Biophysical Journal
Article

Biophysical Society

Stochastic dynamics and ribosome-RNAP interactions in transcription-translation coupling

Xiangting Li[1] and Tom Chou[1,2,*]
[1]Department of Computational Medicine, University of California, Los Angeles, Los Angeles, California and [2]Department of Mathematics, University of California, Los Angeles, Los Angeles, California

ABSTRACT Under certain cellular conditions, transcription and mRNA translation in prokaryotes appear to be "coupled," in which the formation of mRNA transcript and production of its associated protein are temporally correlated. Such transcription-translation coupling (TTC) has been evoked as a mechanism that speeds up the overall process, provides protection against premature termination, and/or regulates the timing of transcript and protein formation. What molecular mechanisms underlie ribosome-RNAP coupling and how they can perform these functions have not been explicitly modeled. We develop and analyze a continuous-time stochastic model that incorporates ribosome and RNAP elongation rates, initiation and termination rates, RNAP pausing, and direct ribosome and RNAP interactions (exclusion and binding). Our model predicts how *distributions* of delay times depend on these molecular features of transcription and translation. We also propose additional measures for TTC: a direct ribosome-RNAP binding probability and the fraction of time the translation-transcription process is "protected" from attack by transcription-terminating proteins. These metrics quantify different aspects of TTC and differentially depend on parameters of known molecular processes. We use our metrics to reveal how and when our model can exhibit either acceleration or deceleration of transcription, as well as protection from termination. Our detailed mechanistic model provides a basis for designing new experimental assays that can better elucidate the mechanisms of TTC.

SIGNIFICANCE Transcription-translation coupling (TTC) in prokaryotes is thought to control the timing of protein production relative to transcript formation. The marker for such coupling has typically been the measured time delay between the first completion of transcript and protein. We formulate a stochastic model for ribosome and RNAP elongation that also includes RNAP pausing and ribosome-RNAP binding. The model is able to predict how these processes control the distribution of delay times and the level of protection against premature termination. We find relative speed conditions under which ribosome-RNAP interactions can accelerate or decelerate transcription. Our analysis provides insight on the viability of potential TTC mechanisms under different conditions and suggests measurements that may be potentially informative.

INTRODUCTION

In prokaryotic cells, transcription and translation of the same genes are sometimes "coupled" in that the first mRNA transcript is detected coincidentally with the first protein associated with that transcript. This observation suggests the proximity of the ribosome to the RNA polymerase (RNAP), as was directly observed by Miller et al. (1). These findings contributed to the idea that the ribosome and RNAP may interact directly or indirectly. Ribosome-RNAP interactions in prokaryotes are thought to maintain the processivity of RNAP by physically pushing it out of the paused, backtracking state (2). Paused RNAP has also been associated with transcriptional error correction and premature mRNA cleavage (3–6); thus, ribosome-RNAP interactions that catalyze RNAP unstalling and speeding up the overall process might do so at the expense of more errors in the transcript. Transcription-translation coupling (TTC) may also play an important role in protecting mRNA from premature transcription termination (7–9). This protection might arise from steric shielding of the elongation complex by the leading ribosome, preventing attack by Rho (9,10).

Evidence for TTC has come from two types of experiments. The first is "time-of-flight" experiments that quantify the time delay between first detection of a complete transcript and a complete protein. For example,

Submitted June 23, 2022, and accepted for publication September 27, 2022.
*Correspondence: tomchou@ucla.edu
Editor: Jason Kahn.
https://doi.org/10.1016/j.bpj.2022.09.041

IPTG-induced LacZ completion experiments (LacZ assays) measure the mean time of mRNA completion $\mathbb{E}[T_{\mathrm{RNAP}}]$ and the mean time of protein completion by the leading ribosome $\mathbb{E}[T_{\mathrm{rib}}]$, with the latter measured from the time of first RNAP engagement (11–13). Since the transcript length L is known, the effective velocities of the RNAP and ribosome over the entire transcript can be estimated by

$$\overline{V}_{\mathrm{RNAP}} = \frac{L}{\mathbb{E}[T_{\mathrm{RNAP}}]}, \; \overline{V}_{\mathrm{rib}} = \frac{L}{\mathbb{E}[T_{\mathrm{rib}}]}. \tag{1}$$

These measurements are performed at the population level, averaging the time-dependent signal from many newly formed transcripts and corresponding proteins. Thus, the individual molecular coupling mechanisms between RNAP and ribosomes cannot be resolved by the time delay unless single-molecule time-of-flight experiments can be designed.

Another class of experiments uses a variety of in vitro and in vivo assays to probe direct and indirect molecular interactions between RNAPs and ribosomes (9,14–16). A recent structural study identified active transcription-translation complexes and confirmed an indirect molecular interaction in situ (17).

Two modes of coupling between the leading ribosome and the RNAP have been proposed. One mode of coupling occurs through a "collided expressome" in which the ribosome and RNAP are held in close proximity (9,14) by direct association, which may be mediated by entropic effects of the intervening mRNA (18). The second coupling mode occurs through a larger complex in which ribosome-RNAP interactions are mediated by the proteins NusG and NusA (15–18). There have been no reports that this mode alters elongation speeds or RNAP processivity, although both are possible. It has been shown that the NusG-coupled expressome can inhibit Rho-induced premature transcription termination (19).

Both potential coupling mechanisms require at least some moments of close proximity between the RNAP and the leading ribosome during the simultaneous transcription-translation process (see Fig. 1), followed by recruitment of NusG for the NusG-coupled expressome mechanism. The ribosome-RNAP proximity requirement can be met if the ribosome elongation speed is, on average, faster than that of the RNAP. Even if the ribosome is fast, proximity also depends on initial condition (ribosome initiation delay after RNAP initiation) and the length of the transcript L. Moreover, both RNAPs and ribosomes are known to experience, respectively, pausing through backtracking (6) and through "slow codons" for which the associated tRNA is scarce (20).

A number of open questions remain. In the "strong coupling" picture, the ribosome and RNAP are nearly always in contact and the ribosome is thought to maintain RNAP processivity during mRNA transcription. Without the ribosome, mRNA transcription can be particularly slow. Therefore, the ribosome elongation rate dictates the RNAP elongation rate. Administration of antibiotics to slow down translation also slowed down transcription (11).

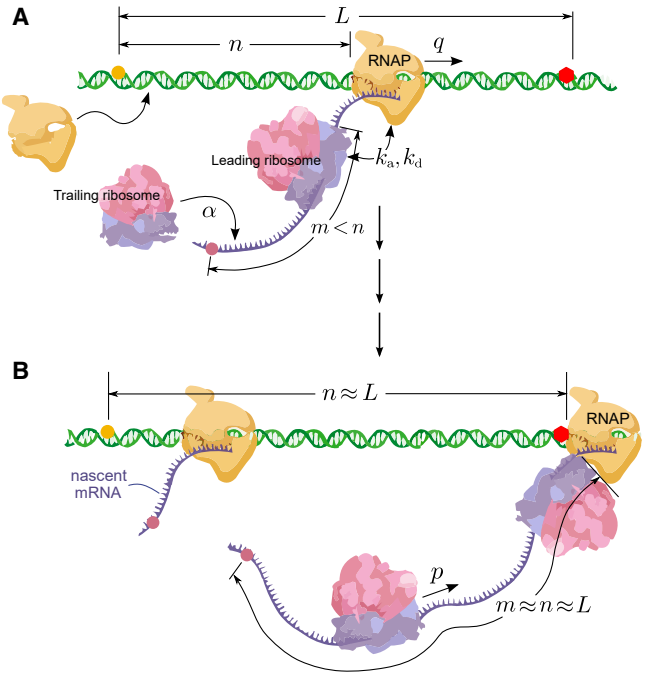

FIGURE 1 Schematic of translation of a nascent mRNA transcript (polypeptide not shown). Ribosomes translocate at rate p while RNAPs elongate at rate q. (*A*) The transcript associated with the RNAP at position n along the gene is shown with a leading ribosome at position m along the mRNA. Ribosomes attach to open initiation sites at rate α. (*B*) A nearly complete transcript is shown. If the leading ribosome has caught up to the RNAP and m is close to n, the two may bind with rate k_a to form a "coupled expressome." Ribosome-RNAP complexes can spontaneously dissociate with rate k_d. We assume that the leading ribosome "terminates" upon reaching the stop codon (not shown). Protein-mediated expressomes (not shown) form larger complexes that can accommodate longer mRNA segments within it.

However, other experiments have shown that the distance between the ribosome and RNAP can be large most of the time, leading to a "weak coupling" picture (21,22). The biological role of weak coupling is unclear since any shielding provided by the ribosome would be limited and ribosome and RNAP speeds could be independently modulated. Even though direct ribosome-RNAP interactions may still arise after an RNAP has stalled for a sufficiently long time, any apparent ribosome-RNAP coordination would be largely coincidental.

Besides the strong and weak coupling dichotomy, another unresolved question is whether or not TTC can slow down transcription. Traditionally, TTC has been invoked as a mechanism for maintaining RNAP processivity by rescuing RNAP from paused states. However, in vivo experiments by Kohler et al. (9) reported that, when translation is inhibited, the ΔαCTD mutant in which RNAP *does not* associate with

ribosome exhibited faster proliferation than that of wild-type RNAP that can associate with ribosomes. Coupled transcription through ribosome-RNAP association may give rise to *slower* transcription. Although in situ cryo-EM has also shown protein-mediated complexes of ribosome and RNAP (23), additional studies should be devised to confirm and characterize ribosome-RNAP interactions in vivo. Nonetheless, our mathematical model can provide possible mechanistic insights into TTC.

To help resolve the puzzles discussed above, provide a quantitative way to explore different molecular mechanisms that may contribute to TTC, and generate predictions that can be compared with experimental observations, we formulate a stochastic model that combines a number of known molecular mechanisms from transcription, keeping track of ribosome and RNAP states and positions along the gene. While an earlier model combined transcription and translation in prokaryotes (24), it did not explicitly incorporate mechanisms of direct transcription and translation coupling and only assumed simple volume exclusion between the RNAP and the leading ribosome.

Here, we explicitly allow for RNAP pausing and formal association and dissociation of the ribosome-RNAP complex. Here, "association" will be generally used to denote specific or nonspecific, direct or indirect molecular interactions. The typical assay used to probe TTC involves measuring the time delay $\Delta T = T_{\mathrm{rib}} - T_{\mathrm{RNAP}}$ between the completion of mRNA and its associated protein. Although time delays can be used as a metric for defining TTC, absence of delay is a necessary but not sufficient condition for direct ribosome-RNAP coupling. Even a direct observation of an interacting expressome does not imply functional consequences of TTC. After formulating our model, we construct additional metrics that better define and reflect possible attributes of TTC. However, since the time delay is the most experimentally measurable quantity, we will still derive and compute the full probability density of delay times $\rho(\Delta T)$.

MODEL AND METHODS

Based on existing structural and interaction information, we formulate a continuous-time Markov chain to model ribosome and RNAP kinetics. As shown in Fig. 1, we describe the position of the head of the leading ribosome along the nascent transcript by $m = 0, 1, \ldots, L$, where 0 denotes a ribosome-free transcript. We also track the length of the nascent mRNA transcript that has cleared the exit channel of the RNAP through the discrete variable $n = 1, 2, \ldots, L$. The positions are described in terms of triplets of nucleotides corresponding to codons, the fundamental step size during ribosome elongation.

Here, L is the length of the coding region of the transcript, typically about $L \sim 300$ codons. Since we are interested in the elongation phase of coupled transcription and translation, the length L is measured from the ribosome binding site immediately upstream of the start codon to the stop codon. We carefully choose the definition of m and n so that ribosome and RNAP sizes are irrelevant and that the difference $d \equiv n - m$ precisely describes the length of the free intervening mRNA between them. Effects of TTC on transcription termination have been previously discussed in (25,26). Therefore, $0 \leq m \leq L$, $1 \leq n \leq L$, and $n = L$ is interpreted as a completed mRNA while $m = L$ indicates a completed polypeptide. Thus triangular state-space defined by (m, n) often arises in related stochastic models of interacting particles in one dimension (6,27,28). Thus far, we assumed monocistronic mRNAs. Polycistronic mRNA with multiple internal ribosome entry sites can be modeled as successive processes with shared RNAP coordinates, independent ribosomes, and L being the respective cistron length.

Within each positional state (m, n), the leading ribosome and RNAP can exist in different internal configurations describing their molecular states. The RNAP at site n can switch between two states, a processive state and a paused state. In the processive state, the RNAP can move forward by one codon at rate q_n or it can transition to a paused or "backtracking" state with stalling rate k_-. The RNAP elongation rate can also depend on its position n through different abundances of corresponding nucleotides. For simplicity, we assume that RNAPs in the backtracking state are fixed and do not elongate ($q_n = 0$) but may transition back to the processive state with unstalling rate k_+. The waiting time distributions in the processive and paused states are exponential with mean $1/k_-$ and $1/k_+$, respectively. The leading ribosome at site m will be assumed to always be in a processive state with forward hopping rate p_m if and only if the next site $m + 1$ is empty (not occupied by the downstream RNAP). In general, the ribosome translation rate can depend on the position m through the codon usage at that site.

When the distance between the leading ribosome and the RNAP is within an interaction range ℓ, ($d \equiv n - m \leq \ell$), they may associate or "bind" with rate k_{a} to form a coupled expressome and dissociate with rate k_{d} (Eqs. 6 and 7). Note that $d \leq \ell$ simply describes proximity and not necessarily molecular coupling. In cases where ribosome-RNAP complex formation is mediated by other proteins and cofactors, such as NusG, etc., k_{a} would represent an effective complex formation rate that could depend on NusG availability. To enumerate coupled and paused-RNAP internal states, we define $(a, b) \in \{0, 1\}^2$ such that $a = 1$ refers to an associated, or "bound" ribosome-RNAP complex, and $b = 1$ refers to an RNAP in a backtracking, or a "paused" or "stalled" state. When $a = 0$, the ribosome is not bound to the RNAP, and when $b = 0$, the RNAP is in the processive state. The state space of our discrete stochastic model is given by $\{(m, n, a, b) : 1 \leq m \leq n \leq L, a, b \in \{0, 1\}\}$, with $\{(0, n, 0, b) : 1 \leq n \leq L, b \in \{0, 1\}\}$ representing

TABLE 1 Model parameters

Parameters	Description	Typical values[a]	Refs.
α	translation initiation rate	$\sim 0.01 - 10.0\ s^{-1}$	(25,30–32) [b]
L	gene and transcript length	$L \in \mathbb{Z}^+$, $L \sim 300$	(33)
m	ribosome position from mRNA 5′	$m \in \mathbb{Z}_{\geq 0}$, $0 \leq m \leq L$	–
n	RNAP position from mRNA 5′	$n \in \mathbb{Z}_+$, $m \leq n \leq L$	–
p	free ribosome translocation rate	~ 15 codons/s	(11,25,34,35)
q	free processing RNAP transcription rate	~ 30 codons/s	(11–13,36) [c]
k_-	processive RNAP $\rightarrow$ paused RNAP rate	$\sim 0.4\ s^{-1}$	(37) [d]
k_+	paused RNAP $\rightarrow$ processive RNAP rate	$\sim 0.3\ s^{-1}$	(37)
k_+^*	paused RNAP $\rightarrow$ processive RNAP rate (pushed)	$k_+^* = k_+ \exp(E_+), E_+ \geq 0$	estimated
k_a, k_d	ribosome-RNAP association, dissociation rates	$k_d = k_a e^{-E_a}$, $E_a \sim 3 - 7$	(14)
ℓ	maximum mRNA length in bound complex	$\sim 4 - 6$ codons	(29)

[a]The unit of length assumed throughout the paper will be nucleotide triplets (codons).
[b]Translation initiation is highly variable. In accordance with LacZ completion assays (25,30), we initially set $\alpha = 1/s$ in our simulations and later discuss the effects of varying translation initiation rates.
[c]Typical noninteracting RNAP transcription rates are $\overline{q} \equiv q k_+/(k_- + k_+) \sim 15$ codons/s. Since typically $k_+/(k_+ + k_-) \sim 1/2$, we use typical values $q \sim 30$ codons/s for the unimpeded transcription rate of processing RNAP.
[d]The pausing probability along an RNAP trajectory has been measured as ~ 0.87 per 100 nucleotides. By using the estimated mean RNAP velocity of ~ 15 codons/s, we convert this probability to a pausing rate $k_- \approx 0.4$/s.

ribosome-free configurations. By appropriately choosing ℓ, k_a, and k_d, our model can effectively describe both the collided expressome and the NusG-mediated expressome. Experimentally, the collided expressome is typically observed in a stalled configuration (9,17,18,29). This observation is consistent with a small k_a and a large k_d, which suggests that the coupled, processive state a is transient. We will define $\log(k_a/k_d) \equiv E_a$ which is not restricted to positive values.

Other than steric exclusion (which constrains $m \leq n$) and ribosome-RNAP association and dissociation, we incorporate a contact-based RNAP "pushing" mechanism. The processing ribosome can directly push (powerstroke) against a stalled RNAP and/or reduce the entropy of a backtracking RNAP to bias it toward a processive state. A similar mechanism arises in RNAP-RNAP interactions, as discussed in (6). To quantify this pushing mechanism, we simply modify the paused-to-processive RNAP ($b = 1 \rightarrow b = 0$) transition rate from k_+ to $k_+^* \equiv k_+ e^{E_+} > k_+$ whenever the ribosome abuts the RNAP ($d \equiv n - m = 0$). The enhanced rate arises from a reduction E_+ in the total transition free energy barrier provided by the adjacent ribosome. Typical model parameters relevant to prokaryotic transcription and translation are listed in Table 1.

The length ℓ may influence direct molecular coupling and stochastic dynamics of transcription. In vitro studies of ribosome and RNAP structure provide constraints on the configuration space accessible to coupled expressomes. Wang et al. (29) and Webster et al. (18) found that collided expressomes are stable only when the spacer mRNA between the ribosome and the RNAP is $\sim 12 - 24$ nucleotides ($\sim 4 - 8$ codons). Because the intervening mRNA must be at least 12 nucleotides to extend beyond the RNA exit channel of the RNAP, the free intervening RNA within an intact collided expressome can vary between 0 and 12 nucleotides. In contrast, the NusG-mediated expressome can accommodate $\sim 24 - 30$ free mRNA nucleotides. RNA looping might allow for even longer spacer mRNA, but there has so far been no in vivo evidence that collided expressomes exist with mRNA loops.

Since mRNA is flexible, we can also assume that k_d is constant for $d \equiv n - m \leq \ell$. The association rate k_a may be dependent on the distance $d = n - m$ between the ribosome and the RNAP; for example, a distance-dependent association rate might take the form $k_a(n - m) \approx k_a(\ell)[(\ell + \xi)/(n - m + \xi)]^3$, where $[(n - m) + \xi]^{-3}$ represents the effective volume fraction of the leading ribosome and ξ is the configuration flexibility of ribosome-RNAP binding when they are close. If we adopt such a distance-dependent k_a, we would also have to let the ratio p/q be dependent on $(n - m)$ in order to conserve free energy during approach and binding steps. To simplify matters, we will assume that $\xi \gg \ell$ and take k_a to be a constant for $d \equiv n - m \leq \ell$ and zero for $d \equiv n - m > \ell$.

The overall kinetics of the internal states pictured in the insets of Fig. 2 can be explicitly summarized by considering the intervening mRNA length $d = n - m$ between RNAP and the ribosome. Fig. S1 in Appendix S1 explicitly depicts the transitions for different d. Since, in our model, the maximum length of mRNA that can fit within the complex is ℓ, a processing ribosome-bound RNAP at $n = m + \ell$ cannot advance to lengthen the already compressed transcript (see Fig. 2 *E*). The only way a coupled state $a = 1$ with $d = \ell$ can reach any state where $d > \ell$ is for the ribosome and RNAP to first dissociate (we assume dissociation rates in all $d = n - m$ states remain constant at k_d). Molecular coupling effectively slows down transcription by preventing RNAP elongation in the $a = 1$, $d = \ell$ state. Such ribosome-mediated slowing down of transcription has been proposed in previous studies (2,9).

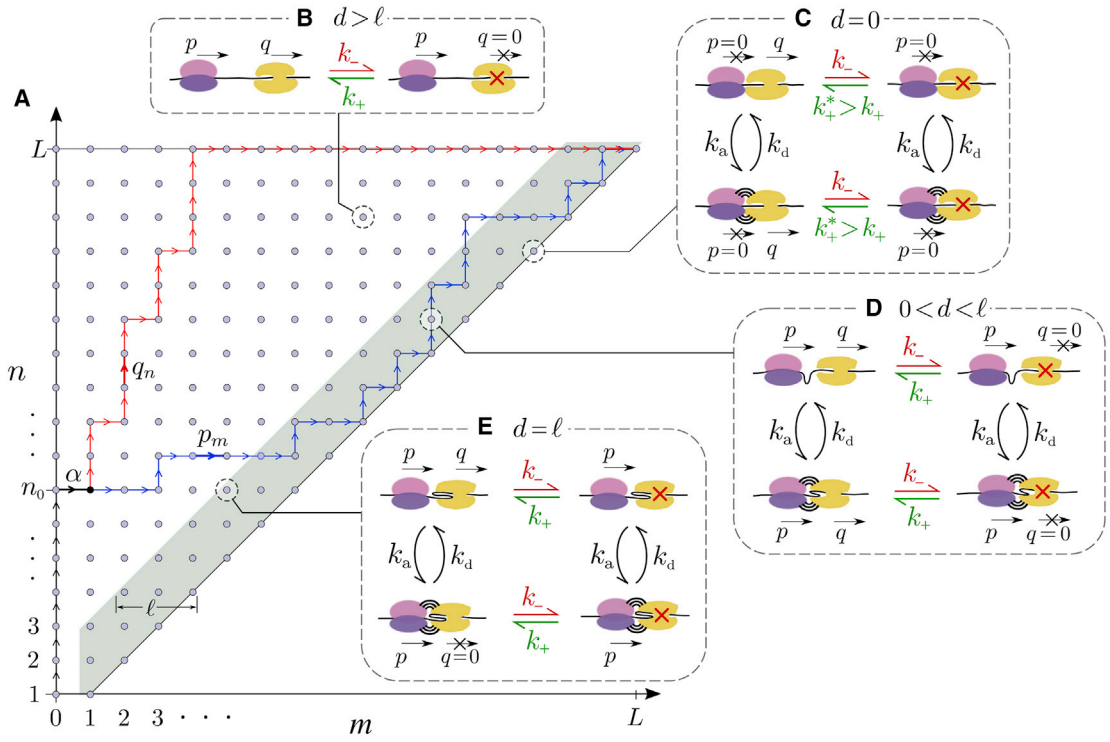

FIGURE 2 (*A*) State space of the stochastic model defined in terms of the leading ribosome and RNAP positions (m, n). The initial time $t = 0$ is defined as the time RNAP first produces a ribosome initiation site, starting the system in $(m = 0, n = 1)$. For $t > 0$, as the RNAP is elongating, the first ribosome binds at rate α. Here, a ribosome binds after the RNAP first reaches position $n = n_0$. Red and blue trajectories indicate scenarios in which the RNAP is relatively fast and slow, respectively. Within each position (m, n) exist internal molecular microstates. (*B*) In the "interior" states $n - m > \ell$ ($\ell = 2$ in this example), the ribosome and RNAP are too distant to be bound, and only stalled and processing RNAP states arise, with transition rates $k_\pm$ between them. (*C*) When $d = n - m = 0$, the ribosome and RNAP are adjacent without any intervening mRNA, allowing them to associate with rate k_a. The RNAP can be in either stalled or processive states. In the stalled state, whether associated or not, the adjacent volume-excluding ribosome entropically "pushes" the stalled RNAP, catalyzing its transition to a processive state so that $k_+^* > k_+$. (*D*) When $0 < n - m < \ell$, the ribosome and the RNAP are close enough to bind with rate k_a. Here, the intervening mRNA dissipates the entropic pushing (so that the stalled RNAP $\rightarrow$ processing RNAP transition rate is k_+) and also allows an RNAP in the processive state to elongate with rate q, regardless of whether it is bound to the ribosome. (*E*) Only when the ribosome and the RNAP are separated by $d = \ell$ is a bound RNAP prevented from processing as this would reel in more mRNA than can be fit inside the complex, either a collided or NusG-mediated expressome. Molecular binding prevents complexed ribosome and RNAP to be separated by more than ℓ mRNA codons.

We now list all allowed transitions in the $\omega := \{m, n, a, b\}$ state space of our continuous-time stochastic Markov model. The probability that an allowable transition from state ω to state ω' occurs in time increment dt is $r(\omega'|\omega)$dt, where the complete set of rates is given by

$$r(1, n, a, b \mid 0, n, a, b) = \alpha, \qquad 1 \le n \le L, \tag{2}$$

$$r(m+1, n, a, b \mid m, n, a, b) = p_m, \quad 1 \le m \le n-1, \tag{3}$$

$$r(m, n+1, 0, 0 \mid m, n, 0, 0) = q_n, \quad m \le n \le L-1, \tag{4}$$

$$r(m, n+1, 1, 0 \mid m, n, 1, 0) = q_n, \quad 0 \le d \le \ell - 1, \tag{5}$$

$$r(m, n, 1, b \mid m, n, 0, b) = k_a, \qquad 0 \le d \le \ell, \tag{6}$$

$$r(m, n, 0, b \mid m, n, 1, b) = k_d, \tag{7}$$

$$r(m, n, a, 1 \mid m, n, a, 0) = k_-, \tag{8}$$

$$r(m, n, 0, 0 \mid m, n, 0, 1) = k_+, \tag{9}$$

$$r(m, n, 1, 0 \mid m, n, 1, 1) = k_+^*, \qquad a = 1,\ m = n. \tag{10}$$

Using these rules, we performed event-based stochastic simulations (38,39) of the model as detailed in Appendix S1 of the supporting material. For completeness, the master equation associated with our model is also formally given in Appendix S2.

Construction of time-delay distribution

Our stochastic molecular model allows for the explicit calculation of the probability density $\rho(\Delta T)$ of time delay $\Delta T = T_{\text{rib}} - T_{\text{RNAP}}$ between mRNA and polypeptide completion. To find $\rho(\Delta T)$, we first find the distribution of ribosome positions $m(T_i)$ at the moment $T_i \equiv T(n = i)$ the RNAP first reaches site i. $T_L \equiv T_{\text{RNAP}}$ denotes the

instant the mRNA is completed. The initial value $m(T_1) = 0$ is known because, immediately after initiation of RNAP at site $n = 1$, the ribosome is not yet present but is trying to bind at a rate of α. As detailed in Appendix S3, we can iteratively find the distribution of $m(T_{i+1})$ given that of $m(T_i)$. By the same method, the distribution of association values $a(T_{\mathrm{RNAP}})$ at the instant of RNAP completion can be computed. After constructing the probability distribution $\mathbb{P}(m,a,b \mid t = T_{\mathrm{RNAP}})$, we evolve it to find the distribution of the additional amount of time $\Delta T = T_{\mathrm{rib}} - T_{\mathrm{RNAP}}$ required for m to first reach L.

Although we are able to construct the whole distribution of delay times that might provide a more resolved metric, especially if single-molecule or single-cell assays can be developed, a short time delay is a necessary but not sufficient condition for TTC. To provide direct information on molecular ribosome-RNAP interactions, we construct additional metrics.

Analytic approximation

Although our stochastic model contains a complex state space with many parameters, simple physical limits are immediately apparent. If the free ribosome translocation rate p is much greater than the free RNAP transcription rate q and $\alpha \gg L/\overline{q}$, the ribosome, for much of the time, abuts the RNAP, inducing it to transcribe at an average rate $qk_+^*/(k_+^* + k_-) \equiv \overline{q}^*$. Here, we predict an expected delay $\mathbb{E}[\Delta T] \approx 0$, $C \approx k_{\mathrm{a}}/(k_{\mathrm{a}} + k_{\mathrm{d}})$, and $F_T \approx 1$. If the ribosome is slow and $p \ll \overline{q}$, the ribosome and RNAP are nearly always free, $\Delta\mathbb{E}[T] \approx 1/\alpha + L(1/p - 1/\overline{q})$, $C \approx 0$, and $F_T \approx 0$. However, when p is intermediate, more intricate kinetics can arise, including tethered elongation and transcription slowdown.

To gain insight into this intermediate case, we first assume the large system limit $L \to \infty$ and divide the states into those with $d > \ell$ and those with $d \leq \ell$. Within each metastate, the behavior of the ribosome and RNAP is relatively homogeneous. States with $d > \ell$ are completely decoupled, while the proximal states $d \leq \ell$ can include associated $(a = 1)$ or unassociated $(a = 0)$ ribosome/RNAP substates.

To characterize the transition between the two metastates, we define the mean dwell time in metastate $d > \ell$ by T_0 and the mean dwell time in the metastate $d \leq \ell$ by T_1. The variability in the $d > \ell$ state dwell time arises mainly from RNAP stalling and unstalling. When translation is slow, the RNAP may undergo multiple stalling-unstalling cycles before leaving the $d > \ell$ metastate. For faster translation, one stalling event will typically be enough to leave the $d > \ell$ metastate. Let $P_0 = p/[p + (q-p)k_+/k_-]$ be the probability of leaving $d > \ell$ following one RNAP stalling event. With details provided in Appendix S4 of the supporting material, we find analytic approximations to T_0 and T_1 in the $q > p > \overline{q}$ and small ℓ limit:

$$T_0 \approx \frac{q}{pk_-} + \left(\frac{1}{k_-} + \frac{1}{k_+}\right)\left[\frac{(q-p)k_+}{pk_-} + \frac{(1-P_0)(q-p)/k_-}{p/k_+ - (q-p)/k_-}\right], \tag{11}$$

$$T_1 \approx \frac{k_{\mathrm{a}} + k_{\mathrm{d}}}{k_{\mathrm{d}}}\left(1 + \frac{k_-}{k_+^*} + \frac{k_-\ell}{p}\right)\left(\frac{1}{p} + \frac{1}{q}\right)\left(\frac{p}{q}\right)^{\ell}\sum_{j=1}^{\ell} j\left(\frac{q}{p}\right)^{j}. \tag{12}$$

As $p \downarrow \overline{q}$, the ribosome and RNAP are less likely to be proximal and $T_0 \to +\infty$. Once $p \leq \overline{q}$, the RNAP and the ribosome remain separated. On the other hand, as $p \uparrow q$, $1 - P_0 \to 1$, $T_0 \to 1/k_-$, and T_1 is always monotonically increasing with p. Kinetics between these two metastates qualitatively capture the behavior of the model in the translation-invariant, infinite length limit. The approximations, T_0 and T_1 will used in some of our subsequent results.

Metrics of TTC

We now define additional metrics that help characterize TTC. To more explicitly quantify direct *molecular* coupling, we define the *coupling coefficient* C by

$$C \equiv \mathbb{P}(a = 1 \mid t = T_{\mathrm{RNAP}}), \tag{13}$$

the probability that the ribosome is associated with the RNAP $(a = 1)$ at the moment the mRNA transcript is completed. The coupling parameter C provides a more direct measure of molecular coupling and further resolves configurations that have short or negligible delays. While delay time distributions do not directly quantify ribosome-RNAP contact, the coupling coefficient C does not directly probe the trajectories or history of ribosome-RNAP dynamics.

To also characterize the history of ribosome-RNAP interactions, we quantify TTC by the fraction of time F_T that the ribosome "protects" the RNAP across the entire transcription process. There are different ways of defining how the transcript is protected. While both modes of TTC are proposed to shield the mRNA from premature termination, neither has been directly observed in vivo. We assume that a termination protein has size $\sim \ell_{\mathrm{p}}$ and that, if the ribosome and RNAP are closer than ℓ_{p}, the termination factor is excluded. Thus, we define the protected time as the total time that $d < \ell_{\mathrm{p}}$ codons, divided by the time to complete transcription:

$$F_T = \frac{\| \{t : (n_t - m_t < \ell_{\mathrm{p}})\} \|}{T_{\mathrm{RNAP}}}. \tag{14}$$

Since the transcription-termination protein Rho has an mRNA footprint of about 80 nt, $\ell_{\mathrm{p}} \approx 27$ codons (40). A high F_T reflects transcription that has been protected against exposure to Rho-mediated termination. The fraction of transcription events that is prematurely

terminated is then proportional to $1 - F_T$. Exploration of this fraction under different conditions might be a way to experimentally estimate F_T.

Using these metrics, including the effective velocities $\overline{V}_{\text{rib}}$ and $\overline{V}_{\text{RNAP}}$, we will explore properties and predictions of our model.

RESULTS AND DISCUSSION

Here, we present analyses of solutions to our model obtained from numerical recursion and Gillespie-type kinetic Monte Carlo simulations detailed in Appendices S1, S2, and S3 of the supporting material. Predictions derived from using different parameter sets are compared, and mechanistic interpretations are provided.

Comparison of TTC metrics

We evaluate our stochastic model to provide quantitative predictions for the coupling indices, $\rho(\Delta T)$, C, and $\mathbb{E}[F_T]$. The results are summarized in Fig. 3.

Limitations of mean delay times

Fig. 3 *A* shows delay-time distributions for various parameter sets and reveals subtle differences in the kinetic consequences of coupling. Without molecular coupling ($k_a = 0$), the distribution has a single peak around the mean delay time. With molecular coupling, the distribution can exhibit two peaks with one at $\Delta T = 0$. This short-time peak reflects trajectories that terminate as a bound ribosome-RNAP complex. These finer structures in $\rho(\Delta T)$ cannot be resolved by evaluating only the mean delay time. Fig. 3 *B* plots the mean delay $\mathbb{E}[\Delta T]$ as a function of p and q. For our chosen parameters, in particular $k_a = 100$/s and $k_d = k_a e^{-3}$, we see that $\mathbb{E}[\Delta T]$ is rather featureless, with a significant delay arising only for small p. Thus, the mean delay time provides little information about the details of TTC.

Coupling coefficient

The dimensionless ratio $p/\overline{q}$ is a key indicator of the overall level of coupling possible. If $p/\overline{q} > 1$, the speed of the ribosome exceeds the average speed of the RNAP, allowing them to approach each other and potentially form a coupled expressome. If $p/\overline{q} < 1$, the ribosome speed is slower than the average RNAP speed and the system can at most be only transiently coupled. It turns out that the coupling coefficient C is mostly determined by $p/\overline{q}$ alone, particularly if the effects of translation initiation vanishes for large L (see "effects of translation initiation rates" section later). Essentially, the transition to a coupled system (large C) is predicted when $p/\overline{q} \gtrsim 1$. In Fig. 3 *C*, we find the values of C for multiple values of p and q (each dot corresponds to each (p, q) pair), and plot them as a function of $p/\overline{q}$, with $k_+/(k_+ + k_-) \approx 0.43$. The mean values of C as a function of p and q are plotted in Fig. 3 *D* and are qualitatively distinct from the mean times shown in (*B*). Given Eqs. 11 and 12, the steady-state coupling coefficient C for $L \to \infty$ can be approximated as

$$C \approx \frac{k_a}{k_a + k_d} \frac{T_1}{T_1 + T_0}, \tag{15}$$

which qualitatively agrees with numerical results as shown in Fig. 3 *C*.

Fraction of time protected

Each point in Fig. 3 *E* indicates the mean value F_T, $\mathbb{E}[F_T]$, for different values of p and q, arranged along values of $p/\overline{q}$. Each value $\mathbb{E}[F_T(p,q)]$ was computed by averaging protected-time fractions F_T (Eq. 14) from 1000 simulated trajectories. As expected, $\mathbb{E}[F_T(p,q)]$ increases with ribosome translation rate p until saturation to above $\mathbb{E}[F_T] \gtrsim 0.9$ for $p \gtrsim 22$ codons/s.

Comparing C and $\mathbb{E}[F_T]$ from Fig. 3 *D* and *F* we find that C and $\mathbb{E}[F_T]$ are qualitatively similar across various values of p and q, although in general we find $\mathbb{E}[F_T] \gtrsim C$. The transition from low to high values occurs at lower values of p for $\mathbb{E}[F_T]$ since the condition for protection ($d \le \ell_p$) is not as stringent as that for $C = 1$ ($d \le \ell$ and binding). Thus, there can be value of (p, q) for which $C(p,q)$ is small but $\mathbb{E}[F_T(p,q)]$ is close to one.

The similarity between $\mathbb{E}[F_T]$ and C is restricted to the dependence on p and q. The coupling and the protection time fraction may respond to changes in other parameters in drastically different ways. For example, C is nonzero only if molecular binding is present, rendering it sensitive to k_a, k_d. However, F_T directly measures the dynamics of TTC and does not depend on actual molecular coupling, so it will be relatively insensitive to k_a, k_d, particularly when p is large. Thus, F_T may be a better index if we wish to quantify functional consequences of TTC. The standard deviation of the simulated F_T values is large and approximately $\sqrt{\mathbb{E}[F_T](1 - \mathbb{E}[F_T])}$ (shown in Appendix S5 of the supporting material), limiting the suitability of the mean protected-time fraction as a robust metric.

As with the coupling coefficient C, F_T can be estimated under our heuristic approximations:

$$F_T \approx \frac{T_1}{T_1 + T_0} + \frac{\frac{q}{pk_-} + \left(\frac{1}{k_-} + \frac{1}{k_+}\right)\frac{(q-p)k_+}{pk_-}}{T_0 + T_1}\left(1 - e^{-\frac{k_-}{q-p}(\ell_p - \ell)}\right). \tag{16}$$

From Eq. 16, we see that F_T is comprised of two terms, the protection provided by coupling $\frac{T_1}{T_1+T_0} \approx \frac{k_a+k_d}{k_a} C$, and the protection provided by ribosome elongation when $d > \ell$. Although Eq. 16 is valid mainly

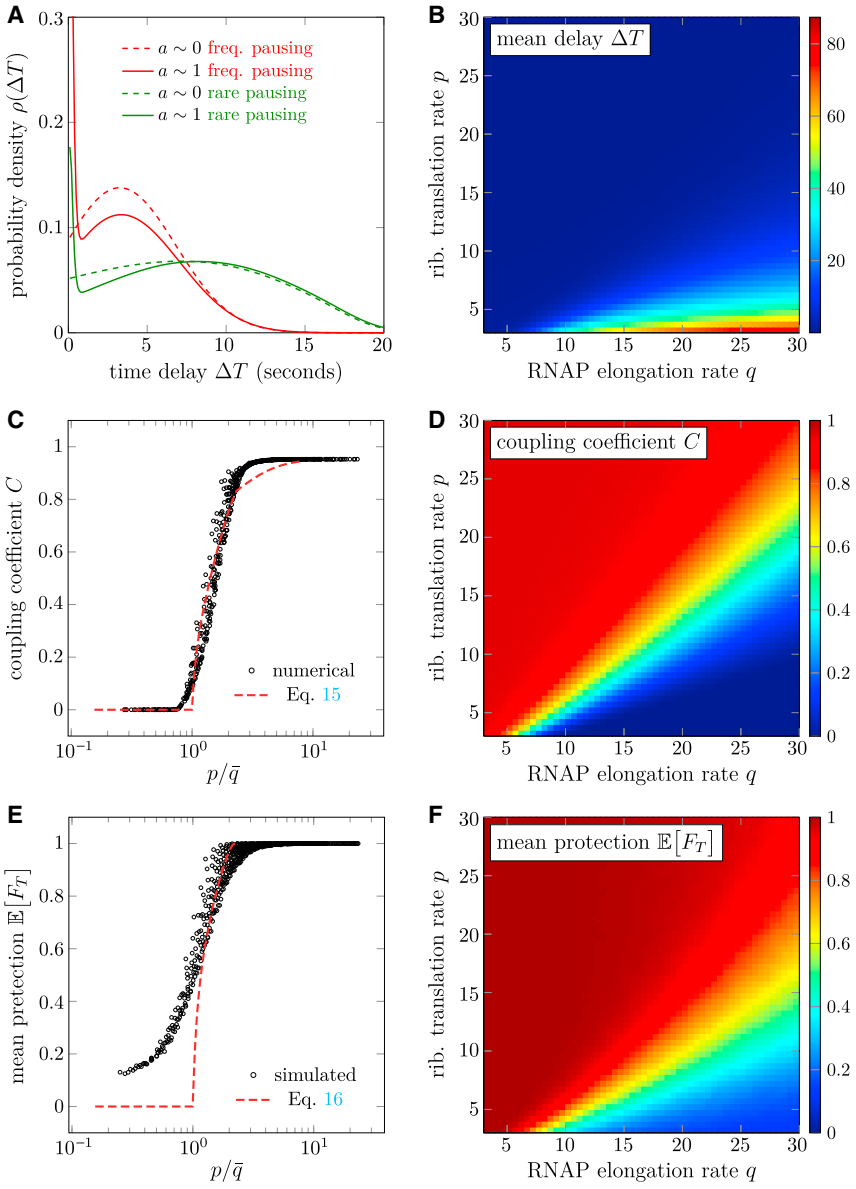

FIGURE 3 Comparison of different TTC indices. Common parameters for all these plots are $\alpha = 1$/s, $E_+ = 2$, $E_a = 3$, $k_d = k_a e^{-E_a}$, $\ell = 4$, $L = 335$, and $k_a = 100$/s, unless stated otherwise. (*A*) The delay-time density $\rho(\Delta T)$ plotted for $p = 12$ codons/s and $q = 30$ codons/s. Densities for $k_+ = 0.4$/s, $k_- = 0.3$/s (rarely pausing RNAPs, *green curves*), and $k_+ = 4.0$/s, $k_- = 3.0$/s (frequently pausing RNAPs, *red curves*) are shown. Within these cases, strong-binding ($a \sim 1$, $k_a = 100$/s, $k_d = k_a e^{-3}$) and no-binding ($a = 0$, $k_a = 0$) subcases are indicated by solid and dashed curves, respectively. (*B*) Mean delay $\mathbb{E}[\Delta T]$ as a function of p and q. (*C*) The direct coupling coefficient C as a function of the relative velocity $p/\overline{q}$. Each point represents C evaluated at specific values of (p,q), each chosen from all integers between 3 and 27 codons/s. The dashed curve represents the analytic approximation given by Eq. 15. (*D*) Heatmap of $C(p,q)$. (*E*) Values of $\mathbb{E}[F_T]$, each derived from 1000 kinetic Monte-Carlo (kMC) trajectories, plotted against $p/\overline{q}$. The analytic approximation given by Eq. 16 is shown by the dashed curve. (*F*) The heatmap of $\mathbb{E}[F_T(p,q)]$.

for positive E_a and $\ell_p > \ell$, the protection from physical proximity exists even if molecular association is completely absent. This formula confirms our observation that $\mathbb{E}[F_T] \gtrsim C$.

Binding-induced slowdown

We now use our model to reveal the major factors contributing to TTC-induced slowdown of transcription and discuss limits this slowdown.

Unstalling rate k_+^* dictates ribosome efficiency

The principal factor that influences the overall velocity $\overline{V}$ of a coupled expressome is the interplay between two antagonistic mechanisms: ribosome-mediated dislodging of an adjacent stalled RNAP and bound-ribosome deceleration of the RNAP. When the reduction in activation free energy of unstalling, $E_+ = \log(k_+^*/k_+)$, is large, the ribosome is less likely to be impeded by a stalled RNAP. Fig. 4 *A* plots the effective velocities $\overline{V}_{rib}$ and $\overline{V}_{RNAP}$ as a function of E_+ and illustrates the increases in overall speed when the ribosome is more effective at dislodging a stalled RNAP (higher E_+).

The decrease in the velocity of a coupled processing RNAP is primarily determined by the ribosome translation speed p. For different values of $E_+ = \log(k_+^*/k_+)$, the dependence of $\overline{V}_{RNAP}$ on p can be quite different, as is shown in Fig. 4 *B*. For large E_+, when the ribosome efficiently pushes stalled RNAPs, increasing p allows the ribosome to more frequently abut the RNAP and dislodge

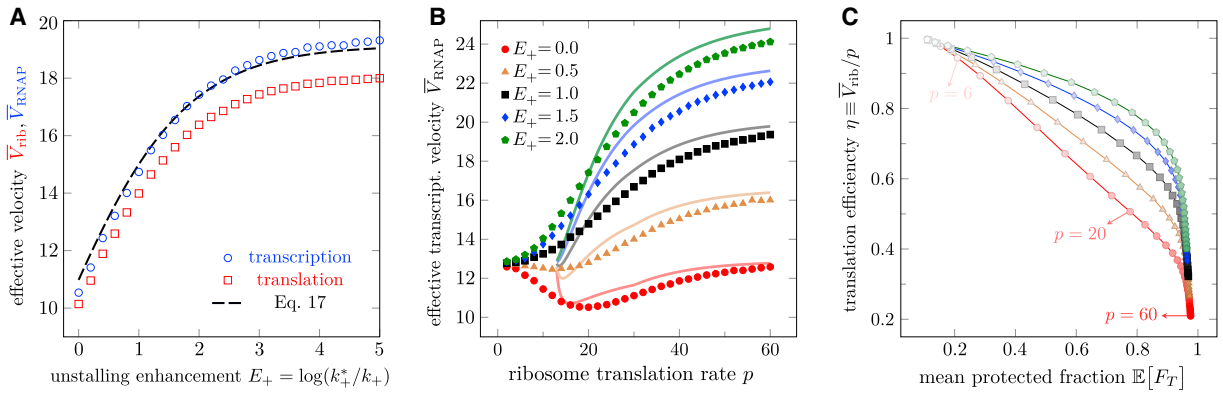

FIGURE 4 Slowdown induced by molecular coupling. The common parameters used are the same as those in Fig. 3. (*A*) Effective velocities as a function activation energy reduction $E_+ = \log(k_+^*/k_+)$ in ribosome-induced RNAP unstalling, $p = 20$ codons/s and $q = 30$ codons/s. (*B*) Effective velocity of RNAP as a function of the free ribosome translation rate p. (*C*) The trade-off between translation efficiency and mean fraction of time protected $\mathbb{E}[F_T]$. The efficiency $\eta \equiv \overline{V}_{\rm rib}/p$ is defined by the ratio of the mean ribosome speed to the translation rate of an isolated ribosome. In (*B*) and (*C*), $q = 30$ codons/s. The variances (not shown) for the plotted quantities are large, typically overlapping the mean-value curves in (*A* and *B*). The dashed curve in (*A*) and solid curves in (*B*) are the analytical predictions of the effective velocities using Eq. 17. To see this figure in color, go online.

it, leading to faster overall transcription. However, for inefficient unstalling (small E_+), we see that faster ribosomes can *decrease* $\overline{V}_{\rm RNAP}$. This feature arises because for inefficient unstalling, a larger p increases the fraction of time the ribosome and RNAP are bound ($a = 1$), allowing a binding-induced slowdown to arise more often. Besides E_+, the emergence of a decreasing transcription velocity $\overline{V}_{\rm RNAP}$ with increasing ribosome translation rate p depends intricately on factors, such as ℓ, $k_\pm$, $k_{\rm a}, k_{\rm d}$, and arises only if $k_{\rm a}/k_{\rm d}$ is sufficiently large and ℓ is not too large.

Although the decrease in $\overline{V}_{\rm RNAP}$ is not large, it certainly suggests that increasing p under small $E_+ \leq 0.5$ is not advantageous. This observation motivates us to define a translation efficiency as the ratio of the effective ribosome speed $\overline{V}_{\rm rib}$ to its unimpeded translation speed p: $\eta \equiv \overline{V}_{\rm rib}/p$. The loss $1 - \eta$ measures how much a ribosome is impeded due to its interactions with the RNAP. As p is increased, we find trajectories that display a trade-off between translation efficiency and protected time. Higher p leads to more proximal ribosomes and protected RNAP at the expense of translation efficiency η. Fig. 4 *C* shows that the decrease in unstalling activation energy E_+ affects this level of trade-off. For large E_+, increasing p can speed up ribosomes beyond the velocity determined by $\overline{q}$ so that η decreases more slowly than $\sim 1/p$. At the same time, the system is only slightly less coupled, leading to a subtle decrease in $\mathbb{E}[F_T]$. In the end, larger E_+ leads to a higher η versus $\mathbb{E}[F_T]$ curve.

Low η values are likely selected against since a cell would be expending more resources than necessary to maintain high levels of tRNA and other translation factors. A potentially optimal setting may be to maintain $p \simeq \overline{q}$, which is the minimally sufficient velocity to keep the RNAP protected. This intermediate choice of p for the ribosome may explain the recent observations that slower ribosomes did not appreciably slow down transcription (22) or prevent folding of specific mRNA segments (21).

When $p > \overline{q}$ and $L \to \infty$, the ribosome and the RNAP intermittently close on each other and share a common effective velocity. In states $d > \ell$, the velocity $v_0 = \min\{p, q\}$ is determined by the slower of the ribosome or the processive RNAP. While the system is in states $d \leq \ell$, its velocity is given by $v_1 \approx q\left[\frac{(q/p)^\ell - 1}{(q/p)^{\ell+1} - 1}\right] \frac{k_+^*}{k_+^* + k_-}$, as detailed in Appendix S4. Finally, the overall effective velocity is estimated by

$$\overline{V} \approx \frac{T_0}{T_0 + T_1} v_0 + \frac{T_1}{T_0 + T_1} v_1. \tag{17}$$

Eq. 17 agrees qualitatively well with simulation and reproduces the slowdown, as shown in Fig. 4 *A* and *B*.

Limits of binding-induced slowdown

When $\overline{q} < p < q$ and $L \to \infty$, we explore the lower bound for the effective velocity of RNAP by examining v_1. For sufficiently large $\ell \geq 1$ and q/p, the term $\left[\frac{(q/p)^\ell - 1}{(q/p)^{\ell+1} - 1}\right]$ is approximately p/q, and lower bounds for $\overline{V}_{\rm RNAP}(a = 1)$ are

$$\overline{V}_{\rm RNAP} \geq \frac{p k_+}{k_+ + k_-} \geq q\left(\frac{k_+}{k_+ + k_-}\right)^2. \tag{18}$$

The first equality holds when $k_+^* = k_+$, and the second equality holds when $p = \overline{q}$. We conclude that the maximum slowdown induced by binding is essentially limited by the slowdown of RNAP due to transcriptional road blocks. The latter plays a fundamental role in the significantly slower rate of mRNA transcription ~ 45 nt/s relative to rRNA transcription ~ 90 nt/s.

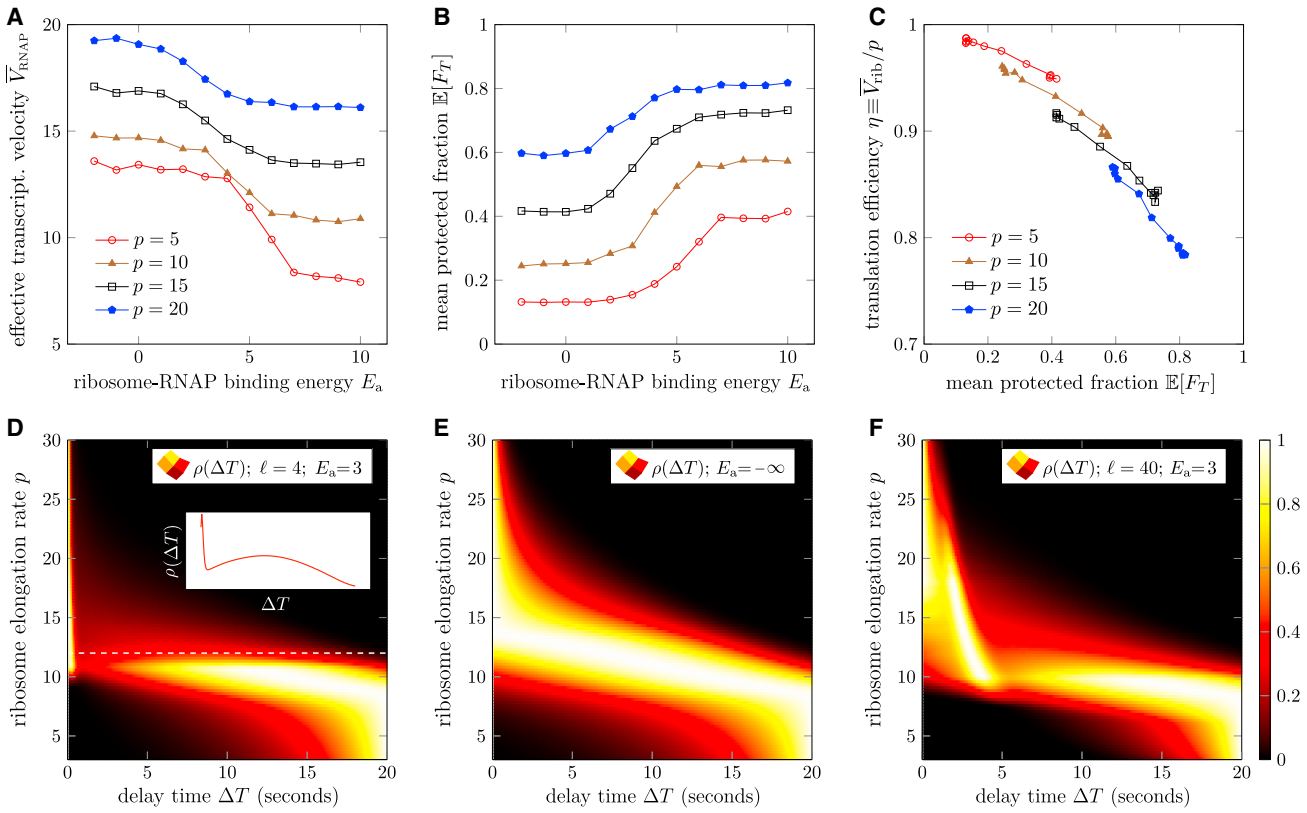

FIGURE 5 Effects of molecular coupling. For those parameters not varied, we use the same values used to generate Figs. 3 and 4. (*A*) For $\ell = 4$, the effective transcription velocity $\overline{V}_{\rm RNAP}$ as a function of binding-energy depth between the ribosome and RNAP E_a. (*B*) Mean protected-time fraction $\mathbb{E}[F_T]$ as a function of binding energy depth E_a ($\ell = 4$). (*C*) The trade-off between efficiency and protection for $\ell = 4$. (*D*) Rescaled heatmap of the delay-time distribution $\rho(\Delta T)$ as a function of ribosome translocation rate p. The brightness indicates the relative probability, and the inset shows the probability distribution at $p = 12$ codons/s indicated by the dashed white line. Here, the binding energy $E_a = 3$ and $\ell = 4$. For $p \approx 9$ codons/s, $\rho(\Delta T)$ is bimodal in ΔT. (*E*) Delay-time distribution in the absence of ribosome-RNAP binding ($k_a = 0$). Here, the ℓ dependence disappears and $\rho(\Delta T)$ is mono-modal. (*F*) Delay-time distribution for $\ell = 40$ and $E_a = 3$. Bimodality arises in more than one regime of p. To see this figure in color, go online.

Testing the molecular coupling hypothesis

It is informative to use our proposed indices to compare scenarios that predict molecular coupling to those that do not. We now vary the binding energy E_a for different velocity ratios $p/\overline{q}$. For $\ell = 4$, Fig. 5 *A* shows $\overline{V}_{\rm RNAP}$ as a function of E_a for various values of p. Although higher p leads to increased $\overline{V}_{\rm RNAP}$, for each value of p, increasing the binding energy increases coupling and leads to RNAP slowdown. Since the most significant slowdown arises when translation is slow, association-induced slowdown is indicative of strong coupling. Both p and E_a increase $\mathbb{E}[F_T]$, as shown in Fig. 5 *B*. The variation in $\overline{V}_{\rm RNAP}$ and $\mathbb{E}[F_T]$ with E_a also suggests that, within the context of our model, one can use experimental measurements of these indices to estimate E_a. As E_a is tuned in our model we also predict a trade-off between ribosome efficiency and protection, as shown in Fig. 5 *A*. Since our analytic approximation is most accurate in the $p \geq \overline{q}$, $L \to \infty$ regime, deviations in $\overline{V}_{\rm RNAP}$ and $\mathbb{E}[F_T]$ are mainly due to finite L effects when $p < \overline{q}$; thus, we forgo plotting our analytic approximations in Fig. 5 *A–C*. In Appendix S6, we provide additional simulation results that confirm the ℓ dependence in Eq. 17 and in $\mathbb{E}[F_T]$.

Additional parallel experiments may show consistent evidence of coupling; thus, we investigate the distribution of delay times $\rho(\Delta T)$ as p is varied. Fig. 5 *D* shows $\rho(\Delta T)$ rescaled so that the largest value is set to unity for easier visualization. We see that for intermediate values of $10 \lesssim p \lesssim 13$ (unit codons/s, omitted henceforth), $\rho(\Delta T)$ can be bimodal. Fig. 5 *E* depicts a single peaked $\rho(\Delta T)$ when coupling is completely turned off by setting $k_a = 0, (E_a = -\infty)$. In this case, ℓ is irrelevant. Fig. 5 *F* shows the rescaled $\rho(\Delta T)$ in the presence of coupling ($E_a = 3$) for $\ell = 40$. Here, there are two regimes, $8 \lesssim p \approx 12$ and $18 \lesssim p \lesssim 24$, that exhibit bimodality. The delay-time density would be ideally measured via single-molecule or single-cell experiments combined with proper fluorescent imaging. However, in Appendix S7 we also consider how improved bulk LacZ completion assays can be used to estimate features of $\rho(\Delta T)$. We find that the distribution of ribosome completion times $T_{\rm rib}$ is always mono-modal and that the bimodality of ΔT is mainly the result of bimodality in $T_{\rm RNAP}$.

Genome-wide variability of coupling

We have so far assumed all parameters are homogeneous along the transcript and time independent. However, a cell is able to dynamically regulate the transcription and translation of different genes by exploiting the transcript sequence or other factors that mediate the process. Such regulation can be effectively described within our model by varying its parameters in the appropriate way.

Regulation of RNAP pausing

The RNAP pausing rate k_- is one parameter that can be modulated by specific DNA sequences and other roadblocks along the gene (36,41–43). There is evidence that consensus pause sequences are enriched at the beginning of genes (44,45). In addition to leading ribosomes, a trailing RNAP can also push the leading RNAP out of a paused state by increasing k_+, much like ribosomes (6,36). Even if k_+ and k_- are varied in our model, the overall predicted performance regimes of the system are still delineated by values of $p/\overline{q}$, and the effective transcription velocity can still be predicted by Eq. 17.

Effects of translation initiation rates

Translation initiation is another process that can be altered by the cell through, e.g., initiation factors that modulate the initiation rate α (46). Genome-wide analysis reveals that translation initiation times in *E. coli* are highly variable, ranging from less than 1 s to more than 500 s (31,47).

As shown in Figs. S8 *A–C* of Appendix S8, varying the translation initiation rate α straightforwardly affects TTC. As indicated in (*A*), the peak in $\overline{V}_{\mathrm{RNAP}}$ at $\alpha \simeq 1/\mathrm{s}$ persists across different values of p. Slower translation initiation results in larger initial separations n_0, decreasing the overall fraction of protected times, as shown in (*B*) and (*C*). To mitigate large initial distances n_0 and lower likelihood coupling due to slow initiation, RNAP pausing occurs more often at the start of the gene to allow time for a slow-initiating ribosome to catch up. Thus, delayed ribosome initiation and early RNAP pausing are two "opposing" processes that can regulate coupling and efficiency, especially for short genes.

A simple analytic description of how the interplay between length L and transcription initiation rate α affects our TTC metrics could not be found. However, large/small α generally leads to higher/lower C and $\mathbb{E}[F_T]$ compared with those associated with the $L \to \infty$ approximations (Eqs. 15 and 16), as shown in Fig. S9 in Appendix S8. For sufficiently small α, a rough estimate of the deviation from steady state is $p\overline{q}/[(p - \overline{q})\alpha L]$, the typical time needed for the ribosome to catch up to the RNAP divided by the time needed for the RNAP to finish the transcription.

Ribosome translocation rate profiles

Although we have thus far assumed uniform ribosome translocation rates, it is known that codon bias and tRNA/amino acid availability can locally affect ribosome translocation (20,48). Snapshots of ribosome positions along transcripts have been inferred from ribosome profiling experiments. After imposing a stochastic exclusion model (49), Khanh and coworkers (50) reconstructed position-dependent ribosome translocation rates p_m. They found that hopping rates p_m are larger near the 5′ end and decrease toward the 3′ end. Although they reconstructed the entire genome-wide p_m profile, translocation rates are gene dependent, so we will propose and test simple profiles p_m.

To qualitatively match the inferred profile (50), we define profile 1 by increasing p by 50% for the first 40 codons, and decreasing it by 50% for the second 40 codons. The rest of the transcript retains the constant baseline value of p. Profile 2 is similarly defined except that, instead of being the second group of 40 codons, the speed across the last 40 codons is decreased. We compared the performance of the three different profiles in Fig. S8 *D–F* as a function of the mean translation rate $\overline{p} \equiv L^{-1}\sum_{m=1}^{L} p_m$. We conclude that the differences in the effective velocities and protection fraction are subtle, suggesting that these performance statistics are relatively insensitive to different speed profiles.

CONCLUSIONS

We have presented a detailed stochastic model of TTC. The continuous-time discrete-state model tracks the distance between the leading ribosome and the RNAP and assumes they sterically exclude each other along the nascent mRNA transcript. All current experimental understanding of interactions between RNAP and ribosome, including ribosome initiation, RNAP pausing, and direct ribosome-RNAP association, have also been incorporated. Our model exhibits a number of rich features that depend on the interplay of these intermediate mechanisms.

To quantitatively investigate the predictions of our model, we constructed three different metrics to quantify TTC, the delay-time probability distribution $\rho(\Delta T)$, the probability C that the ribosome and the RNAP are in a bound state $(a = 1)$ at termination, and the fraction of time F_T that the ribosome and the RNAP are proximal over the entire process. F_T is a measure of protection against binding of termination proteins. These metrics were computed or simulated under different model parameters. Specifically, since a bound RNAP at distance ℓ from the trailing ribosome needs to first detach before $d = \ell$ can be increased, the $d \leq \ell$ states shown in Figs. 2 and S1 form an effective attractive well that tethers RNAP to ribosome. By allowing direct ribosome-RNAP binding, we find that this effective attraction zone can allow a slower ribosome to dynamically hold back bound RNAP, leading to decreased $\overline{V}_{\mathrm{RNAP}}$. Indirect coupling mediated by Nus proteins provides a longer interaction range, which may improve translation efficiency η compared with direct coupling where $\ell \sim 4$ codons.

Qualitatively, our model predicts two different regimes of TTC that appear to be consistent with observations. One limit can arise when E_a is large, resulting in close proximity and strong molecular coupling that may lead to slowdown of RNAP, while the other arises when E_a is small leading to intermittent contacts and perhaps modest speed up of pausing RNAPs. Besides E_a, our model suggests that ℓ, α, and $p/\overline{q}$ also control which type of TTC arises. Across different genes, E_a and ℓ are expected to be unchanged, but variations in $p/\overline{q}$ (and to some degree α) can affect the balance between these qualitative models of TTC. Our analysis suggests that the unstalling enhancement E_+ is a key factor determining how TTC is manifested. For inefficient unstalling, increased coupling slows down the whole expressome, which might be disadvantageous under growth conditions. Gene-dependent kinetic parameters and signaling pathways involving ppGpp (12,22) and other factors can also be incorporated into our model to provide a more complete picture of how TTC influences cellular regulation.

Our model reveals that a bimodal time-delay distribution when $p \approx \overline{q}$ is a hallmark of molecular association. We provide a discussion of possible measurements of this distribution through traditional bulk assays in the supporting material. By quantitatively controlling ribosome speed p and measuring the corresponding effective velocity measured in the LacZ-completion assay or other single-molecule assays. such as DNA curtains, important kinetic information associated with the delay-time density may be revealed by fitting our model to data.

Measurement of important parameters and indices, such as E_+ and F_T, could be obtained from quantitative analysis of previous experiments (2,22). Cryo-EM approaches used in (17) might also be used to obtain a genome-wide estimate of the coupling coefficient C. Finally, FRET experiments or super-resolution imaging may shed light on macromolecular-level ribosome and RNAP dynamics (51). Our model can guide how in vitro measurements can be designed and used to inform delay-time distributions $\rho(\Delta T)$, coupling coefficients C, protected-time fractions F_T, and efficiencies η.

Modeling of transcription and translation coupling may also inform the stochastic modeling of gene transcription dynamics. Simple kinetic models of transcription that incorporate the active and inactive states of the gene, the transcription initiation, elongation, pausing, premature termination, and the degradation of mRNA products have been developed to understand cell-to-cell variability of gene-specific mRNA levels (52–54). Such models could be merged with our modeling approach to accommodate TTC.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.09.041.

AUTHOR CONTRIBUTIONS

X.L. and T.C. devised and analyzed the model and wrote the paper. X.L. developed the computational algorithms and performed the numerical calculations and kinetic Monte Carlo simulations.

ACKNOWLEDGMENTS

This work was supported by grants from the NIH through grant R01HL146552 and the NSF through grant DMS-1814364 (to T.C.).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Miller, O. L., B. A. Hamkalo, and C. A. Thomas. 1970. Visualization of bacterial genes in action. *Science.* 169:392–395.
2. Stevenson-Jones, F., J. Woodgate, …, N. Zenkin. 2020. Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest. *Proc. Natl. Acad. Sci. USA.* 117:8462–8467.
3. Sydow, J. F., and P. Cramer. 2009. RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* 19:732–739, Catalysis and regulation/Proteins.
4. Knippa, K., and D. O. Peterson. 2013. Fidelity of RNA polymerase II transcription: role of Rbp9 in error detection and proofreading. *Biochemistry.* 52:7807–7817.
5. Sahoo, M., and S. Klumpp. 2013. Backtracking dynamics of RNA polymerase: pausing and error correction. *J. Phys. Condens. Matter.* 25:374104.
6. Zuo, X., and T. Chou. 2022. Density- and elongation speed-dependent error correction in RNA polymerization. *Phys. Biol.* 19:026001.
7. Chalissery, J., G. Muteeb, …, R. Sen. 2011. Interaction surface of the transcription terminator Rho required to form a complex with the C-terminal domain of the antiterminator NusG. *J. Mol. Biol.* 405:49–64.
8. Lawson, M. R., W. Ma, …, J. M. Berger. 2018. Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. *Mol. Cell.* 71:911–922.e4.
9. Kohler, R., R. A. Mooney, …, P. Cramer. 2017. Architecture of a transcribing-translating expressome. *Science.* 356:194–197.
10. Ma, C., M. Mobli, …, P. J. Lewis. 2015. RNA polymerase-induced remodelling of NusA produces a pause enhancement complex. *Nucleic Acids Res.* 43:2829–2840.
11. Proshkin, S., A. R. Rahmouni, …, E. Nudler. 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science.* 328:504–508.
12. Iyer, S., D. Le, B. R. Park, and M. Kim. 2018. Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in Escherichia coli. *Nat. Microbiol.* 3:741–748.
13. Vogel, U., and K. F. Jensen. 1994. The RNA chain elongation rate in Escherichia coli depends on the growth rate. *J. Bacteriol.* 176:2807–2813.
14. Fan, H., A. B. Conn, …, G. M. Blaha. 2017. Transcription–translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res.* 45:11043–11055.
15. Mooney, R. A., S. E. Davis, …, R. Landick. 2009. Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell.* 33:97–108.
16. Saxena, S., K. K. Myka, …, M. E. Gottesman. 2018. Escherichia coli transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* 108:495–504.
17. O'Reilly, F. J., L. Xue, A. Graziadei, L. Sinn, S. Lenz, D. Tegunov, C. Blötz, N. Singh, W. J. H. Hagen, P. Cramer, J. Stülke, J. Mahamid, and

J. Rappsilber. 2020. In-cell architecture of an actively transcribing-translating expressome. *Science.* 369:554–557.

18. Webster, M. W., M. Takacs, …, A. Weixlbaumer. 2020. Structural basis of transcription-translation coupling and collision in bacteria. *Science.* 369:1355–1359.
19. Burmann, B. M., K. Schweimer, …, P. Rösch. 2010. A NusE:NusG complex links transcription and translation. *Science.* 328:501–504.
20. Chou, T., and G. Lakatos. 2004. Clustered bottlenecks in mRNA translation and protein synthesis. *Phys. Rev. Lett.* 93:198101.
21. Chen, M., and K. Fredrick. 2018. Measures of single- versus multiple-round translation argue against a mechanism to ensure coupling of transcription and translation. *Proc. Natl. Acad. Sci. USA.* 115:10774–10779.
22. Zhu, M., M. Mori, …, X. Dai. 2019. Disruption of transcription–translation coordination in Escherichia coli leads to premature transcriptional termination. *Nat. Microbiol.* 4:2347–2356.
23. Washburn, R. S., P. K. Zuber, …, J. Frank. 2020. Escherichia coli NusG links the lead ribosome with the transcription elongation complex. *iScience.* 23:101352.
24. Mäkelä, J., J. Lloyd-Price, …, A. S. Ribeiro. 2011. Stochastic sequence-level model of coupled transcription and translation in prokaryotes. *BMC Bioinf.* 12:121–213.
25. Johnson, G. E., J.-B. Lalanne, …, G.-W. Li. 2020. Functionally uncoupled transcription–translation in Bacillus subtilis. *Nature.* 585:124–128.
26. Li, R., Q. Zhang, …, H. Shi. 2016. Effects of cooperation between translating ribosome and RNA polymerase on termination efficiency of the Rho-independent terminator. *Nucleic Acids Res.* 44:2554–2563.
27. Chou, T. 2007. Peeling and sliding in nucleosome repositioning. *Phys. Rev. Lett.* 99:058105.
28. Teimouri, H., C. Spaulding, and A. B. Kolomeisky. 2022. Optimal pathways control fixation of multiple mutations during cancer initiation. *Biophys. J.* 121:3698–3705.
29. Wang, C., V. Molodtsov, …, R. H. Ebright. 2020. Structural basis of transcription-translation coupling. *Science.* 369:1359–1365.
30. Dai, X., M. Zhu, M. Warren, …, T. Hwa. 2016. Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. *Nat. Microbiol.* 2:16231.
31. Shaham, G., and T. Tuller. 2018. Genome scale analysis of Escherichia coli with a comprehensive prokaryotic sequence-based biophysical model of translation initiation and elongation. *DNA Res.* 25:195–205.
32. Kennell, D., and H. Riezman. 1977. Transcription and translation initiation frequencies of the Escherichia coli lac operon. *J. Mol. Biol.* 114:1–21.
33. Xu, L., H. Chen, …, Z. W. Luo. 2006. Average gene length is highly conserved in prokaryotes and eukaryotes and diverges only between the two kingdoms. *Mol. Biol. Evol.* 23:1107–1108.
34. Young, R., and H. Bremer. 1976. Polypeptide-chain-elongation rate in Escherichia coli B/r as a function of growth rate. *Biochem. J.* 160:185–194.
35. Zhu, M., X. Dai, and Y.-P. Wang. 2016. Real time determination of bacterial in vivo ribosome translation elongation speed based on LacZα complementation system. *Nucleic Acids Res.* 44:e155.
36. Epshtein, V., and E. Nudler. 2003. Cooperation between RNA polymerase molecules in transcription elongation. *Science.* 300:801–805.
37. Neuman, K. C., E. A. Abbondanzieri, …, S. M. Block. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell.* 115:437–447.
38. Bortz, A. B., M. H. Kalos, and J. L. Lebowitz. 1975. A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.* 17:10–18.
39. Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem. A.* 81:2340–2361.
40. Koslover, D. J., F. M. Fazal, …, S. M. Block. 2012. Binding and translocation of termination factor Rho studied at the single-molecule level. *J. Mol. Biol.* 423:664–676.
41. Komissarova, N., and M. Kashlev. 1997. RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* 272:15329–15338.
42. Nudler, E. 2009. RNA polymerase active center: the molecular engine of transcription. *Annu. Rev. Biochem.* 78:335–361.
43. Davenport, R. J., G. J. Wuite, …, C. Bustamante. 2000. Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase. *Science.* 287:2497–2500.
44. Hatoum, A., and J. Roberts. 2008. Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. *Mol. Microbiol.* 68:17–28.
45. Larson, M. H., R. A. Mooney, …, J. S. Weissman. 2014. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science.* 344:1042–1047.
46. Chou, T. 2003. Ribosome recycling, diffusion, and mRNA loop formation in translational regulation. *Biophys. J.* 85:755–773.
47. Siwiak, M., and P. Zielenkiewicz. 2013. Transimulation - protein biosynthesis web service. *PLoS One.* 8:e73943.
48. Klumpp, S., J. Dong, and T. Hwa. 2012. On ribosome load, codon bias and protein abundance. *PLoS One.* 7:e48542.
49. MacDonald, C. T., and J. H. Gibbs. 1969. Concerning the kinetics of polypeptide synthesis on polyribosomes. Biopolymers. *Biopolymers.* 7:707–725.
50. Dao Duc, K., and Y. S. Song. 2018. The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. *PLoS Genet.* 14:e1007166.
51. Morisaki, T., K. Lyon, …, T. J. Stasevich. 2016. Real-time quantification of single RNA translation dynamics in living cells. *Science.* 352:1425–1429.
52. Choubey, S., J. Kondev, and A. Sanchez. 2015. Deciphering transcriptional dynamics in vivo by counting nascent RNA molecules. *PLoS Comput. Biol.* 11:e1004345.
53. Xu, H., S. O. Skinner, …, I. Golding. 2016. Stochastic kinetics of nascent RNA. *Phys. Rev. Lett.* 117:128101.
54. Filatova, T., N. Popovic, and R. Grima. 2020. Statistics of nascent and mature RNA fluctuations in a stochastic model of transcriptional initiation, elongation, pausing, and termination. *Bull. Math. Biol.* 83:3.