The hidden function of nonspecifically DNA-bound RNA polymerases: a buffer of resource competition in bacteria

Yichen Yan,[1] Tairan Li,[2] and Jie Lin[1,2]

[1] *Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China*
[2] *Center for Quantitative Biology, Peking University, Beijing, China*

(Dated: November 4, 2023)

RNA polymerase (RNAP) is the workhorse of bacterial gene expression, transcribing rRNA and mRNA. Experiments found that a significant fraction of RNAPs in bacteria are nonspecifically bound to DNA, which is puzzling as these idle RNAPs could have produced more RNAs. Whether nonspecifically DNA-bound RNAPs have any function or are merely a consequence of passive interaction between RNAP and DNA is unclear. In this work, we propose that nonspecifically DNA-bound RNAPs mitigate the crosstalk between rRNA and mRNA transcription. They buffer the free RNAP concentration change upon the regulation of rRNA genes, thus reducing its impact on the mRNA transcription and vice versa. We verify our theory using mean-field models and an agent-based model of transcription. We show that the buffering effects are robust against the finite number of sigma factors and spatial fluctuations of RNAP concentration. We analyze the relevant data of *Escherichia coli* and show that its number and affinity of nonspecific binding sites are in the range where the buffering effects are significant, in concert with our theories.

INTRODUCTION

RNA polymerases (RNAPs) and ribosomes are finite resources that limit the overall rate of gene expression in bacteria [1–10]. As a result, genes compete for these resources, leading to inevitable crosstalk of gene expression [11–13]. This competition is particularly significant in bacteria as all RNAs are transcribed using the same type of RNAP. Thus, increasing the transcription of rRNA, will inevitably reduce the availability of RNAPs for transcribing mRNA. However, despite this competition, a significant proportion of RNAPs in bacteria (30% to 50%) are nonspecifically bound to DNA and do not participate in transcription [14–18]. This phenomenon is puzzling as these idle RNAPs could have contributed to transcription, producing more mRNAs. Furthermore, producing these non-transcribing RNAPs is a significant investment in the proteome, and bacterial cells could have avoided producing these idle RNAPs and allocated the finite resources to other essential proteins to enhance their fitness. Indeed, eukaryotes have been found to suppress the nonspecific binding of RNAPs to DNA actively [19]. Therefore, a puzzle emerges: why do bacteria keep so many RNAPs nonspecifically bound to DNA?

In this study, we propose that nonspecifically DNA-bound (ND) RNAPs in bacteria are crucial in regulating gene expression. These RNAPs mitigate the unwanted crosstalk between the expressions of different genes due to the limited availability of RNAPs. For example, when a subset of genes, such as the genes of rRNAs, are upregulated, more RNAPs participate in their transcription, reducing the concentration of free RNAPs. This reduction in free RNAPs decreases the initiation rates of other genes since the probability of an RNAP binding to a promoter depends on the free RNAP concentration [14, 15, 17]. Nevertheless, ND RNAPs help buffer the free RNAP concentration reduction and minimize the unwanted impact of resource competition (see the schematic in Figure 1a). Similarly, when the expression of a group of genes is suppressed, ND RNAPs buffer the increase in the free RNAP concentration. In this way, ND RNAPs act like a pH buffer made of weak acid-base pairs, helping maintain the pH of an aqueous solution after adding an acid or base [20].

To assess the impact of ND RNAPs on gene expression, we use a mean-field model of transcription neglecting spatial information [14, 15]. RNAPs can bind to promoters and nonspecific binding sites on DNA, and the probabilities of RNAP binding to these sites depend on the concentration of free RNAPs through the Michaelis-Menten (MM) mechanism. Using both analytical and numerical calculations, we demonstrate that ND RNAPs act as buffers for changes in free RNAP concentration caused by the regulation of specific genes, such as genes of rRNAs. The buffering effects help suppress crosstalk between different genes, promoting robust and stable gene expression. In particular, the physiological parameters of *Escherichia coli* are in the parameter space where the buffering benefits of ND RNAPs are significant, suggesting that an appropriate amount of ND RNAPs may be under evolution selection. We further extend our mean-field model by including sigma factors such that RNAP can bind to promoters and initiate transcription only if they are bound by a sigma factor. Our conclusions on the buffering effects of ND RNAPs remain robust against a finite number of sigma factors.

To strengthen our hypothesis, we also develop an agent-based model that explicitly incorporates spatial information, including the diffusion of RNA polymerases and their interactions with multiple binding sites, including promoters and nonspecific binding sites. In this model, free RNAPs diffuse until they collide with a bind-

ing site. For an RNAP that just leaves a binding site, we relocate it to a random position within a finite distance from the binding site. The agent-based model provides further evidence supporting our predictions. Importantly, our conclusions regarding the buffering effects of ND RNAPs remain robust in the presence of spatial fluctuations in free RNAP concentration beyond the mean-field model.

ND RNAPs benefit the cell by minimizing unwanted crosstalk between genes; nevertheless, they also require the cell to spend additional resources to establish a reservoir of RNAPs. Therefore, we hypothesize that the number of nonspecific binding sites and the nonspecific binding affinity of RNAPs to DNA are under evolutionary selection. To test our hypothesis, we examine the number of nonspecific binding sites and the nonspecific binding affinity in *Escherichia coli* and find that they fall within a parameter space where the buffering benefits of ND RNAPs are significant, in concert with our theory. On the application side, our findings suggest a strategy for reducing the interference between the expressions of exogenous and endogenous genes in synthetic biology [21, 22] by adding non-coding DNA sequences to the bacterial chromosome to increase the number of nonspecific binding sites.

RESULTS

Mean-field model

In the mean-field model, the probability of a promoter bound by an RNAP is an MM function of the free RNAP concentration [14, 15, 17],

$$\mathrm{Pb}_i = \frac{[n]_{\mathrm{free}}}{[n]_{\mathrm{free}} + K_i}, \tag{1}$$

where $[n]_{\mathrm{free}}$ is the concentration of free RNAPs, and K_i is the MM constant of the promoter i. Regulation affects the expression level of a gene by changing the kinetic on and off rates of RNAP upon its promoters. In the mean-field model, these kinetic rate changes are modeled through the change in the MM constant K_i. The smaller K_i is, the stronger the promoter's binding affinity to RNAP. Once an RNAP binds to a promoter, it transitions to the elongating state with an initiation rate $k_{\mathrm{cat},i}$. The number of elongating RNAPs ($n_{\mathrm{el,i}}$) on the operon following the promoter is proportional to the probability of the promoter bound by an RNAP:

$$n_{\mathrm{el},i} = \mathrm{Pb}_i \Lambda_i, \tag{2}$$

where $\Lambda_i \equiv k_{\mathrm{cat},i} L_i / c_i$, the maximum possible number of RNAPs working on the operon, which we name as capacity in the following. Here, L_i is the length of the operon following the promoter i in base pairs, and c_i is the elongation speed of RNAP.

We categorize RNAPs into three types: (1) active RNAPs bound to promoters or engaged in transcription, (2) free RNAPs diffusing within the nucleoid, (3) and nonspecifically DNA-bound RNAPs [14–17]. For simplicity, we consider a bacterial cell with a fixed nucleoid volume V and a fixed total number of RNAPs n. This simplification is reasonable as the time required for RNAPs to transcribe an operon (around one minute) is much shorter than the typical doubling time of bacteria (around 30 minutes or longer) [23]. Therefore, RNAP allocation to the different categories can be considered as instantaneous.

In this work, we aim to explore the buffering effects of ND RNAPs such that specific genetic details are not critical to our conclusions. Therefore, we employ a coarse-grained approach by dividing the genome into two parts: promoters of rRNA genes (g_{r}) and promoters of mRNA genes (g_{m}), each with their respective MM constants (K_{r} and K_{m}) and capacities (Λ_{r} and Λ_{m}). By conserving the RNAP copy number, we obtain a self-consistent equation to determine the fraction of free RNAPs (f_{free}) as follows:

$$n_{\mathrm{t}}(1 - f_{\mathrm{free}}) = \underbrace{g_{\mathrm{r}} \frac{[n]_{\mathrm{t}} f_{\mathrm{free}}}{[n]_{\mathrm{t}} f_{\mathrm{free}} + K_{\mathrm{r}}} (1 + \Lambda_r)}_{\text{RNAPs on rRNA genes}} + \underbrace{g_{\mathrm{m}} \frac{[n]_{\mathrm{t}} f_{\mathrm{free}}}{[n]_{\mathrm{t}} f_{\mathrm{free}} + K_{\mathrm{m}}} (1 + \Lambda_m)}_{\text{RNAPs on mRNA genes}} + \underbrace{g_{\mathrm{ns}} \frac{[n]_{\mathrm{t}} f_{\mathrm{free}}}{[n]_{\mathrm{t}} f_{\mathrm{free}} + K_{\mathrm{ns}}}}_{\text{ND RNAPs}}. \tag{3}$$

Here, $[n]_{\mathrm{t}} = n_{\mathrm{t}}/V$ represents the total RNAP concentration, and $[n]_{\mathrm{free}} = [n]_{\mathrm{t}} f_{\mathrm{free}}$. We use the nucleoid volume to compute the concentration because RNAPs are almost entirely within the nucleoid [24–26]. On the right side of the above equation, the three terms represent the number of RNAPs binding to rRNA genes, mRNA genes, and nonspecific binding sites, respectively. g_{ns} refers to the number of nonspecific binding sites on DNA, and K_{ns} represents the MM constant of nonspecific binding.

Linear analysis of the mean-field model

To make our model biologically transparent, we adopt the estimated parameters of the model organism, *E. coli* with a doubling time of 30 minutes (see Table 1 with

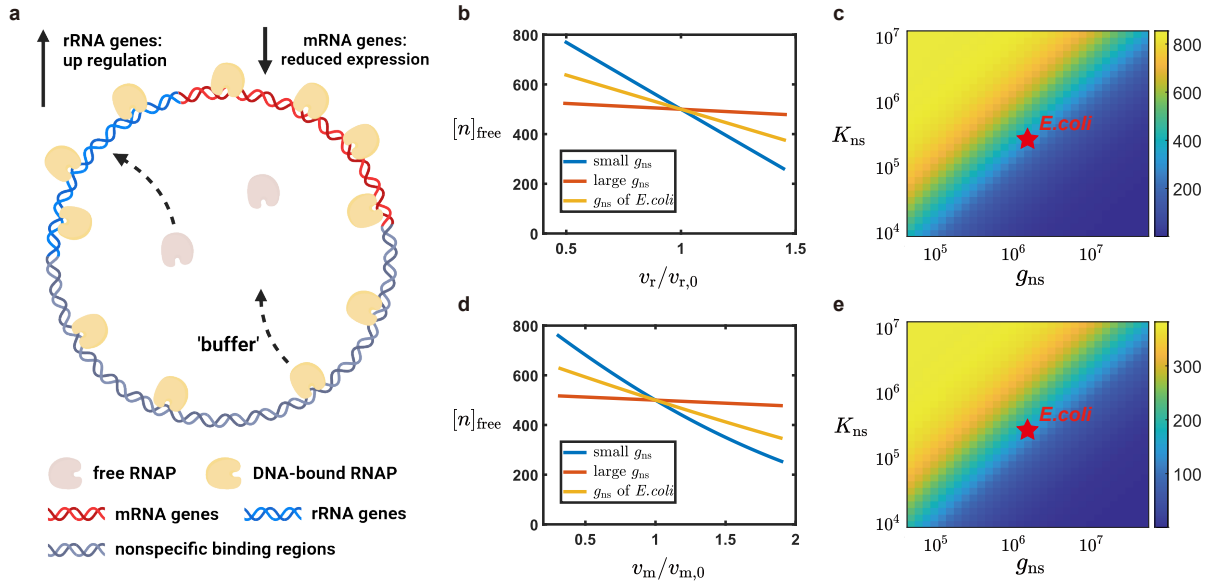

FIG. 1. **Nonspecifically DNA-bound RNAPs buffer the change of free RNAP concentration**. (a) The ND RNAPs replenish the pool of free RNAPs and buffer the decrease of free RNAP concentration upon the up-regulation of rRNA genes. Therefore, they reduce the competition between mRNA and rRNA for the resource of RNAPs. (b) The free RNAP concentration *vs.* the relative change of rRNA production rate $v_{\rm r}/v_{\rm r,0}$. For *E. coli*, $g_{\rm ns} = 1.5 \times 10^6$. The small $g_{\rm ns}$ line corresponds to $g_{\rm ns} = 1.5 \times 10^5$ and the large $g_{\rm ns}$ line corresponds to $g_{\rm ns} = 1.5 \times 10^7$. The more nonspecific binding sites, the less sensitive $[n]_{\rm free}$ is to the change of $v_{\rm r}$. (c) The sensitivity of free RNAP concentration to the relative changes in rRNA production rate $\partial[n]_{\rm free}/\partial(v_{\rm r}/v_{\rm r,0})$ as a function of $g_{\rm ns}$ and $K_{\rm ns}$. The star marks the parameters of *E. coli*. (d) The free RNAP concentration *vs.* the relative change of mRNA production rate $v_{\rm m}/v_{\rm m,0}$. The values of $g_{\rm ns}$ are same as (b). (e) The sensitivity of free RNAP concentration to the relative changes in mRNA production rate $\partial[n]_{\rm free}/\partial(v_{\rm m}/v_{\rm m,0})$ as a function of $g_{\rm ns}$ and $K_{\rm ns}$. The star marks the parameters of *E. coli*. In (b) and (d), we use the $K_{\rm ns}$ value of *E. coli*.

the detailed calculations and references included in the caption). As expected, the promoter of the mRNA genes is much weaker than that of rRNA genes, and the binding affinity of nonspecific binding sites is even weaker: $K_{\rm r} = 250\ \mu{\rm m}^{-3}$, $K_{\rm m} = 4.5 \times 10^4\ \mu{\rm m}^{-3}$, and $K_{\rm ns} = 2.6 \times 10^5\ \mu{\rm m}^{-3}$. Meanwhile, the free RNAP concentration in this condition is about $[n]_{\rm free} = 500\ \mu{\rm m}^{-3}$, suggesting that the rRNA promoter has a 2/3 probability of being occupied while the binding probability of mRNA promoter is much less than 1. In the following of this work, we set the parameters as the physiological values of *E. coli* (Table 1) unless otherwise mentioned.

We next study the impact of rRNA regulation on the free RNAP concentration by changing the MM constant of the rRNA promoter K_r from its unregulated value $K_{\rm r,0} = 500\ \mu{\rm m}^{-3}$. We define the production rate of rRNA as $v_r = g_r n_{\rm el,r} c_r$, namely, the number of RNAPs transcribing rRNA genes multiplied by the corresponding elongation speed. Interestingly, using the parameters of *E. coli*: $K_{\rm ns} = 4.5 \times 10^4\ \mu{\rm m}^{-3}$ and $g_{\rm ns} = 1.5 \times 10^6$ (Table 1), the free RNAP concentration is weakly dependent on the rRNA production rate (Figure 1b). If we take a smaller $g_{\rm ns}$ with the same $[n]_{\rm free}$ before regulation, $[n]_{\rm free}$ becomes much more sensitive to the rRNA production rate: an up-regulation of rRNA production significantly reduces the production rate of mRNA. In contrast, if we make the $g_{\rm ns}$ larger, $[n]_{\rm free}$ can be almost independent of the rRNA production rate.

To quantify this buffering effects, we define the sensitivity of $[n]_{\rm free}$ against the change in the rRNA production rate due to a small change in K_r as $\partial[n]_{\rm free}/\partial v_{\rm r}$ (with the same $[n]_{\rm free}$ before perturbation):

$$\frac{\partial[n]_{\rm free}}{\partial v_{\rm r}} = -\frac{\Lambda_{\rm r}+1}{c_r \Lambda_{\rm r}} \times \frac{1}{A + g_{\rm ns}\frac{K_{\rm ns}}{([n]_{\rm free}+K_{\rm ns})^2}}. \qquad (4)$$

Here, A is a constant independent of $g_{\rm ns}$ and $K_{\rm ns}$ (see details in Methods). Interestingly, Eq. (4) shows that increasing the number of nonspecific binding sites reduces the impact of rRNA production on the free RNAP concentration. We plot the sensitivity as a function of the number of nonspecific binding sites $g_{\rm ns}$ and the nonspecific binding affinity $K_{\rm ns}$ (Figure 1c). Remarkably, the corresponding parameters of *E. coli* sit in the crossover regime at which the buffering effects are significant.

We also apply the same analysis to the change in the mRNA production rate $v_m = g_m n_{\rm el,m} c_m$, i.e., the number of RNAPs transcribing mRNA genes multiplied by the elongation speed. The change of $[n]_{\rm free}$ due to a perturbation of the mRNA production rate $v_{\rm m}$ is similar to Eq. (4): one only needs to replace the lower index r by

TABLE 1. **Parameters for *E. coli* at doubling time 30 min.**
(1) There are two nucleoid lobes in a cell, segregated into two sublobes each. Each sublobe volume is 0.5 μm^3, so the total nucleoid volume at 30 min doubling time is $V = 2$ μm^3 [16].
(2) $[n]_{\rm free} = nf_{\rm free}/V$ where $n = 8400$ is the total number of RNAPs [27]. $f_{\rm free} = 0.12$ is the fraction of free RNAPs [16].
(3) The number of rRNA operons at 30 min doubling time is about 30 [15, 27].
(4) $g_{\rm m}$ is estimated by dividing 4500 genes in a genome by three genes per promoter, and multiplied by three genomes per cell [28, 29].
(5) $g_{\rm ns}$ is estimated as the genome size divided by the size of a single RNAP, which is about ten base pairs.
(6, 7) Λ_r and Λ_m are calculated using Eq. (2), where $k_{\rm cat,m} = 0.5$ s^{-1}, $L_m = 3000$ nt, $c_m = 52$ nt/s, and $k_{\rm cat,r} = 2$ s^{-1}, $L_r = 6500$ nt, $c_r = 85$ nt/s [15, 17]. Here, we approximate $\Lambda_r = 150$ and $\Lambda_m = 30$ for simplicity, which does not affect our results.
(8) $K_{\rm ns}$ is calculated using Eq. (3) from the experimentally estimated fraction of ND RNAPs $f_{\rm ns} = 0.34$ [16]. In Ref. [16], there is a small fraction of RNAPs unable to bind to DNA, which are likely immature RNAPs [15, 17]. We neglect these immature RNAPs for simplicity since they only account for a small part of RNAPs. To ensure that $f_{\rm r} + f_{\rm m} + f_{\rm free} + f_{\rm ns} = 1$, the $f_{\rm ns}$ we estimate is slightly larger than the reported value ($f_{\rm ns} = 0.28 \pm 0.05$).
(9, 10) $K_{\rm m}$ and $K_{\rm r}$ are calculated using Eq. (3) from the experimentally estimated fractions of RNAPs transcribing rRNA ($f_{\rm r}$) and mRNA ($f_{\rm m}$). Given the fraction of transcribing RNAPs (0.54) [16] and the ratio between RNAPs transcribing rRNA and mRNA at 30-min doubling time (2:1) [15], we obtain $f_{\rm r} = 0.36$ and $f_{\rm m} = 0.18$, from which we get $K_{\rm m} = 4.5 \times 10^4$ μm^{-3} and $K_{\rm r} = 250$ μm^{-3}.

Parameters	**Symbols**	**Values**
(1) Nucleoid volume	V	2 μm^3
(2) Free RNAP concentration	$[n]_{\rm free}$	500 μm^{-3}
(3) Number of rRNA promoters	$g_{\rm r}$	30
(4) Number of mRNA promoters	$g_{\rm m}$	4500
(5) Number of nonspecific binding sites	$g_{\rm ns}$	1.5×10^6
(6) Maximum number of transcribing RNAPs per rRNA operon	$\Lambda_{\rm r}$	150
(7) Maximum number of transcribing RNAPs per mRNA operon	$\Lambda_{\rm m}$	30
(8) MM constant for nonspecific binding	$K_{\rm ns}$	2.6×10^5 μm^{-3}
(9) MM constant of mRNA promoters	$K_{\rm m}$	4.5×10^4 μm^{-3}
(10) MM constant of rRNA promoters	$K_{\rm r}$	250 μm^{-3}

m. Interestingly, the conclusions from the regulation of rRNA production are equally applicable to the case of mRNA regulation (Figure 1d, e). Therefore, in the following, we mainly study the cases in which the expression of rRNA genes is regulated.

Numerical analysis of the mean-field model

Regulation on rRNA expression can be significant, e.g., the ratio between rRNA to total protein can vary dramatically as the environment changes, e.g., through the nutrient conditions and application of antibiotics [11], reflected by a large change in K_r in our mean-field model. To test whether our conclusions on the buffering effects of ND RNAPs remain valid given a significant perturbation to rRNA production, we go beyond the linear analysis and numerically calculate the mean-field model to investigate the impact of large perturbations to $K_{\rm r}$. We assess the changes in the free RNAP concentration and the mRNA production rate $v_{\rm m}$ after a ten-fold change of $K_{\rm r}$. We find that the ratio of free RNAP concentration after and before the regulation becomes closer to 1 as $g_{\rm ns}$ increases (Figure 2a, b), and the same for $v_{\rm m}$. We also change the MM constant for nonspecific binding and find that the buffering effects are more pronounced for smaller $K_{\rm ns}$ values (Figure 2c, d). Remarkably, the values of $g_{\rm ns}$ and $K_{\rm ns}$ in *E. coli* lie in the region of the S-shaped curves (dashed lines in Figure 2a-d) where the buffering effects are significant, in agreement with our linear analysis. This observation suggests that *E.coli* may have fine-tuned the number and the binding affinity of nonspecific binding sites to balance the benefits of buffering against the cost of producing extra RNAPs. Similar conclusions apply to the case of mRNA regulation (Figure S3).

We also study the buffering effects against a time-dependent random rRNA production rate to mimic a bacterial cell exposed to a fluctuating environment (see simulation details in Methods). We expect that ND RNAPs can also attenuate the noise propagation from rRNA production to mRNA production. To verify this idea, we introduce a time-dependent noise to the MM constant of rRNA genes, which leads to a stochastic rRNA production rate. Due to the finite resource of RNAPs, the mRNA production rate is also stochastic. We measure the coefficients of variation of the mRNA and rRNA production rates (CV$_{\rm m}$, CV$_{\rm r}$). Indeed, CV$_{\rm m}$

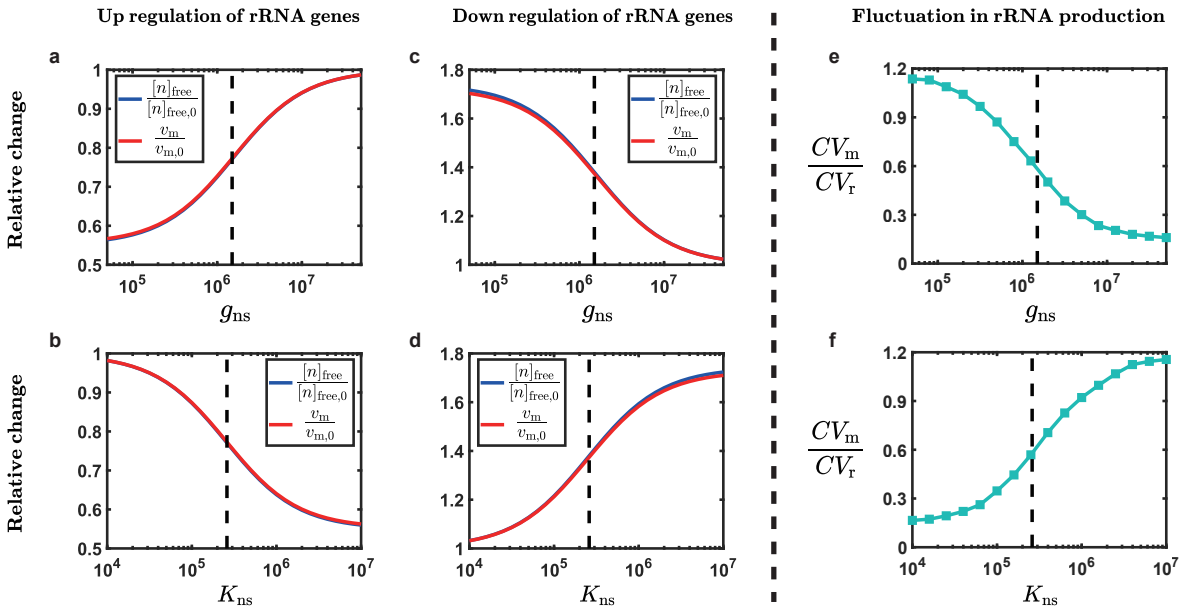

FIG. 2. **Predictions of the mean-field model and the locations of the parameters of** *E. coli* **in the parameter space.** (a-d) The ratio of free RNAP concentrations $[n]_{\rm free}$ and mRNA production rates $v_{\rm m}$ after and before the up/down-regulation of rRNA genes *vs.* $g_{\rm ns}$ and $K_{\rm ns}$. Up-regulation of rRNA expression reduces the free RNAP concentration and mRNA production rate, while down-regulation increases them. A large number of nonspecific binding sites (large $g_{\rm ns}$) and strong nonspecific binding affinity (small $K_{\rm ns}$) result in a relatively stable free RNAP concentration and mRNA production rate. Here, the MM-constant of rRNA promoters in up-regulation (down-regulation) is set to be $K_{\rm r}^{\rm up} = 0.1K_{\rm r}$ ($K_{\rm r}^{\rm down} = 10K_{\rm r}$). The black dashed lines represent the corresponding parameters of *E. coli*. The free RNAP concentration is fixed at $[n]_{\rm free} = 500\ \mu{\rm m}^{-3}$ before perturbation. (e-f) ND RNAPs attenuate the noise propagation from rRNA production rate to mRNA production rate. A small CV ratio indicates that the mRNA production rate is insensitive to the noise in the rRNA production rate.

is substantially smaller than $\mathrm{CV_r}$ when $g_{\rm ns}$ is sufficiently large and $K_{\rm ns}$ is sufficiently small (Figure 2e, f). Notably, the values of $g_{\rm ns}$ and $K_{\rm ns}$ of *E. coli* are again within the parameter range where the buffering against temporal noise is significant, showing the robustness of our results. We note that the initiation rate of the mRNA promoters used in the numerical analysis is from the measurement of $P_{\rm spc}$ promoter, which may not necessarily reflect the typical value of $k_{\rm cat,m}$ across the genome. To confirm the robustness of our conclusions, we also modify the value of $k_{\rm cat,m}$, and our findings remain valid (Supplementary Information and Figure S2).

Effects of sigma factors

In bacteria, sigma factors bind RNAPs and are required for RNAPs to recognize promoters. RNAPs without sigma factor binding are called core RNAPs, and RNAPs bound by a sigma factor are known as holoenzymes [30–32]. While both core RNAPs and holoenzymes nonspecifically bind to DNA, only holoenzymes specifically bind to promoters and initiate transcription. Meanwhile, anti-sigma factors bind to sigma factors and prevent their binding with RNAPs, inhibiting transcription [32–35]. Because the binding affinity between sigma factors and core RNAPs is strong (Table S2), if the number of sigma factors is more than the number of RNAPs in the absence of anti-sigma factors, virtually all core RNAPs are bound by sigma factors and our conclusions in the previous section remain the same.

Nevertheless, it is unclear whether the buffering effects of ND RNAPs are still valid if the number of sigma factors is comparable or smaller than the number of RNAPs. The presence of anti-sigma factors, which compete with core RNAPs for sigma factors, makes the conclusions even more unclear. To test the robustness of our results, we extend our model to include sigma and anti-sigma factors (Figure 3a and Methods). For simplicity, we focus on the common scenario where most genes are under the control of the same housekeeping sigma factors – σ^{70} [31, 36–38]. We assume that the sigma factor quickly dissociates from the holoenzyme early after transcription initiation, although some experiments show that a small fraction of sigma factors can be retained on an RNAP until termination [39, 40]. We use σ to represent the sigma factor, E to represent the core RNAP, and $E\sigma$ to represent the holoenzyme. The total RNAP as the sum of core RNAPs and holoenzyme is still represented by n.

Some experiments showed that the total number of sigma factor is more than the total RNAPs, while others

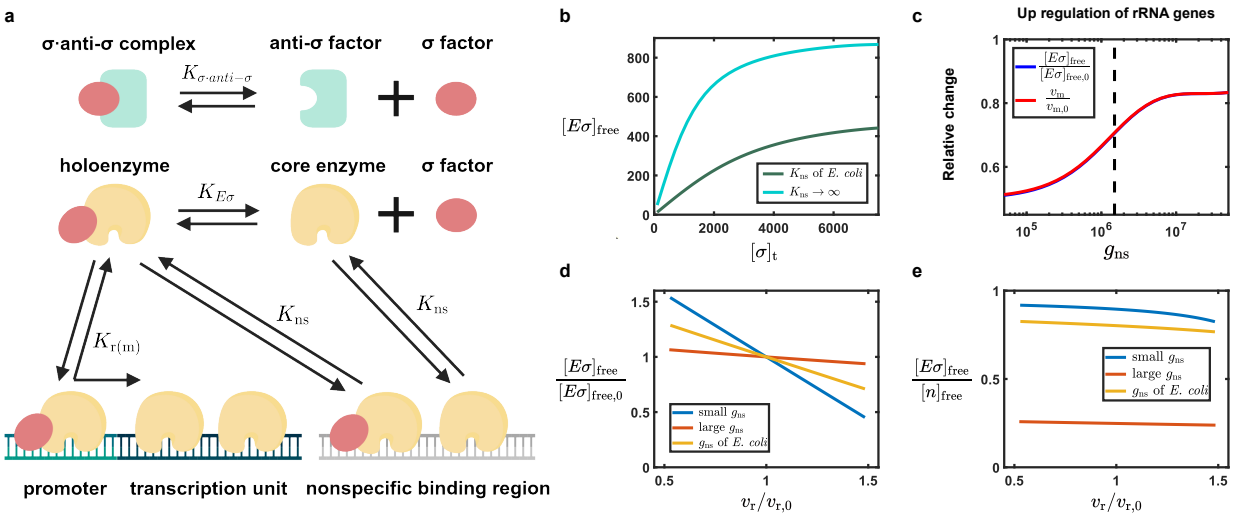

FIG. 3. **The buffering effects of ND RNAPs still hold given a finite number of sigma factors**. (a) The mean-field model with sigma and anti-sigma factors included. The dissociation constants or the MM constants for the corresponding reactions are shown. (b) The concentration of free holoenzymes *vs* the concentration of total sigma factors. $K \to \infty$ means that there is no nonspecific binding. (c) The relative change of free holoenzyme concentration and mRNA production rate after an up-regulation of rRNA genes ($K_{\rm r}^{\rm up} = 0.1K_{\rm r}$) under different numbers of nonspecific binding sites. The black dashed lines represent the corresponding parameters of *E. coli*. (d) The concentration of free holoenzyme *vs* the rRNA production rate given different numbers of nonspecific binding sites. Both the x and y coordinates are expressed in terms of relative changes. (e) The fraction of free holoenzymes in total free RNAPs is nearly constant as the rRNA production rate increases. In (c-e), we keep the free RNAP concentration before regulation to be the same across different g_{ns} values. In this figure, we use the parameters in Table 1 and Table S2 unless otherwise mentioned.

showed the opposite [37, 38, 41–43]. We first study the effects of the copy number of sigma factors on the concentration of free holoenzymes, $[E\sigma]_{\rm free}$, which directly sets the mRNA and rRNA production rates. We find that $[E\sigma]_{\rm free}$ increases with the concentration of total sigma factors, $[\sigma]_{\rm t}$ (Figure 3b). Without nonspecific binding ($K_{\rm ns} \to \infty$), the concentration of free holoenzyme is higher. However, the relative change of $[E\sigma]_{\rm free}$ with nonspecific binding has a larger range: the ratio between the finite $K_{\rm ns}$ and infinite $K_{\rm ns}$ lines increases as $[\sigma]_{\rm t}$ increases in Figure 3b. Intuitively, this is because ND RNAPs replenish the pool of free holoenzymes as transcriptional activity becomes more significant due to the increasing number of sigma factors. In the following, we set the total number of sigma factors to be 10000, slightly larger than the number of total RNAPs at a doubling time of 30 min, and the total number of anti-sigma factors to be about 1.4 times RNAP (see detailed parameters in Table S2).

We next calculate the relative changes of $[E\sigma]_{\rm free}$ and mRNA production rate after an up-regulation of rRNA genes (Figure 3c and Figure S4b) and a down-regulation of rRNA genes (Figure S4a, c). In both cases, we observe robust buffering effects due to the finite number of nonspecific binding sites. In particular, the parameter of *E. coli* is in the regime where the buffering effects are significant. We notice that in the case of up-regulation of rRNA genes, the ratio of free holoenzyme concentrations before and after regulation is smaller than one even when $g_{\rm ns}$ is very large, although the ratio of free RNAP concentration change approaches one in the limit of $g_{\rm ns} \to \infty$ (Figure S4d). This is because the dissociation constant for the binding between core RNAPs and sigma factors, $K_{E\sigma}$ increases as the rRNA production rate increases due to a reduction of $K_{\rm r}$ since transcription activity speeds up the dissociation of holoenzyme (Eq. (18) and Figure S4h). A faster dissociation of holoenzymes to core RNAPs and sigma factors leads to a smaller fraction of free holoenzymes in total free RNAPs, which cannot be avoided even when $g_{\rm ns}$ is large. The relative changes of total free RNAP concentration $[n]_{\rm free}$ and free core enzyme concentration $[E]_{\rm free}$ after an up-regulation or a down regulation of rRNA genes under different $g_{\rm ns}$ and $K_{\rm ns}$ are shown in Figure S4d-g.

We also plot $[E\sigma]_{\rm free}$ *vs.* the rRNA production rate as we continuously change the MM constant of rRNA genes (Figure 3d) and again observe significant buffering effects on the concentration of free holoenzymes similar to the previous results (Figure 1b). ($[n]_{\rm free}$ *vs.* the rRNA production rate is shown in Figure S4i.) Since the mean-field model without including sigma and anti-sigma factors explicitly corresponds to the limit of infinite sigma factors without anti-sigma factors, these observations support that the buffering effects of ND RNAPs are robust against a finite number of sigma factors and anti-sigma factors. To rationalize our observations, we notice that

the fraction of free holoenzymes in the total free RNAPs is nearly constant (Figure 3e). Based on this result, we define α as $[E\sigma]_{\rm free}/[n]_{\rm free}$ and rewrite Eq.(3) as

$$\begin{aligned} n_{\rm t}(1-f_{\rm free}) =& g_{\rm ns}\frac{[n]_{\rm t}f_{\rm free}}{[n]_{\rm t}f_{\rm free}+K_{\rm ns}} \\ &+g_{\rm r}\frac{[n]_{\rm t}f_{\rm free}}{[n]_{\rm t}f_{\rm free}+K_{\rm r}/\alpha}(1+\Lambda_r) \\ &+g_{\rm m}\frac{[n]_{\rm t}f_{\rm free}}{[n]_{\rm t}f_{\rm free}+K_{\rm m}/\alpha}(1+\Lambda_m). \end{aligned} \tag{5}$$

Eq. (5) is identical to Eq.(3) for the model neglecting sigma factors, except that the effective MM constants of rRNA and mRNA promoters become larger since $\alpha < 1$. Therefore, our previous theoretical conclusions for the mean-field model without sigma factors, e.g., Eq. (4), are approximately valid for the extended model as long as we include the approximately constant ratio between $[E\sigma]_{\rm free}$ and $[n]_{\rm free}$. In Methods, we include a detailed discussion rationalizing why $[E\sigma]_{\rm free}/[n]_{\rm free}$ is nearly constant. The main idea is that $[E\sigma]_{\rm free}/[n]_{\rm free}$ is approximately a MM function of the concentration of sigma factors not bound to DNA, $[\sigma]'_{\rm t}$ with an effective MM constant $K'_{E\sigma}$ (Eq. (26) and Figure S4l). As the rRNA production rate increases, the concentration of free holoenzyme $[E\sigma]_{\rm free}$ decreases (Figure 3d); in the meantime, $[\sigma]'_{\rm t}$ increases (Figure S4j) and $K'_{E\sigma}$ increases (Figure S4k), leading to a relatively constant $[E\sigma]_{\rm free}/[n]_{\rm free}$.

To summarize, the free RNAP concentration $[n]_{\rm free}$ can always be buffered by a big reservoir of ND RNAPs (Eq. (5)). Since only free holoenzymes can initiate transcription, to minimize the crosstalk between different genes, it is important that the ratio between $[E\sigma]_{\rm free}$ and $[n]_{\rm free}$ is stable as well. In typical biological scenarios where the numbers of sigma factors and anti-sigma factors are comparable to the total RNAP number, and the binding affinity between sigma factors and core RNAPs is strong (see Table S2 for details about the parameters), we find that $[E\sigma]_{\rm free}/[n]_{\rm free}$ is indeed relatively stable, which means the simplified mean-field model neglecting sigma factors is valid. In the following, we ignore sigma factors for simplicity in the agent-based model.

Agent-based simulations

The mean-field model assumes that once an RNAP leaves a promoter or nonspecific binding site, it enters the well-mixed pool of free RNAPs. However, this model does not capture that the free RNAP concentration is spatially heterogeneous, and RNAPs that have just left a binding site are more likely to rebind the same site. To test whether the buffering effects of ND RNAP are still valid in the presence of the spatial effect, we develop an agent-based model incorporating spatial information (Figure 4a). This model represents the nucleoid as a three-dimensional space with multiple binding sites corresponding to promoters and nonspecific binding sites. We model RNAPs as point particles, diffusing in the nucleoid with a diffusion constant D until they hit any free binding site. If RNAPs enter a promoter, they will either leave with an off-rate of $k_{\rm off,i}$ or transition to the elongation state with a rate of $k_{\rm cat,i}$ (Figure 4b), where i is either "r" or "m", depending on whether the promoter is followed by rRNA or mRNA genes [45]. An elongating RNAP becomes free after a fixed duration, which is the time for an RNAP to finish transcribing an operon, i.e., $\Lambda_{\rm m}/k_{\rm cat,m}$ for mRNA genes (Table 1, 2). If RNAPs enter a free nonspecific binding site, they will leave with an off-rate of $k_{\rm off,ns}$. Once a binding site is occupied, it cannot bind another RNAP until the RNAP hops off or transitions to the elongation state.

We represent the binding sites as spheres with radii $a_{\rm s}$ for promoters and $a_{\rm ns}$ for nonspecific binding sites. Once an RNAP leaves a binding site, we relocate it randomly inside a shell with inner diameter a and outer diameter $R = \gamma \times a$ (Figure 4c), where a is $a_{\rm s}$ for mRNA and rRNA promoter, or $a_{\rm ns}$ for nonspecific binding. Here, $\gamma > 1$ is the relocation parameter. In this work, we simplify the problem by assuming random and fixed positions of binding sites.

In the agent-based model, the free RNAP concentration $[n]_{\rm free}(r,t)$ becomes spatially and temporally dependent. We define $\overline{[n]}_{\rm free}$ as the free RNAP concentration averaged over the space not occupied by binding sites after the system reaches the steady state. The agent-based model can be mapped to the mean-field model by replacing $[n]_{\rm free}$ in Eq. (1) with the averaged free RNAP concentration $\overline{[n]}_{\rm free}$. By balancing the influx and outflux of a binding site, we have

$$k_{\rm on}\overline{[n]}_{\rm free}(1-{\rm Pb}_i) = (k_{{\rm off},i}+k_{{\rm cat},i}){\rm Pb}_i. \tag{6}$$

Here, $k_{\rm on}$ is the on rate, which is the flux entering the binding site per unit free RNAP concentration. Comparing Eq. (6) with Eq. (1), we obtain the relationship between the binding affinity K_i in the mean-field model and the kinetic parameters $k_{{\rm off},i}$, $k_{{\rm cat},i}$ and $k_{\rm on}$ in the agent-based model:

$$K_i = \frac{k_{{\rm off},i}+k_{{\rm cat},i}}{k_{\rm on,s}}. \tag{7}$$

Similarly, the MM constant of nonspecific binding is

$$K_{\rm ns} = \frac{k_{\rm off,ns}}{k_{\rm on,ns}}. \tag{8}$$

We also derive the analytical expression of $k_{\rm on}$ for the agent-based model (Methods):

$$k_{\rm on} = \frac{2(\gamma^3-1)}{2\gamma^3-3\gamma^2+1}4\pi Da \tag{9}$$

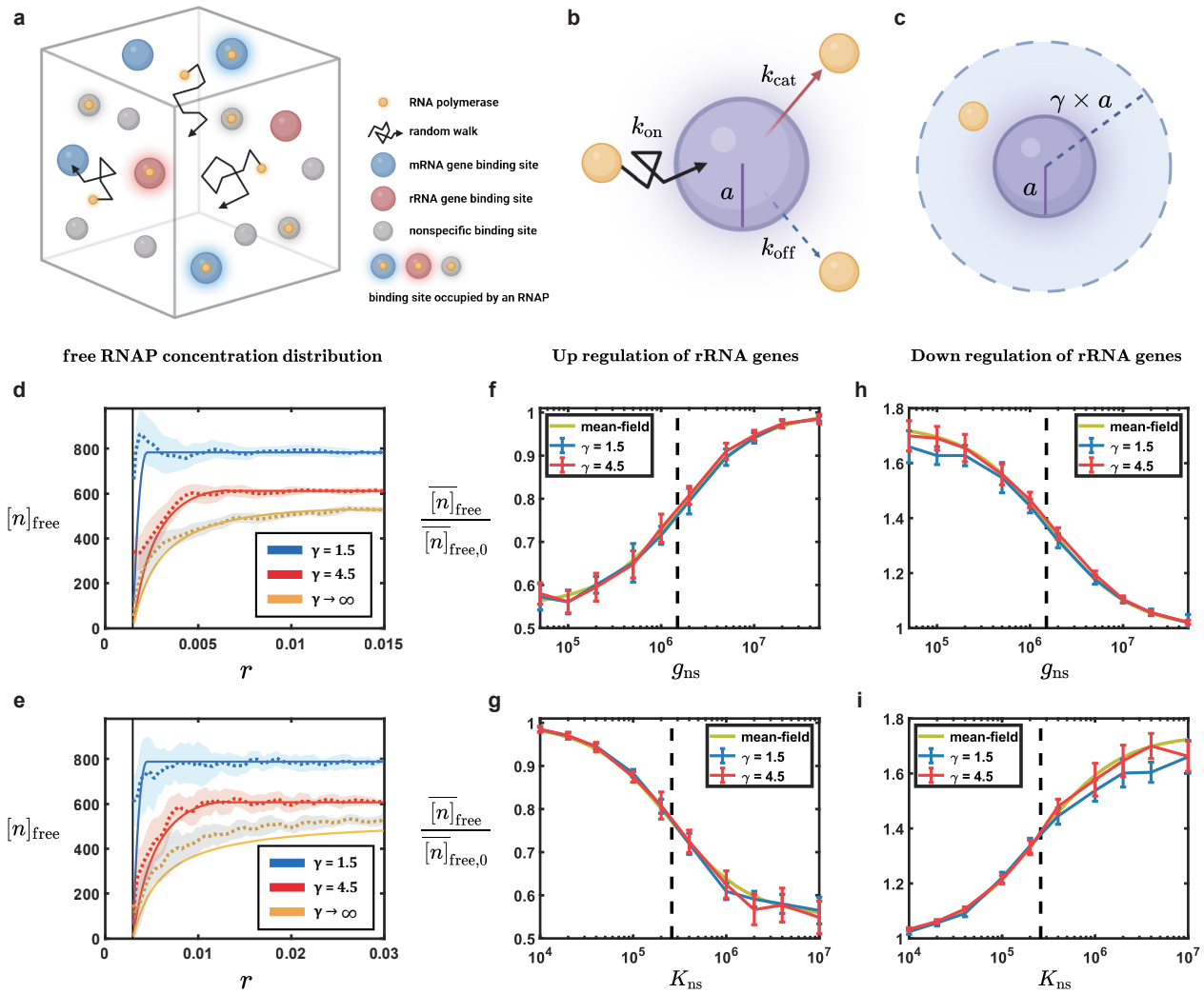

FIG. 4. **Simulations of the agent-based model.** (a) A schematic of the agent-based model of bacterial transcription. We model the three-dimensional nucleoid with multiple binding sites corresponding to promoters and nonspecific binding sites. RNAPs diffuse until they hit any free-binding site. (b) If an RNAP hits a promoter, it will start transcription with a rate $k_{\rm cat,i}$ or hop off the promoter with a rate $k_{\rm off,i}$. Here, i can be "r" or "m", depending on whether rRNA or mRNA genes follow the promoter. If an RNAP hits a nonspecific binding site, it will only hop off with a rate $k_{\rm off,ns}$. (c) We relocate an RNAP that just left a binding site inside a shell with inner diameter a and outer diameter $R = \gamma \times a$. (d-e) The free RNAP concentration vs. the distance to a mRNA promoter (d) and the distance to a nonspecific binding site (e). The dashed lines are averaged over binding sites of the same type. The solid lines are the theoretical predictions given the measured asymptotic free RNAP concentration far from the binding site (Eq. (28)). All the results here are before the regulation of rRNA genes. The shaded area represents the standard deviation. Here, $g_{\rm ns} = 5 \times 10^4$ and $K_{\rm ns} = 2.6 \times 10^5$ μm^{-3}. (f-i) The ratio of the averaged free RNAP concentration after and before the up-regulation (f, g) or down-regulation (h, i) of rRNA genes *vs.* the number of nonspecific binding sites $g_{\rm ns}$ and the MM constant of nonspecific binding $K_{\rm ns}$. The black dashed lines represent the estimated parameters of *E. coli*. In this figure, we use the parameters in Table 1, 2 unless otherwise mentioned. We change $K_{\rm ns}$ by changing the off-rate $k_{\rm off,ns}$. Data are shown as mean ± s.d.

Here $a = a_{\rm s}$ or $a_{\rm ns}$ depending on the nature of the binding site. In the limit of $\gamma \to \infty$, the entering flux reduces to the classical result of diffusion-limited on-rate: $J = 4\pi D a \overline{[n]}_{\rm free}$ [46]. In the agent-based model simulations, $\gamma \to \infty$ corresponds to a random relocation of RNAPs in the whole system.

Given the MM constants of the mean-field model, $K_{\rm m}$, $K_{\rm r}$, along with other parameters including γ and the initiation rates of rRNA and mRNA promoters (Table 2), we infer the values of $k_{\rm off,m}$ and $k_{\rm off,r}$ so that the mean-field model and the agent-based model are one-to-one correspondence. We also adjust the total number of RNAPs so that the agent-based model has the same $\overline{[n]}_{\rm free}$ before regulation as the mean-field model according to Eq. (3).

TABLE 2. **Parameters of agent-based simulations.**
(1) $k_{\mathrm{cat,r}}$ is the initiation rate of rRNA promoter, which is the rate for a promoter-bound RNAP to start transcribing [17, 44].
(2) $k_{\mathrm{cat,m}}$ is the initiation rate of mRNA promoter [17, 44].
(3) The radius of nonspecific binding sites is set to be comparable to the width of double-stranded DNA.
(4) The radius of specific binding sites, i.e., promoters, is set to be larger than the radius of nonspecific binding sites.
(5) $D = 2.6\ \mu\mathrm{m}^2\mathrm{s}^{-1}$ is the diffusion constant of free RNAPs [26].

Parameters	**Symbols**	**Values**
(1) Initiation rate of rRNA promoter-bound RNAPs	$k_{\mathrm{cat,r}}$	$2\ \mathrm{s}^{-1}$
(2) Initiation rate of mRNA promoter-bound RNAPs	$k_{\mathrm{cat,m}}$	$0.5\ \mathrm{s}^{-1}$
(3) Radius of nonspecific binding sites	a_{ns}	$1.5 \times 10^{-3}\ \mu\mathrm{m}$
(4) Radius of specific binding sites	a_s	$3 \times 10^{-3}\ \mu\mathrm{m}$
(5) Diffusion constant of free RNAPs	D	$2.6\ \mu\mathrm{m}^2\mathrm{s}^{-1}$

We apply up-regulation and down-regulation to rRNA promoters by changing the off rate $k_{\mathrm{off,r}}$ and study how the free RNAP concentration relative to its original value before regulation changes (see simulation details in Methods).

We find that the simulated free RNAP concentration (averaged over binding sites) vs. the distance to the binding site center can be well fitted by our analytical predictions for both finite and infinite γ values (Figure 4d-e). Nevertheless, the free RNAP concentration far from a binding site deviates from the preassigned value (which is 500 $\mu\mathrm{m}^{-3}$), particularly for small γ values. We think this deviation is due to the finite time interval of the numerical simulations. We confirm this idea by simulating a simplified model with only a single binding site with a given distribution of free RNAPs based on Eqs. (30). We compute the simulated on rate $k_{\mathrm{on}}^{\mathrm{sim}}$ and find that it is smaller than the theoretical prediction [Eq. (9)]. The difference becomes more significant as the time interval increases (Supplementary Information and Figure S1). We also compare the free RNAP concentration distributions for the full agent-based simulations under different time intervals. As the time interval decreases, the simulated curves get closer to the predicted curves for both finite γ and infinitely large γ (Supplementary Information and Figure S1).

Intriguingly, the agent-based simulations agree well with the mean-field predictions regarding the relative changes of the averaged free RNAP concentration (Figure 4f-i) and the relative changes in the mRNA production rate (Figure S5), despite the deviations of the free RNAP concentrations. Therefore, we conclude that the buffering effects of ND RNAPs are equally applicable to the agent-based model and robust against spatial fluctuation of free RNAP concentration.

DISCUSSION

While the abundance of nonspecifically DNA-bound RNAPs in bacteria is well known, the biological functions of these idle RNAPs are unclear. In this work, we demonstrate that the nonspecifically DNA-bound RNAPs are natural buffers to mitigate the unwanted crosstalk between the transcription of rRNA and mRNA. This evolutionary advantage is crucial for bacteria because rRNA and mRNA use the same type of RNAP. Interestingly, the number of nonspecific binding sites and the nonspecific binding affinity of *E. coli* lead to significant buffering effects. In particular, the number of nonspecific binding sites is much more than that of promoters, and the nonspecific binding affinity is much weaker than those of promoters, similar to pH buffers made of weak acid-base pairs.

Meanwhile, producing ND RNAPs also has a cost because cells have to make more RNAPs to keep the same free RNAP concentration – the total number of RNAPs increases linearly with g_{ns} according to Eq. (3). We propose that the benefit of weaker interference between genes and the cost of making extra RNAPs together set the number of nonspecific binding sites and their binding affinity through evolution. We cannot exclude the possibility that the origin of nonspecific binding may result from passive interaction between RNAP and DNA. But still, whether it is passive or selected actively, one cannot ignore the function of nonspecifically DNA-bound RNA polymerases as a buffer of resource competition.

In growing cells, the number of σ^{70} is significantly larger than other alternative types of sigma factors in *E. coli* [37, 38, 47], and most genes are controlled by σ^{70}. However, in the presence of multiple types of sigma factors [37, 48, 49], different types of sigma factors can compete for the pool of core RNAP such that the up-regulation of one group of genes can result in reduced expressions of genes controlled by other types of sigma factors, which is called sigma competition. It has been suggested that ND RNAPs have mild effects on mitigating the crosstalk between different groups of genes due to sigma competition [47]. This result does not contradict the buffering effects between genes controlled by the same type of sigma factor we propose in this work. We

demonstrate that ND RNAPs help to reduce the crosstalk between genes regulated by the same type of sigma factors, i.e., σ^{70}, which is an important scenario in *E. coli*. In particular, we show that the buffering effects of ND RNAPs are robust against a finite number of sigma factors. Meanwhile, cells can use alternative sigma factors to up-regulate some genes while down-regulating others during the global switch of the gene expression program.

Our model neglects the spatial correlation of promoter locations while spatial and temporal organizations of the nucleoid have been observed [18, 50–52]. In the future, it will be interesting to study how the nucleoid structure influences our conclusions. Furthermore, it is also essential to verify whether the buffering effects due to nonspecifically DNA-bound RNAPs apply to other bacteria. On the application side, our work provides a new avenue to reduce the burden effects of synthetic circuits on bacteria by adding non-coding DNA to the chromosome to increase the number of nonspecific binding sites.

METHODS

Details of the linear analysis of the mean-field model

Directly differentiating Eq.(3) with respect to K_r, we get

$$\begin{aligned} -V\frac{\partial[n]_\mathrm{free}}{\partial K_\mathrm{r}} &= -g_\mathrm{r}(1+\Lambda_r)\frac{[n]_\mathrm{free}}{([n]_\mathrm{free}+K_\mathrm{r})^2} \\ &+ g_\mathrm{r}(1+\Lambda_r)\frac{K_\mathrm{r}}{([n]_\mathrm{free}+K_\mathrm{r})^2}\frac{\partial[n]_\mathrm{free}}{\partial K_\mathrm{r}} \\ &+ g_\mathrm{m}(1+\Lambda_m)\frac{K_\mathrm{m}}{([n]_\mathrm{free}+K_\mathrm{m})^2}\frac{\partial[n]_\mathrm{free}}{\partial K_\mathrm{r}} \\ &+ g_\mathrm{ns}\frac{K_\mathrm{ns}}{([n]_\mathrm{free}+K_\mathrm{ns})^2}\frac{\partial[n]_\mathrm{free}}{\partial K_\mathrm{r}}. \end{aligned} \tag{10}$$

Rearranging the above equation leads to

$$\frac{\partial[n]_\mathrm{free}}{\partial K_\mathrm{r}} = \frac{g_\mathrm{r}(1+\Lambda_r)\frac{[n]_\mathrm{free}}{([n]_\mathrm{free}+K_\mathrm{r})^2}}{A+g_\mathrm{r}(1+\Lambda_r)\frac{K_\mathrm{r}}{([n]_\mathrm{free}+K_\mathrm{r})^2}+g_\mathrm{ns}\frac{K_\mathrm{ns}}{([n]_\mathrm{free}+K_\mathrm{ns})^2}}, \tag{11}$$

where

$$A = V + g_\mathrm{m}(1+\Lambda_m)\frac{K_\mathrm{m}}{([n]_\mathrm{free}+K_\mathrm{m})^2}. \tag{12}$$

We then differentiate the rRNA production rate with respect to $[n]_\mathrm{free}$,

$$v_\mathrm{r} = c_\mathrm{r} g_\mathrm{r}\Lambda_\mathrm{r}\frac{[n]_\mathrm{free}}{[n]_\mathrm{free}+K_\mathrm{r}}, \tag{13}$$

so that

$$\frac{\partial v_\mathrm{r}}{\partial[n]_\mathrm{free}} = \frac{c_\mathrm{r} g_\mathrm{r}\Lambda_\mathrm{r}}{([n]_\mathrm{free}+K_\mathrm{r})^2}(K_\mathrm{r}-[n]_\mathrm{free}\frac{\partial K_\mathrm{r}}{\partial[n]_\mathrm{free}}). \tag{14}$$

Finally we substitute $\frac{\partial K_\mathrm{r}}{\partial[n]_\mathrm{free}}$ using Eq. (11), and obtain Eq. (4).

Simulations of time-dependent noise

We let K_r perform a random walk with the initial value $K_\mathrm{r}^0 = 250$ such that

$$K_\mathrm{r}(t+\Delta t) = K_\mathrm{r}(t) - 0.05(K_\mathrm{r}(t)-K_\mathrm{r}^0)\Delta t + \xi\sqrt{2D_\mathrm{r}\Delta t} \tag{15}$$

Here, the diffusion coefficient $D_\mathrm{r} = 100$. ξ is a Gaussian random variable with zero mean and unit variance. We run this for 1000 steps with each step $\Delta t = 1$ and repeat it 10 times. At each step, we calculate the free RNAP concentration $[n]_\mathrm{free}(t)$ using Eq.(3). Then we calculate $v_\mathrm{m}(t) = c_\mathrm{m} g_\mathrm{m}\Lambda_\mathrm{m}\frac{[n]_\mathrm{free}}{[n]_\mathrm{free}+K_\mathrm{m}}$ and $v_\mathrm{r}(t) = c_\mathrm{r} g_\mathrm{r}\Lambda_\mathrm{r}\frac{[n]_\mathrm{free}}{[n]_\mathrm{free}+K_\mathrm{r}(t)}$. We pool data from all time steps in repeated simulations, and calculate the means and standard deviations of v_m and v_r.

Mean-field model including sigma factors

The following equation describes the equilibrium condition of core RNAP-sigma binding:

$$k_{fE\sigma}[E]_\mathrm{free}[\sigma]_\mathrm{free} = k_{bE\sigma}[E\sigma]_\mathrm{free} + J_\mathrm{cat}. \tag{16}$$

Here, $[\sigma]_\mathrm{free}$ is the concentration of free sigma factors unbound by anti-sigma factors or core RNAPs. $[E]_\mathrm{free}$ is the concentration of free core RNAPs, and $[E\sigma]_\mathrm{free}$ is the concentration of free holoenzymes. $k_{fE\sigma}$ and $k_{bE\sigma}$ are the forward and backward rates of holoenzyme formation. J_cat is the transcription flux, which comes from the dissociation of holoenzymes into core RNAPs and sigma factors during elongation [39, 40, 47]:

$$J_\mathrm{cat} = k_\mathrm{cat,r}[g_r]\frac{[E\sigma]_\mathrm{free}}{[E\sigma]_\mathrm{free}+K_\mathrm{r}} + k_\mathrm{cat,m}[g_m]\frac{[E\sigma]_\mathrm{free}}{[E\sigma]_\mathrm{free}+K_\mathrm{m}} \tag{17}$$

where $[g_r]$ and $[g_m]$ are the concentrations of rRNA and mRNA promoters, respectively. Therefore, the dissociation constant of core RNAP-sigma binding becomes

$$\begin{aligned} K_{E\sigma} = \frac{[E]_\mathrm{free}[\sigma]_\mathrm{free}}{[E\sigma]_\mathrm{free}} &= \frac{k_{bE\sigma}}{k_{fE\sigma}} + \frac{k_\mathrm{cat,r}}{k_{fE\sigma}}\frac{[g_r]}{[E\sigma]_\mathrm{free}+K_\mathrm{r}} \\ &+ \frac{k_\mathrm{cat,m}}{k_{fE\sigma}}\frac{[g_m]}{[E\sigma]_\mathrm{free}+K_\mathrm{m}}. \end{aligned} \tag{18}$$

Here, the last two terms on the right side of Eq. (18) make $K_{E\sigma}$ larger than $k_{bE\sigma}/k_{fE\sigma}$.

The equilibrium equation of sigma and anti-sigma factor binding is

$$[\sigma]_\mathrm{free}[Anti]_\mathrm{free} = K_{\sigma\cdot\mathrm{anti}-\sigma}[\sigma\cdot Anti]. \tag{19}$$

Here, $[Anti]_\mathrm{free}$ is the concentration of free anti-sigma factors, and $[\sigma\cdot Anti]$ is the concentration of sigma-anti-sigma complexes. $K_{\sigma\cdot\mathrm{anti}-\sigma}$ is the corresponding dissociation constant.

The concentrations of free core RNAPs and free holoenzymes add up to the concentration of all free RNAPs, $[n]_{\text{free}}$,

$$[E]_{\text{free}} + [E\sigma]_{\text{free}} = [n]_{\text{free}}. \tag{20}$$

The concentrations of free anti-sigma factors and sigma-anti-sigma complexes add up to the concentration of total anti-sigma factors,

$$[Anti]_{\text{free}} + [\sigma \cdot Anti] = [Anti]_{\text{t}}. \tag{21}$$

The concentrations of free sigma factors, sigma-anti-sigma complexes, free holoenzymes, and holoenzymes bound to DNA add up to the concentration of total sigma factors,

$$[\sigma]_{\text{free}} + [\sigma \cdot Anti] + [E\sigma]_{\text{free}} + [E\sigma]_{\text{bound}} = [\sigma]_{\text{t}} \tag{22}$$

where $[\sigma]_{\text{t}}$ is the concentration of total sigma factors. $[E\sigma]_{\text{bound}}$ is the concentration of holoenzymes bound to DNA (both specifically and nonspecifically):

$$\begin{aligned}[E\sigma]_{\text{bound}} =& [g_{\text{r}}]\frac{[E\sigma]_{\text{free}}}{[E\sigma]_{\text{free}} + K_{\text{r}}} + [g_{\text{m}}]\frac{[E\sigma]_{\text{free}}}{[E\sigma]_{\text{free}} + K_{\text{m}}} \\ &+ [g_{\text{ns}}]\frac{[E\sigma]_{\text{free}}}{[n]_{\text{free}} + K_{\text{ns}}}\end{aligned} \tag{23}$$

Here, we assume the core RNAPs and holoenzymes have equal nonspecific binding affinities with DNA, i.e., K_{ns}.

Since only holoenzymes can specifically bind to promoters, we rewrite Eq.(3) as

$$\begin{aligned}n_{\text{t}} - [n]_{\text{free}}V =& g_{\text{ns}}\frac{[n]_{\text{free}}}{[n]_{\text{free}} + K_{\text{ns}}} \\ &+ g_{\text{r}}\frac{[E\sigma]_{\text{free}}}{[E\sigma]_{\text{free}} + K_{\text{r}}}(1 + \Lambda_r) \\ &+ g_{\text{m}}\frac{[E\sigma]_{\text{free}}}{[E\sigma]_{\text{free}} + K_{\text{m}}}(1 + \Lambda_m).\end{aligned} \tag{24}$$

In the following, we rationalize the observation that the fraction of free holoenzymes in total free RNAPs is nearly constant as the rRNA production rate increases (Figure 3e). Using Eqs. (18-22), we find the relationship between $[E\sigma]_{\text{free}}$ and $[n]_{\text{free}}$:

$$\begin{aligned}[E\sigma]_{\text{free}} = \frac{1}{2}\Big([\sigma]'_{\text{t}} + K'_{E\sigma} + [n]_{\text{free}} \\ - \sqrt{([\sigma]'_{\text{t}} + K'_{E\sigma} + [n]_{\text{free}})^2 - 4[\sigma]'_{\text{t}}[n]_{\text{free}}}\Big).\end{aligned} \tag{25}$$

Here, $[\sigma]'_{\text{t}} = [\sigma]_{\text{t}} - [E\sigma]_{\text{bound}}$ is the concentration of sigma factors not bound to DNA. $K'_{E\sigma} = (1 + \frac{[Anti]_{\text{t}}}{K_{\sigma\cdot\text{anti}-\sigma} + [\sigma]_{\text{free}}})K_{E\sigma}$, where $\frac{[Anti]_{\text{t}}}{K_{\sigma\cdot\text{anti}-\sigma} + [\sigma]_{\text{free}}}$ is equal to $[\sigma \cdot Anti]/[\sigma]_{\text{free}}$, such that anti-sigma factors make the dissociation constant $K_{E\sigma}$ effectively larger. Note that $[\sigma]'_{\text{t}}$ and $K'_{E\sigma}$ are functions of $[E\sigma]_{\text{free}}$ as well. In the biological scenario where $[n]_{\text{free}} \ll ([\sigma]'_{\text{t}} + K'_{E\sigma})$, the concentration of free holoenzyme is simplified as

$$[E\sigma]_{\text{free}} = [n]_{\text{free}}\frac{[\sigma]'_{\text{t}}}{[\sigma]'_{\text{t}} + K'_{E\sigma}}. \tag{26}$$

In Figure 3e, the concentration of free holoenzymes decreases as we increase the rRNA production rate, leading to an overall increased $[\sigma]'_{\text{t}}$ (Figure S4j). Meanwhile, the reduction of $[E\sigma]_{\text{free}}$ increases the effective dissociation constant $K'_{E\sigma}$ (Eq. (18) and Figure S4k). Interestingly, the synergistic changes of $[\sigma]'_{\text{t}}$ and $K'_{E\sigma}$ lead to a relatively constant $[E\sigma]_{\text{free}}/[n]_{\text{free}}$ (Figure 3e and see Figure S4l for predictions based on Eq. (26)).

Derivations of the free RNAP concentration distribution and k_{on} in the agent-based model

Here, we consider that RNAPs are randomly relocated in a shell with inner diameter a and outer diameter $R = \gamma \times a$, where a could be a_{s} for mRNA and rRNA promoter, or a_{ns} for nonspecific binding site. Let $[n]_{\text{free}}(r, t)$ be the free RNAP concentration distribution and it satisfies

$$\frac{\partial [n]_{\text{free}}}{\partial t} = D\nabla^2[n]_{\text{free}} + j = 0 \tag{27}$$

where j is the RNAP flux per unit volume from RNAPs randomly relocated near the binding site. For $r > R$, $j = 0$. The time derivative is zero in the steady state, and j is constant. The solution of Eq. (27) is

$$\begin{aligned}[n]_{\text{free}} &= A + \frac{B}{r} - \frac{j}{6D}r^2, \quad a < r \leqslant R \\ [n]_{\text{free}} &= A' + \frac{B'}{r}, \quad r > R\end{aligned} \tag{28}$$

where A, B, A', B' are constants defined by boundary conditions, which are

$$\begin{aligned}&[n]_{\text{free}}|_{r=a} = 0; \\ &[n]_{\text{free}}|_{r=R^-} = [n]_{\text{free}}|_{r=R^+}; \\ &[n]_{\text{free}}|_{r\to\infty} = [n]_{\text{free},\infty}; \\ &\left.\frac{\partial [n]_{\text{free}}}{\partial r}\right|_{r=R^-} = \left.\frac{\partial [n]_{\text{free}}}{\partial r}\right|_{r=R^+}.\end{aligned} \tag{29}$$

Here, $[n]_{\text{free},\infty}$ is the asymptotic free RNAP concentration. These lead to the solutions (notice $\gamma = R/a$)

$$\begin{aligned}A &= \frac{2\gamma^3 + 1}{2\gamma^3 - 3\gamma^2 + 1}[n]_{\text{free},\infty}, \\ B &= -\frac{2\gamma^3}{2\gamma^3 - 3\gamma^2 + 1}a[n]_{\text{free},\infty}, \\ A' &= [n]_{\text{free},\infty}, \\ B' &= 0.\end{aligned} \tag{30}$$

In the steady state, the flux j is related to the total flux

entering a binding site J through

$$J = \frac{4\pi}{3}(R^3 - r^3)j \tag{31}$$

Meanwhile, J is also proportional to the gradient of the free RNAP concentration near the binding site surface, $J = 4\pi a^2 D \frac{\partial [n]_{\text{free}}}{\partial r}|_{r=a}$. So from Eq.(28), we obtain

$$J = -4\pi DB - \frac{4\pi}{3}a^3 j \tag{32}$$

Combining Eq.(31), Eq.(32) and Eq.(30), we get

$$\begin{aligned} j &= \frac{1}{2\gamma^3 - 3\gamma^2 + 1}\frac{6D[n]_{\text{free},\infty}}{a^2}, \\ J &= \frac{2(\gamma^3 - 1)}{2\gamma^3 - 3\gamma^2 + 1}4\pi Da[n]_{\text{free},\infty}. \end{aligned} \tag{33}$$

Finally, according to the definition $k_{\text{on}} = J/[n]_{\text{free},\infty}$, we obtain Eq. (9). In Eq.(6), we use $\overline{[n]}_{\text{free}}$ to replace $[n]_{\text{free},\infty}$, because the binding sites only occupy a small fraction of the whole space, so the values of $\overline{[n]}_{\text{free}}$ and $[n]_{\text{free},\infty}$ are virtually the same.

Agent-based simulation

Since our system is extensive, scaling the volume, gene copy number, nonspecific binding site number, and the RNAP number by the same factor does not change intensive variables such as the averaged free RNAP concentration $\overline{[n]}_{\text{free}}$. Therefore, we simulate 1/10 of the whole system to speed up the simulation. In our simulation, the nucleoid is a cube; promoters and nonspecific binding sites are randomly distributed. Promoters are modeled as spheres with radius a_{s} while nonspecific binding sites are modeled as spheres with radius a_{ns}. For the simulation of each set of g_{ns} and K_{ns}, we first use the mean-field model to estimate the total number of RNAPs given $\overline{[n]}_{\text{free}} = 500\mu m^3$. We model RNAPs as point particles, which are randomly distributed initially. We simulate the diffusion of RNAPs so that the x coordinate of an RNAP particle follows

$$x(t + \Delta t) = x(t) + \sqrt{2D\Delta t}\xi. \tag{34}$$

Here, ξ is a Gaussian random variable with zero mean and unit variance. The y and z coordinates follow similar dynamics. We take the time interval $\Delta t = 10^{-7}$ s and simulate for a total duration from 300 to 600 s depending on the values of g_{ns} and K_{ns} to ensure the system reaches the steady state. We calculate the time-averaged free RNAP concentration, excluding the data of the first quarter simulation. To apply the regulation of rRNA genes, we change $k_{\text{off,r}}$. We repeat the simulations ten times for each g_{ns} and K_{ns} and calculate the mean and standard deviation across different simulations.

ACKNOWLEDGMENTS

The research was funded by the National Key R&D Program of China (2021YFF1200500) and supported by grants from Peking-Tsinghua Center for Life Sciences. Figure 1a, 3a, and Figure 4a-c are created with BioRender.com.

[1] G. Churchward, H. Bremer, and R. Young, Transcription in bacteria at different dna concentrations, Journal of Bacteriology **150**, 572 (1982).

[2] J. Zhurinsky, K. Leonhard, S. Watt, S. Marguerat, J. Bähler, and P. Nurse, A coordinated global control over cellular transcription, Current Biology **20** (2010).

[3] O. Padovan-Merhar, G. P. Nair, A. G. Biaesch, A. Mayer, S. Scarfone, S. W. Foley, A. R. Wu, L. S. Churchman, A. Singh, and A. Raj, Single mammalian cells compensate for differences in cellular volume and dna copy number through independent global transcriptional mechanisms, Molecular Cell **58**, 339 (2015).

[4] M. Kafri, E. Metzl-Raz, F. Jonas, and N. Barkai, Rethinking cell growth models, FEMS yeast research **16**, fow081 (2016).

[5] M. Kafri, E. Metzl-Raz, G. Jona, and N. Barkai, The cost of protein production, Cell Reports **14**, 22 (2016).

[6] G. Neurohr, R. Terry, J. Lengefeld, M. Bonney, G. Brittingham, F. Moretto, T. Miettinen, L. Vaites, L. Soares, J. Paulo, J. Harper, S. Buratowski, S. Manalis, F. van Werven, L. Holt, and A. Amon, Excessive cell growth causes cytoplasm dilution and contributes to senescence, Cell **176**, 1083 (2019).

[7] X.-M. Sun, A. Bowman, M. Priestman, F. Bertaux, A. Martinez-Segura, W. Tang, C. Whilding, D. Dormann, V. Shahrezaei, and S. Marguerat, Size-dependent increase in rna polymerase ii initiation rates mediates gene expression scaling with cell size, Current Biology **30**, 1217 (2020).

[8] M. P. Swaffer, G. K. Marinov, H. Zheng, C. Y. Tsui, A. W. Jones, J. Greenwood, A. Kundaje, W. J. Greenleaf, R. Reyes-Lamothe, and J. M. Skotheim, Rna polymerase ii dynamics and mrna stability feedback scale mrna in proportion to cell size, bioRxiv 10.1101/2021.09.20.461005 (2022).

[9] P. Chakraborty and S. Ghosh, Emergent correlations in gene expression dynamics as footprints of resource competition, The European Physical Journal E **44**, 1 (2021).

[10] N. M. Belliveau, G. Chure, C. L. Hueschen, H. G. Garcia, J. Kondev, D. S. Fisher, J. A. Theriot, and R. Phillips, Fundamental limits on the rate of bacterial growth and their influence on proteomic composition, Cell Systems **12**, 924 (2021).

[11] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, Interdependence of Cell Growth and Gene Expression: Origins and Consequences, Science **330**, 1099 (2010).

[12] J. Lin and A. Amir, Homeostasis of protein and mrna concentrations in growing cells, Nature Communications **9**, 4496 (2018).

[13] Q. Wang and J. Lin, Heterogeneous recruitment abilities to rna polymerases generate nonlinear scaling of gene expression with cell volume, Nature Communications **12**, 6852 (2021).

[14] H. Bremer, P. Dennis, and M. Ehrenberg, Free rna polymerase and modeling global transcription in escherichia coli, Biochimie **85**, 597 (2003).

[15] S. Klumpp and T. Hwa, Growth-rate-dependent partitioning of rna polymerases in bacteria, Proceedings of the National Academy of Sciences **105**, 20245 (2008).

[16] S. Bakshi, R. M. Dalrymple, W. Li, H. Choi, and J. C. Weisshaar, Partitioning of rna polymerase activity in live escherichia coli from analysis of single-molecule diffusive trajectories, Biophysical journal **105**, 2676 (2013).

[17] M. Patrick, P. P. Dennis, M. Ehrenberg, and H. Bremer, Free rna polymerase in escherichia coli, Biochimie **119**, 80 (2015).

[18] M. Stracy, C. Lesterlin, F. G. de Leon, S. Uphoff, P. Zawadzki, and A. N. Kapanidis, Live-cell superresolution microscopy reveals the organization of rna polymerase in the bacterial nucleoid, Proceedings of the National Academy of Sciences **112**, E4390 (2015).

[19] M. T. Killeen and J. F. Greenblatt, The general transcription factor rap30 binds to rna polymerase ii and prevents it from binding nonspecifically to dna., Molecular and Cellular Biology **12**, 30 (1992).

[20] D. L. Nelson, A. L. Lehninger, and M. M. Cox, Lehninger principles of biochemistry (Macmillan, 2008).

[21] O. Borkowski, F. Ceroni, G.-B. Stan, and T. Ellis, Overloaded and stressed: whole-cell considerations for bacterial synthetic biology, Current Opinion in Microbiology **33**, 123 (2016), antimicrobials • Microbial systems biology.

[22] A. Grob, R. Di Blasi, and F. Ceroni, Experimental tools to reduce the burden of bacterial synthetic biology, Current Opinion in Systems Biology **28**, 100393 (2021).

[23] R. Milo and R. Phillips, Cell biology by the numbers (Garland Science, 2015).

[24] S. Bakshi, A. Siryaporn, M. Goulian, and J. C. Weisshaar, Superresolution imaging of ribosomes and rna polymerase in live escherichia coli cells, Molecular Microbiology **85**, 21 (2012).

[25] C. J. Dorman, Structure and function of the bacterial genome (John Wiley & Sons, 2020).

[26] M. Stracy, C. Lesterlin, F. Garza de Leon, S. Uphoff, P. Zawadzki, and A. N. Kapanidis, Live-cell superresolution microscopy reveals the organization of rna polymerase in the bacterial nucleoid, Proceedings of the National Academy of Sciences **112**, E4390 (2015).

[27] H. Bremer and P. P. Dennis, Modulation of chemical composition and other parameters of the cell at different exponential growth rates, EcoSal Plus **3** (2008).

[28] H. Salgado, G. Moreno-Hagelsieb, T. F. Smith, and J. Collado-Vides, Operons in escherichia coli: genomic analyses and predictions, Proceedings of the National Academy of Sciences **97**, 6652 (2000).

[29] I. M. Keseler, J. Collado-Vides, A. Santos-Zavaleta, M. Peralta-Gil, S. Gama-Castro, L. Muñiz-Rascado, C. Bonavides-Martinez, S. Paley, M. Krummenacker, T. Altman, et al., Ecocyc: a comprehensive database of escherichia coli biology, Nucleic acids research **39**, D583 (2010).

[30] J.-G. Kang, M.-Y. Hahn, A. Ishihama, and J.-H. Roe, Identification of sigma factors for growth phase-related promoter selectivity of rna polymerases from streptomyces coelicolor a3 (2), Nucleic acids research **25**, 2566 (1997).

[31] T. M. Gruber and C. A. Gross, Multiple sigma subunits and the partitioning of bacterial transcription space, Annual Reviews in Microbiology **57**, 441 (2003).

[32] M. S. Paget, Bacterial sigma factors and anti-sigma factors: structure, function and distribution, Biomolecules **5**, 1245 (2015).

[33] K. T. Hughes and K. Mathee, The anti-sigma factors, Annual review of microbiology **52**, 231 (1998).

[34] N. Hofmann, R. Wurm, and R. Wagner, The e. coli anti-sigma factor rsd: studies on the specificity and regulation of its expression, PLoS One **6**, e19235 (2011).

[35] L. Gerardo Trevino-Quintanilla, J. Augusto Freyre-Gonzalez, and I. Martínez-Flores, Anti-sigma factors in e. coli: common regulatory mechanisms controlling sigma factors availability, Current genomics **14**, 378 (2013).

[36] M. S. Paget and J. D. Helmann, The σ70family of sigma factors, Genome biology **4**, 1 (2003).

[37] I. L. Grigorova, N. J. Phleger, V. K. Mutalik, and C. A. Gross, Insights into transcriptional regulation and σ competition from an equilibrium model of rna polymerase binding to dna, Proceedings of the National Academy of Sciences **103**, 5332 (2006).

[38] S. E. Piper, J. E. Mitchell, D. J. Lee, and S. J. Busby, A global view of escherichia coli rsd protein and its interactions, Molecular BioSystems **5**, 1943 (2009).

[39] A. N. Kapanidis, E. Margeat, T. A. Laurence, S. Doose, S. O. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R. H. Ebright, and S. Weiss, Retention of transcription initiation factor σ70 in transcription elongation: single-molecule analysis, Molecular cell **20**, 347 (2005).

[40] T. T. Harden, C. D. Wells, L. J. Friedman, R. Landick, A. Hochschild, J. Kondev, and J. Gelles, Bacterial rna polymerase can retain σ70 throughout transcription, Proceedings of the National Academy of Sciences **113**, 602 (2016).

[41] H. Maeda, M. Jishage, T. Nomura, N. Fujita, and A. Ishihama, Two extracytoplasmic function sigma subunits, ςe and ςfeci, of escherichia coli: Promoter selectivity and intracellular levels, Journal of Bacteriology **182**, 1181 (2000).

[42] A. Ishihama and R. Fukuda, Autogenous and post-transcriptional regulation of rna polymerase synthesis, Molecular and cellular biochemistry **31**, 177 (1980).

[43] M. Jishage, A. Iwata, S. Ueda, and A. Ishihama, Regulation of rna polymerase sigma subunit synthesis in escherichia coli: intracellular levels of four species of sigma subunit under various growth conditions, Journal of bacteriology **178**, 5447 (1996).

[44] S.-T. Liang, M. Bipatnath, Y.-C. Xu, S.-L. Chen, P. Dennis, M. Ehrenberg, and H. Bremer, Activities of constitutive promoters in escherichia coli, Journal of molecular biology **292**, 19 (1999).

[45] D. Jensen and E. A. Galburt, The context-dependent influence of promoter sequence motifs on transcription initiation kinetics and regulation, Journal of Bacteriology **203**, 10.1128/jb.00512 (2021), https://journals.asm.org/doi/pdf/10.1128/jb.00512-20.

[46] R. Phillips, J. Kondev, J. Theriot, and H. Garcia, Physical biology of the cell (Garland Science, 2012).

[47] M. Mauri and S. Klumpp, A model for sigma factor competition in bacterial cells, PLoS computational biology

10, e1003845 (2014).

[48] M. Jishage and A. Ishihama, Regulation of rna polymerase sigma subunit synthesis in escherichia coli: intracellular levels of sigma 70 and sigma 38, Journal of Bacteriology **177**, 6832 (1995).

[49] V. K. Kandavalli, H. Tran, and A. S. Ribeiro, Effects of σ factor competition are promoter initiation kinetics dependent, Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms **1859**, 1281 (2016).

[50] X. Wang, P. M. Llopis, and D. Z. Rudner, Organization and segregation of bacterial chromosomes, Nature Reviews Genetics **14**, 191 (2013).

[51] E. Toro and L. Shapiro, Bacterial chromosome organization and segregation, Cold Spring Harbor perspectives in biology **2**, a000349 (2010).

[52] B. P. Bratton, R. A. Mooney, and J. C. Weisshaar, Spatial distribution and diffusive motion of rna polymerase in live escherichia coli, Journal of bacteriology **193**, 5138 (2011).