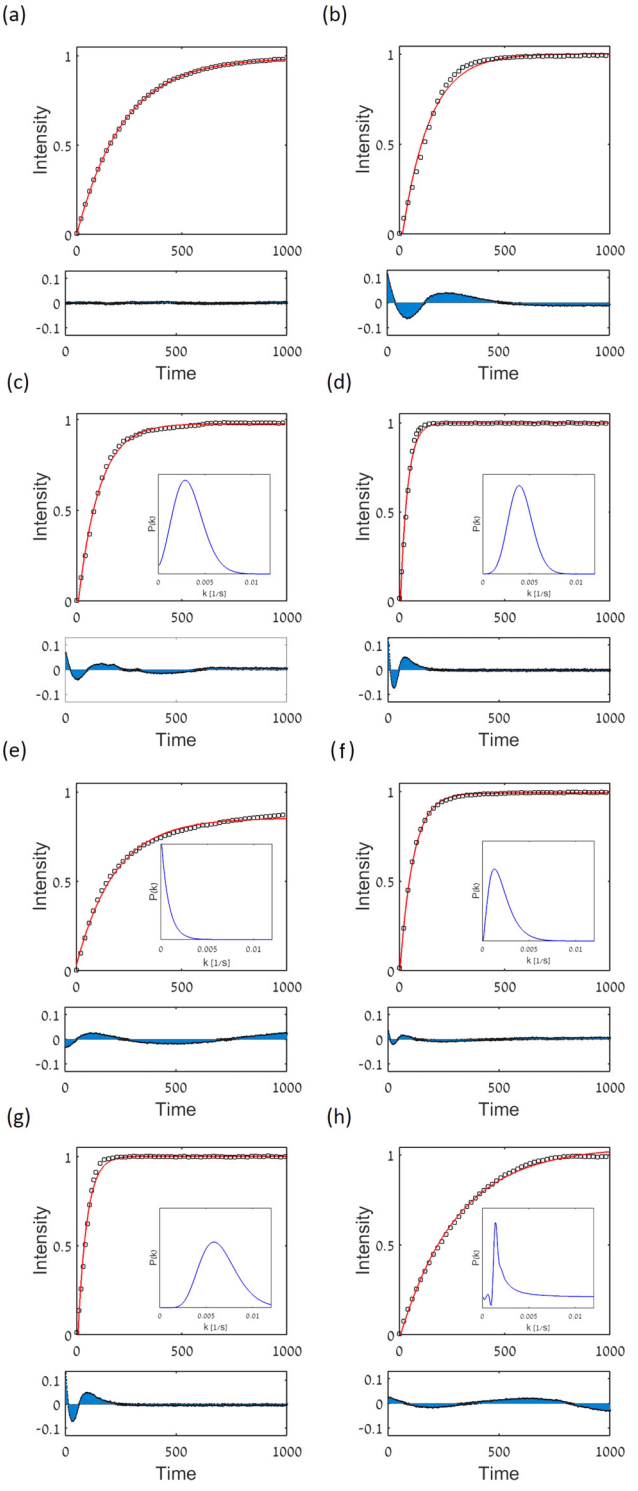

FIG. 6. (Color online) Simulation results of average recovery slopes according to different distributions. The average of 1500 trapezoidal intensity curves with slopes distributed according to Eq. (5) (a) and according to normal distribution (b), Poisson with $\lambda = 4$ (c) and $\lambda = 10$ (d). χ^2 distribution with $k = 2,5,20$ degrees of freedom (e–g) and (h) Pareto distribution with finite mean and infinite variance. Noise was added to each single trapezoidlike curve in order to imitate experimental data. All panels present data in black squares, fit with solid line. The bottom of each figure shows the residuals plot. Inset figures show the distribution curves according to the distributions discussed in (c–h).

for extreme rate values and can therefore explain the wide variability in published *in vivo* rates [17,18].

To validate this distribution we simulated a model for transcription in MATLAB [24]. We assumed constant elongation rates for each single gene, illustrated as a trapezoidlike recovery curve. To mimic experimental data we also added random noise to each curve. We draw the elongation rate for each single-gene simulation experiment randomly, from the distribution we found [Eq. (5)] and calculated the average FRAP curve. This average produced an exponential growth curve that resembles experimental results [Fig. 6(a)]. A good fit was found even after taking only ten randomly selected different rates from the distribution. In contrast, when applying a normal distribution to the single-gene elongation rates the resulting average FRAP curve did not result in an exponential curve [Fig. 6(b)], but rather had a bilinear time dependency with linear recovery at short times and a roundoff to the steady state.

We have also tested other types of distributions to estimate the accuracy and uniqueness of the distribution we have found, and identify the necessary characteristics in order to recreate the experimental results. We concentrated on studying the significance of the distribution's shape (symmetric or asymmetric) and moments (finite mean and infinite variance). We first tested how the shape of a distribution affects the shape of the calculated average curve. We therefore chose two known distributions that the dependence on their parameters can show a transition from asymmetric to symmetric shape. One is a Poisson distribution and the other is a χ^2 distribution. For the Poisson distribution we set $\lambda = 4$ which presents an asymmetric distribution [Fig. 6(c), inset], and $\lambda = 10$ which resembles the properties of normal distribution [Fig. 6(d), inset]. χ^2 distribution is another widely used distribution with finite mean and variance; it is highly asymmetric for a low number of degrees of freedom, and it converges again to normal distribution at a large number of degrees of freedom [Fig. 6(g), inset]. Finally we overview the weight of the moments of a distribution to the ensemble average curve. We tested the Pareto distribution which is a skewed heavy-tailed distribution [Fig. 6(h), inset]. For suitable parameter values, the Pareto distribution will have finite mean and infinite variance.

All results can be viewed in Fig. 6. These clearly show that the shape of the distribution has a strong effect on the recovery curve. Asymmetric distributions give better recovery behavior. When examining the Poisson and χ^2 distribution it is easy to see that the fit of the curve is getting worse as these distributions shift to a more symmetric shape distribution [Figs. 6(c)–6(g)]. Moreover, the residuals become larger and unevenly distributed. The next natural step was to examine whether an asymmetric distribution that has similar features will improve the fit. The Pareto distribution has the same features as our distribution regarding its shape and moments, but when looking at Fig. 6(h) one can see that the fit quality and residuals did not present better results. χ^2 distribution with five degrees of freedom [Fig. 6(f)] presents the best fit results and the minimal residuals among the additional tested distributions. Therefore, we conclude that the characteristic of the shape of the asymmetric distribution of transcription rates has the biggest influence on the average curve shape. Even so, the solution by our analytic distribution gives the best results with $R^2 = 0.99$ and smallest residuals [Fig. 6(a)].

In order to support our findings and to link them to additional transcription FRAP data, we applied our analysis and distribution to reconstruct previous experimental data performed on the same gene with two different promoters [13]. For most of the experiments, noise posed a problem as it dominated the nature of the recovery curve; nevertheless, many of the recovery curves showed a linear increase in the intensity. To obtain a quantitative measure of transcription rates, we tried fitting the single experiment's data with both a linear fit and an exponential fit following Eq. (2). About 10% of the data could not be fitted based on a poor R^2 ($\leqslant 0.7$). For the rest of the data sets, about 70% showed a better linear fit. These findings support the simple representation of a recovery curve of a single gene (Fig. 3), and suggest that it can be modeled with a constant transcription rate.

Next, we turned to examine the rates distribution. To test our distribution on more than one population experimental data

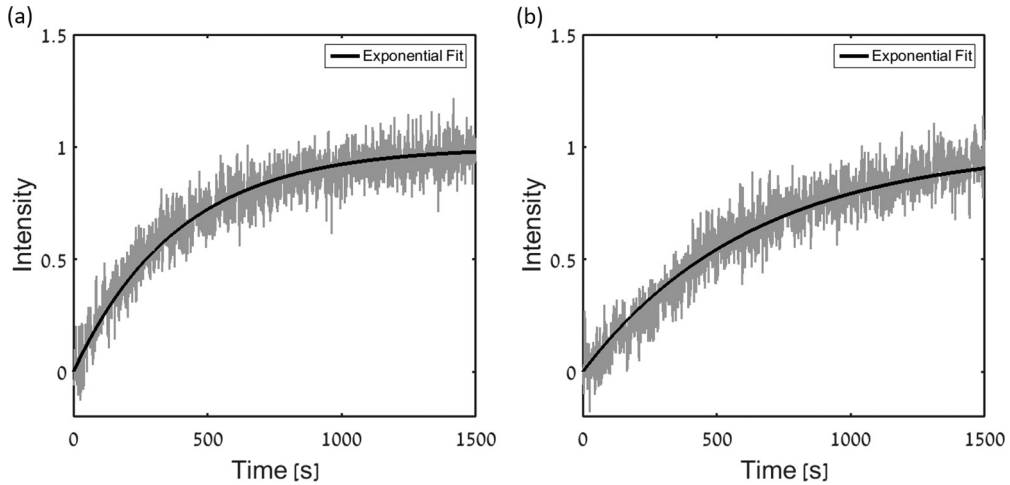

FIG. 7. Simulation results for rates distribution according to experimental results [13] for the CMV promoter (a) and endogenous cyclin D1 promoter (b). Noisy signal is the simulation data and solid line is an exponential fit according to Eq. (2). Fit results: for CMV $\gamma = (2.8 \pm 0.6) \times 10^{-3}$ s^{-1} and for cyclin D1 $\gamma = (1.6 \pm 0.2) \times 10^{-3}$ s^{-1}.

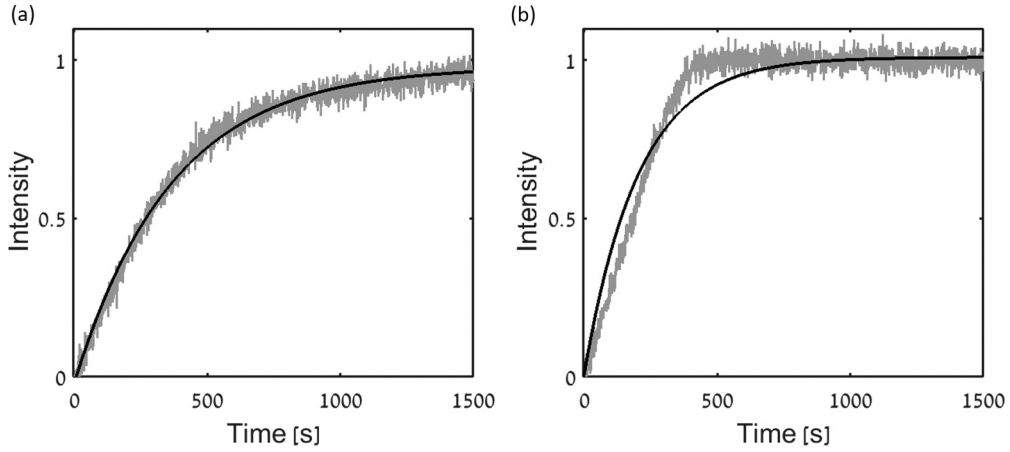

FIG. 8. Reconstruction of experimental data with rate distribution according to Eq. (5) (a) and according to normal distribution (b). For both distributions the average was set with $\gamma = 2.3 \times 10^{-3}$ s^{-1}. Data are shown as noisy signal with fitting according to Eq. (2) in a solid line.

we monitored transcription kinetics of the same gene under the control of two promoters that differ in their expression strength [13], the endogenous cyclin D1 promoter (D1) and a more efficient viral CMV promoter. We used the exponential parameter $\gamma = 2.8 \times 10^{-3}$ s^{-1} or $\gamma = 1.4 \times 10^{-3}$ s^{-1} as was found from FRAP results [13] for the two promoter types (CMV and cyclin D1 promoter, respectively), and inserted them into Eq. (5). Using these values, we simulated a FRAP curve that is the average of ten possible elongation rates according to the distribution we found. We fitted these curves with Eq. (2) and found dissociation rates that match the experimental results with excellent fit $R^2 = 0.89$ for the CMV promoter and $R^2 = 0.9$ for the cyclin D1 promoter. Figure 7 shows our fit parameters for the two promoters CMV, $\gamma = (2.8 \pm 0.6) \times 10^{-3}$ s^{-1}, and cyclin D1, $\gamma = (1.6 \pm 0.2) \times 10^{-3}$ s^{-1}. They are in good agreement with previous fit parameters.

We repeated this analysis for an average of more than ten rates and examined the whole data set of 27 measurements of the CMV promoter, and again got good agreement between experimental results and our distribution simulation with $\gamma = (2.3 \pm 0.2) \times 10^{-3}$ s^{-1}. Attempting to reconstruct the results with a normal distribution with the same average rate did not succeed (Fig. 8).

According to our findings, which are based on real-time live-cell experimental results, we conclude that transcription dynamics can only be described through a wide distribution of single-cell kinetic rates. Single-gene transcription exhibits constant average elongation rates but these rates vary widely between different cells. Even if the single polymerase steps have stochastic distribution, the net outcome of multiple such steps would constantly converge to an effective mean. Thus, elongation can be investigated as a single effective rate process.

IV. CONCLUSIONS

We have developed an analysis approach for ensemble-averaged transcription dynamics and extracted the distribution for single-site transcription rates. Our distribution analysis predicts not only the single-cell average transcription rate but also the diversity of transcription rates arising from inhomogeneity between cells. This distribution provides an explanation to the wide variability of previously published transcription rates as it supports the emergence of extreme rate values. The significant differences in rates between cells can be related to the availability of different transcription factors and nucleotides, chromatin state, and even to the stage of the cell cycle. We suggest that this distribution of kinetic rates is not only a signature of the cell state, but also a vital part of genetic activity regulation and diversification. Future studies should look into the biological factors that affect the elongation rate distribution or even change its type altogether.

ACKNOWLEDGMENTS

This work was supported in part by the Israel Centers of Research Excellence (ICORE) No. 1902/12 and Israel Science Foundation No. 51/12 (YG); and the European Research Council (YST).

[1] B. Munsky, G. Neuert, and A. van Oudenaarden, Science **336**, 183 (2012).
[2] O. G. Berg, J. Theor. Biol. **71**, 587 (1978).
[3] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, PLoS Biol. **4**, e309 (2006).
[4] D. R. Larson, R. H. Singer, and D. Zenklusen, Trends Cell Biol. **19**, 630 (2009).
[5] A. Raj and A. van Oudenaarden, Annual Rev. Biophys. **38**, 255 (2009).
[6] E. Bertrand, P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long, Mol. Cell **2**, 437 (1998).
[7] D. Fusco, N. Accornero, B. Lavoie, S. M. Shenoy, J.-M. Blanchard, R. H. Singer, and E. Bertrand, Curr Biol. **13**, 161 (2003).
[8] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025 (2005).
[9] K. A. LeCuyer, L. S. Behlen, and O. C. Uhlenbeck, EMBO J. **15**, 6847 (1996).

[10] Y. Brody, N. Neufeld, N. Bieberstein, S. Z. Causse, E.-M. Böhnlein, K. M. Neugebauer, X. Darzacq, and Y. Shav-Tal, PLoS Biology **9**, e1000573 (2011).
[11] X. Darzacq, Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer, Nature Struct. Mol. Biol. **14**, 796 (2007).
[12] S. Boireau, P. Maiuri, E. Basyuk, M. de la Mata, A. Knezevich, B. Pradet-Balade, V. Bäcker, A. Kornblihtt, A. Marcello, and E. Bertrand, J. Cell Biol. **179**, 291 (2007).
[13] S. Yunger, L. Rosenfeld, Y. Garini, and Y. Shav-Tal, Nat. Methods **7**, 631 (2010).
[14] T. Lionnet, K. Czaplinski, X. Darzacq, Y. Shav-Tal, A. L. Wells, J. A. Chao, H. Y. Park, V. de Turris, M. Lopez-Jones, and R. H. Singer, Nat. Methods **8**, 165 (2011).
[15] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer, Science **332**, 475 (2011).
[16] P. Maiuri, A. Knezevich, A. De Marco, D. Mazza, A. Kula, J. G. McNally, and A. Marcello, EMBO Rep. **12**, 1280 (2011).
[17] A. Marcello, Transcription **3**, 29 (2012).
[18] D. Cannon and J. R. Chubb, EMBO Rep. **12**, 1208 (2011).
[19] S. Yunger, L. Rosenfeld, Y. Garini, and Y. Shav-Tal, Nat. Protoc. **8**, 393 (2013).
[20] http://bigwww.epfl.ch/sage/soft/spottracker.
[21] B. L. Sprague, R. L. Pego, D. A. Stavreva, and J. G. McNally, Biophys. J. **86**, 3473 (2004).
[22] T. Lionnet, B. Wu, D. Grünwald, R. H. Singer, and D. R. Larson, Cold Spring Harbor Symp. Quant. Biol. **75**, 113 (2010).
[23] R. D. De Veaux, P. F. Velleman, and D. E. Bock, *Stats: Data and Models* (Pearson Addison Wesley, Boston, 2005).
[24] MATLAB, The MathWorks Inc., 2014.