Biophysical Journal, Volume *114*

Supplemental Information

Effects of mRNA Degradation and Site-Specific Transcriptional Pausing on Protein Expression Noise

Sangjin Kim and Christine Jacobs-Wagner

Supporting Material

Effects of mRNA degradation and site-specific transcriptional pausing on protein expression noise

Sangjin Kim and Christine Jacobs-Wagner

Contents

Supporting Text

Guidelines for using our simulation code

Our MATLAB-based code can be used to reproduce the data presented in this paper and to produce new data using different input variables. Below, we explain (i) the input variables that need to be changed to model the expression of genes of interest (*lacZ* and others) and (ii) the types of simulation results (output) that can be obtained.

Universal parameters (no need to change):

Parameter name as in the code	Value (unit)	Description
`dx`	1 (nt)	Size of the lattice for TASEP. RNAP and ribosome make an one-directional step every 1 nt.
`polWidth`	35 (bp)	Footprint size of RNAPs
`riboWidth`	30 (nt)	Footprint size of ribosomes

Gene-specific parameters (we used values relevant for the *E. coli lacZ* gene):

Parameter name as in the code	Value (unit)	Description
`geneLength`	3075 (bp)	Length of the gene of interest. The first base is assumed as the promoter, transcription start site, as well as the ribosome-binding site.
`mRNALL`	90 (sec)	Mean mRNA lifetime
`proteinLL`	1200 (sec)	Mean protein lifetime. See `simTime`.
`kRiboLoading`	0.2 (sec^{-1})	Translation initiation rate
`kOn`/ `kOff`/ `kLoading`	0.007/ varied/ 0.45 (sec^{-1})	Transcription initiation rate for bursty promoters. The inverse of `kOn` and `kOff` are average duration of OFF and ON periods, respectively. `kLoading` is the RNAP loading rate during ON periods.
`kLoading`	Varied (sec^{-1})	Transcription initiation rate for nonbursty promoters.
`avgSpeed`	30 (nt/sec)	Mean RNAP or ribosome speed
`pauseProfile`	`'flat'` or `'OnepauseAbs'` or `'MultipauseAbs'` (capital letters are important)	Profile of pauses along the gene. `'flat'` is for no pause. `'OnepauseAbs'` is for one pause site. `'MultipauseAbs'` is for two pause sites along the gene.
`pauseSite`	1500 (nt)	A pause site at a single position.

	1500 and 2500	Two pause sites at two distinct locations
pauseDuration	10 (sec) 10 and 15	Pause duration at a single site. Pause durations specified for two pause sites (corresponding to the location specified in pauseSite).
pauseProb	80 (%)	Probability of pausing at the pause site specified in pauseSite.
specificDwelltime		Dwell time of an RNAP at each nucleotide position along the template. Generally, it is calculated based on input (avgSpeed, pauseProfile, pauseSite, pauseDuration). However, one can specify dwell time at every nucleotide position by modifying specificDwelltime array in the code.
simTime	0:1:40*60 (sec)	Simulation time is from 0 to 40 min (2400 sec). The simulation time is set to achieve steady state in mRNA distribution. Longer simulation time would be needed for long mRNA lifetimes. To achieve steady state in protein levels with protein lifetime of 20 min, we set the simulation time from 0 to 200 min.

Technical parameters that users may want to change:

Parameter name as in the code	Value (unit)	Description
N_totalLoci	100	Total copy number of DNA templates to be simulated. This equals the number of iterations. To get distributions of mRNA and proteins per DNA template, we performed many iterations. Default value for N_totalLoci is 100 and we repeat 10 times, for a total of 1000 iterations (this helps more efficient PC core usage than setting N_totalLoci as 1000).

code_for_gene_expression_dist_no_elongation:

This is the folder to check if one wants to test an elongation-free model (i.e., that does not include transcription or translation elongation). Based on the transcription initiation, translation initiation, and mRNA lifetime input values, one can obtain distributions of mRNA and protein numbers per DNA template. mRNA and proteins are counted without any length information, so the results can be directly compared with analytical models that do not consider elongation

processes (e.g. (1-4). The mRNA distributions are calculated at multiple time points, so that users can see the temporal progression of mRNA levels (e.g., as in the case of transcription induction experiments). The same is true for protein levels, when protein degradation was considered. Additionally, the code calculates the average RNAP loading rate, the RNAP headway distribution (at the time of initiation), and the mRNA lifetime distribution. See `mastertranscript.m` for example command lines to use. The output parameters of the analysis are described below:

- `tsxInitiationrate`: Effective RNAP loading rate (sec^{-1}) and RNAP loading interval (sec) by calculating the number of RNAPs loaded during the sample time (default: 15-30 min) after the start of simulation.
- `RNAPheadway`: Distribution of RNAP headway at the start of transcription.
- `mRNAdist`: Distribution of mRNA copy numbers per DNA template at steady state (or at a specific time of interest).
- `mRNAstat`: mean, Fano factor, CV and CV^2 of `mRNAdist`.
- `proteindist`: Distribution of the number of proteins accumulated during 10 min between 20-30 min of the simulation time.
- `proteinstat`: mean, Fano factor, CV and CV^2 of `proteindist`.
- `steadystateprotein`: mean, Fano factor, CV and CV^2 of protein abundance considering protein degradation.

`code_for_gene_expression_distribution`:

This is the folder to check if one wants to test our integrated model, as described in Fig. 1B. The outputs are in the form of distributions or average values, calculated from many iterations. The output parameters of "par" functions are described below:

- `loadingSample`: The number of RNAPs successfully loaded during the sample time (default: 15-30 min after the start of simulation) on each DNA template.
- `presentonDNA`: The number of elongating RNAPs on each DNA template at a given time (default: 20 min after the start of simulation).
- `fishSignal1, fishSignal2, fishSignal3`: The number of mRNAs produced (and not yet decayed) from each DNA template at a given time (default: `fishTime` = every 1 min in the first 30 min of the simulation). mRNAs are counted either by 5'-end mRNA nucleotide (`fishSignal1`), or 3'-end mRNA nucleotide (`fishSignal2`), or using tiling probes (`fishSignal3`), as mentioned in Fig. S2B.
- `riboNumHistFine`: Distribution of the number of ribosomes successfully loaded per mRNA during its lifetime. This indicates translation efficiency. It is calculated from transcripts whose RNAPs initiated during the sample time (default: 15-30 min after the start of simulation).

- `rEndStampA`: The number of ribosomes that finish translation elongation from each DNA template during a certain time period (default: every 10 sec during 10-40 min after the start of simulation).
- `proteinLoci2`: The number of proteins produced during a certain time period (default: 20-30 min after the start of simulation).
- `proteinSS`: The number of proteins produced (and not yet degraded) from each DNA template at a given time (default: measured every 1 min during `fishTime`)..
- `lifeTimeHist1, lifeTimeHist2, lifeTimeAvg`: Distributions of the lifetime of mRNA at the 5' and 3' ends are `lifeTimeHist1` and `lifeTimeHist2`, respectively. Average lifetimes of 5'- and 3'-end mRNA are `lifeTimeAvg`. Lifetimes were calculated from transcripts whose RNAPs were loaded during the sample time (default: 15-30 min after the start of simulation).
- `tDiffStartHist, tDiffEndHist, tDiffDHist`: Distributions of RNAP headway at the first base of the template and at the end of the template are `tDiffStartHist` and `tDiffEndHist`, respectively. `tDiffDHist` is the distribution of headway changes (Δheadway) during transcription elongation.

In order to run the code, one should start with `masterscript.m`.

Note: Since the simulation is based on the MATLAB function *parfor*, it is important to specify the number of workers (CPU core) inside the code. The default is set as 12 workers. If the available CPU number is different from 12, one needs to change the *parpool* line in the code accordingly.

`code_for_RNAP_traffic`:

This is the folder to check if one wants to examine only the RNAP traffic (without mRNA degradation and translation). Code in this folder calculates transcription initiation and elongation under various initiation and elongation conditions. The output is RNAP trajectories or time points when an RNAP exits each nucleotide position (`exitTime`). Codes do not use the *parfor* function, as in the previous case, and therefore, the variables are more trackable.

Supporting Materials and Methods

Miller's β-galactosidase assay (related to Fig. S1A)

A wild-type *E. coli* strain (MG1655) was grown in liquid cultures of M9 minimal medium supplemented with 0.2% glycerol, 0.1% casamino acids and 1 mg/L thiamine at 30ºC. Expression of *lacZ* was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) when the cultures reached early exponential growth phase (OD_{600} ~0.2 from >10^3 dilution from overnight culture). IPTG (0.5 mM) was added at t = 0 and 300 µL of culture was taken every 20-60 sec and pipetted into a 1.5-mL tube containing cold chloramphenicol (5 mg/mL; to be diluted 10 fold by the cell culture). The tubes were kept in an ice bucket throughout the remaining procedure, following a standard Miller assay protocol (5). In short, 190 µL cell cultures from each tube were transferred to a 96-well plate to measure the optical density of the culture at 600 nm (OD600). Next, the remaining cells in the tubes were lysed by adding 50 µL of lysis buffer (60 mM disodium phosphate, 40 mM monosodium phosphate, 10 mM potassium chloride, 1 mM magnesium sulfate, 0.0075% sodium dodecyl sulfate, plus 188 mM β-mercaptoethanol) and 20 µL of chloroform. After lysis, 150 µL of the supernatant was transferred to a new 96-well plate, containing 40 µL of ortho-nitrophenyl-β-galactoside (ONPG) solution (4 mg/mL in the lysis buffer). Hydrolysis of ONPG by LacZ in the cell lysate produces *o*-nitrophenol, which is yellow. The color change was monitored by measuring absorbance at 420 nm with a microplate reader (Synergy 2, Biotek). Plus, absorbance at 550 nm was measured to account for cell debris. LacZ activity was calculated as:

$$LacZ\ activity = 1000 \times \frac{OD_{420} - 1.75 \times OD_{550}}{t \times OD_{600}} \times dilution\ factor,$$

where t is the ONPG reaction time in minutes and OD indicates optical density measurements. The dilution factor is the fraction of cell lysates in the final volume of the ONPG reaction.

Fluorescence *in situ* hybridization (FISH) microscopy (related to Fig. S1B)

To determine the lifetime of *lacZ* mRNAs, we grew wild-type *E. coli* strain (MG1655) in M9 minimal medium supplemented with 0.2% glycerol, 0.1% casamino acids and 1 mg/L thiamine at 30ºC. A culture grown to early exponential growth phase (OD_{600} ~0.2 from >10^3 dilution from overnight culture) was split into two separate cultures. One, the "repressed" sample, was used

directly to examine *lacZ* mRNA expression under the repressed condition. The other, the "induced" sample, was first used to induce *lacZ* expression by IPTG (0.2 mM). In this induced sample, glucose (500 mM) was added at t = 3.5 min to turn off the promoter and to observe degradation of the *lacZ* mRNA. Aliquots (750 µL) were taken from the induced culture every 1 min after restoring repression, and were immediately put into a 1.5-mL tube containing a 4x fixing solution made of 16% formaldehyde in sodium phosphate buffer at pH 7.4. Cells were fixed for 15 min at room temperature and 30 min on ice. For the repressed sample, an aliquot was taken only once and cells were fixed in the same manner. Next, the fixed cells were washed with diethyl pyrocarbonate (DEPC)-treated phosphate buffered saline (PBS) buffer 3 times and were applied onto coverslips coated with poly-L-lysine. All the remaining procedures were performed while cells adhered on the coverslip.

Cells were lysed with 70% ethanol for 5 min. Next, prehybridization was performed in a solution contatining 20% formamide, 2x saline-sodium citrate buffer (2x SSC; 300 mM sodium chloride, 30 mM sodium citrate, pH 7.0), and 0.2 mM vanadyl ribonucleoside complex (VRC) for 30 min at 37ºC. Hybridization was performed with 24 20-nt-long oligonucleotides that have sequences complementary to the first 1000-nt region of the *lacZ* mRNA (see Table S1 for sequence data). The probes were labeled with Cy5 at the 5' end (6) and purified by ultra performance liquid chromatography (Acquity UPLC system, Waters) on an Acquity BEH C18 column. The eluents were A: 100 mM triethylammonium acetate in Milli-Q water and B: 100% acetonitrile. The gradient was set as follows: 0-5 min with 0% B, 5-35 min with a 0-30% linear gradient of B, 35-37 min with a 30-100% linear gradient of B, and 37-40 min with 0% B. The flow rate was kept at 0.1 mL/min. Chromatograms were recorded at 260 and 650 nm.

The hybridization solution consists of 4 nM of probes in 20% formamide, 2x SSC, 0.2 mM VRC, 10% dextran sulfate, 0.1% bovine serum albumin, and 0.4 mg/mL *E. coli* tRNA. The solution was applied to cells for 2 h at 37ºC. After washing with wash solutions (25% formamide and 2x SSC) and DEPC-PBS, DEPC-PBS was applied on the sample, and the coverslip was mounted on a glass slide. Imaging by phase contrast and fluorescence microscopy was performed on an Eclipse 80i microscope (Nikon) equipped with a phase-contrast objective Plan Apochromat 100×/1.40 NA (Nikon) and an Orca-II-ER CCD camera (Hamamatsu Photonics).

To determine the average fluorescent intensity of a single *lacZ* mRNA, we measured the fluorescence spot intensity of *lacZ* mRNA signal under the repressed condition (7, 8) using the

spotFinder tool in MicrobeTracker (9). The number of mRNAs per cell was obtained by dividing the total fluorescence signal per cell by the single-mRNA signal.

Supporting Table

Table S1. Oligonucleotide probes used for the *lacZ* mRNA FISH microscopy experiment

Oligonucleotide name	Sequence (from 5')
lacZ1	GTGAATCCGTAATCATGGTC
lacZ2	TCACGACGTTGTAAAACGAC
lacZ3	ATTAAGTTGGGTAACGCCAG
lacZ4	TATTACGCCAGCTGGCGAAA
lacZ5	ATTCAGGCTGCGCAACTGTT
lacZ6	AAACCAGGCAAAGCGCCATT
lacZ7	AGTATCGGCCTCAGGAAGAT
lacZ8	AACCGTGCATCTGCCAGTTT
lacZ9	TAGGTCACGTTGGTGTAGAT
lacZ10	AATGTGAGCGAGTAACAACC
lacZ11	GTAGCCAGCTTTCATCAACA
lacZ12	AATAATTCGCGTCTGGCCTT
lacZ13	AGATGAAACGCCGAGTTAAC
lacZ14	AATTCAGACGGCAAACGACT
lacZ15	TTTCTCCGGCGCGTAAAAAT
lacZ16	ATCTTCCAGATAACTGCCGT
lacZ17	AACGAGACGTCACGGAAAAT
lacZ18	GCTGATTTGTGTAGTCGGTT
lacZ19	TTAAAGCGAGTGGCAACATG
lacZ20	AACTGTTACCCGTAGGTAGT
lacZ21	ATAATTTCACCGCCGAAAGG
lacZ22	TTTCGACGTTCAGACGTAGT
lacZ23	ATAGAGATTCGGGATTTCGG
lacZ24	TTCTGCTTCAATCAGCGTGC

The probes were labeled with Cy5 at the 5' end and were used for the FISH microscopy experiment shown in Fig. S1B.

Supporting Figures

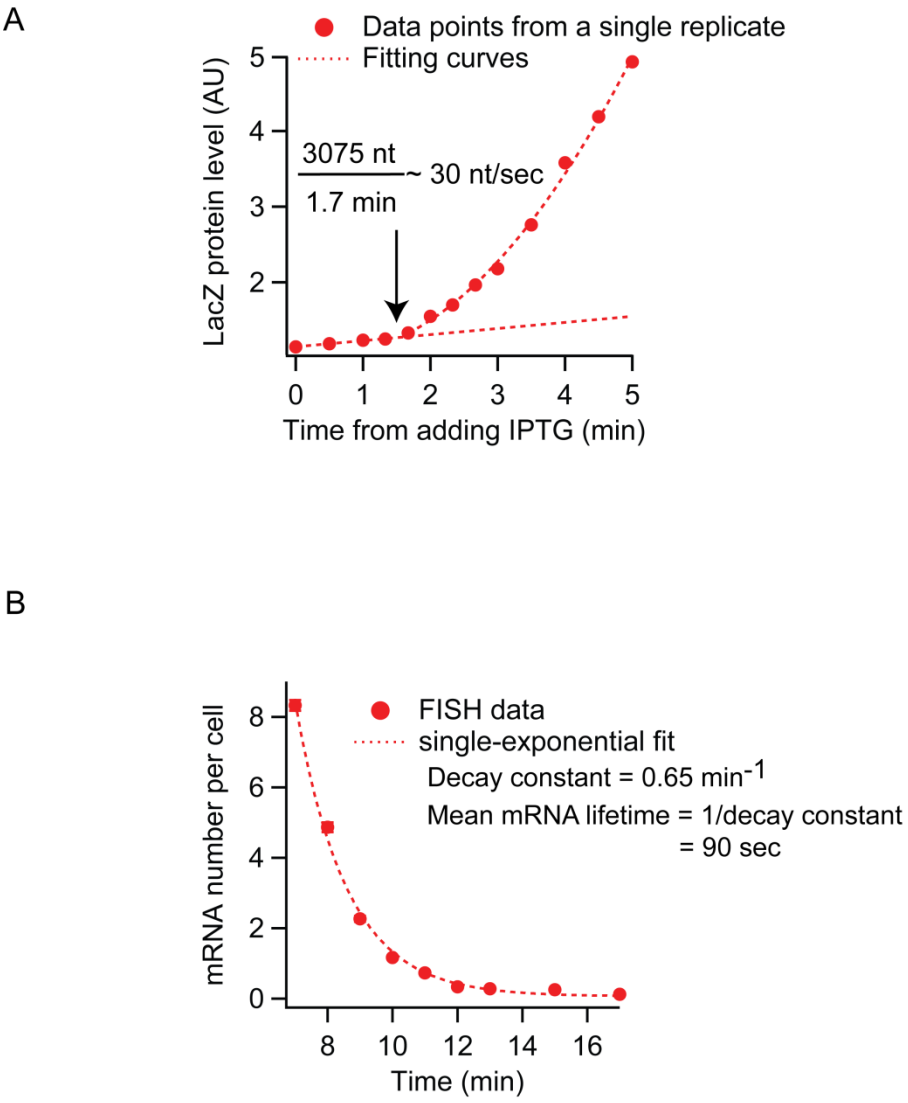

FIGURE S1. Experimental determination of RNAP translocation speed and *lacZ* mRNA lifetime. (A) *In vivo* determination of RNAP speed on the *lacZ* DNA template. The overall speed of RNAP during transcription elongation was calculated from the time lag of the first LacZ protein appearance after addition of the inducer IPTG to wild-type *E. coli* MG1655 cell culture (10). Red dotted lines are a line fit for baseline and a quadratic fit for the increase of LacZ level over time, respectively (11). From the two fits, we found that it took ~1.7 min for the synthesis of 3075-nt *lacZ* mRNA. By performing this procedure in multiple replicates ($n > 3$), we obtained ~30 nt/sec of average RNAP speed. (B) Determination of *lacZ* mRNA lifetime *in vivo*. Mean lifetime of *lacZ* mRNA was measured from wild-type *E. coli* cells growing under the same condition as in (A). We fit an exponential function to the decaying mRNA signal to obtain the mean mRNA lifetime. Error bars are the standard error of mean, estimated from bootstrapping.

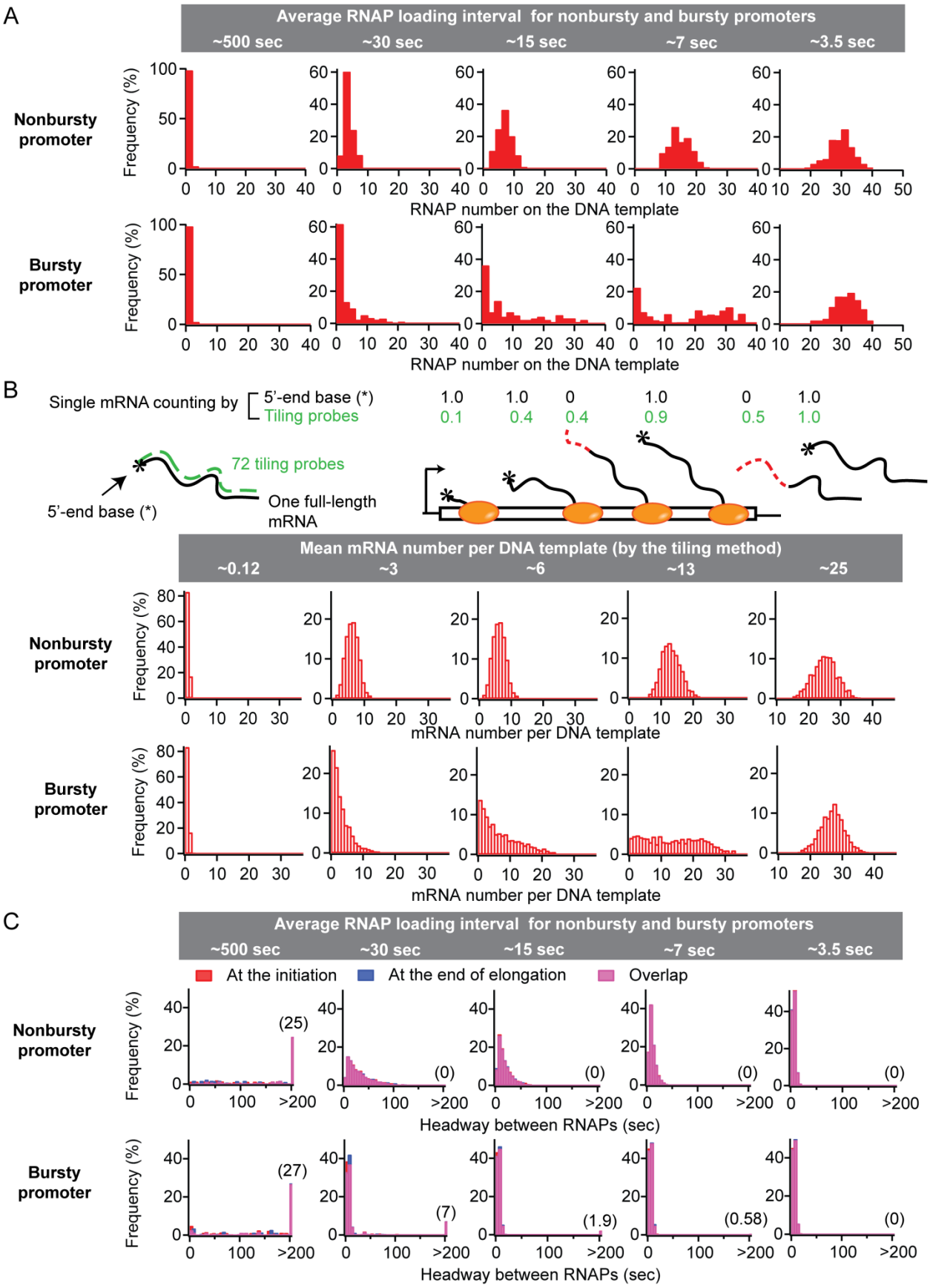
A
Average RNAP loading interval for nonbursty and bursty promoters
~500 sec
~30 sec
~15 sec
~7 sec
~3.5 sec
Nonbursty promoter
Bursty promoter
Frequency (%)
RNAP number on the DNA template
B
Single mRNA counting by
5'-end base (*)
Tiling probes
1.0 1.0 0 1.0 0 1.0
0.1 0.4 0.4 0.9 0.5 1.0
72 tiling probes
5'-end base (*)
One full-length mRNA
Mean mRNA number per DNA template (by the tiling method)
~0.12
~3
~6
~13
~25
mRNA number per DNA template
C
Average RNAP loading interval for nonbursty and bursty promoters
~500 sec
~30 sec
~15 sec
~7 sec
~3.5 sec
At the initiation
At the end of elongation
Overlap
(25)
(0)
(27)
(7)
(1.9)
(0.58)
Headway between RNAPs (sec)

FIGURE S2. Effect of nonbursty and bursty promoters on RNAP traffic and mRNA number distributions. (A) Number of RNAPs on the DNA template under various conditions of transcription initiation used in Fig. 2. We plotted the number of RNAPs on a DNA template at a given moment (at t = 1200 sec of the simulation time) from 1000 simulations. The distinction between nonbursty and bursty transcription initiations was apparent in the intermediate range of initiation intervals, as shown in Fig. 2B. (B) Distributions of mRNA numbers determined by a method that better reflects the counting method used in experiments. (Top) Schematics of different mRNA counting methods. In general, we counted mRNAs by considering their first base (5' end, marked as *), as in Fig. 2C. However, in experiments, the number of mRNA molecules per cell is generally obtained by mRNA fluorescence *in situ* hybridization (FISH) microscopy using multiple probes that tile the mRNA (e.g., 72 independent probes of 20 nt for *lacZ* mRNA (7, 8)). Since our modeling results contain information about which nucleotides are present in each mRNA at a given time, we were able to examine the effect of using multiple tiling probes on the mRNA number distributions. In this case, mRNA was counted as one mRNA when all the nucleotides complementary to the 72 probes, equally spaced throughout 3000 nt, were present on the mRNA. Using the first-base counting method, mRNAs are counted as one whenever the 5'-end base (*) is present. Using the tiling method, nascent or partially degraded mRNAs are counted as a fraction based on their length. Red dotted lines indicate degraded mRNA bases. (Bottom) Steady-state distributions of mRNA numbers per DNA template (from n = 1000 simulations) when mRNAs were counted using 72 virtual tiling probes to mimic FISH experiments (7, 8). Each distribution is from the corresponding simulation in Fig. 2C. (C) Distributions of headways between two subsequently loaded RNAPs at initiation and completion of transcription under various transcription initiation conditions used in Fig. 2B. The probability of seeing RNAP pairs with large (>200) headways is indicated in parentheses. The results were obtained from simulations described in Fig. 2.

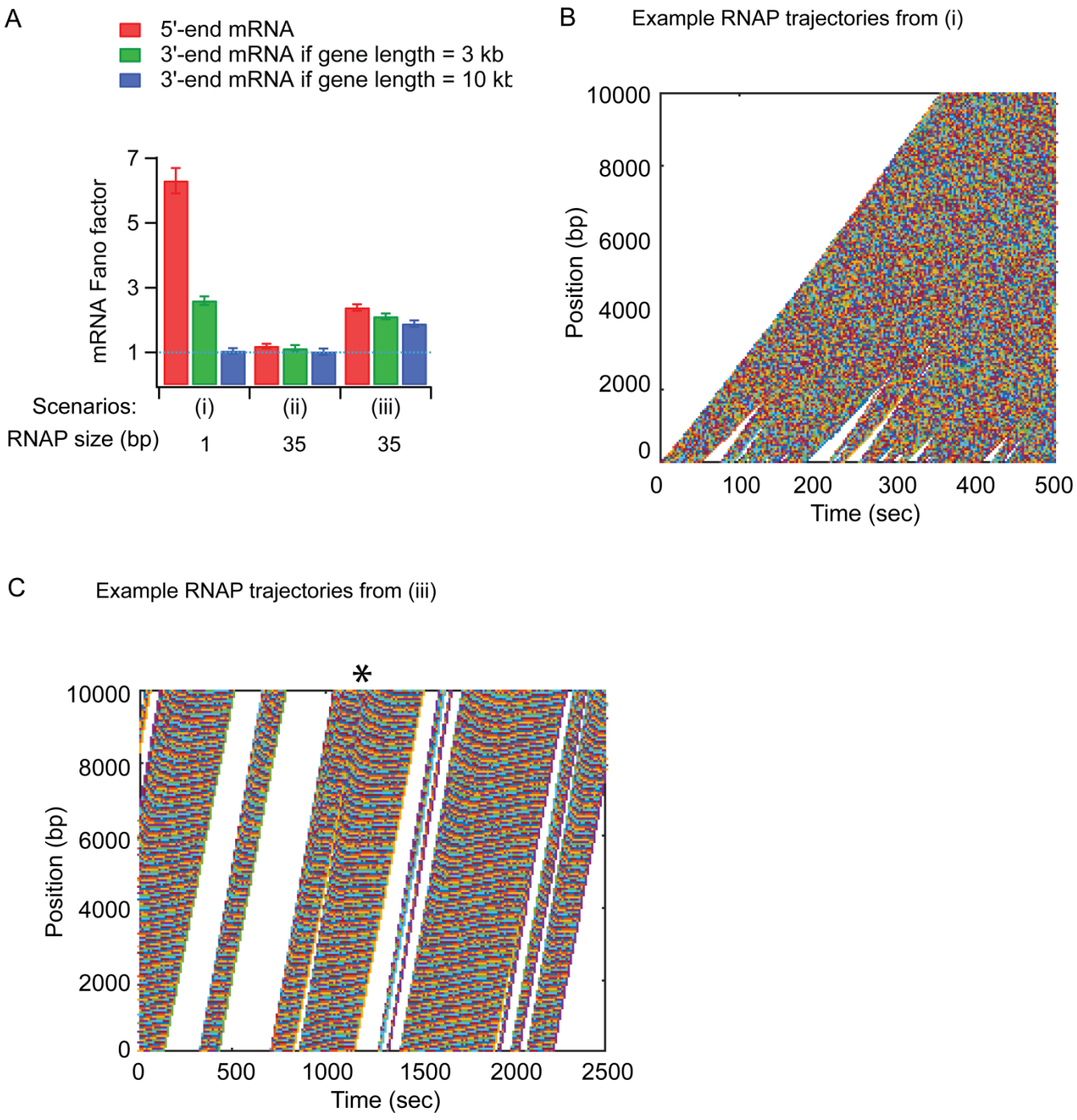

FIGURE S3. RNAP bursts from a bursty promoter can disappear during transcription elongation under special conditions (related to Fig. 2D). As in Dobrzyński and Bruggeman (2009), RNAPs were continually loaded back-to-back during ON periods. This was achieved by setting $k_{\text{loading}} = 10 \times k_{\text{elongation}}$. This condition ensures that successive RNAP loading occurs as soon as the promoter is cleared. (A) mRNA Fano factors at 5'

and 3' ends under various scenarios. The blue dotted line denotes an mRNA Fano factor of 1. (i) Case of RNAP size of 1 bp, as in Dobrzyński and Bruggeman (2009). τ_{ON} = 5 sec, f_{ON} = 0.5, $k_{elongation}$ = 30 nt/sec. Under this condition, we expect ~150 RNAP loading events per ON period. mRNA Fano factors calculated from the 5'-end and the 3'-end indicate that the burstiness from the promoter is completely lost at the end of 10-kb transcription (red vs. blue bar). (ii) Case of an RNAP size of 35 bp. τ_{ON} = 5 sec, f_{ON} = 0.5, $k_{elongation}$ = 30 nt/sec (same as (i) except RNAP size). Under this condition, we expect ~4 RNAP loading events per ON period. Under this condition, the 5'-end mRNA Fano factor is close to 1, indicating that the promoter is not bursty. There were not enough loading events during well-separated ON periods. (iii) Case of an RNAP size of 35 bp. τ_{ON} = 150 sec, f_{ON} = 0.5, $k_{elongation}$ = 30 nt/sec. τ_{ON} was increased to match the number of RNAPs loaded during ON periods as in (i). Under this condition, we expected ~130 RNAP loading events per ON period. Although the promoter became slightly burstier than (ii), as shown by the increased 5'-end mRNA Fano factor, the burstiness set by the promoter was not completely lost at the end of a 10-kb gene (i.e., the 3'-end mRNA Fano factor is still greater than 1). (B) RNAP trajectories from an example simulation under the scenario (i). (C) RNAP trajectories from an example simulation under the scenario (iii). While we saw the expansion of RNAP bursts set by the promoter and the disappearance of certain OFF periods (indicated with *), the burstiness from the promoter is mostly conserved during transcription elongation, indicated by the presence of OFF periods at both the beginning and end of transcription.

The results shown in (A-C) indicate that a 1-bp, and not 35-bp, RNAP size results in a complete loss of RNAP bursts from the promoter. This is because the small 1-bp RNAP footprint allows many RNAPs to load as a convoy for a given τ_{ON} (i.e., RNAPs rapidly clear the RNAP binding site for the next initiation event), resulting in a large separation between subsequent RNAPs during transcription elongation. We tried testing similar numbers of RNAPs loaded back-to-back during ON periods by increasing τ_{ON} for a larger RNAP size (scenario (i) vs (iii)). However, longer τ_{ON} had to be accompanied by longer τ_{OFF} to keep the promoter relatively bursty, and longer τ_{OFF} prevented a loss of RNAP bursts by the end of transcription elongation through RNAP-RNAP separation. Overall, we were unable to find conditions for which we observed a complete loss of RNAP bursts from a bursty promoter when the RNAP footprint was 35 bp.

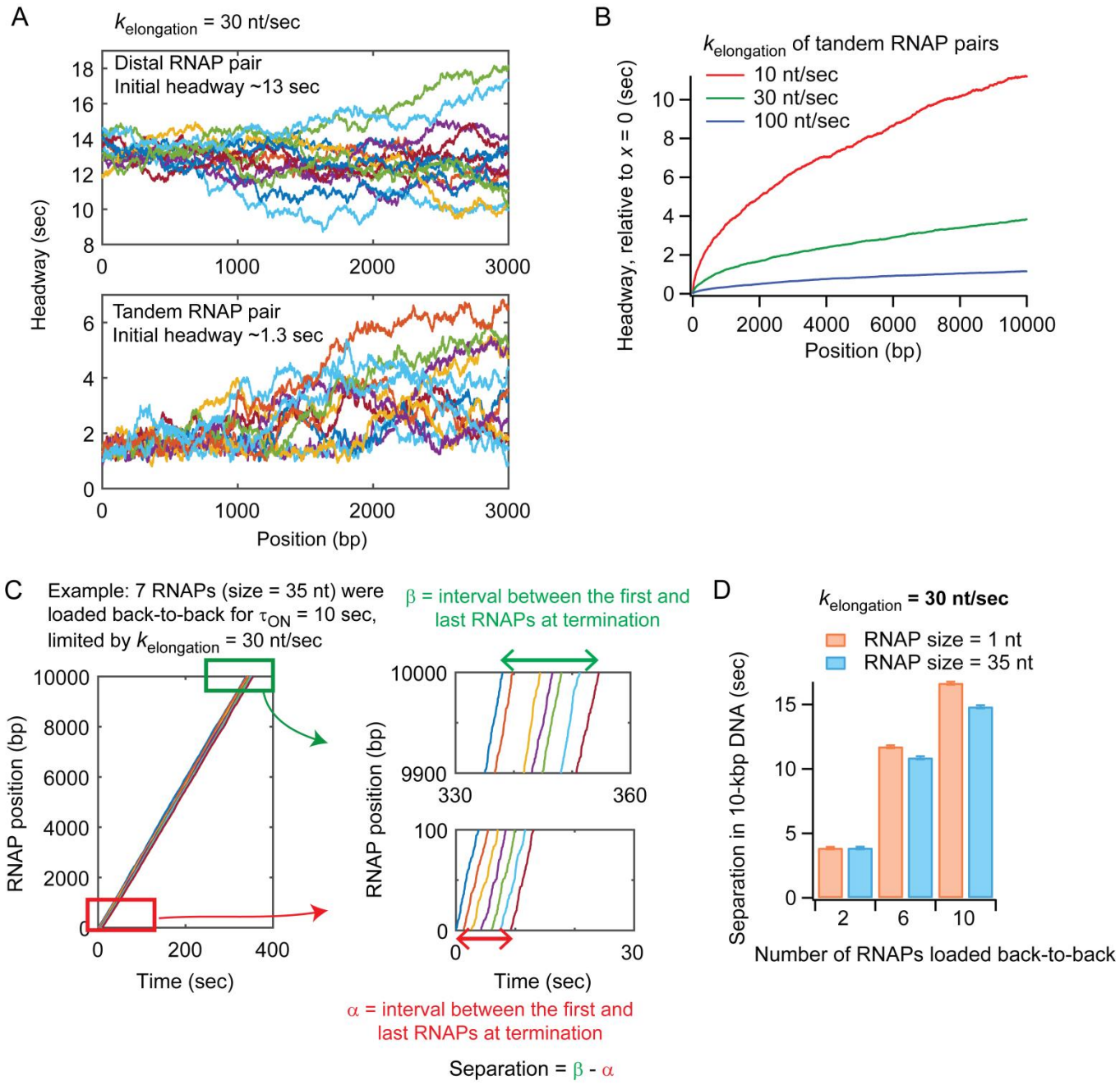

FIGURE S4. Increase in the headway between adjacent RNAPs upon back-to-back RNAP loading. (A) Example headway trajectories showing how the headway between two adjacent RNAPs changed during transcription elongation as a function of the trailing RNAP position. Average RNAP speed was 30 nt/sec. (Top) In the case of distal RNAP pairs (two RNAPs loaded ~13 sec apart), the headway could either increase or decrease during transcription elongation. (Bottom) In the case of tandem RNAP pairs (two RNAPs loaded back-to-back at an

interval of ~1.3 sec), the headway between them could only increase during transcription elongation due to the steric hindrance between RNAPs. (B) Effect of RNAP speed on the average change in headway between two tandem RNAPs as a function of the trailing RNAP position. A larger headway change was observed under slower RNAP speed. In addition, the headway keeps increasing during transcription elongation, suggesting a larger headway change for longer genes. (C) Example RNAP trajectories showing RNAPs separating from each other during elongation. We considered 7 RNAPs loaded back-to-back and travelling at an average speed of 30 nt/sec along a 10-kbp DNA template. The time intervals between the first and last RNAPs in a convoy at transcription initiation (α, red arrow) and termination (β, green) are illustrated in the zoomed trajectories. We defined "separation" as the difference between the red and green arrows. (D) Separation in an RNAP convoy as a function of the number of RNAPs in the convoy (convoy size). Average RNAP speed and gene length were 30 nt/sec and 10 kbp, respectively. Larger separations were observed with greater numbers of RNAPs per convoy. Changing the RNAP size (1 nt vs. 35 nt) had little effect.

The results in (A-D) indicate that the time separation between adjacent RNAPs can be maximized with (i) back-to-back RNAP loading, (ii) slower elongation speed, (iii) longer gene, and (iv) a large number of RNAPs loaded as a convoy.

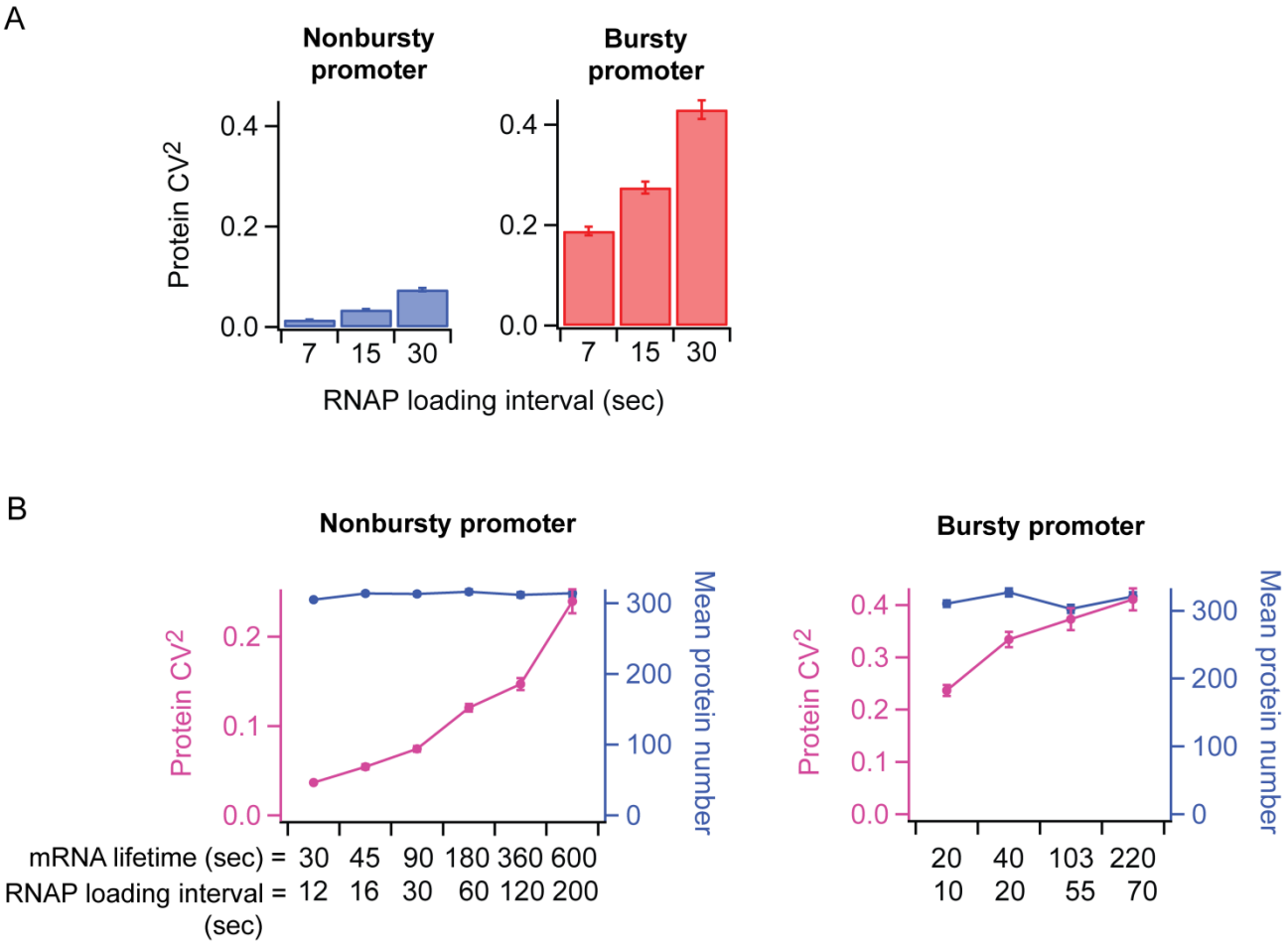

FIGURE S5. Effect of the RNAP loading interval and the mRNA lifetime on protein expression noise. (A) Effect of the RNAP loading interval on protein expression noise (CV^2) when the mRNA lifetime was 90 sec. (B) Effect of simultaneous change in mRNA lifetime and RNAP loading interval on protein expression noise. We tested various conditions of mRNA lifetime and RNAP loading interval that maintain the same protein production (i.e., an increase in mRNA lifetime was compensated by an increase in RNAP loading interval to keep the same mean in protein levels). For the nonbursty-promoter case, we tested input RNAP loading intervals of 10, 15, 30, 60, 120, and 200 sec. For the bursty-promoter case, we tested input f_{ON} of 0.05, 0.1, 0.25, and 0.45. The RNAP loading intervals indicated in the plots was calculated from the output of the simulations.

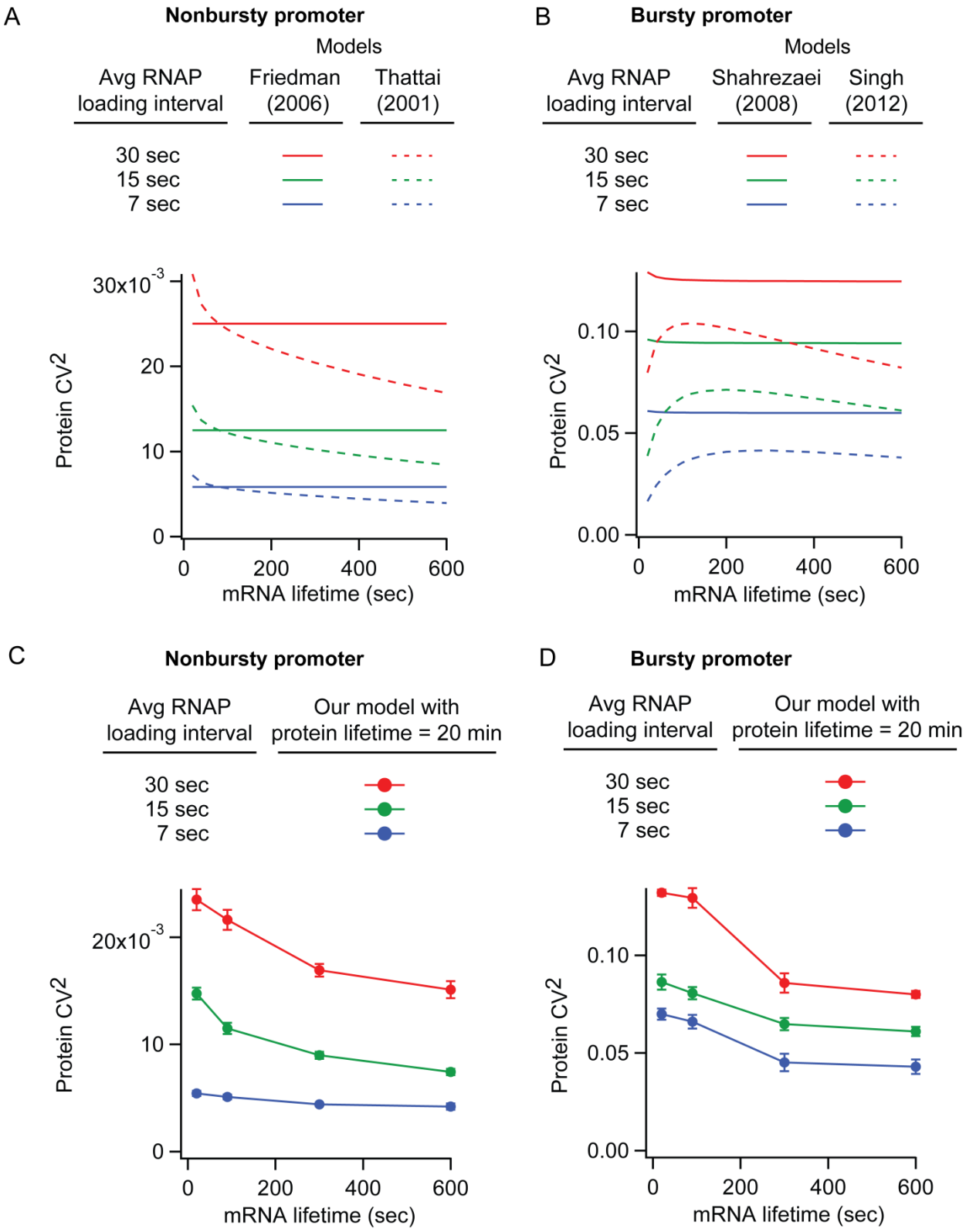

FIGURE S6. Effect of mRNA lifetime on the noise in protein levels. (A) Calculations of protein CV^2 as a function of mRNA lifetime using analytical models of nonbursty

transcription initiations by Friedman et al (2006) and Thattai et al (2001), using different average RNAP loading intervals ($1/k_{loading}$) as input parameters. The model by Friedman et al (2006) assumes that proteins are produced as a random uncorrelated event from each transcript and expects that protein CV^2 does not depend on the mRNA lifetime. In contrast, the model by Thattai et al (2001) predicts a negative effect of mRNA lifetime on CV^2. (B) Calculations of protein CV^2 as a function of mRNA lifetime using analytical models of bursty transcription initiations by Shahrezaei and Swain (2008) and Singh, et al (2012), using different average RNAP loading intervals as input parameters. Average RNAP loading interval, $1/(k_{loading} \cdot f_{ON})$, was varied by changing f_{ON} while keeping $k_{loading}$ and k_{ON} the same as in the main text. The model by Shahrezaei and Swain (2008) expects that protein CV^2 is independent of mRNA lifetime under the parameters we used. In contrast, the model by Singh et al (2012) predicts a negative effect of mRNA lifetime on CV^2. (C) Results from our model of nonbursty transcription initiation for different average RNAP loading intervals. We used $\tau_{loading}$ = 30, 15, 6 sec as input values to achieve average loading intervals of 30 (red), 15 (green), and 7 (blue) sec, respectively. We calculated protein CV^2 from the protein number distribution at steady state (simulations ran for 60-200 min, instead of 40 min to achieve steady state in protein levels). (D) The same as (C) except for bursty transcription initiations. To obtain the indicated average RNAP loading intervals, we used f_{ON} = 0.1, 0.25, 0.5. In all cases (A-D), protein degradation was considered with an arbitrary protein lifetime of 20 min.

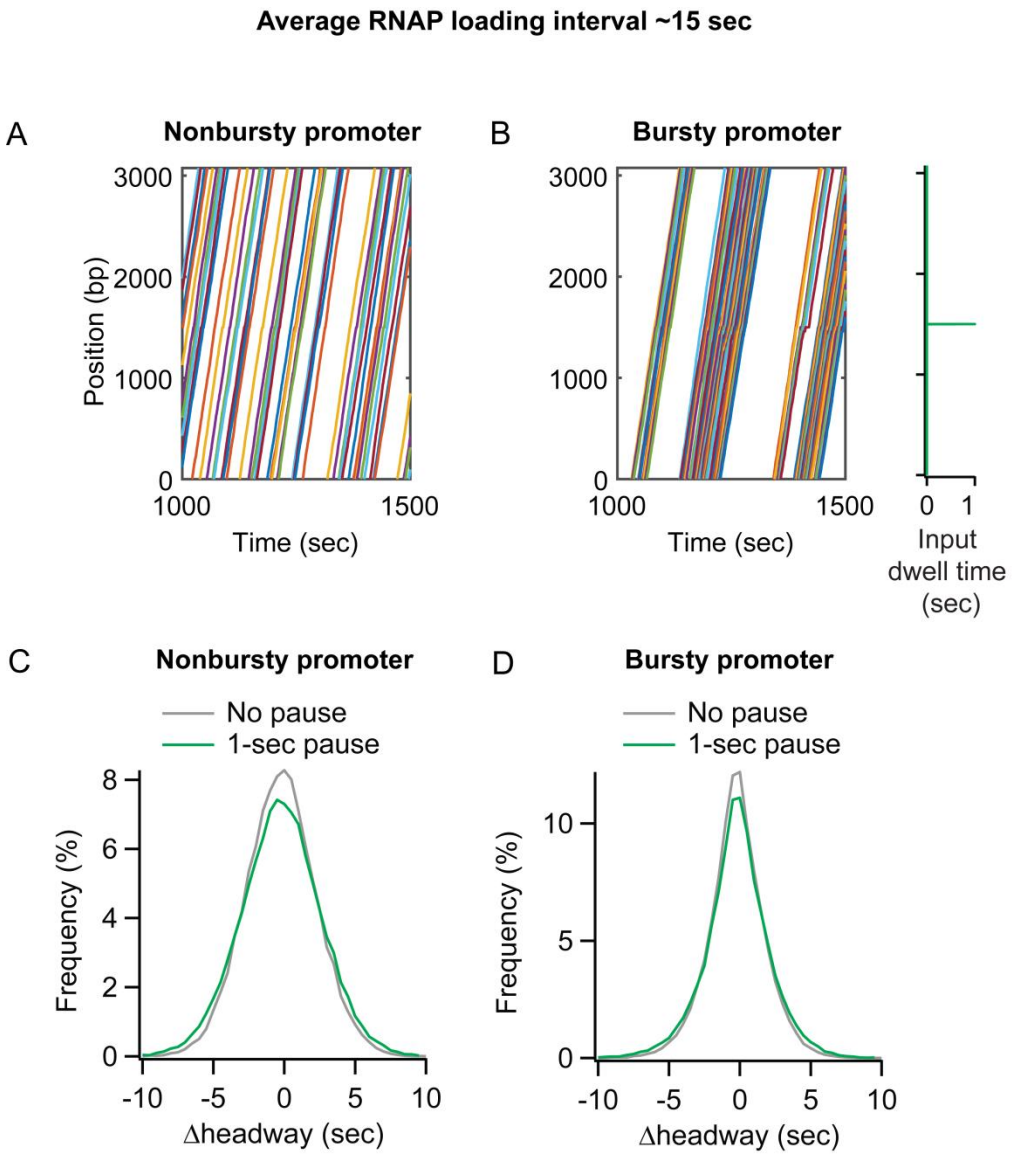

FIGURE S7. RNAP pauses much shorter than the RNAP loading interval have virtually no effect on RNAP traffic. (A and B) Trajectories of RNAPs from a representative simulation with a 1-sec pause at x_{pause} = 1500 nt (with p_{pause} = 100%). We tested both nonbursty (A) and bursty (B) initiations with the same average RNAP loading interval of ~15 sec measured. Individual RNAP trajectories were parallel to each other and burstiness (or lack thereof) from the promoter was preserved until the end of transcription elongation. This indicates that 1-sec pause did not change RNAP traffic. (C and D) Distributions of headway changes between transcription initiation and termination.

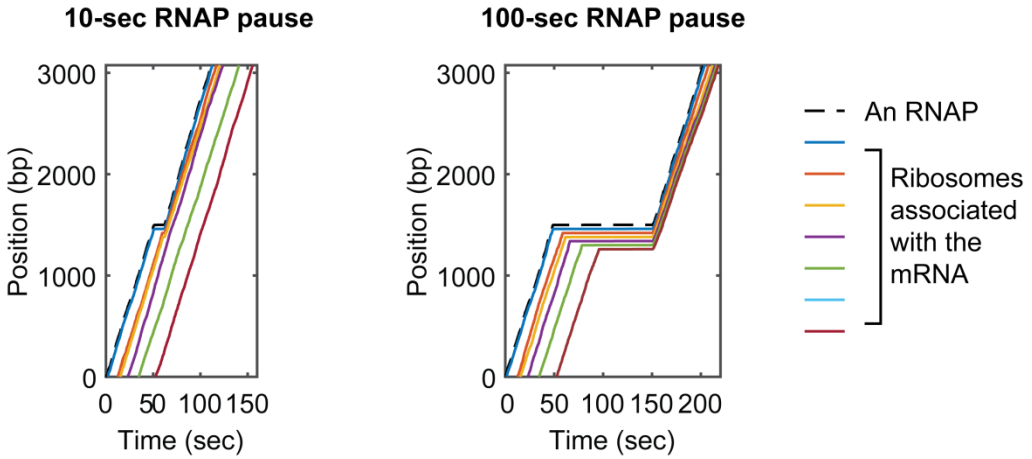

FIGURE S8. Effect of a very long-lived RNAP pause on ribosome traffic. The location of ribosomes traveling on a given nascent transcript was plotted over time, together with the location of the RNAP on the DNA, to showcase how an RNAP pause of 10 or 100 sec affects the ribosome traffic. When the RNAP pause duration was ~10 sec (left), only a few leading ribosomes piled behind the paused RNAP (dotted line), resulting in a minimal effect on ribosome traffic. If the pause lasted much longer, such as ~100 sec (right), most of the ribosomes on the transcript piled behind the paused RNAP and then travelled together as a convoy when the RNAP was released from the pause site, resulting in a protein burst. For illustration purposes, we used the same set of ribosome loading times for both pause scenarios. Five ribosomes loaded on a transcript during its lifetime (mean mRNA lifetime = 90 sec).

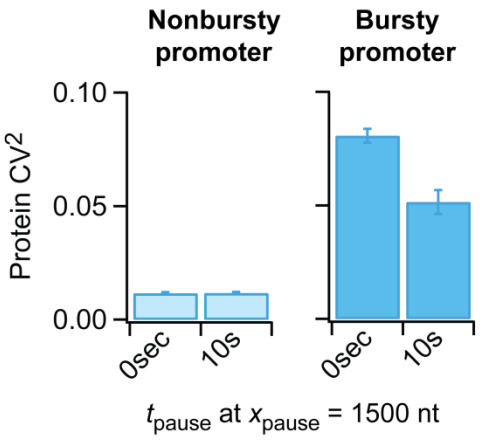

FIGURE S9. Effect of a high-probability RNAP pause on steady-state protein expression noise. A high-probability (p_{pause} = 80%) pause site attenuates steady-state protein noise (CV^2) dictated by a bursty transcription initiation when protein degradation (or dilution occurring upon cell division) was included in the model. We chose an arbitrary protein lifetime of 20 min. We used the same input conditions as shown in Fig. 5D and H: an average RNAP loading interval of ~15 sec for both nonbursty and bursty promoters.

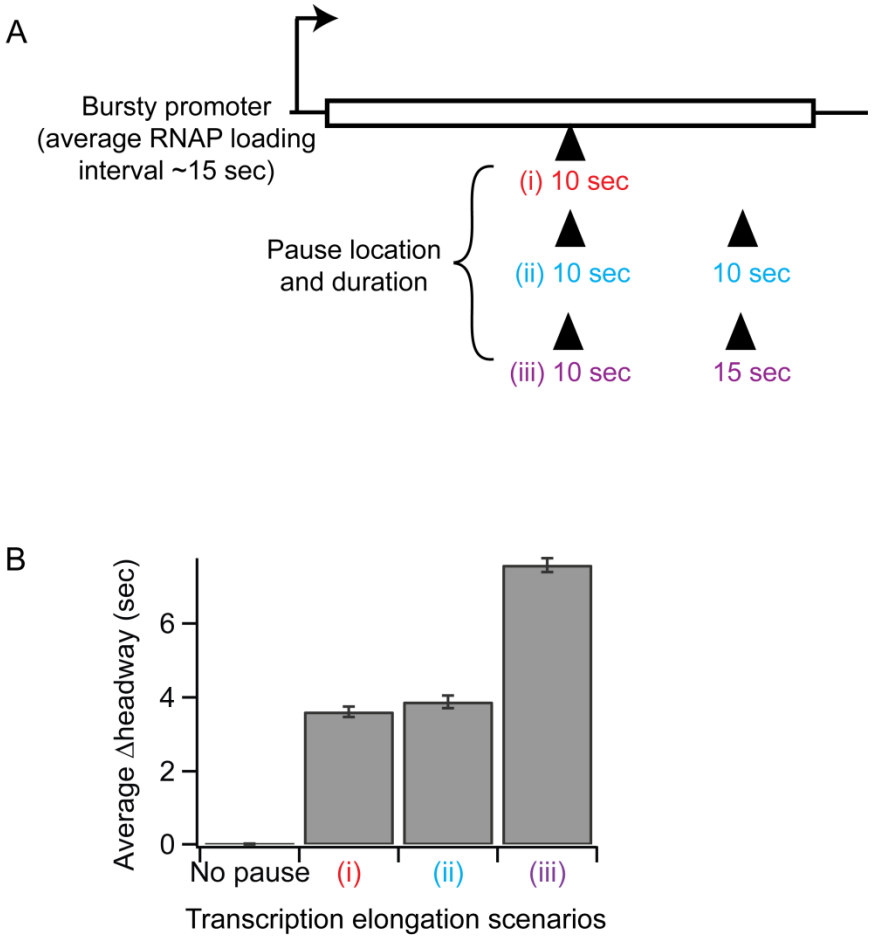

FIGURE S10. Two pause sites can have additive effects on RNAP traffic depending on their durations. (A) Relevant simulation conditions. We tested the effects of different pause scenarios on RNAP traffic from a bursty promoter with an average RNAP loading interval of ~15 sec. The pause scenarios tested were (i) a single 10-sec pause at x_{pause} = 1500 nt, (ii) a 10-sec pause at x_{pause} = 1500 nt followed by another 10-sec pause at x_{pause} = 2500 nt, (iii) a 10-sec pause at x_{pause} = 1500 nt followed by a 15-sec pause at x_{pause} = 2500 nt (all pause sites with p_{pause} = 100%). (B) Effect of these pause scenarios on RNAP traffic was measured by the average headway change between initiation and termination of transcription.

A

Scenarios	(i)	(ii)	(iii)
Transcription initiation rate	same	same	same
Translation initiation rate	same	same	same
RNAP pause?	no	σ^{70}	σ^{70}
RNAP-ribosome coupling	yes	yes	no

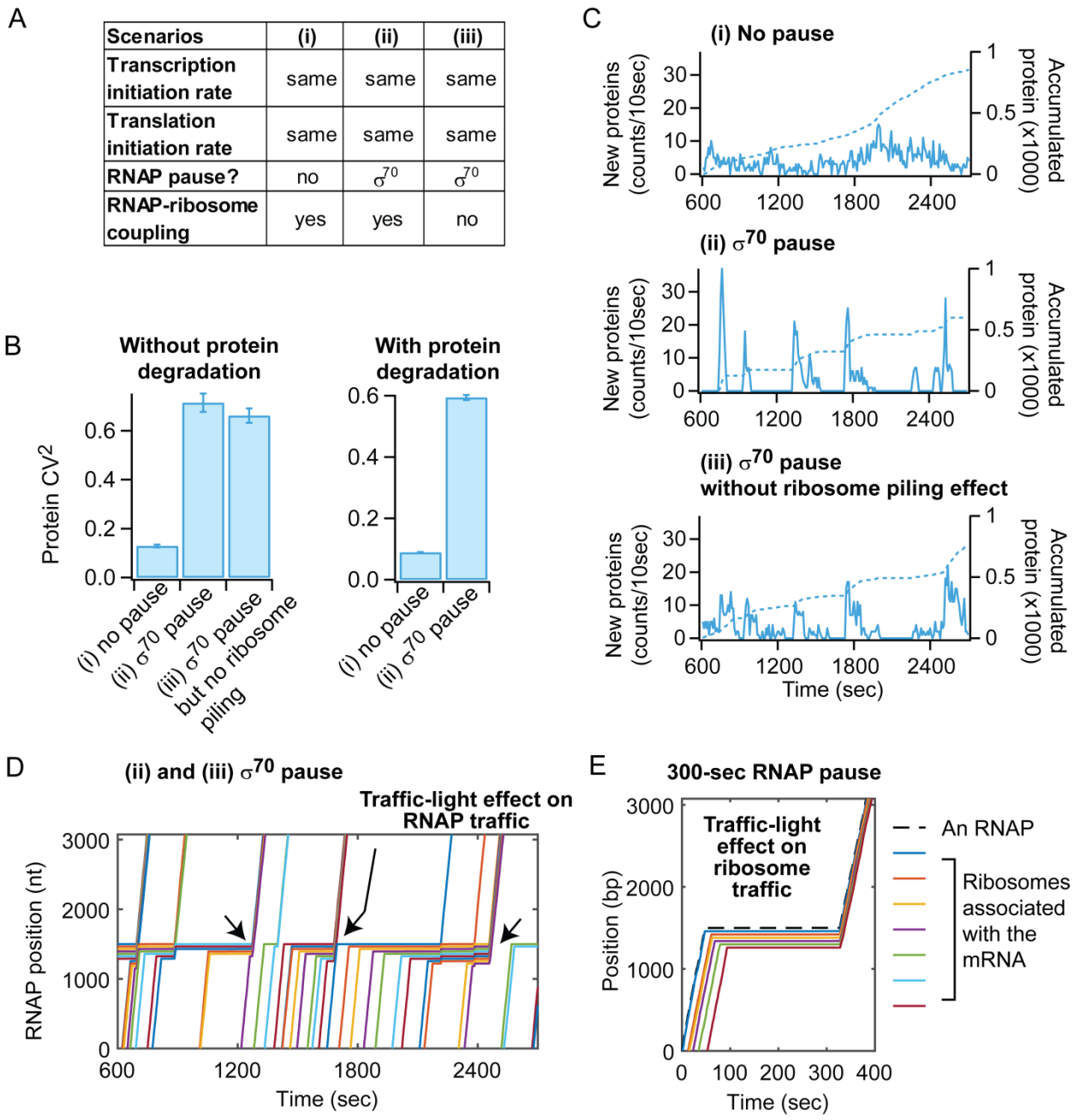

FIGURE S11. RNAP pause due to σ^{70} retention can increase protein expression noise by creating RNAP convoys. (A) Three scenarios were considered to examine the potential effect of σ^{70}-RNAP pausing: (i) No pause. (ii) σ^{70}-RNAP pausing: we simulated a pause site for σ^{70}-associated RNAP (and not for RNAPs free of σ^{70}) at x = 1,500 bp. To match *in vitro* experimental observations (12), we chose t_{pause} = 300 sec and p_{pause} = 30% (the latter because only 30% of the RNAP population retained σ^{70}). (iii) Same as (ii) except that the ribosome piling effect was ignored. This was done by ignoring ribosome elongation and modeling translation

after each mRNA is made (i.e., at the end of transcription elongation). This method uncouples transcription and translation and avoids ribosome piling behind paused RNAP. In all three scenarios, we considered a nonbursty transcription initiation and an mRNA lifetime of 90 sec. The input values of transcription and translation initiations were tuned to obtain an average RNAP loading interval of ~ 50 sec and an average ribosome loading rate of ~0.15 sec^{-1} as output in all scenarios. (B) Effect of σ^{70}-induced RNAP pausing on protein noise (CV^2) for all three scenarios. Comparison between (i) and (iii) shows the contribution of RNAP convoy formation on the CV^2 increase, whereas the comparison between (ii) and (iii) shows the contribution from ribosome convoy formation (ribosome piling behind the paused RNAP). We also found that the σ^{70}-induced RNAP pausing increases protein CV^2 from nonbursty transcription initiation ((i) vs. (ii)) when protein degradation was considered, using an arbitrary protein lifetime of 20 min. (C) Temporal fluctuations in protein production from an example DNA template for each scenario shown in (A). The dotted lines show protein accumulation. Scenario (ii) creates bursts of protein production. To illustrate the effect of ribosome piling on the protein bursts (ii vs. iii), the fluctuations in protein production in (ii) and (iii) were generated using the RNAP trajectories shown in (D). (D) Example RNAP trajectories from scenarios (ii) and (iii) showing a "traffic-light" effect in which back-to-back RNAPs accumulated behind the pause site travel together as a convoy once the pausing ends. (E) Representative ribosome trajectories from scenario (ii). The location of ribosomes for a given nascent transcript was plotted over time to showcase how the long pause by a σ^{70}-associated RNAP causes a "traffic-light" effect on ribosome traffic, in which back-to-back ribosomes accumulated behind the paused RNAP travel together after the pause.

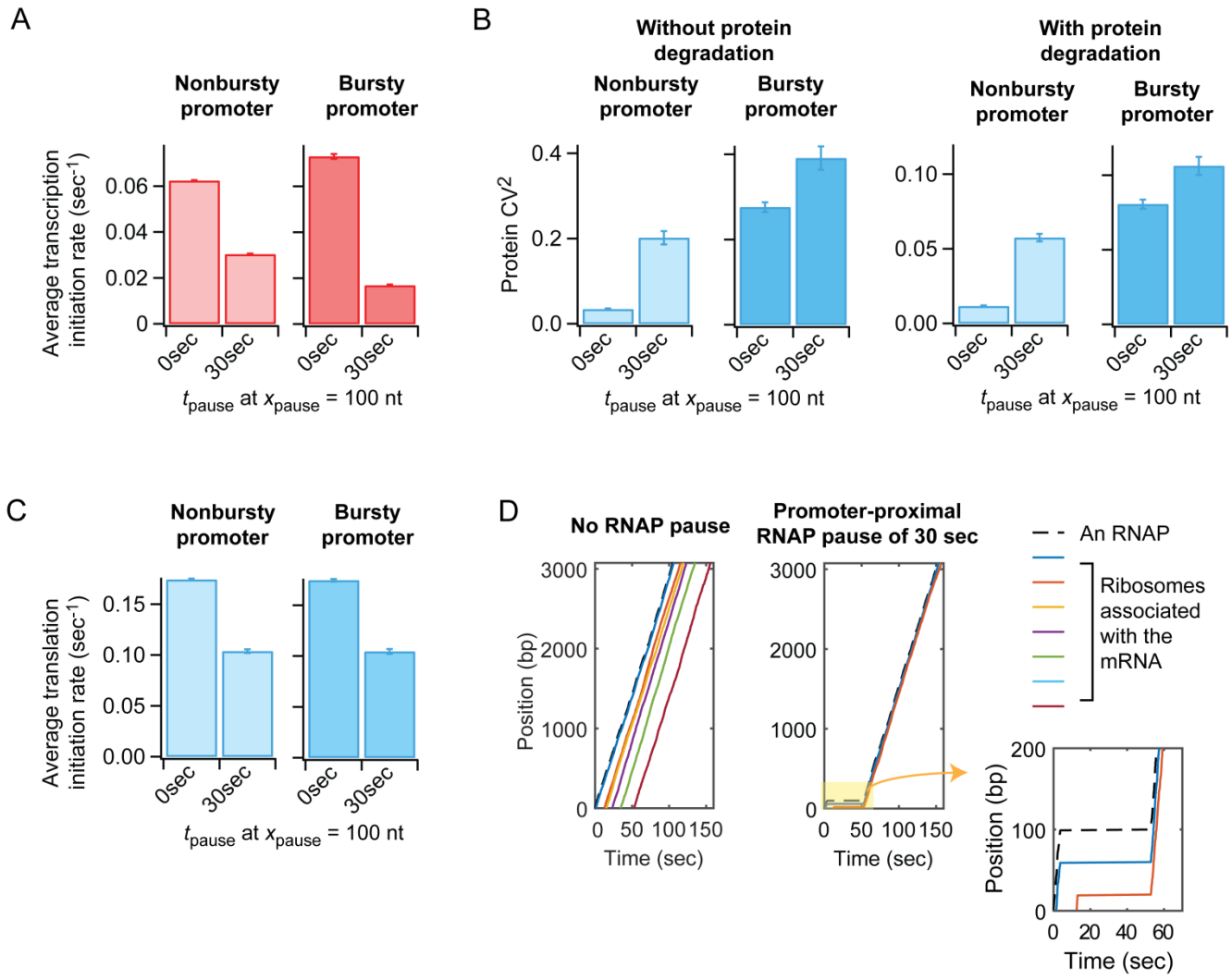

FIGURE S12. A promoter-proximal RNAP pause can reduce both transcription and translation rates. (A) A promoter-proximal pause of t_{pause} = 30 sec (x_{pause} = 100 nt and p_{pause} = 100%) lowers the effective transcription initiation rate from both nonbursty and bursty promoters with an input average transcription initiation of ~1/15 sec^{-1}. The effect is greater for the bursty promoter, because of the frequent RNAP loading attempts during ON periods. (B) Effect of a promoter-proximal pause of t_{pause} = 30 sec (x_{pause} = 100 nt and p_{pause} = 100%) on protein expression noise from both nonbursty and bursty promoters shown in (A). As a result of lower transcription initiation rate shown in (A), the noise in protein levels (CV^2) increases when a promoter-proximal pause site exists. Similar results were obtained when we considered protein degradation (arbitrary protein lifetime of 20 min). (C) Promoter-proximal pausing also lowers the translation initiation rate (promoter and pause properties are the same as noted in (A)). (D)

Promoter-proximal pausing of an RNAP results in piling of ribosomes that can easily reach the RBS and prevent new ribosome loading. (Left) No RNAP pause case: Five ribosomes were loaded on to the transcript and traveled without pausing. (Right) If an RNAP pauses at x_{pause} = 100 nt for ~30 sec, only 2 out of 5 ribosomes were able to load during the mRNA lifetime (see the zoomed trajectories on the right). Once the RNAP pause ends, these two ribosomes form a convoy, as also shown in Fig. S8.

Supporting References

1. 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.
2. Shahrezaei, V., and P. S. Swain. 2008. Analytical distributions for stochastic gene expression. Proc. Natl. Acad. Sci. USA. 105:17256-17261.
3. 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.
4. Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. USA. 98:8614-8619.
5. 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.
6. Joo, C., and T. Ha. 2012. Labeling DNA (or RNA) for single-molecule FRET. Cold Spring Harbor Protocols 2012:1005-1008.
7. Jones, D. L., R. C. Brewster, and R. Phillips. 2014. Promoter architecture dictates cell-to-cell variability in gene expression. Science 346:1533-1536.
8. 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.
9. 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.
10. 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.
11. 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.
12. Harden, T. T., C. D. Wells, L. J. Friedman, R. Landick, A. Hochschild, J. Kondev, and J. Gelles. 2016. Bacterial RNA polymerase can retain σ70 throughout transcription. Proc. Natl. Acad. Sci. USA. 113:602-607.