Combinatorial effect of promoter activity, mRNA degradation and site-specific transcriptional pausing in modulating protein expression noise

Sangjin Kim [1,2,3], Christine Jacobs-Wagner [1,2,3,4*]

Affiliations:

[1] Microbial Sciences Institute, Yale University, West Haven, CT 06516 USA

[2] Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06511 USA

[3] Howard Hughes Medical Institute, Yale University, New Haven, CT 06536 USA

[4] Department of Microbial Pathogenesis, Yale School of Medicine, New Haven, CT 06536 USA

*Corresponding author and lead contact: christine.jacobs-wagner@yale.edu, tel: +1-203-737-6778, fax: +1-203-737-6715

Competing interests: The authors declare that no competing interests exist.

Keywords: Transcriptional pausing, stochastic modeling, protein expression noise, mRNA degradation

ABSTRACT

Genetically identical cells exhibit diverse phenotypes, even when experiencing the same environment. This phenomenon, in part, originates from cell-to-cell variability (noise) in protein expression. While various kinetic schemes of stochastic transcription initiation are known to yield transcriptional noise, how post-transcription initiation events may affect noise at the protein level remains incompletely understood. To address this question, we developed an integrated model of bacterial gene expression. By exploring different scenarios of transcription initiation, transcription elongation dynamics, mRNA degradation and gene copy number, we found that the mRNA lifetime, gene dosage and sequence-dependent transcriptional pausing modulate the protein expression noise that was set by the promoter properties. Our findings highlight the interplay between transcription initiation, transcription elongation, translation and mRNA degradation in shaping protein number distributions in clonal populations. They also have implications for our understanding of genome evolution, and suggest combinatorial strategies for modulating phenotypic variability by genetic engineering.

INTRODUCTION

One of the most important tasks cells do is regulate the levels of gene expression—the conversion of genetic information written in DNA into a certain amount of functional proteins. Interestingly, isogenic cells in the same environment produce variable amounts of mRNA and protein due to the stochastic nature of gene expression processes (1-7). Population-wide homogeneity in protein numbers (low noise) may be expected for proteins that are needed in all cells, such as house-keeping proteins. Consistent with this idea, experiments in *Escherichia coli* and budding yeast have shown that genes essential for viability tend to exhibit lower noise than nonessential genes (8-10). On the other hand, variability in protein expression can lead to beneficial traits for the population by generating variability in cell phenotypes. In bacterial populations, phenotypic heterogeneity can provide a "bet-hedging" strategy for survival in fluctuating and stressful environments (11-14). It can also be beneficial for cooperative social adaptations through "division of labor" within the cell population (15).

What is the origin of gene expression noise? Previous experimental and theoretical studies have shown that an important source of gene expression noise comes from transcription initiation, i.e., the loading of RNA polymerase (RNAP) onto the promoter region (6, 16-20). Specifically, if transcription initiation occurs stochastically at a constant frequency, a mode known as "nonbursty" initiation, the mRNA number at steady state will follow a Poisson distribution, which characteristically shows a Fano factor (variance/mean) equal to 1. In contrast, if the rate of transcription initiation varies over time, such as in pulsatile transcription from a promoter that cycles between active and inactive states, the mRNA levels become much more variable among cells (Fano factor > 1). This ON/OFF model of transcription, which is often referred to as "bursty" initiation, has been supported by the observation of pulsatile transcription events in live *E. coli* cells (21) and by the grouping of RNAPs along the rRNA operon in electron micrographs (22). The mRNA Fano factors measured in *E. coli* span from 1 to ~10, suggesting that both nonbursty and bursty promoters are likely in operation *in vivo* (9, 18, 19).

Cell phenotypes are generally dictated at the protein, rather than mRNA, level. While transcription initiation sets the initial conditions for protein expression noise, it is not fully understood how post-transcription initiation processes, such as transcription elongation, mRNA decay and translation, may affect protein distributions in a cell population. Modeling studies have already suggested a potential role for transcription elongation in noise generation by

showing that spurious sequence-independent RNAP pauses can create mRNA and protein bursts from nonbursty promoters depending on pause duration and frequency (23-26).

These studies motivated us to develop an integrated stochastic model for bacterial gene expression to examine the interplay between transcription initiation, transcription elongation dynamics, mRNA degradation and translation in shaping protein expression noise. The model includes well-established dependencies, such as the coupling between transcription and translation (27, 28), co-transcriptional mRNA degradation (29, 30), and the ribosome effect on mRNA degradation (31) (Fig. 1A).

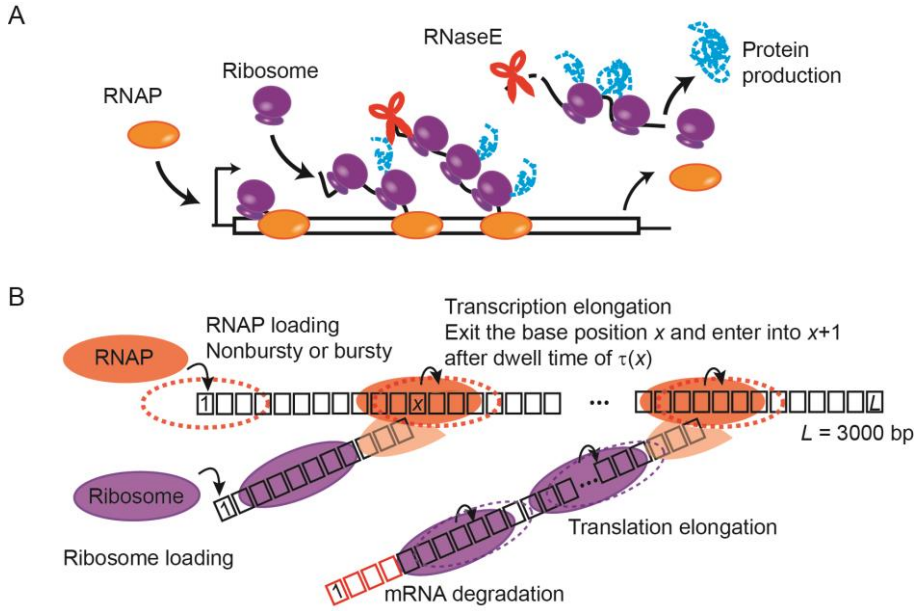

FIGURE 1 Integrated model of gene expression. (A) Schematic showing the current view on the temporal coordination between transcription, translation, and mRNA degradation in *E. coli*. RNAP loads onto the promoter. Once the ribosome binding site (RBS) is transcribed, ribosomes load and translocate. mRNA degradation by RNase E can start on the nascent mRNA. The first ribosome maintains contact with the RNAP throughout elongation. The first full-length protein is released almost immediately after the RNAP reaches the end of the DNA template. (B) Schematic showing the different steps included in our TASEP-based model of bacterial gene expression. Transcription starts at the first base pair of the template, and transcription elongation occurs by RNAP stepping along the template (3,000 bp in length). Translation initiates at the first base of the mRNA template (considered here as the RBS), and translation elongation occurs by ribosome stepping on each nucleotide of the mRNA template. mRNA degradation starts from the 5' end of the mRNA and continues following the last ribosome on the transcript. Degraded ribonucleotides are shown in red.

METHODS

Modeling transcription, translation, and mRNA degradation

Transcription was simulated in two steps. (1) Stochastic transcription initiation: In the case of a bursty promoter, we generated a series of time points when the promoter was ON or OFF, assuming that τ_{ON} and τ_{OFF} follow exponential waiting time distributions. In the case of a nonbursty promoter, the promoter was always ON. Next, we determined a series of time points for RNAP loading attempts during ON periods assuming exponentially distributed waiting times between loading attempts. (2) Transcription elongation by stochastic 1-nt stepping with an average speed of 30 nt/sec (Fig. S1A in the Supporting Material): We assumed an exponential dwell time distribution at each nucleotide position based on experimental observations made by others (32-34). Steric hindrance between RNAPs was checked before each stepping, assuming an RNAP footprint of 35 nt (35). After RNAP trajectories were simulated, the headways were calculated as the time interval between two consecutive RNAP exit events at every nucleotide position.

To model the coupling between transcription and translation (28), the first ribosome was loaded when the first 33-nt mRNA was transcribed due to the footprint sizes of the RNAP and the ribosome, 35 and 30 nt, respectively (35, 36). The first ribosome then moved on the nascent mRNA in concert with the RNAP to maintain contact throughout transcription elongation (27, 37). Additional ribosomes were stochastically loaded based an exponential waiting time distribution. These ribosomes made stochastic 1-nt steps at the same average speed as the RNAP (27). During ribosome translocation, the same steric hindrance principle as for RNAP translocation was used: ribosomes were unable to bypass each other on an mRNA and the first ribosome was unable to bypass the RNAP on the nascent mRNA.

In our simulations, the 5' end of mRNAs was inactivated following an exponential waiting time distribution from its synthesis. Specifically, once the 5'-end nucleotide was transcribed by an RNAP, it took on average 90 sec before it was inactivated to reflect experimental results (Fig. S1B). Then, the rest of the nucleotides on the mRNA were removed following the last ribosome (38). This model is consistent with previous experimental observations for the 5'-to-3' net directionality of mRNA degradation (29, 39, 40) and for the ribosome shielding effect (31). Each run of simulation calculated transcription, translation and

mRNA degradation occurring on a DNA template for 40 min. We performed a total of 1000 simulations for each scenario.

Analysis of the simulated data

We counted 1 mRNA when the first nucleotide (5' end) was present, unless noted otherwise (see Supporting Text). To obtain steady-state distributions of mRNA numbers, we counted the number of mRNAs made from a DNA template (total n = 1000 simulations), but not yet degraded at t = 30 min of simulation time (to ensure that the system had reached steady-state). We used this distribution to calculate the mean value and the Fano factor.

Protein number increased by 1 every time a ribosome reached the end of a transcript. To obtain distributions of protein numbers, we summed all proteins made from a DNA template between t = 20-30 min of simulation time. Mean and CV^2 of protein numbers were calculated from these distributions.

To calculate gene expression per cell with two gene copies, we added simulation results from two different DNA templates, assuming that gene copies within a cell act independently (41). All error bars indicate an estimation of the standard error of the mean calculated by bootstrap resampling the original sample size (1,000 simulations) 3,000 times.

RESULTS

Integrated modeling of bacterial gene expression

Our integrated model of gene expression has three components. The first two components, transcription and translation, were based on the totally asymmetric exclusion process algorithm (TASEP) (42-44). The DNA and mRNA templates were considered as one-dimensional lattices, where each lattice site corresponds to 1 base pair (bp) or nucleotide (nt). RNAPs and ribosomes with respective footprint sizes of 35 nt (35) and 30 nt (36) were modeled to stochastically load onto their DNA and mRNA templates, respectively. Once loaded, RNAPs and ribosomes stepped each lattice site from 5' to 3' (Fig. 1B). Because of steric hindrance, loading was allowed only when the loading region was cleared. During translocation, RNAPs and ribosomes kept a minimum distance of 35 nt and 30 nt, respectively. The first ribosome maintained contact with

the RNAP throughout transcription elongation to account for the coupling between transcription and translation in bacteria (28).

We modeled the third component, mRNA degradation, as the initial 5'-end mRNA inactivation followed by mRNA nucleotide removal in the 5'-to-3' direction (29, 39, 40). Protein degradation was considered negligible because most bacterial proteins are very stable (45).

While our model can be applied to any gene, we modeled the expression of the 3075-nt *lacZ* mRNA of *E. coli*, which is a popular model in quantitative gene expression studies. Given an average RNAP speed ($k_{\mathrm{ELONGATION}}$) of 30 nt/sec on the *lacZ* region (Fig. S1A) (27), the input average RNAP dwell time at each base position x was $\tau(x) = 1\ \mathrm{nt}/k_{\mathrm{ELONGATION}} = 1/30$ sec. We used the experimentally determined mean *lacZ* mRNA lifetime of 90 sec (Fig. S1B) as the first-order rate constant for 5'-end inactivation. For transcription initiation, we varied the RNAP loading rate on the DNA template to achieve a range of expression levels seen in experiments (18). For translation initiation, we used an average rate of ribosome loading of 0.2 sec^{-1} (46, 47). The average ribosome speed was the same as that of RNAPs (27). Stochastic simulations computed transcription, translation and mRNA degradation events based on exponential waiting time distributions (see Methods).

Effects of nonbursty and bursty transcription initiation on mRNA and protein expression

To build on previously known promoter properties, we considered two different types of promoters: nonbursty and bursty. Transcription initiation from nonbursty promoters was modeled as a Poisson process with RNAP loading rate, k_{LOADING}, which is the inverse of the loading interval, τ_{LOADING} (Fig. 2A). This parameter was varied to reflect the strength of a constitutive promoter. For the bursty case, the promoters cycled between ON and OFF states, with rate constants k_{ON} (rate of switching from OFF to ON state, $1/\tau_{\mathrm{OFF}}$) and k_{OFF} (rate of switching from ON to OFF state, $1/\tau_{\mathrm{ON}}$). RNAPs loaded only during the ON state at an interval of τ_{LOADING} (RNAP loading interval during the ON state) (Fig. 2A) (48). This model resulted in multiple RNAP loading events bunched in time in a pulsatile manner. We used $\tau_{\mathrm{OFF}} = 143$ sec and $\tau_{\mathrm{LOADING}} = 2.2$ sec (18), and varied the duration of the ON state (τ_{ON}) to achieve different expression levels (18).

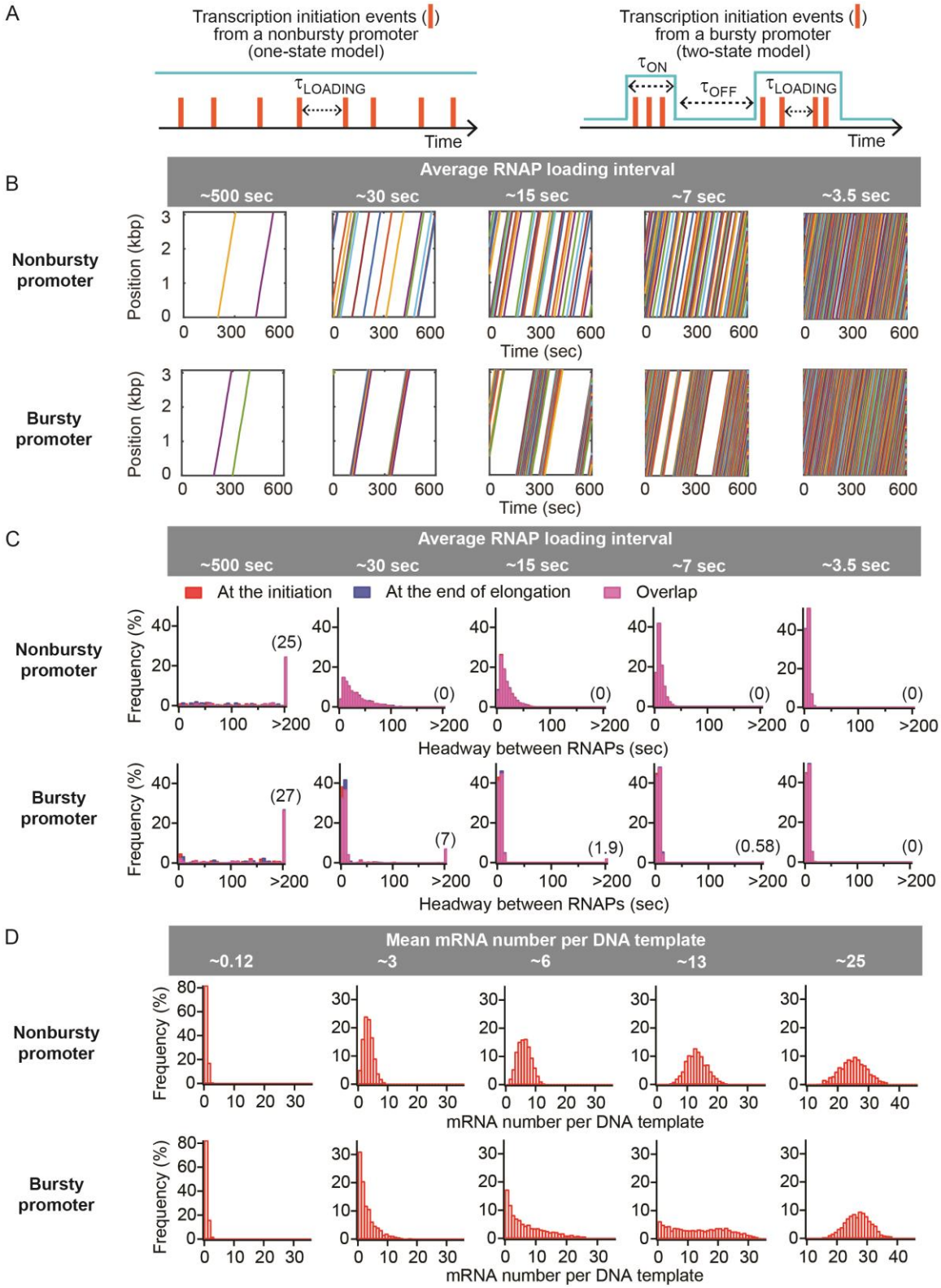

FIGURE 2 Effect of nonbursty and bursty promoters on RNAP trafficking and mRNA number distribution. (A) Schematics for two modes of transcription initiation. The promoter states are shown in cyan, and RNAP loading events are represented by orange bars. (B) Example

trajectories of RNAP loading and translocating on individual DNA templates over 10 min based on different transcription initiation conditions. Specific parameter values used are provided in Table S1. (C) Distributions of headways between two subsequently loaded RNAPs at initiation and completion of transcription. The probability of seeing RNAP pairs with large (>200) headways is indicated in parenthesis. (D) Steady-state distributions of mRNA numbers per DNA template. Each distribution is from the corresponding simulation in (B) and (C). mRNAs were counted based on the presence of the first base.

When the input RNAP loading interval is shorter than the time it takes for the previous RNAP to clear from the promoter region, the rate of RNAP loading is expected to be governed by the promoter clearance time (49). Our input average RNAP loading interval (varying between 2 and 500 sec; Table S1) was always longer than the average promoter clearance time (1.2 sec). However, the stochasticity in RNAP loading and translocation occasionally resulted in failure in RNAP loading due to the promoter being already occupied, yielding a lower *effective* transcription initiation rate (see Supporting Text, Fig. S2-S3, and Table S1).

When we compared nonbursty and bursty promoters of similar strength (i.e., yielding similar effective initiation rates and numbers of mRNAs at steady state), we found expected differences at intermediate average RNAP loading intervals (e.g., 7, 15 and 30 sec in Fig. 2B, 2C and 2D). First, bursty promoters showed pronounced bursts of RNAP loading events followed by notable OFF periods, resulting in temporal profiles of RNAP trajectories that were very different from those obtained from a nonbursty promoter of similar strength (Fig. 2B). Second, the distribution of time intervals (headways) between subsequently loaded RNAPs appeared largely exponential for nonbursty promoters. In contrast, we observed two populations for bursty promoters: small headways arising from loading events within an ON period and large headways arising from loading events separated by an OFF period (Fig. 2C). In both promoter cases, the headway set upon transcription initiation was conserved to the end of transcription, as shown by the near-perfect overlap in distributions between headways at initiation and at termination (Fig. 2C). Third, the distribution of the number of RNAPs on the DNA templates was wider for bursty promoters than for nonbursty promoters, with a noticeable peak close to zero due to the stochastic occurrence of OFF and ON states (Fig. S4). Lastly, the steady-state distributions of mRNA numbers for bursty initiations were broader than those for nonbursty ones despite having

similar mean mRNA numbers (Fig. 2D and see Fig. S5 for results using a different mRNA counting method).

In contrast, at initiation frequencies that were either very low or very high (e.g., average loading interval ≈ 500 sec or 3.5 sec), bursty promoters were virtually indistinguishable from nonbursty promoters since they effectively reproduced the characteristics of a nonbursty initiation (Fig. 2B, 2C, 2D and S4). At very low initiation frequencies, the ON period was too short to accommodate enough RNAP loading events to show transcriptional bursts. This is consistent with the experimentally determined mRNA Fano factor of 1 for the repressed *lacZ* promoter (18). At very high initiation frequencies, the OFF period was so short that it became negligible (1). These results suggest that transcriptional bursts are unlikely for genes at either side of the expression spectrum.

Importantly, our simulations identified a dynamic range of initiation rates for which our model produced a clear difference between nonbursty and bursty promoters (Fig. 2 and S4). Hereafter, we exploit this range to identify characteristics that distinguish between nonbursty and bursty transcription initiations. When we examined protein production under this range of transcription initiation rates, we found that the temporal profile of protein production was largely dictated by the temporal profile of RNAP loading onto the promoter. Nonbursty transcription initiations yielded a relatively constant number of proteins made from a DNA template over time (Fig. 3A, S6A and S6B). In contrast, bursty transcription initiations resulted in bursty protein productions, showing periods of time without any new protein production from the DNA template (Fig. 3A, S6A and S6B). As a result, bursty promoters produced a broader protein number distribution in comparison to nonbursty promoters despite both yielding the same mean protein numbers (Fig. 3B, S6C and S6D).

Short mRNA lifetimes facilitate production of protein bursts

The finding that a bursty promoter led to bursty protein production was dependent on the use of a 90-sec mRNA lifetime. When the lifetime of the mRNA was increased, protein bursts generated over time from bursty promoters became less apparent, as illustrated with a 10-min mRNA lifetime (Fig. 3C). To quantify the noise in protein production, we calculated the squared coefficient of variation (CV^2, $(\text{standard deviation/mean})^2$) of the protein number distribution (4).

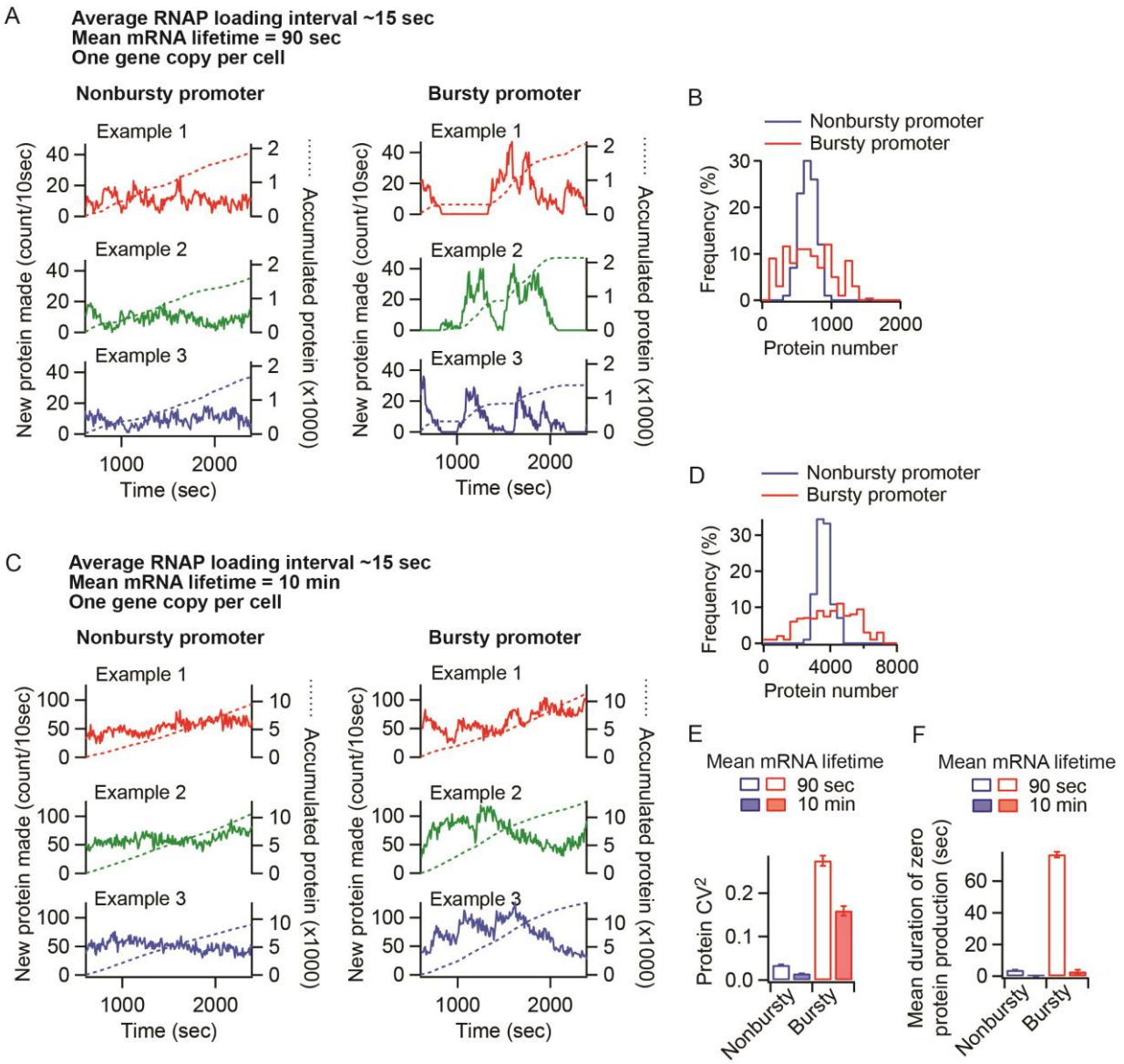

FIGURE 3 Bursty promoters result in bursty protein synthesis and greater noise in protein levels when the mRNA lifetime is short. (A) Temporal profiles of protein production. The number of new proteins made per DNA template was quantified every 10 sec. Example trajectories were from three independent simulations for nonbursty and bursty promoters with average RNAP loading intervals of ~15 sec. The mean mRNA lifetime was 90 sec. The number of proteins accumulated from expression of the DNA template is shown in dotted lines. Only the accumulation of proteins occurring within the time frame is shown. (B) Distributions of the number of proteins made per DNA template over 10 min. (C) Temporal profiles of protein production under the same conditions as in (A), except that the mRNA lifetime was 10 min. (D) Distributions of the number of proteins made per DNA template over 10 min when the mRNA lifetime was 10 min. (E) Variability in protein numbers for a nonbursty and bursty promoter, depending on the mRNA lifetime. (F) Mean durations of zero protein production in the temporal profiles shown in (A) and (C).

While bursty promoters still produced broader protein number distributions than nonbursty promoters (Fig. 3D), the noise from both types of promoters was reduced by the increase in mRNA lifetime (Fig. 3E). In both cases, the reduction in protein expression noise was correlated with an overall attenuation in temporal fluctuations of protein synthesis (Fig. 3C vs. Fig. 3A). Periods of no protein production virtually disappeared with the 10-min mRNA lifetime (Fig. 3F).

We reasoned that the reduced temporal fluctuations in protein synthesis stems from the mRNA lifetime being much longer than the RNAP loading interval, resulting in increased protein production between transcription events. Consistent with this idea, the noise in protein expression for both nonbursty and bursty promoters increased either by shortening the mRNA lifetime for a given average RNAP loading interval or by increasing the average RNAP loading interval for a given mRNA lifetime (Fig. 4A and 4B). In fact, if the mRNA lifetime was much smaller than the average RNAP loading interval (e.g., 90 sec vs. 500 sec), a nonbursty promoter was able to produce protein bursts (Fig. 4C and 4D), resulting in much higher protein production noise than in the opposite case when the mRNA lifetime was longer than the average RNAP loading interval (e.g., 90 sec vs. 15 sec) (Fig. 4E). This is consistent with *in vivo* observations that the occasional firing of the repressed *lac* promoter (average RNAP loading interval ≈ 40 min under the experimental condition used in the cited studies) causes spikes of protein production (46, 50). This is because each mRNA is degraded before the next one is made, resulting in well-separated bursts of protein production.

These results suggest that short mRNA lifetimes (in the minute time scale), a common characteristic of bacterial mRNAs (29, 39, 51), facilitate bursty protein synthesis and increase the variability in protein levels across the population, both for bursty and nonbursty promoters.

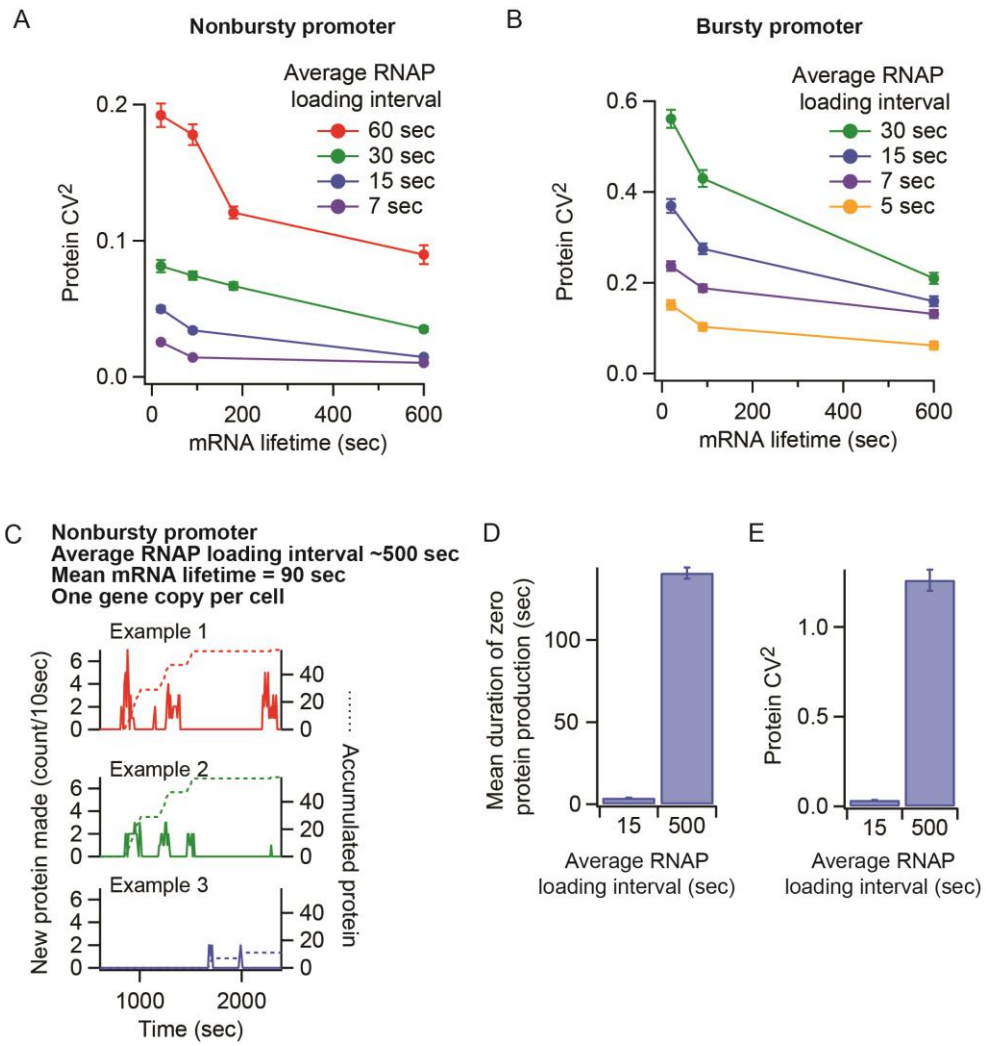

FIGURE 4 Temporal fluctuations in protein production and variability in protein number are modulated by the RNAP loading interval and the mRNA lifetime. (A and B) Variability in protein numbers as a function of mRNA lifetimes and RNAP loading intervals from nonbursty (A) and bursty (B) promoters. (C) Example temporal profiles of protein production from a nonbursty promoter with an average RNAP loading interval of 500 sec, which is much longer than the 90-sec mRNA lifetime. (D) Mean durations of zero protein production calculated from the temporal profiles of protein production shown in (C) and Fig. 3A. (E) Variability in protein numbers for the conditions described in (D).

Gene dosage has opposing effects on protein expression noise at the cell-population level

Our simulations quantified gene expression per DNA template. In real cells, a chromosomal gene can coexist in multiple copies due to DNA replication. We found that gene dosage has opposing

effects on protein expression noise at the cell-population level for both nonbursty and bursty promoters. First, having more than one gene copy smoothed the temporal fluctuations in mRNA production in a cell, because fluctuations were not synchronized among the gene copies (25). Due to this effect, virtual cells with two gene copies had smoother protein accumulation over time (Fig. 5A, dotted line) than cells with a single copy (Fig. 3A), resulting in reduced protein synthesis noise (Fig. 5B). Indeed, if *all* cells in a population contain *n* independent copies of a gene, the protein noise at the cell-population level (CV^2_{cell}) should decrease by 1/*n* from the protein noise at the gene-population level (CV^2_{gene}). However, in reality, a population consists of cells with different gene copy numbers depending on their stage of DNA replication. For example, *E. coli* cells under slow growth conditions (i.e., undergoing only one round of DNA replication) will contain either 1 or 2 copies of a gene of interest. A mixture of cells with different gene copy numbers will increase CV^2_{cell}, opposing the attenuating effect of gene copy number on CV^2_{cell}. For a population containing a mixture of cells with 1 or 2 gene copies, CV^2_{cell} can be predicted by the following equation:

$$CV^2_{cell} = \frac{1}{1+f} CV^2_{gene} + \frac{f(1-f)}{(1+f)^2},$$

where *f* is the fraction of cells with 2 gene copies (see Supporting Text for derivation). Accordingly, CV^2_{cell} will become higher or lower than CV^2_{gene} depending on the CV^2_{gene} and *f* values (Fig. 5C and S7). As discussed above, CV^2_{gene} is determined by several factors, including the temporal profile of RNAP loading and the mRNA lifetime. *f* is determined by cell cycle parameters and the chromosomal position of the gene of interest (52). The equation suggests that CV^2_{gene} for a gene located near the chromosomal terminus (lower *f*) is more likely to increase by a gene-dosage effect than CV^2_{gene} for a gene close to the origin of replication. In addition, since nonbursty promoters produce lower CV^2_{gene} than bursty promoters, the CV^2_{gene} of a nonbursty promoter is more likely to increase at the cell-population level ($CV^2_{cell} > CV^2_{gene}$) due to gene dosage while the opposite is true for the bursty promoters. When $CV^2_{gene} > 1$, as seen in minimally expressed promoters (Fig. 4C) such as the repressed Lac promoter (46, 50), the equation predicts that the gene-dosage effect will decrease protein expression noise at the cell-population level ($CV^2_{cell} < CV^2_{gene}$) (Fig. S7).

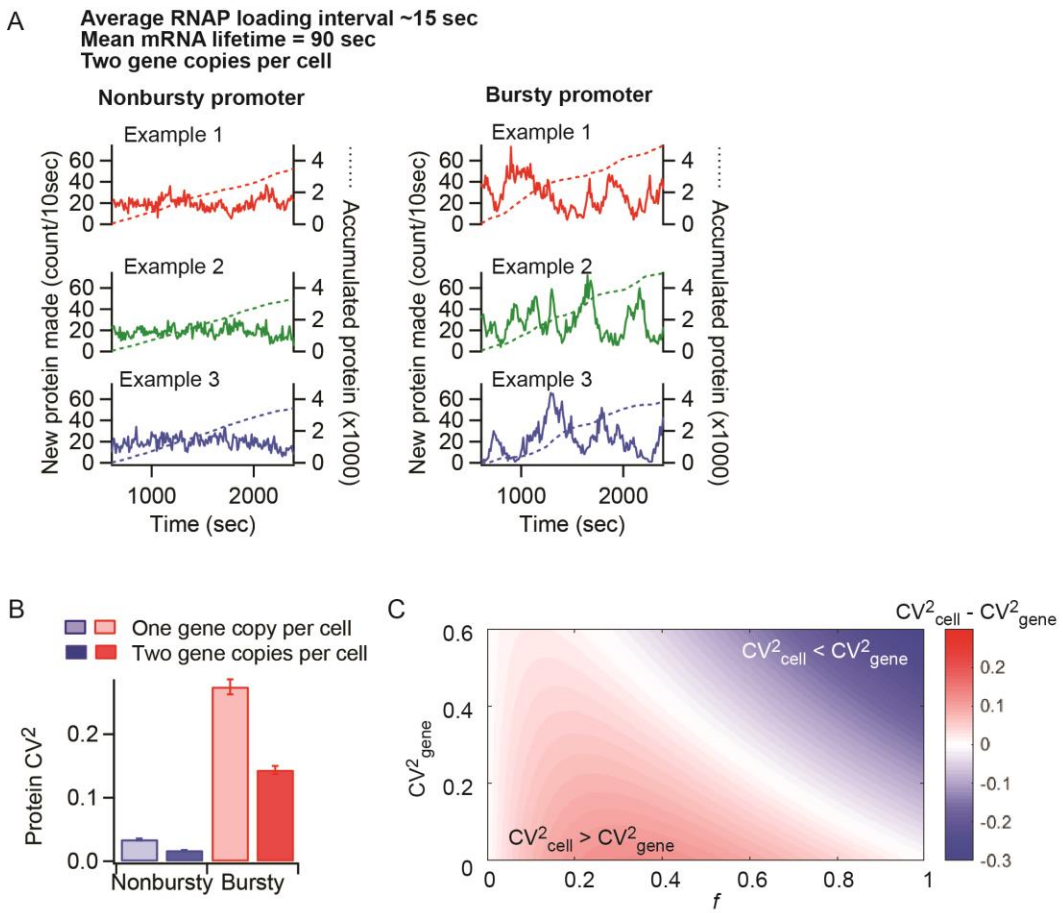

FIGURE 5 Gene dosage has opposing effects on gene expression noise at the cell-population level. (A) Temporal profiles of protein production under the conditions used in Fig. 3A, but for cells with two gene copies. (B) Effect of the number of gene copies (1 vs. 2) per cell on the variability in protein numbers from a nonbursty and bursty promoter with an average loading interval of ~15 sec. (C) Relationship between noise at the cell-population level (CV^2_{cell}) and at the gene-population level (CV^2_{gene}). The color map indicates whether CV^2_{cell} is greater or lower than CV^2_{gene}, due to opposing effects of the gene dosage. *f* indicates the fraction of cells with 2 gene copies when the population consists of cells with either 1 or 2 gene copies.

In the absence of RNAP pauses, RNAP bursts at initiation are largely preserved throughout transcription elongation

The gene expression scenarios described above showed that RNAP bursts are maintained from the promoter to the end of the template, as evidenced by the profile of RNAP trajectories (Fig. 2B) and the negligible changes in RNAP headway during transcription elongation (Fig. 2C). Furthermore, the Fano factor calculated based on the 5' end of the mRNA was similar to the Fano factor calculated based on the 3' end of the mRNA (Fig. S8), indicating that burstiness (or lack thereof) set at initiation did not change by the end of transcription. We wondered whether

these results were generalizable because a previous modeling study showed that RNAP congestion can cause RNAP bursts at initiation to completely disappear by the end of elongation (25). Complete loss of RNAP bursts means that the 3'-end mRNA Fano factor for a bursty promoter should be 1, a characteristic of Poissonian processes. We observed this phenomenon only if RNAPs were continually loaded back-to-back during the ON state, and even then, the loss of burstiness between transcription initiation and termination was most dramatic if the RNAP footprint was small (e.g., 1 nt) (see Supporting Text and Fig. S9, S10 and S11). We concluded that, under realistic parameter values, the memory of a bursty promoter's ON/OFF switch is largely maintained throughout transcription elongation.

Sequence-dependent pauses can attenuate RNAP bursts and affect gene expression

So far, our simulations have considered pause-free elongation. In *in vitro* experiments, RNAPs can pause spuriously along the DNA template (32, 53, 54). Modeling studies have shown that stochastic (sequence-independent) pauses can produce RNAP bursts from nonbursty promoters (23-26). It is also well-known that RNAPs pause at specific DNA sites for durations that range from seconds to ~1 minute (32, 34, 55-63). RNAP profiling experiments in *E. coli* have demonstrated that pause sites are common *in vivo* (64, 65).

The demonstrated prevalence of sequence-dependent pauses *in vivo* and their potentially long lifetimes motivated us to investigate their effects on RNAP trafficking and gene expression noise. For illustration purposes, we first modeled a single pause site in the middle of a 3-kbp gene, driven by either a nonbursty or bursty promoter with an average RNAP loading interval of ~15 sec (Fig. 6A). In this elongation profile, RNAPs resided at each nucleotide on average for 1/30 sec as before, except at the pause site (x_p), where the average RNAP dwell time (t_p) was varied. Dwell time of an RNAP at the pause site was randomly chosen from an exponential distribution with a mean dwell time of t_p (32-34). When the pause was much shorter than the RNAP loading interval, such as t_p = 1 sec, the effect on RNAP trafficking was negligible regardless of the promoter type, as shown by the near-perfectly overlapping distributions of Δheadways (headway at the end - headway at the start for two consecutive RNAPs) between the 1-sec pause and the no-pause cases (Fig. S12).

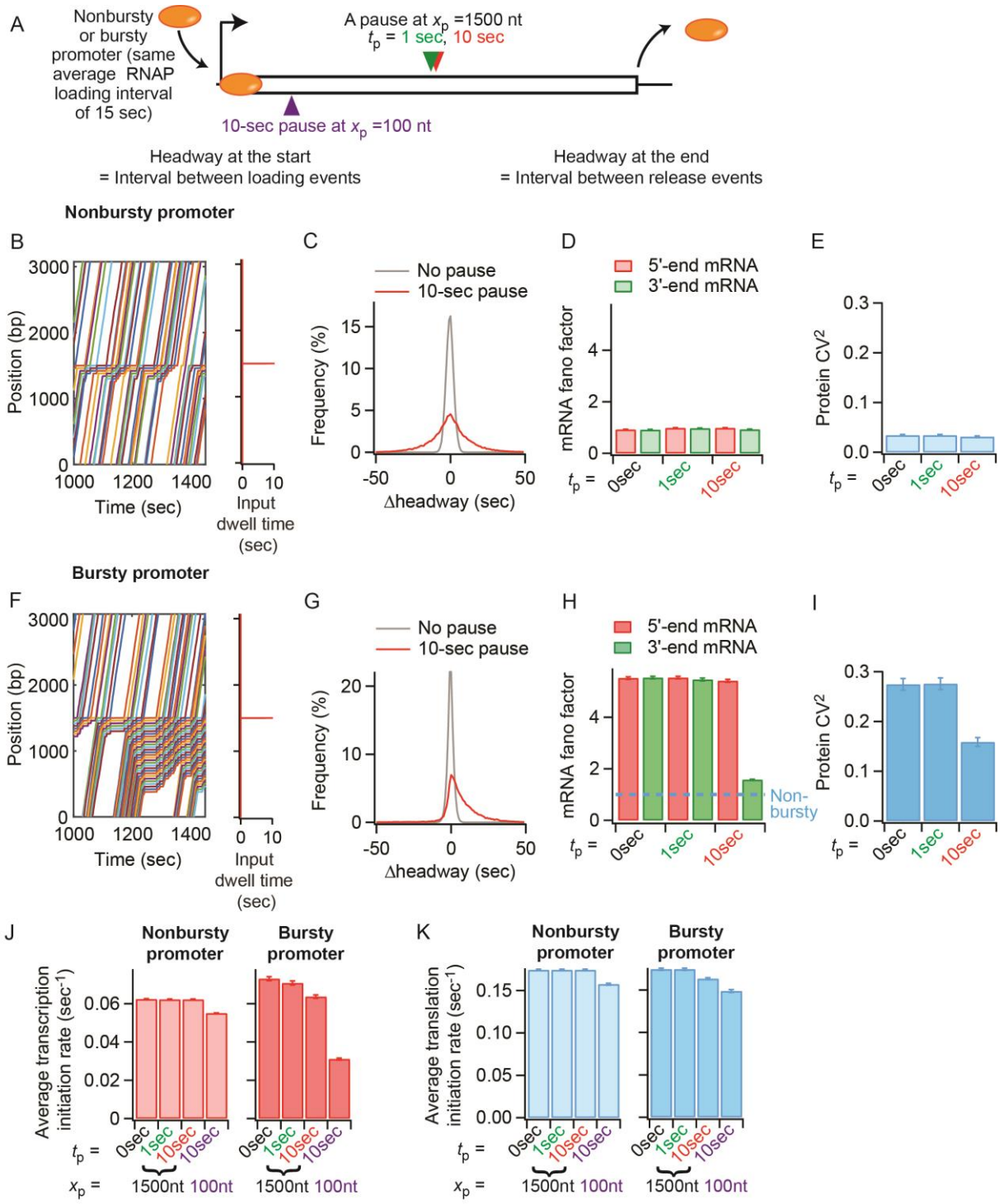

FIGURE 6 Sequence-dependent RNAP pauses affect gene expression dynamics and steady-state distributions of mRNA and protein numbers. (A) Schematics showing scenarios of a single RNAP pause site with different durations and locations: t_p = 1 sec and x_p = 1500 nt (green), t_p = 10 sec and x_p = 1500 nt (red) and t_p = 10 sec and x_p = 100 nt (purple). For each scenario, we considered either a nonbursty or bursty promoter with a similar average loading interval of ~15

sec (measured in the absence of any pause) and a mean mRNA lifetime of 90 sec. The consequence of the pause was analyzed in several ways, including RNAP headway change from initiation to the completion of transcription elongation (Δheadway). (B) Example trajectories of RNAPs along a DNA template from a simulation with a pause duration t_p = 10 sec for the nonbursty promoter. (C) Distributions of Δheadways for all simulated RNAPs (n = 1000 simulations) using the same conditions as in (B). (D) Fano factors calculated based on either the 5' or 3' end of the mRNA, considering three different pause durations for the nonbursty promoter. (E) Protein noise (CV^2) of the nonbursty promoter, considering three different pause durations at x_p = 1500 nt. (F) Example trajectories of RNAPs along a DNA template with t_p = 10 sec for the bursty promoter. (G) Distributions of Δheadways from all simulated RNAPs (n = 1000 simulations) using the same conditions as in (F). (H) 5'- and 3'-end mRNA Fano factors for the bursty promoter with different pause durations. The blue dotted line denotes a Fano factor equal to 1, representing a nonbursty process. (I) Protein noise (CV^2) for the bursty promoter, considering three different pause durations at x_p = 1500 nt. (J) Effect of sequence-dependent pauses of different durations and locations on transcription initiation rates. (K) Effect of these same pauses on translation initiation rates.

When the pause was similar to (e.g., t_p = 10 sec) or longer than the RNAP loading intervals, the pause resulted in two effects: (1) RNAP piling upstream of the pause site and (2) a change in RNAP headway downstream of the pause. In the case of a nonbursty initiation, these effects (Fig. 6B) resulted in a broader distribution of Δheadways between subsequent RNAPs, but the Δheadway distribution remained centered around zero, with no net change (Fig. 6C). The nonbursty conditions remained until the end of the template because the headway between subsequent RNAPs could either increase or decrease with similar probabilities. As a consequence, the Fano factor was maintained at ~1 at both the 5' and 3' ends of the mRNA (Fig. 6D) and the protein level noise was not affected (Fig. 6E). However, in the case of a bursty initiation, the frequent back-to-back loading of RNAPs (i.e., small initial headway) caused the majority of RNAPs to catch up with each other at the pause site (Fig. 6F). As a result, the headway between RNAPs showed net increase after the pause site (Fig. 6G). This situation is analogous to car traffic near a stop sign at rush hour. Cars pile up back-to-back before the stop sign, but as each car pauses at the intersection, they resume traveling with a new headway dictated by the pause duration. The pause-dependent increase in headway led initial RNAP bursts to largely dissipate into a nonbursty-like situation after the pause, as seen in the example RNAP trajectories (Fig. 6F) and in the large decrease in mRNA Fano factor between 5' and 3' ends (Fig. 6H). In other words, pauses similar to or longer than RNAP loading intervals diminished the effect from the bursty promoter's ON/OFF switch by increasing the RNAP headway after the

pause. This memory loss of initial conditions smoothed the temporal profile of protein production (Fig. S13) and lowered protein expression noise (Fig. 6I).

When we examined the effect of having two pause sites on a gene driven by a bursty promoter, we found that the second pause did not affect the RNAP headway as much as the first pause (Fig. S14, scenarios (ii) vs. (i)). We reasoned that the headway increase generated by the first pause produced fewer RNAPs to pile up at the second pause site. However, when the second pause was longer than the first one, it further increased RNAP headways (Fig. S14, scenarios (iii) vs. (i)), suggesting that multiple long-lived pauses can have additive effects.

In the *E. coli* genome, RNAP pause sites are often found near promoters (64, 66, 67). When we modeled a promoter-proximal pause, such as a 10-sec pause at x_p = 100 nt (Fig. 6A), we found that the RNAP piling caused by the pause could easily extend to the promoter, blocking the loading of additional RNAPs and thereby reducing the effective rate of transcription initiation for both nonbursty and bursty promoters (Fig. 6J). In addition, RNAP stalling near the promoter resulted in reduced translation initiation rates (Fig. 6K) because ribosomes piling on the short nascent transcript could quickly reach the RBS. A smaller reduction in transcription and translation initiation rates was also observed for more distal pauses, such as a 10-sec pause at x_p = 1,500 nt (Fig. 6J and 6K) because RNAP piles occasionally reached the promoter, preventing some RNAPs to load.

We concluded that sequence-dependent pauses can impact gene expression by attenuating RNAP bursts and/or by affecting transcription and translation initiation rates.

DISCUSSION

Our study shows that the conditions set by transcription initiation play a critical role in modulating protein expression noise, consistent with previous works (16, 17, 20, 68, 69). It also highlights the importance of post-transcription initiation processes, such as mRNA degradation and RNAP pausing, in shaping the steady-state distribution of protein production.

The lifetime of mRNA impacts the amount of proteins produced, and larger protein abundance is expected to be associated with lower variability in protein numbers (17, 20, 69). We showed that mRNA lifetimes longer than the OFF period dampen the temporal fluctuations of protein synthesis (Fig. 3A vs. 3C). Because one mRNA typically generates more than one

protein, the mRNA lifetime reduces the effect of the bursty promoter's ON/OFF switch (Fig. 4B). The mRNA lifetime also smooths temporal fluctuations in protein production and hence its noise in the case of nonbursty promoters (Fig. 4A) because mRNA expression from these promoters still fluctuates over time due to the stochastic nature of transcription initiation.

Altogether, this suggests that mRNA degradation is an important factor to consider when studying noise in gene expression. The difference among genes can be considerable given that the lifetimes of bacterial mRNAs can vary over an order of magnitude (29, 39, 51). Furthermore, mRNA lifetimes have been shown to vary with environmental cues in some cases (70), suggesting a unappreciated means by which environmental changes may affect the expression noise of certain genes.

Our results identified sequence-dependent RNAP pausing as another important post-transcription initiation factor that can attenuate noise in protein synthesis. RNAP pauses affect protein expression noise by modulating RNAP trafficking downstream of the pause (Fig. 6F and 6I). The headway between RNAPs upon transcription termination, which shapes the temporal fluctuations in mRNA and protein production, is altered by the pause (Fig. 6F and 6G) to the point that the memory of a bursty promoter's ON/OFF switch can be lost (Fig. 6H). This effect depends on the duration of the pause relative to the loading intervals of RNAPs (Fig. 6G vs. Fig. S12D).

We also note that the noise-attenuating effect of RNAP pauses occurred when the pause duration was shorter than the mRNA lifetime, which we expect to be the general case *in vivo* based on reported mRNA lifetimes and pause durations. However, if a pause lasted longer than the mRNA lifetime, we observed bursty protein production, even in the case of a nonbursty promoter (data not shown). Protein bursts arose because the majority of ribosomes piled behind the paused RNAP on the nascent mRNA and then traveled together once the RNAP escaped the pause site, as reported before (71). Certain mechanisms, such as premature termination (72) and physical push between RNAPs (73), may help mitigate these pause-dependent effects on protein expression noise.

In addition to the duration of the pause, its location relative to the promoter matters. These pause properties determine the probability that trailing RNAPs reach the promoter and prevent new initiation events (Fig. 6J). Pauses close to promoters can quickly result in RNAP piles blocking the promoter. Promoter blockage by pause-induced RNAP queuing has been

discussed in the literature in the context of the highly transcribed ribosomal RNAs in *E. coli* (74, 75). Given the prevalence of sequence-dependent pauses near promoters (64, 67, 76), the occurrence of RNAP queuing up to the promoters may be more common than anticipated, beyond highly transcribed genes. Promoter-proximal pauses have indeed been shown to reduce mRNA production in *E. coli* (66). Our modeling also shows that RNAP pauses can reduce not only the rate of transcription initiation, but also that of translation initiation (Fig. 6K) when ribosomes catch up to the stalled RNAP and pile up to the RBS. This double negative effect on protein production stems from the temporal coupling between transcription and translation in bacteria. Altogether, our work suggests potential roles for RNAP pauses in modulating protein expression level and noise, on top of their known functions in transcription-translation coupling, transcriptional termination, transcriptional fidelity, sigma factor association at the promoter, and elongation speed in bacteria (77-80).

In our simulations, the probability of RNAP pausing at a particular site was assumed to be 100%. In reality, it is likely to be lower because the probability varies 30-80% in *in vitro* experiments depending on the pause site (34). While a lower probability does not qualitatively alter our conclusions, an accurate quantitative prediction of mRNA and protein number distributions will require knowledge of not only the location and mean duration of pauses, but also the probability of RNAPs to pause at those sites. It may also be important to consider whether the duration of a pause is affected by the environment. For instance, a high concentration of magnesium has been shown to increase pause duration in a gene encoding a magnesium transporter in *Salmonella enterica* serovar Typhimurium (63). This example raises the possibility of environmental control of protein expression noise through modulation of pause duration.

mRNA lifetimes and RNAP pauses are evolvable features at the gene-specific level because they are sequence-dependent and can change through mutations. Therefore, our work predicts that mutations altering pause duration or mRNA lifetime (e.g., by affecting mRNA secondary structure at the pause site or at the 5'-untranslated region) will affect protein expression noise at the level of individual genes. Our work also suggests possible ways by which protein expression noise may change globally. For example, mutations that render RNAP less prone to pausing (e.g., *rpoB2* mutation in *E. coli* (81)) are expected to affect the protein expression noise of pause-sensitive genes. In addition, mutations decreasing the activity of the

mRNA degradation machinery (RNase E in *E. coli*) (82) are predicted to decrease protein expression noise at the genome-wide level.

CONCLUSION

Our work underscores the combinatorial origin of protein expression noise, which includes not only promoter properties, but also mRNA lifetime, transcriptional pausing, and gene location relative to the origin of replication. This combinatorial effect may affect how genome sequences evolve by modulating phenotypic variability. It could also be exploited for genetic engineering in synthetic biology. For instance, our findings suggest conditions to maximize protein expression noise and phenotypic diversity: a bursty promoter, a short mRNA lifetime, an absence of long pauses, and gene location near the chromosomal terminus. The opposite conditions are expected to minimize phenotypic heterogeneity.

SUPPORTING MATERIAL

Supporting Material contains Supporting Text, Supporting Materials and Methods, fourteen figures and two tables. Source code can be found in http://github.com/JacobsWagnerLab/TASEP noise.

SUPPORTING CITATIONS

References (82-86) appear in the Supporting Material.

AUTHOR CONTRIBUTIONS

C.J.-W and S.K. designed the study and wrote the manuscript. S.K. performed the experiments and modeling. S.K. and C.J.-W. analyzed the data. C.J.-W supervised the project.

ACKNOWLEDGEMENTS

We thank Bruno Beltran and Drs. Damon Clark, Thierry Emonet and Alvaro Sanchez for helpful discussions. We also thank the members of the Jacobs-Wagner Laboratory for discussions and for critical reading of the manuscript. C.J.-W. is an investigator of the Howard Hughes Medical Institute. The authors declare no conflict of interest.

REFERENCES

1. Sanchez, A., and I. Golding. 2013. Genetic determinants and cellular constraints in noisy gene expression. Science 342:1188-1193.
2. McAdams, H. H., and A. Arkin. 1997. Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. USA. 94:814-819.
3. Raser, J. M., and E. K. O'Shea. 2005. Noise in gene expression: origins, consequences, and control. Science 309:2010-2013.
4. Paulsson, J. 2004. Summing up the noise in gene networks. Nature 427:415.
5. Kaern, M., T. C. Elston, W. J. Blake, and J. J. Collins. 2005. Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. 6:451-464.
6. Elowitz, M. B., A. J. Levine, E. D. Siggia, and P. S. Swain. 2002. Stochastic gene expression in a single cell. Science 297:1183-1186.
7. Raj, A., and A. van Oudenaarden. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell 135:216-226.
8. Newman, J. R. S., S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman. 2006. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature 441:840-846.
9. Taniguchi, Y., P. Choi, G. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. Xie. 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science 329:533-538.
10. Silander, O. K., N. Nikolic, A. Zaslaver, A. Bren, I. Kikoin, U. Alon, and M. Ackermann. 2012. A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. PLoS Genet. 8:e1002443.
11. Balaban, N. Q., J. Merrin, R. Chait, L. Kowalik, and S. Leibler. 2004. Bacterial persistence as a phenotypic switch. Science 305:1622-1625.
12. Veening, J.-W., W. K. Smits, and O. P. Kuipers. 2008. Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. 62:193-210.
13. Thattai, M., and A. van Oudenaarden. 2004. Stochastic gene expression in fluctuating environments. Genetics 167:523-530.
14. Kussell, E., and S. Leibler. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. Science 309:2075-2078.
15. West, S. A., and G. A. Cooper. 2016. Division of labour in microorganisms: an evolutionary perspective. Nat. Rev. Microbiol. 14:716-723.
16. Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. USA. 98:8614-8619.
17. Shahrezaei, V., and P. S. Swain. 2008. Analytical distributions for stochastic gene expression. Proc. Natl. Acad. Sci. USA. 105:17256-17261.
18. So, L.-h., A. Ghosh, C. Zong, L. A. Sepulveda, R. Segev, and I. Golding. 2011. General properties of transcriptional time series in Escherichia coli. Nat. Genet. 43:554-560.
19. Jones, D. L., R. C. Brewster, and R. Phillips. 2014. Promoter architecture dictates cell-to-cell variability in gene expression. Science 346:1533-1536.
20. Singh, A., B. S. Razooky, R. D. Dar, and L. S. Weinberger. 2012. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. Mol. Syst. Biol. 8:607.

21. Golding, I., J. Paulsson, S. M. Zawilski, and E. C. Cox. 2005. Real-time kinetics of gene activity in individual bacteria. Cell 123:1025-1036.

22. Voulgaris, J., S. French, R. L. Gourse, C. Squires, and C. L. Squires. 1999. Increased rrn gene dosage causes intermittent transcription of rRNA in Escherichia coli. J. Bacteriol. 181:4170-4175.

23. Rajala, T., A. Häkkinen, S. Healy, O. Yli-Harja, and A. S. Ribeiro. 2010. Effects of transcriptional pausing on gene expression dynamics. PLoS Comput. Biol. 6:e1000704.

24. Voliotis, M., N. Cohen, C. Molina-París, and T. B. Liverpool. 2008. Fluctuations, pauses, and backtracking in DNA transcription. Biophys. J. 94:334-348.

25. Dobrzyński, M., and F. J. Bruggeman. 2009. Elongation dynamics shape bursty transcription and translation. Proc. Natl. Acad. Sci. USA. 106:2583-2588.

26. Fujita, K., M. Iwaki, and T. Yanagida. 2016. Transcriptional bursting is intrinsically caused by interplay between RNA polymerases on DNA. Nat. Commun. 7:13788.

27. Proshkin, S., A. R. Rahmouni, A. Mironov, and E. Nudler. 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science 328:504-508.

28. Miller, O. L., B. A. Hamkalo, and C. A. Thomas. 1970. Visualization of bacterial genes in action. Science 169:392-395.

29. Chen, H., K. Shiroguchi, H. Ge, and X. S. Xie. 2015. Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. Mol. Syst. Biol. 11:781.

30. Yarchuk, O., N. Jacques, J. Guillerez, and M. Dreyfus. 1992. Interdependence of translation, transcription and mRNA degradation in the lacZ gene. J. Mol. Biol. 226:581-596.

31. Deana, A., and J. G. Belasco. 2005. Lost in translation: the influence of ribosomes on bacterial mRNA decay. Genes Dev. 19:2526-2533.

32. Adelman, K., A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts, and M. D. Wang. 2002. Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. Proc. Natl. Acad. Sci. USA. 99:13538-13543.

33. Abbondanzieri, E. A., W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block. 2005. Direct observation of base-pair stepping by RNA polymerase. Nature 438:460-465.

34. Herbert, K. M., A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick, and S. M. Block. 2006. Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell 125:1083-1094.

35. Gelles, J., and R. Landick. 1998. RNA polymerase as a molecular motor. Cell 93:13-16.

36. Steitz, J. A. 1969. Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. Nature 224:957-964.

37. Kohler, R., R. A. Mooney, D. J. Mills, R. Landick, and P. Cramer. 2017. Architecture of a transcribing-translating expressome. Science 356:194-197.

38. Carrier, T. A., and J. D. Keasling. 1997. Mechanistic modeling of prokaryotic mRNA decay. J. Theor. Biol. 189:195-209.

39. Selinger, D. W., R. M. Saxena, K. J. Cheung, G. M. Church, and C. Rosenow. 2003. Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. Genome Res. 13:216-223.

40. Cannistraro, V. J., and D. Kennell. 1985. Evidence that the 5' end of lac mRNA starts to decay as soon as it is synthesized. J. Bacteriol. 161:820-822.

41. Sepúlveda, L. A., H. Xu, J. Zhang, M. Wang, and I. Golding. 2016. Measurement of gene regulation in individual cells reveals rapid switching between promoter states. Science 351:1218-1222.

42. MacDonald, C. T., and J. H. Gibbs. 1969. Concerning the kinetics of polypeptide synthesis on polyribosomes. Biopolymers 7:707-725.

43. MacDonald, C. T., J. H. Gibbs, and A. C. Pipkin. 1968. Kinetics of biopolymerization on nucleic acid templates. Biopolymers 6:1-25.

44. Tripathi, T., and D. Chowdhury. 2008. Interacting RNA polymerase motors on a DNA track: Effects of traffic congestion and intrinsic noise on RNA synthesis. Phys. Rev. E 77:011921.

45. Nath, K., and A. L. Koch. 1970. Protein degradation in Escherichia coli: I. Measurement of rapidly and slowly decaying components. J. Biol. Chem. 245:2889-2900.

46. Cai, L., N. Friedman, and X. S. Xie. 2006. Stochastic protein expression in individual cells at the single molecule level. Nature 440:358-362.

47. Li, G.-W., D. Burkhardt, C. A. Gross, and J. S. Weissman. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 157:624-635.

48. Peccoud, J., and B. Ycart. 1995. Markovian modeling of gene-product synthesis. Theor. Popul. Biol. 48:222-234.

49. Ehrensberger, Andreas H., Gavin P. Kelly, and Jesper Q. Svejstrup. 2013. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. Cell 154:713-715.

50. Yu, J., J. Xiao, X. Ren, K. Lao, and X. S. Xie. 2006. Probing gene expression in live cells, one protein molecule at a time. Science 311:1600-1603.

51. Bernstein, J., A. Khodursky, P. Lin, S. Lin-Chao, and S. Cohen. 2002. Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. USA. 99:9697-9702.

52. Bremer, H., and P. P. Dennis. 1996. Modulation of chemical composition and other parameters of the cell by growth rate. Escherichia coli and Salmonella: cellular and molecular biology 2:1553-1569.

53. Neuman, K. C., E. A. Abbondanzieri, R. Landick, J. Gelles, and S. M. Block. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell 115:437-447.

54. Shaevitz, J. W., E. A. Abbondanzieri, R. Landick, and S. M. Block. 2003. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature 426:684-687.

55. Shundrovsky, A., T. J. Santangelo, J. W. Roberts, and M. D. Wang. 2004. A single-molecule technique to study sequence-dependent transcription pausing. Biophys. J. 87:3945-3953.

56. Hein, P. P., K. E. Kolb, T. Windgassen, M. J. Bellecourt, S. A. Darst, R. A. Mooney, and R. Landick. 2014. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. Nat. Struct. Mol. Biol. 21:794-802.

57. Duchi, D., David L. V. Bauer, L. Fernandez, G. Evans, N. Robb, Ling C. Hwang, K. Gryte, A. Tomescu, P. Zawadzki, Z. Morichaud, K. Brodolin, and Achillefs N. Kapanidis. 2016. RNA polymerase pausing during initial transcription. Mol. Cell 63:939-950.

58. Davenport, J., R., G. J. L. Wuite, R. Landick, and C. Bustamante. 2000. Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase. Science 287:2497-2500.

59. Landick, R., J. Carey, and C. Yanofsky. 1985. Translation activates the paused transcription complex and restores transcription of the trp operon leader region. Proc. Natl. Acad. Sci. USA. 82:4663-4667.

60. Artsimovitch, I., and R. Landick. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc. Natl. Acad. Sci. USA. 97:7090-7095.

61. Lee, D. N., L. Phung, J. Stewart, and R. Landick. 1990. Transcription pausing by Escherichia coli RNA polymerase is modulated by downstream DNA sequences. J. Biol. Chem. 265:15145-15153.

62. Forde, N. R., D. Izhaky, G. R. Woodcock, G. J. L. Wuite, and C. Bustamante. 2002. Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by Escherichia coli RNA polymerase. Proc. Natl. Acad. Sci. USA. 99:11682-11687.

63. Hollands, K., A. Sevostiyanova, and E. A. Groisman. 2014. Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. Proc. Natl. Acad. Sci. USA. 111:E1999-E2007.

64. Larson, M. H., 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. 2014. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. Science 344:1042-1047.

65. Vvedenskaya, I. O., H. Vahedian-Movahed, J. G. Bird, J. G. Knoblauch, S. R. Goldman, Y. Zhang, R. H. Ebright, and B. E. Nickels. 2014. Interactions between RNA polymerase and the “core recognition element” counteract pausing. Science 344:1285-1289.

66. Hatoum, A., and J. Roberts. 2008. Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. Mol. Microbiol. 68:17-28.

67. Mooney, R. A., S. E. Davis, J. M. Peters, J. L. Rowland, A. Z. Ansari, and R. Landick. 2009. Regulator trafficking on bacterial transcription units in vivo. Mol. Cell 33:97-108.

68. Friedman, N., L. Cai, and X. S. Xie. 2006. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys. Rev. Lett. 97:168302.

69. Bar-Even, A., J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai. 2006. Noise in protein expression scales with natural protein abundance. Nat. Genet. 38:636-643.

70. Caron, M.-P., L. Bastet, A. Lussier, M. Simoneau-Roy, E. Massé, and D. A. Lafontaine. 2012. Dual-acting riboswitch control of translation initiation and mRNA decay. Proc. Natl. Acad. Sci. USA. 109:E3444–E3453.

71. Mäkelä, J., J. Lloyd-Price, O. Yli-Harja, and A. S. Ribeiro. 2011. Stochastic sequence-level model of coupled transcription and translation in prokaryotes. BMC Bioinformatics 12:121.

72. Ray-Soni, A., M. J. Bellecourt, and R. Landick. 2016. Mechanisms of bacterial transcription termination: all good things must end. Annu. Rev. Biochem. 85:319-347.

73. Epshtein, V., and E. Nudler. 2003. Cooperation between RNA polymerase molecules in transcription elongation. Science 300:801-805.

74. Klumpp, S., and T. Hwa. 2008. Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. Proc. Natl. Acad. Sci. USA. 105:18159-18164.

75. Ehrenberg, M., P. P. Dennis, and H. Bremer. 2010. Maximum rrn promoter activity in Escherichia coli at saturating concentrations of free RNA polymerase. Biochimie 92:12-20.

76. Adelman, K., and J. T. Lis. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. 13:720-731.

77. Landick, R. 2006. The regulatory roles and mechanism of transcriptional pausing. Biochem. Soc. Trans. 34:1062-1066.

78. Imashimizu, M., N. Shimamoto, T. Oshima, and M. Kashlev. 2014. Transcription elongation. Transcription 5:e28285.

79. Nudler, E. 2012. RNA polymerase backtracking in gene regulation and genome instability. Cell 149:1438-1445.

80. Perdue, S. A., and J. W. Roberts. 2011. σ70-dependent transcription pausing in Escherichia coli. J. Mol. Biol. 412:782-792.

81. McDowell, J., J. Roberts, D. Jin, and C. Gross. 1994. Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. Science 266:822-825.

82. Moffitt, J. R., S. Pandey, A. N. Boettiger, S. Wang, and X. Zhuang. 2016. Spatial organization shapes the turnover of a bacterial transcriptome. eLife 5:e13065.

83. Griffith, K. L., and R. E. Wolf. 2002. Measuring β-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. Biochem. Bioph. Res. Co. 290:397-402.

84. Joo, C., and T. Ha. 2012. Labeling DNA (or RNA) for single-molecule FRET. Cold Spring Harbor Protocols 2012:1005-1008.

85. Sliusarenko, O., J. Heinritz, T. Emonet, and C. Jacobs-Wagner. 2011. High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. Mol. Microbiol. 80:612-627.

86. Kepes, A. 1969. Transcription and translation in the lactose operon of Escherichia coli studied by in vivo kinetics. Prog. Biophys. Mol. Biol. 19:199-236.