RNA polymerase sliding on DNA can couple the transcription of nearby bacterial operons.

Debora Tenenbaum,[1,2,3] Koe Inlow,[1] Larry Friedman, [1] Anthony Cai, [1] Jeff Gelles, [1,*] Jane Kondev [2,*]
[1]Department of Biochemistry, Brandeis University, Waltham, MA, United States
[2]Department of Physics, Brandeis University, Waltham, MA, United States
[3]Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
[*]Correspondence: gelles@brandeis.com, kondev@brandeis.com

SIGNIFICANCE STATEMENT

After transcribing an operon, a bacterial RNA polymerase can stay bound to DNA, slide along it, and reinitiate transcription of the same or a different operon. Quantitative single-molecule biophysics experiments combined with mathematical theory demonstrate that this reinitiation process can be quick and efficient over gene spacings typical of a bacterial genome. Reinitiation may provide a mechanism to orchestrate the transcriptional activities of groups of nearby operons.

ABSTRACT

DNA transcription initiates after an RNA polymerase (RNAP) molecule binds to the promoter of a gene. In bacteria, the canonical picture is that RNAP comes from the cytoplasmic pool of freely diffusing RNAP molecules. Recent experiments suggest the possible existence of a separate pool of polymerases, competent for initiation, which freely slide on the DNA after having terminated one round of transcription. Promoter-dependent transcription reinitiation from this pool of post-termination RNAP may lead to coupled initiation at nearby operons, but it is unclear whether this can occur over the distance- and time-scales needed for it to function widely on a bacterial genome in vivo. Here, we mathematically model the hypothesized reinitiation mechanism as a diffusion-to-capture process and compute the distances over which significant inter-operon coupling can occur and the time required. These quantities depend on previously uncharacterized molecular association and dissociation rate constants between DNA, RNAP and the transcription initiation factor σ^{70}; we measure these rate constants using single-molecule experiments in vitro. Our combined theory/experimental results demonstrate that efficient coupling can occur at physiologically relevant σ^{70} concentrations and on timescales appropriate for transcript synthesis. Coupling is efficient over terminator-promoter distances up to $\sim 1,000$ bp, which includes the majority of terminator-promoter nearest neighbor pairs in the *E. coli* genome. The results suggest a generalized mechanism that couples the transcription of nearby operons and breaks the paradigm that each binding of RNAP to DNA can produce at most one messenger RNA.

Keywords Reaction-diffusion model | Single-molecule fluorescence | Gene coupling

INTRODUCTION

The core RNA polymerase (RNAP) in bacteria, which is composed of five subunits ($\beta, \beta', \alpha^I, \alpha^{II}$, and ω), can catalyze the synthesis of RNA, but cannot recognize specific promoter sequences. To recognize promoters, RNAP must first bind an initiation factor such as the *E. coli* housekeeping σ^{70}, forming an RNAP holoenzyme (σ^{70}RNAP)[1–3]. DNA transcription initiates when σ^{70}RNAP is recruited from cytoplasm to bind to the promoter region of a gene. Controlling this process is thought to be the principal means through which transcription repressors and activators modulate gene transcription.

Bacterial metabolism requires the coordinated expression of multiple genes[4]. A basic way in which this coordination is achieved in bacterial cells is by the organization of functionally related genes into operons, which are groups of consecutive genes that can be transcribed from the same promoter[5,6]. Functionally related operons often reside in contiguous regions of the bacterial genome[7]. Proximally located operons show higher levels of correlated expression than distant operons in *E. coli*[8–10]. The same is true of closely spaced genes in eukaryotes[11,12].

While specific groups of bacterial operons may have correlated activities simply because they have common regulatory proteins (e.g., alternative sigma factors), there are also proximity-based mechanisms that can couple transcription of adjacent operons. Terminator readthrough, in which the RNAP fails to read a terminator signal and keeps elongating the mRNA molecule, can generate the joint transcription of co-directional neighboring operons. Transcription-coupled DNA supercoiling[13] can induce coupled transcription of divergently transcribed genes[14,15].

Recently, a new mechanism of proximity-based transcription coupling was observed. Using single-molecule microscopy, Harden et al.[16] and Kang et al.[17,18] observed that RNAP can remain bound to DNA after termination for at least hundreds of seconds in vitro. This post-termination RNAP-DNA complex may retain a partially open bubble in the DNA[19]. The retained RNAP exhibits one-dimensional diffusive sliding over hundreds or thousands of base pairs along the DNA. In the presence of σ^{70} in solution, the sliding RNAP can re-initiate transcription at a nearby promoter. This post-termination behavior of bacterial RNAP may couple transcription of nearby operons in a way that is dependent on both the distance between the two transcription units and the available concentration of σ^{70}. Genome-wide transcription measurements are consistent with this mechanism, but do not prove that it operates in vivo in both *E. coli* and *B. subtilis*[16].

In this work, we test whether RNAP post-termination sliding followed by σ^{70} rebinding can efficiently couple the transcription of nearby operons. First, we mathematically model the mechanism as a diffusion-to-capture process, in which the association of a σ^{70} molecule with the sliding RNAP is required for re-initiation at a nearby promoter sequence. Next, we use single-molecule microscopy experiments under conditions designed to mimic the ionic composition of bacterial cytoplasm to measure the values of the model kinetic parameters. Finally, we input the measured values into the model to predict the distances

and times over which post-termination sliding of RNAP could couple expression of neighboring genes.

RESULTS

Model of operon coupling by sliding RNAP

We model transcriptional coupling between proximal operons using the sliding RNAP mechanism depicted in Fig. 1. The distance between the terminator of the first operon and the promoter of the second operon is d. Upon reaching the terminator sequence T of the primary operon (at time $t = 0$), the RNAP releases an RNA transcript but remains non-specifically bound, enabling it to diffuse along the DNA with a diffusion coefficient D. During this time interval of sliding, the RNAP can either dissociate from the DNA with rate k_{off}, or bind a σ^{70} molecule from solution with a rate $k_{\text{b}}[\sigma^{70}]$, where $[\sigma^{70}]$ denotes the free σ^{70} solution concentration. In the latter case, the RNAP-σ^{70} complex continues diffusing along the DNA molecule, and can dissociate with a rate $k_{\text{off,s}}$, or can encounter and be captured by the promoter for the secondary operon. We define the time it takes the captured RNAP-σ^{70} complexes to find the secondary promoter as the search time t_{f}. In this mechanism, σ^{70} can have conflicting effects because it can stimulate RNAP dissociation from DNA via the $k_{\text{off,s}}$ step and yet is also required for secondary promoter capture.

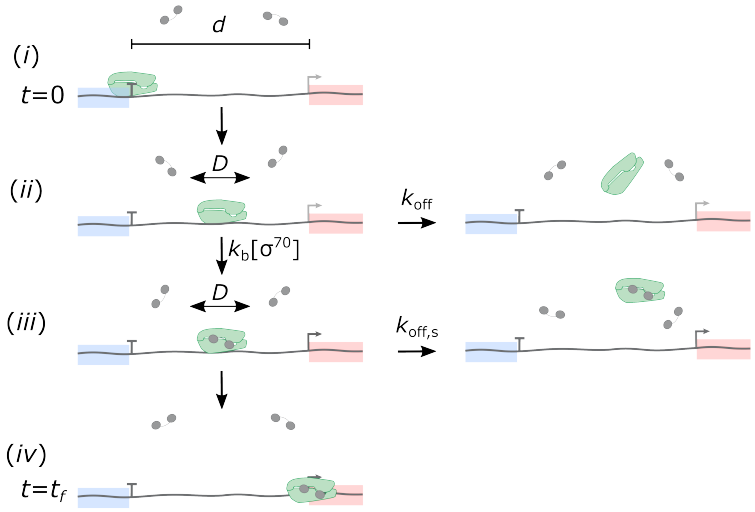

Figure 1: Model of operon coupling by sliding RNAP. (i) An RNAP molecule (green) terminates transcription of the primary operon (blue), and (ii) starts sliding along the DNA molecule with a diffusion constant D. (iii) While sliding, the RNAP can either dissociate from the DNA with rate k_{off}, or bind σ^{70} (gray) with rate $k_{\text{b}}[\sigma^{70}]$. (iv) After binding of σ^{70}, the RNAP-σ^{70} complex can either dissociate with a rate $k_{\text{off,s}}$, or find the promoter (bent arrow) for the secondary operon (pink), which is located is at a distance d along the DNA from the primary operon terminator (T).

For simplicity, we assume that the binding of σ^{70} to sliding RNAP is irreversible. This is equivalent to assuming that the unbinding of σ^{70} from the sliding RNAP is significantly slower than the dissociation of the RNAP-σ^{70} complex from the DNA. This assumption is reasonable, given previous measurements of σ^{70} dissociation from free[20,21] and DNA-bound RNAP[17]. Also for simplicity, we assume that the diffusion

coefficient on DNA of the RNAP-σ^{70} complex and RNAP are the same.

In this work we will refer to the complex that is formed by binding of σ^{70} to DNA-bound RNAP as RNAP-σ^{70}-DNA and to the complex formed by binding σ^{70}RNAP holoenzyme to DNA as holoenzyme-DNA. It is not currently known whether these different orders of assembly produce complexes with the same structure (see Discussion).

Calculation of coupling efficiency

To quantify how transcriptional coupling by sliding RNAP changes with varying distance between operons and with $[\sigma^{70}]$, we define the coupling efficiency (E) as the probability that an RNAP molecule, which terminates transcription of the primary operon, reaches the promoter of the secondary promoter by the sliding RNAP mechanism. To reach the secondary promoter, (i) the sliding RNAP has to bind a σ^{70} molecule from solution before falling off the DNA, and (ii) the RNAP-σ^{70} complex then has to reach the secondary promoter before falling off the DNA. The probability of (i), P_{bind}, is given by the partition ratio

$$P_{\text{bind}} = \frac{k_{\text{b}}\,[\sigma^{70}]}{k_{\text{b}}\,[\sigma^{70}] + k_{\text{off}}}, \tag{1}$$

while the probability of (ii), P_{find}, can be computed as the probability that a RNAP-σ^{70} complex remains bound to DNA for at least the time needed to find the secondary promoter before dissociating.

To determine P_{find}, we combine the distribution of times it takes RNAP-σ^{70} complexes to encounter the secondary promoter for the first time, $p_{\text{first passage}}(t)$, and the probability that the complex will stay bound on the DNA long enough for the encounter to happen. Thus, P_{find} is given by

$$P_{\text{find}} = \int_0^\infty p_{\text{first passage}}(t') \exp\left(-k_{\text{off,s}} t'\right) dt'. \tag{2}$$

Here we use uppercase P to refer to probabilities, and lowercase p to represent probability density functions (PDFs).

RNAP terminates transcription of the primary operon at $x = 0$ and starts performing a one-dimensional random walk along the DNA with diffusion coefficient D. How long it takes for a RNAP-σ^{70} complex to first encounter the secondary promoter will depend on its position on the DNA (x_{b}) when it starts the search, i.e., when σ^{70} binds the sliding RNAP molecule. x_{b} in turn depends on how long after termination at the primary terminator σ^{70} binds (t_{b}). At a time t_{b} drawn at random from the exponential distribution

$$p_{\text{bind,t}}(t_{\text{b}}) = (k_{\text{b}}[\sigma^{70}] + k_{\text{off}}) \exp\left(-(k_{\text{b}}[\sigma^{70}] + k_{\text{off}}) t_{\text{b}}\right), \tag{3}$$

the sliding RNAP will either bind a σ^{70} molecule or dissociate from the DNA. If it binds a σ^{70} molecule, the

binding position will be a random value drawn from

$$p^{\text{cond}}_{\text{bind,x}}(x_{\text{b}}|t_{\text{b}}) = \mathcal{N}(0, 2Dt_{\text{b}}), \tag{4}$$

where $\mathcal{N}(\mu, var)$ describes a normal distribution with mean μ and variance *var*. Eq. 4 represents the conditional probability distribution for x_{b} given the binding time t_{b}. Now we can calculate the unconditional distribution of binding positions

$$\begin{aligned} p_{\text{bind,x}}(x_{\text{b}}) &= \int_0^{\infty} p^{\text{cond}}_{\text{bind,x}}(x_{\text{b}}|t')p_{\text{bind,t}}(t')dt' \\ &= \sqrt{\frac{k_{\text{b}}[\sigma] + k_{\text{off}}}{4D}} \exp\left(-\sqrt{\frac{k_{\text{b}}[\sigma] + k_{\text{off}}}{D}}|x_{\text{b}}|\right). \end{aligned} \tag{5}$$

The distribution of times t_{fp} it would take the RNAP-σ^{70} complex to reach the secondary promoter at $x = d$ for the first time is then given by the first-passage-time density for a one-dimensional random walk[22]

$$p^{\text{cond}}_{\text{first passage}}(t_{\text{fp}}|x_{\text{b}}) = \frac{|d - x_{\text{b}}|}{\sqrt{4\pi D t_{\text{fp}}^3}} \exp\left(-\frac{(d - x_{\text{b}})^2}{4Dt_{\text{fp}}}\right), \tag{6}$$

which is conditional on x_{b}. The unconditional distribution of first passage times is then calculated as

$$p_{\text{first passage}}(t_{\text{fp}}) = \int_{-\infty}^{+\infty} p^{\text{cond}}_{\text{first passage}}(t_{\text{fp}}|x')p_{\text{bind,x}}(x')dx'. \tag{7}$$

Finally, substituting Eq. 7 into Eq. 2 to get P_{find}, and multiplying by P_{bind} (Eq. 1), we get the expressions in Eq. 8.

$$E(d, [\sigma^{70}]) = \begin{cases} \frac{k_{\text{b}}[\sigma^{70}]}{k_{\text{b}}[\sigma^{70}] + k_{\text{off}} - k_{\text{off,s}}} \left(\exp\left(-\sqrt{\frac{k_{\text{off,s}}}{D}}d\right) - \sqrt{\frac{k_{\text{off,s}}}{k_{\text{b}}[\sigma^{70}] + k_{\text{off}}}} \exp\left(-\sqrt{\frac{k_{\text{b}}[\sigma^{70}] + k_{\text{off}}}{D}}d\right)\right), & \text{if } k_{\text{b}}[\sigma^{70}] + k_{\text{off}} \neq k_{\text{off,s}} \\ \frac{k_{\text{b}}[\sigma^{70}]}{2} \exp\left(-\sqrt{\frac{k_{\text{off,s}}}{D}}d\right)\left(\frac{d}{\sqrt{k_{\text{off,s}}D}} + \frac{1}{k_{\text{off,s}}}\right), & \text{if } k_{\text{b}}[\sigma^{70}] + k_{\text{off}} = k_{\text{off,s}} \end{cases} \tag{8}$$

Coupling regimes and calculation of coupling distance

We can distinguish three different coupling regimes depending on the availability of free σ^{70} molecules. For this, we define a critical σ^{70} concentration at which the diffusion time intervals available before and after σ^{70} binding to RNAP are equal, $[\sigma^{70}]_c = \frac{k_{\text{off,s}} - k_{\text{off}}}{k_{\text{b}}} \approx \frac{k_{\text{off,s}}}{k_{\text{b}}}$. Here we used $k_{\text{off}} << k_{\text{off,s}}$ based on prior studies[23] and our experimental results (see below).

1. When $[\sigma^{70}] \gg [\sigma^{70}]_c$, the coupling efficiency at small d is given by P_{bind} (Eq. 1) since any RNAP that

binds σ^{70} will subsequently encounter the promoter. For larger d the efficiency decays exponentially,

$$E \approx P_{\text{bind}} \exp\left(-\sqrt{\frac{k_{\text{off,s}}}{D}}d\right). \tag{9}$$

In this regime, we can define the *coupling distance* as the characteristic decay distance of the coupling, $d_c^{\gg} \approx \sqrt{\frac{D}{k_{\text{off,s}}}}$.

2. When $[\sigma^{70}] \ll [\sigma^{70}]_c$, the decay is also exponential but in this case,

$$E \approx P_{\text{bind}} \sqrt{\frac{k_b[\sigma^{70}] + k_{\text{off}}}{k_{\text{off,s}}}} \exp\left(-\sqrt{\frac{k_b[\sigma^{70}] + k_{\text{off}}}{D}}d\right), \tag{10}$$

and the characteristic distance is $d_c^{\ll} \approx \sqrt{\frac{D}{k_b[\sigma^{70}]+k_{\text{off}}}}$. However, in this case $\sqrt{\frac{k_b[\sigma^{70}]+k_{\text{off}}}{k_{\text{off,s}}}} \ll 1$, which means that there is no significant coupling between adjacent transcription units no matter the distance between them.

3. For $[\sigma^{70}] \approx [\sigma^{70}]_c$, we get the bottom expression in Eq. 8. Even though it is not exponential, we can still define a coupling distance $d_c^{\approx}$ over which the coupling efficiency decays by a factor of e. Following the calculations in Appendix S1, we get

$$\begin{aligned} d_c^{\approx} \approx 2.15\sqrt{\frac{D}{k_{\text{off,s}}}} &= 2.15\sqrt{\frac{D}{k_b[\sigma^{70}] + k_{\text{off}}}} \\ &= 1.08\left(\sqrt{\frac{D}{k_{\text{off,s}}}} + \sqrt{\frac{D}{k_b[\sigma^{70}] + k_{\text{off}}}}\right). \end{aligned} \tag{11}$$

In simple terms, the regimes differ by whether most of the diffusional search for the secondary promoter takes place after (regime 1) or before (regime 2) the binding of σ^{70}. In addition, given that $\sqrt{\frac{D}{k_b[\sigma^{70}]+k_{\text{off}}}} \ll \sqrt{\frac{D}{k_{\text{off,s}}}}$ in regime 1, and $\sqrt{\frac{D}{k_{\text{off,s}}}} \ll \sqrt{\frac{D}{k_b[\sigma^{70}]+k_{\text{off}}}}$ in regime 2, for all three regimes we can then approximate the coupling distance as $d_c \approx \sqrt{\frac{D}{k_b[\sigma]+k_{\text{off}}}} + \sqrt{\frac{D}{k_{\text{off,s}}}}$, which is roughly the sum of the root mean squared displacements of the RNAP before and after binding a σ^{70} molecule (Fig. S1). Efficiency curves as a function of distance scaled by the critical distance are shown for all three regimes in Fig. 2. As expected, when the concentration of σ^{70} is well below its critical concentration, the efficiency is small for any distance between the two promoters, while the efficiency can be of order one when $[\sigma^{70}]$ is well above $[\sigma^{70}]_c$.

Single-molecule microscopy experiments to measure model parameters

To estimate the coupling efficiency E, coupling distance d_c, and search times t_f that can be achieved via the proposed mechanism (Fig. 1), we need values for the model parameters D, k_b, k_{off}, and $k_{\text{off,s}}$.

The diffusion coefficient of RNAP post termination, $D = (3.9 \pm 0.5) \times 10^4$ bp^{2}s^{-1}, was experimentally measured in ref.[16]. Those investigators also sometimes observed a non-diffusing post termination RNAP-

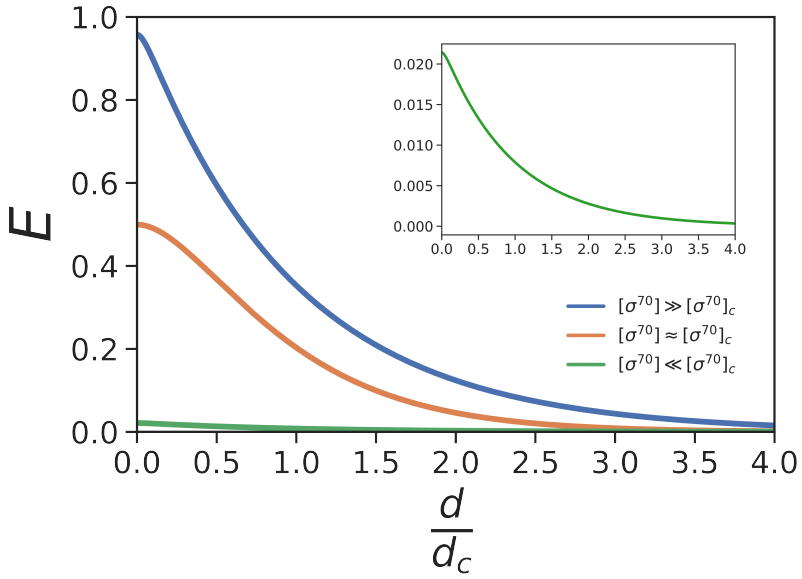

Figure 2: Predicted relationship of Coupling efficiency E to the distance between operons. The coupling efficiency is the probability of an RNAP that terminated transcription at the end of the primary operon reaching the secondary promoter. Efficiency curves are shown for the three regimes of σ^{70} concentration described in the text, for a chosen set of parameter values: $k_{\mathrm{b}} = 10^7$ M^{-1}s^{-1}, $k_{\mathrm{off}} = 10^{-3}$ s^{-1}, $k_{\mathrm{off,s}} = 1$ s^{-1}, $D = 4 \times 10^4$ bp^{2}s^{-1}. Curves were calculated using Eq. 8. Distance is represented in units of the coupling distance $d_c \approx \sqrt{D(k_{\mathrm{b}}[\sigma^{70}] + k_{\mathrm{off}})^{-1}} + \sqrt{D(k_{\mathrm{off,s}})^{-1}}$. Inset: enlarged plot of the curve for $[\sigma^{70}] \ll [\sigma^{70}]_c$.

DNA complex in their experiments, but attributed this to RNAP binding to the ends of the linear DNA molecules they used.

We performed single-molecule experiments to measure k_{b}, k_{off} and $k_{\mathrm{off,s}}$. Specifically, we quantified the dwell times of RNAP on promoterless DNA templates in the presence of different concentrations of σ^{70} (Fig. 3A). These experiments allow us to measure all three rate constants. This is because at low σ^{70} concentrations the measured dwell times are limited by the rate of RNAP dissociation from DNA, at intermediate σ^{70} concentrations they are limited by the rate of σ^{70} binding to the RNAP-DNA complex, and at high σ^{70} concentrations they are limited by the rate of RNAP-σ^{70} complex dissociation from DNA.

For experimental convenience, we did not use core RNAP-DNA complexes that were formed after termination of transcription. Instead, we directly formed sequence non-specific RNAP-DNA complexes by adding core RNAP to a DNA that lacks known promoter sequences. The two types of complexes have the same protein composition and have similar properties: both are long-lived, in both RNAP slides on DNA, both are sensitive to the polyanion heparin[16], and both are rapidly disassembled by the bacterial SNF2 ATPase RapA[24].

To implement these experiments we designed and synthesized a biotinylated, 3,033 bp circular DNA lacking known promoter sequences that was labeled with the red-excited dye Cy5 (we refer to this construct as $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$). Circular DNAs were used to avoid possible binding of RNAP to DNA ends[25], which are largely non-physiological since the *E. coli* chromosome is circular.

We immobilized $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$ molecules on the surface of a glass flow chamber via a biotin-streptavidin link-

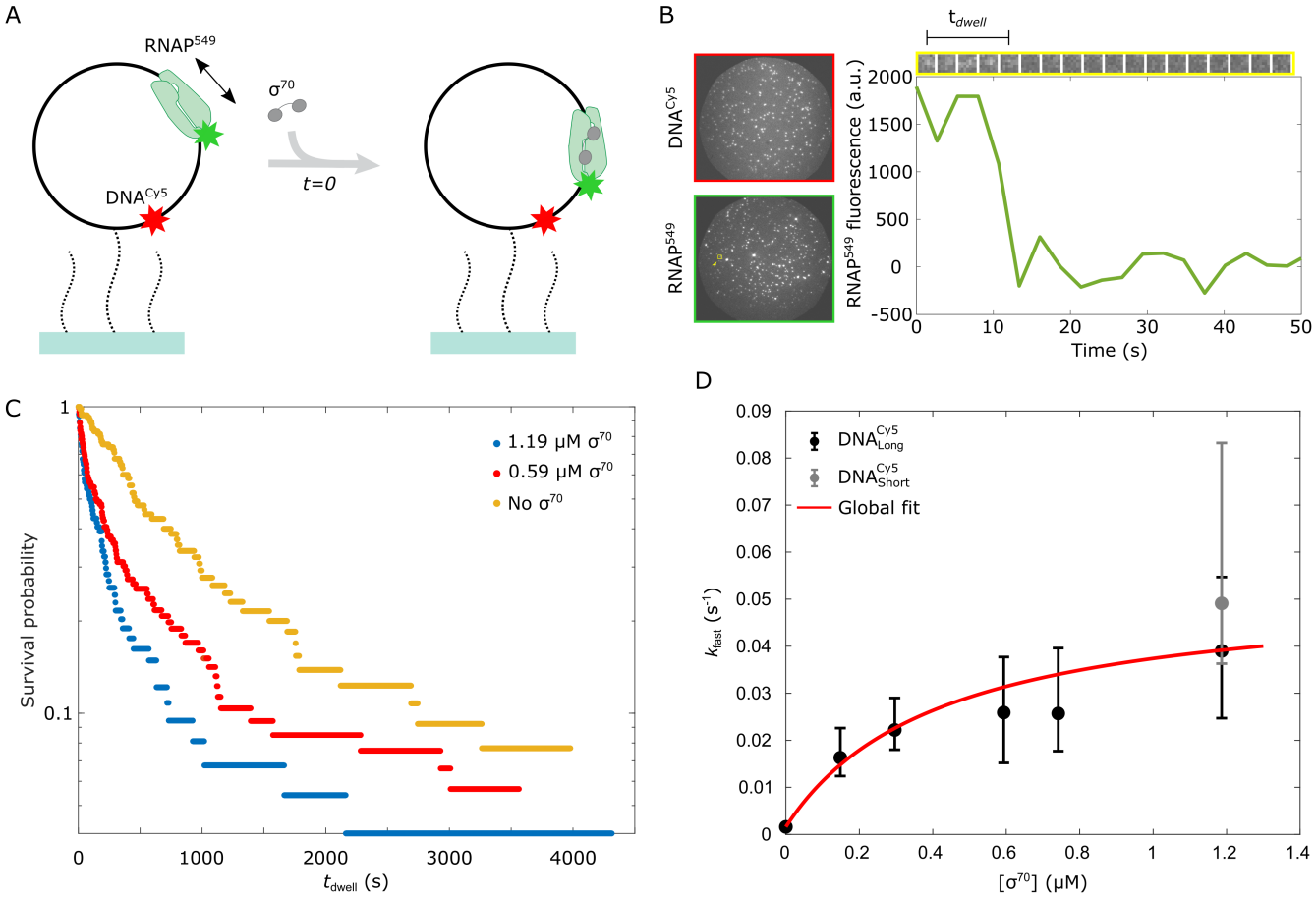

Figure 3: Single-molecule experiments to measure k_b, k_{off}, and $k_{off,s}$. a. Experiment schematic. Fluorescently labeled, promoterless circular DNA templates (DNA^{Cy5}_{Long}; black circles) were tethered to the surface of a glass flow chamber (blue) through polyethylene glycol linkers (dotted black curves). The chamber was then incubated with fluorescently labeled core RNAP ($RNAP^{549}$) to form RNAP-DNA nonspecific complexes, which correspond to the post-termination complex. In each of six experiments, a different concentration of σ^{70} (gray) was introduced at $t = 0$, and the lifetime of each $RNAP^{549}$ that colocalized with a surface-tethered DNA molecule was monitored by single-molecule fluorescence microscopy. **B.** Example of experiment record. Left: DNA^{Cy5}_{Long} and $RNAP^{549}$ fluorescence images of the same field of view ($65\ \mu$m diameter) upon introducing $1.19\ \mu$M σ^{70} at $t = 0$. Right: Time record excerpt of $RNAP^{549}$ fluorescence at the location of a single DNA molecule. Gallery shows 5×5 pixel images centered on the DNA molecule; graph shows the summed, background-corrected intensity of the 3×3 pixels centered on the DNA. t_{dwell} represents the duration of the fluorescent spot. **C.** t_{dwell} survival probability distributions in the presence of 0, 0.59, and $1.19\ \mu$M σ^{70}. **D.** Rates (with 68% C.I.s) of σ^{70}-dependent dissociation of $RNAP^{549}$ from DNA^{Cy5}_{Long} (black) and DNA^{Cy5}_{short} (gray) as a function of σ^{70} concentration, and global fit (red; see eq. (12) and accompanying text).

age (Fig. 3A). We then incubated the chamber with a solution containing *E. coli* core RNAP labeled with a green-excited dye ($RNAP^{549}$) for ~ 10 min, and washed it out at time $t = 0$ with a solution containing σ^{70} in the 0 to 1.2 μM range. Single-molecule total internal reflection microscopy was performed with alternating red and green excitation for observation of DNA^{Cy5}_{Long} and $RNAP^{549}$, respectively. An example of the fluorescence records used for extracting the dwell times of the $RNAP^{549}$ molecules on the DNA^{Cy5}_{Long} template for each experiment is shown in Fig. 3B.

Given that these experiments study sequence-nonspecific interactions between $RNAP^{549}$ and DNA^{Cy5}_{Long}, it was expected that multiple $RNAP^{549}$ molecules could be bound to the same DNA^{Cy5}_{Long} template simultaneously. To reduce complications in the dwell time measurements arising from multiple $RNAP^{549}$ molecules bound to the same template, we restricted the analysis to only those DNA^{Cy5}_{Long} locations with a single colocalized $RNAP^{549}$. The number of $RNAP^{549}$ molecules bound to each DNA^{Cy5}_{Long} template was quantified by counting the number of decreasing steps present in the $RNAP^{549}$ fluorescence intensity records (Fig. S2 and Appendix S3).

Distributions of RNAP dwell times on DNA

Example dwell time probability distributions of $RNAP^{549}$ on promoterless circular DNA templates for different concentrations of σ^{70} are shown in Fig. 3C. Consistent with the results in [23], σ^{70} accelerates the dissociation of RNAP from DNA.

In the absence of promoter sequences in the DNA, the model in Fig. 1 predicts that the dwell time distributions for RNAP obtained in the limits of low and high $[\sigma^{70}]$ are exponential. Theoretically, at intermediate $[\sigma^{70}]$ the dwell time distributions are non-exponential, due to the presence of two sequential steps (k_b and $k_{off,s}$). Still, for reasonable values of the rate constants the distribution is well approximated by an exponential and the effective rate constant has a hyperbolic dependence on $[\sigma^{70}]$,

$$k_{eff} \approx k_{off} + \frac{k_{off,s}\, k_b\, [\sigma^{70}]}{k_b\, [\sigma^{70}] + k_{off,s}}. \tag{12}$$

However, the experimental distributions are in fact described not by a single exponential, but by the sum of multiple exponential components with very different rate constants. Specifically, for the experiments in which $[\sigma^{70}] > 0$ the distributions are well fit by a sum of three exponentials with characteristic rates k_{slow}, k_{inter}, and k_{fast} (Fig. S3). This suggests that in these experiments there are at least three types of RNAP-σ^{70}-DNA complexes. The resulting fits, obtained using a maximum likelihood method, are shown in Fig. S4, and the fit parameters are summarized in Table 1. In all cases, non-specific binding of $RNAP^{549}$ to the chamber surface was minimal (Table S1), and therefore was not considered when fitting the data.

For the experiments with $[\sigma^{70}] > 0$, k_{fast} showed a hyperbolic dependence of $[\sigma^{70}]$ (Fig. 3D); k_{inter} and k_{slow} did not (Table 1). Therefore, we hypothesize that the fastest component corresponds to σ^{70}-induced dissociation of $RNAP^{549}$ bound to DNA, $k_{eff} = k_{fast}$. This suggests that at low σ^{70} concentrations binding of σ^{70}

Table 1: Parameters for fits to dwell time distributions of RNAP^{549}-promoterless DNA complexes at different σ^{70} concentrations.

$[\sigma^{70}]$ (μM)	N	N_d	a_{fast} ($\times 10^{-1}$)	a_{inter} ($\times 10^{-1}$)	k_{fast} ($\times 10^{-2}$ s^{-1})	k_{inter} ($\times 10^{-3}$ s^{-1})	k_{slow} ($\times 10^{-4}$ s^{-1})	DNA preparation
0	65	60	$8.6(6.8-9.4)$	-	$0.16(0.13-0.22)$	-	$1.59(0-3.99)$	DNA^{Cy5}_{Long}, preparation 1
0.15	60	42	$1.2(0.4-2.0)$	$1.6(0.4-2.9)$	$1.63(1.24-2.26)$	$3.11(2.27-4.52)$	$1.90(1.44-2.33)$	DNA^{Cy5}_{Long}, preparation 2
0.30	139	85	$1.8(1.3-2.1)$	$1.3(0.7-3.8)$	$2.22(1.80-2.90)$	$1.23(0.53-2.04)$	$1.27(0.48-1.52)$	DNA^{Cy5}_{Long}, preparation 2
0.59	106	100	$3.4(2.5-4.8)$	$5.4(4.1-6.2)$	$2.59(1.52-3.77)$	$2.61(1.87-3.37)$	$2.18(0.26-3.63)$	DNA^{Cy5}_{Long}, preparation 1
0.74	106	60	$1.8(1.2-2.2)$	$2.0(1.0-5.2)$	$2.57(1.77-3.96)$	$0.81(0.34-1.72)$	$1.05(0-1.54)$	DNA^{Cy5}_{Long}, preparation 3
1.19	74	71	$3.4(2.3-4.9)$	$5.8(4.4-6.8)$	$3.90(2.47-5.47)$	$3.88(2.45-5.26)$	$1.61(0-3.24)$	DNA^{Cy5}_{Long}, preparation 1
1.19	76	68	$5.8(4.0-6.7)$	$3.0(2.0-4.7)$	$4.91(3.63-8.32)$	$5.77(4.03-10.09)$	$0.72(0-1.56)$	DNA^{Cy5}_{Short}

The models used for the fit are described in Methods (Eqs. 13 and 14). N is the number of DNA sites with co-localized RNAP that were used in the analysis. N_d is the number of DNA sites for which the co-localized RNAPs disappeared before the end of the experiment. The values are presented with 68% CI. In some cases, the lower confidence limit on k_{slow} is poorly defined because k_{slow}^{-1} exceeds the duration of the experiment. DNA^{Cy5}_{Long} preparations 1, 2, and 3 were made by the same method on different occasions.

to the sliding RNAP is rate-limiting so that the dissociation rate of RNAP from DNA increases linearly with σ^{70} concentration, while at high concentrations the dissociation of the RNAP-σ^{70} complex from DNA becomes limiting, and the dissociation rate saturates. Possible origins of the longer-lived RNAP-DNA complexes with $[\sigma^{70}]$-independent dissociation rates k_{inter} and k_{slow} are discussed in the Appendix S4.

To confirm that k_{fast} depends on $[\sigma^{70}]$ and not on the DNA template used, we repeated the 1.19 μM σ^{70} experiment using a different DNA template, the promoterless 586 bp circular DNA^{Cy5}_{Short}. Similar values were obtained for k_{fast} for both templates (Table 1, Fig. 3D), supporting the idea that k_{fast} depends on σ^{70} concentration, and not on the length or sequence of the DNA template used.

Two characteristic rates were observed for the experiment with $[\sigma^{70}] = 0$. The slower one is similar to the values of k_{slow} observed in the presence of σ^{70} (Table 1). The faster one is similar to the mean dissociation rate observed for the post-termination RNAP-DNA complex in the absence of σ^{70}, as well as to the mean dissociation rate observed for core RNAP sequence-nonspecifically bound to DNA[16]. Therefore, we assume that the faster rate corresponds to the dissociation rate of RNAP^{549} from DNA in the limit where $[\sigma^{70}] = 0$.

Extraction of model parameters k_b, k_{off}, and $k_{off,s}$

Having established the hyperbolic dependence of the σ^{70}-induced dissociation rate k_{fast} on σ^{70} concentration, we can now determine the values for the Fig. 1 model parameters k_b, k_{off}, and $k_{off,s}$. For this, we jointly fit the data from experiments at different σ^{70} concentrations to a global model that incorporates our conclusions about the origins of the different components of the dwell time distributions (see Appendix S4). The σ^{70}-independent rates k_{inter} and k_{slow} were globally fit for all six experiments performed with DNA^{Cy5}_{Long} (Table S3 and Fig. S8). A separate set of parameters k'_{inter} and k'_{slow} were obtained by fitting the dwell time distribution from the experiment with DNA^{Cy5}_{Short}. The global model explicitly included the $[\sigma^{70}]$ dependency of k_{fast} (Eq. 12, where $k_{eff} = k_{fast}$).

The model fit well to the data (Fig. S8) and gave well-constrained values for the rate constants (Table 2). The rate constants, together with the diffusion coefficient D measured in ref.[16] provide the information needed to calculate the extent and kinetics of operon coupling.

Table 2: Global model parameters.

Parameter	Description	Value (68% C.I.)	Source
D	Diffusion coefficient of sliding RNAP	3.9×10^4 bp^{2}s^{-1}†	[16]
k_{b}	Binding rate of σ^{70} to sliding RNAP	$1.2(0.7-2.7) \times 10^5$ M^{-1}s^{-1}	Fit
k_{off}	Dissociation rate of RNAP before binding σ^{70}	$1.6(1.3-1.9) \times 10^{-3}$ s^{-1}	Fit
$k_{\text{off,s}}$	Dissociation rate of RNAP-σ^{70} complex	$5.1(3.4-12.2) \times 10^{-2}$ s^{-1}	Fit

†Kang et al [17] measured a somewhat lower value corresponding to 0.8×10^4bp^{2}s^{-1} at similar ionic strength but in a different buffer.

Extent of operon coupling by sliding RNAP

Operon coupling cannot be biologically functional if it takes an infeasibly long time for RNAP to find the secondary promoter after terminating transcription at the terminator of the primary operon. To compute the distribution of search times, we used the experimental results for D, k_{b}, k_{off}, and $k_{\text{off,s}}$ to simulate the mechanism for a realistic σ^{70} concentration and terminator-promoter spacing (Fig. 4A). The distribution of the search times is roughly exponential with a mean $\langle t_{\text{f}} \rangle \sim 7$ s, which is comparable to the time for transcription initiation at well-studied promoters (a few seconds to a few minutes[26,27]). This indicates that transcription re-initiation by sliding RNAP is capable of effectively increasing expression of the secondary operon.

Inputting the experimental results for D, k_{b}, k_{off}, and $k_{\text{off,s}}$ into Eq. 8, we can predict the value of the efficiency as a function of the distance between operons and the σ^{70} concentration. The total concentration of σ^{70} in *E. coli* is on order 10 μM[28], but its availability is highly regulated through sequestration by anti-σ factors, whose activity is also tightly regulated. This means that at any time, the free σ^{70} concentration could be anything below roughly 10 μM. Thus, the free σ^{70} concentration in the cell could be either above or below $[\sigma^{70}]_c$, which we calculate to be 0.4 μM. To test whether the model predicts appreciable coupling at typical operon spacings, we calculated the predicted coupling efficiency as a function of the distance d between the primary terminator and the secondary promoter at high and low σ^{70} concentrations (Figure 4B). At 5 μM σ^{70}, this calculation predicts efficient coupling at distances d up to $1,000$ bp. At a much lower free σ^{70} concentration of 50 nM, the model predicts a smaller but still significant amount of coupling on this distance scale, with less dependence on operon spacing. Regulation of the free σ^{70} concentration would therefore allow the cell to vary the amount of coupling in response to internal and environmental conditions.

The spacing between the final terminator of an operon and the nearest operon initial promoter has a broad distribution in the *E. coli* genome (Figure 4C). Nevertheless, based on our calculations, a large fraction of these pairs are capable of efficient coupling by sliding RNAP. For example, 52% of the terminator-promoter pairs are at distances where the coupling efficiency is at least 50% at $[\sigma^{70}] = 5$ μM. In other words, at this σ^{70} concentration (and in general when $[\sigma^{70}] >> [\sigma^{70}]_c \approx 0.4$ μM) the predicted critical distance $d_c \approx 1,000$ bp is of the same order of magnitude as the typical inter-operon distance (median 600 bp). This could allow many pairs of adjacent operons in the genome to be coupled, while at the same time enabling other operons to be transcribed independently of their neighbors, depending on the terminator-promoter spacing.

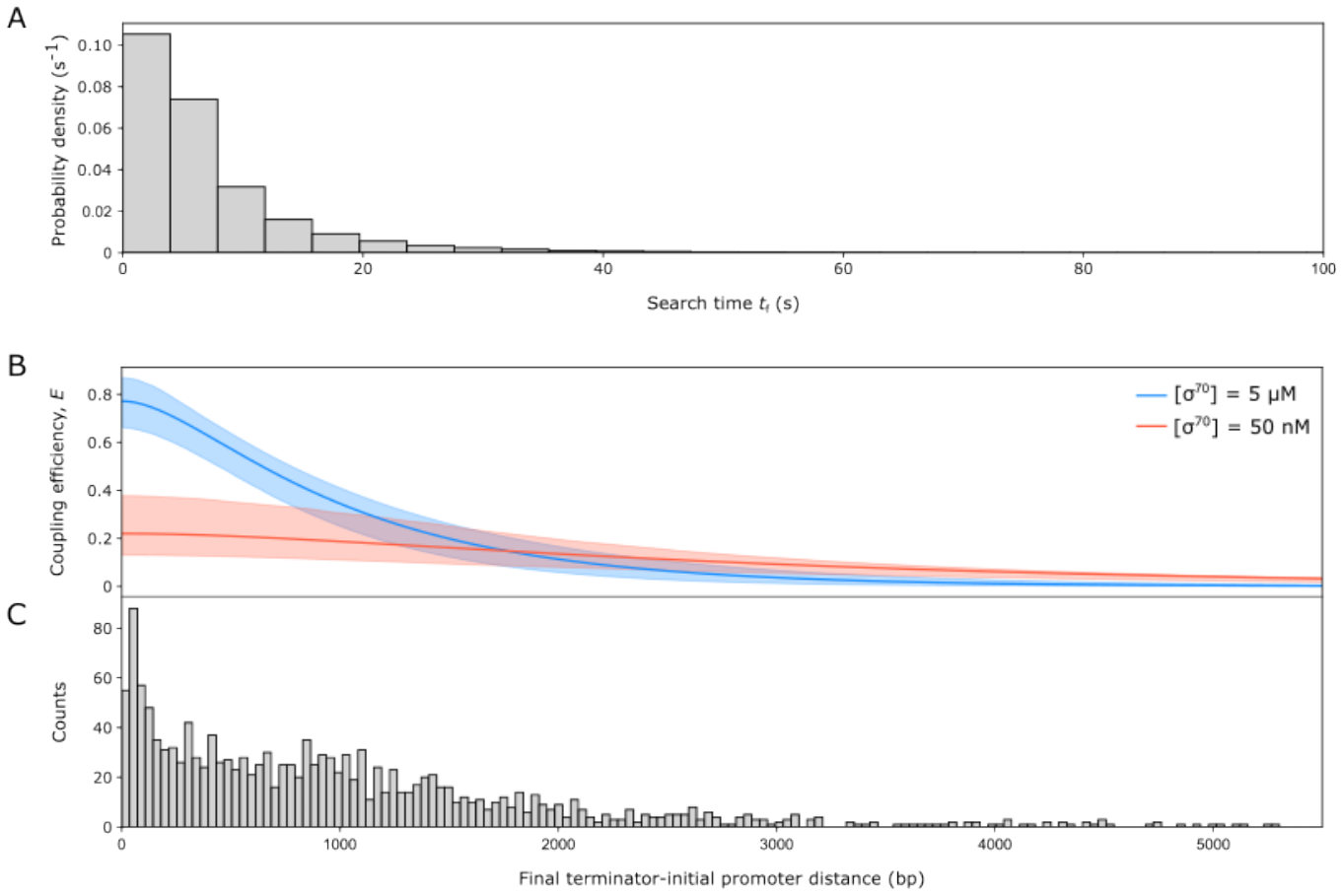

Figure 4: Extent of operon coupling predicted by the sliding RNAP model, using the kinetic parameter values from Table 2. a. Distribution of search times that end in a promoter encounter t_f obtained by simulating the model in Figure 1 for $d = 600$ bp and $[\sigma^{70}] = 5\ \mu$M. **b.** Coupling efficiency dependence on the distance d between primary operon terminator and secondary operon promoter, for two possible free σ^{70} concentrations. Shaded areas show the 68% C.I.s. **c.** Distribution of distances between operon final terminators and the nearest operon initial promoter in the *E. coli* genome determined from data in ref.[29].

Thus, the model predicts significant coupling under relevant cellular conditions and predicts that coupling can be regulated by tuning these conditions.

The model makes the simplifying assumption that every encounter of the RNAP-σ^{70} complex with a promoter is productive and leads to synthesis of a transcript. To the extent that not all encounters are productive, the model will overestimate efficiency (Fig. 4B) and underestimate search time (Fig. 4A).

DISCUSSION

Using a combination of theory, stochastic simulations, and single-molecule microscopy experiments, we characterized the potential spatial and temporal reach of transcriptional coupling between adjacent operons mediated by diffusive sliding of RNAP that remains bound to DNA following transcription termination. We predict that σ^{70} has both stimulatory and inhibitory effects on reinitiation. The stimulatory effect arises from the fact that only RNAP with bound σ^{70} can recognize the secondary promoter. On the other hand, we show that RNAP-σ^{70} has only a short lifetime on DNA, during which the sliding σ^{70}RNAP must find a pro-

moter on the fly to reinitiate transcription. Despite the latter difficulty, we show that reinitiation is expected to be common and efficient for physiological ranges of terminator-promoter spacings and σ^{70} concentrations. Thus, our results show that the proposed reinitiation mechanism is consistent with experiments that demonstrate reinitiation in vitro [16,17] and in vivo [16].

To quantitatively define the reinitiation process, we measured three previously uncharacterized rate constants: the second-order rate constant for binding of σ^{70} to the RNAP-DNA complex, $k_{\text{b}} = 1.2 \times 10^5$ $\text{M}^{-1}\text{s}^{-1}$, the rate constant for the dissociation of the RNAP-DNA complex, $k_{\text{off}} = 1.6 \times 10^{-3}$ s^{-1}, and the rate constant for the dissociation of the RNAP-σ^{70}-DNA complex, $k_{\text{off,s}} = 5.1 \times 10^{-2}$ s^{-1}. The value obtained for k_{b} is an order of magnitude smaller than the rate constant of formation of a stable σ^{70}-RNAP complex in the absence of DNA, 1.5×10^6 $\text{M}^{-1}\text{s}^{-1}$ [21], which suggests that the presence of bound DNA significantly impedes σ^{70} association with RNAP. The value obtained for $k_{\text{off,s}}$ is an order of magnitude smaller than the value obtained for k_1^{D}, the dissociation rate of the RNAP holoenzyme-DNA complex (see Appendix S4 and Table S2). This suggests that the RNAP-σ^{70}-DNA complex (formed by RNAP-DNA binding σ^{70} from solution) and the holoenzyme-DNA complex (formed by mixing σ^{70}RNAP with non-promoter DNA) have different conformations, despite them having the same protein and DNA constituents. It is possible that in the two complexes different subsets of σ^{70} subregions interact with RNAP and/or DNA. More information, kinetic and structural, will be required to understand these differences.

The search for target sequences by proteins sliding on DNA has been demonstrated both in vitro and in vivo (e.g., [30,31]). Post-termination sliding of core RNAP on DNA is atypically slow compared to a sample of other DNA binding proteins [16,32], possibly because RNAP maintains an open bubble of non-base-paired DNA in the post-termination RNAP-DNA complex [19]. The presence in cells of sliding RNAP molecules that may take on order of 10 s after termination to reinitiate transcription (Fig. 4A) is consistent with demonstration of a substantial population in vivo of slowly diffusing RNAP molecules that are neither bound to a fixed site on DNA nor freely diffusing in solution [33,34].

Rapid, efficient reinitiation of transcription through sliding of post-termination RNAP over relevant genomic distances may have significant implications for transcription homeostasis and regulation in both natural and engineered genomes. Under particular growth conditions, transcription activity is often concentrated in clusters of genes or operons in confined genomic regions [4,6,7,35,36]. Sliding-mediated reinitiation may help to maintain a localized pool of RNAP molecules that are efficiently reused in these transcriptionally active regions. Indeed, the efficiency of reinitiation by sliding core RNAP compared to conventional initiation by RNAP holoenzyme from solution may be one of the factors that confers a selective advantage to the clustering of functionally related operons. In the context of synthetic biology, reinitiation by sliding might cause problems by giving rise to non-intended connectivity between transcription units that are intended to act modularly, but conversely could be a used as a tool to introduce correlations in designed genetic circuits.

MATERIALS AND METHODS

Plasmids

Plasmid pDT4 is identical to pCDW116[16] except for mutation of CTGGAGTGCG to CTGGAGACCG to introduce a second BsaI site.

DNA templates

Circular DNA templates DNA^{Cy5}_{Long} and DNA^{Cy5}_{Short} were built by Golden Gate Assembly[37] using a plasmid or PCR product and a synthetic 'ligator" duplex oligonucleotide containing both dye and biotin modifications. Ligator was made by annealing two complementary oligonucleotides: 5′-CGATTAGGTCTCGGGCTAGTAC TGGTTTCTAGAG/iCy5/GTTCCAAGCC/iBiodTCACGGCGGCCGCCCATCGAGACCGGTTAACC-3′ and 5′-GGTTA ACCGGTCTCGATGGGCGGCCGCCGTGAGGCTTGGAACCTCTAGAAACCAGTACTAGCCCGAGACCTAATCG-3′ (IDT).

For making template DNA^{Cy5}_{Long}, two identical Golden Gate Assembly reactions were carried out by mixing 7 μl of each $\sim$ 20 nM DT4 plasmid and $\sim$ 20 nM ligator fragment with 1 μl of Golden Gate Mix (New England Biolabs) in T4 DNA Ligase Buffer (New England Biolabs), in total volumes of 20 μl. The mixtures were incubated for alternating cycles of 5 min at 37^{o}C and 10 min at 16^{o}C 35 times, followed by 5 min at 55^{o}C. After the reaction, the ligase was inactivated for 10 min at 65^{o}C. The resulting 40 μl of reaction product was mixed with 4 μl of T5 Exonuclease (New England Biolabs) in NEB Buffer 4, to a total of 50 μl and incubated at 37^{o}C for 30 min. The digestion was stopped by adding 15 mM of EDTA, and a Qiagen PCR Cleanup Kit was used to remove the cleaved nucleotides and enzymes.

For making template DNA^{Cy5}_{Short}, a linear DNA fragment was first amplified by PCR from plasmid pCDW114 ([16], Addgene #70061), using primers 5′-GAAGGTCTCCAGCCGTACCAACCAGCGGCTTATC-3′ and 5′-CCGGG TCTCACCATACCCGCTGTCTGAGATTACG-3′. A Golden Gate Assembly reaction was carried out by mixing 2 μl of 343 nM PCR product, 0.6 μl of 1 μM ligator and 1 μl of Golden Gate Mix in T4 DNA Ligase Buffer, in a total volume of 20 μl. The mixture was incubated for alternating cycles of 5 min at 37^{o}C and 10 min at 16^{o}C 35 times, followed by 5 min 55^{o}C. After the reaction, the ligase was inactivated for 10 min at 65^{o}C. The resulting reaction product was mixed with 1 μl Exonuclease V (New England Biolabs) and 3 μl of 10 mM ATP in NEB Buffer 4, in a total volume of 30 μl, and incubated at 37^{o}C for 30 min. Exonuclease V was then inactivated for 30 min at 70^{o}C. Finally, a Qiagen PCR Cleanup Kit was used to remove the cleaved nucleotides and enzymes.

Proteins

Fluorescent labeling of core RNAP

E. coli core RNAP with a SNAP tag on the C-terminus of β'[38] (RNAP-SNAP, gift from the Robert Landick lab) was labeled with SNAP-Surface 549, yielding $RNAP^{549}$, as follows: 13.65 μM SNAP-RNAP (core) and 45.5 μM of SNAP-Surface 549 were mixed in a buffer containing 9 mM Tris-Cl^- pH 7.9, 5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, and 90 mM NaCl, and incubated for 30 min at room temperature. The sample was then mixed

with an equivalent amount of Dilution buffer (11 mM Tris-Cl$^-$ pH 8.0, 30% glycerol, 110 mM NaCl, 1 mM DTT), flash-frozen in liquid nitrogen and stored at -80^{o}C.

Expression and purification of His-tagged σ^{70}

His$_6$-tagged *E. coli* σ^{70} (σ^{70}) was overexpressed in T7 Express cells (New England Biolabs) as inclusion bodies from the pET-28a-σ^{70} overexpression plasmid[39] by growing the cells at 37^{o}C to an OD$_{600}$ of ~ 0.8, and then inducing by addition of IPTG to 0.4 mM. The temperature was decreased to 20^{o}C and cells were left shaking at 200 rpm overnight. Cells were then harvested by centrifugation at 4^{o}C, followed by sonication in Lysis buffer (50 mM Tris-Cl$^-$, pH 7.9, 5 mM imidazole, 5% [v/v] glycerol, 233 mM NaCl, 2 mM EDTA, 10 mM β-mercaptoethanol) plus 1× cOmplete™protease inhibitor cocktail (Roche). The lysate was centrifuged at 22,000 × g for 30 min at 4^{o}C and the supernatant was discarded. To remove *E. coli* membrane and cell wall material, the pellet was resuspended in 10 ml of 2 M Urea Cleaning buffer (20 mM Tris-Cl$^-$ pH 8.0, 500 mM NaCl, 2 M urea, 2% Triton X-100, 10 mM β-mercaptoethanol) and sonicated. The resulting sample was centrifuged again at 22,000 × g for 30 min at 4^{o}C and the supernatant was discarded. Four consecutive resuspension-centrifugation cycles were carried out, two of them in 2 M Urea Cleaning buffer, and the other two in Wash buffer (20 mM Tris-Cl$^-$ pH 8.0, 500 mM NaCl, 7% glycerol, 20 mM imidazole, 10 mM β-mercaptoethanol) to remove Triton X-100 from the pellet. To solubilize and denature the protein, the washed pellet was resuspended in 6 M Guanidine Binding buffer (20 mM Tris-Cl$^-$ pH 8.0, 500 mM NaCl, 5 mM imidazole, 6 M guanidine hydrochloride, 2 mM β-mercaptoethanol), stirred for 1 hour, and centrifuged at 22,000 × g for 30 min at 4^{o}C. The supernatant was passed through a 0.22 μm filter and injected into a 1 ml HisTrap column (Cytiva Life Sciences), followed by a wash with Wash buffer supplemented with 6 M urea. Refolding of the bound protein was performed using a linear 1-hour-long 6 M to 0 M urea gradient in Wash buffer. The refolded protein was eluted with 500 mM imidazole in Wash buffer. The purified protein was dialyzed overnight into σ^{70} storage buffer (10 mM Tris-Cl$^-$, pH 8.0, 30% [v/v] glycerol, 0.1 mM EDTA, 100 mM NaCl, 20 μM $ZnCl_2$, 1 mM $MgCl_2$, 0.1 mM DTT) and aliquots were flash-frozen in liquid N_2 and stored at -80^oC.

Fluorescent labeling of σ^{70}

An N-terminal His$_6$-tagged single-cysteine derivative of *E. coli* σ^{70} (C132S C291S C295S S366C)[40] (see Appendix S4) was overexpressed and purified following the same protocol used for σ^{70}. The purified protein was concentrated 5× using an Amicon Ultra-0.5ml 30K filter by centrifuging for 11 minutes at 14,000 × g at 4^{o}C. For fluorescent labeling, the concentrated protein was mixed with Cy5-maleimide dye (Cytiva) (1:15 protein:dye ratio), incubated first for 10 min at room temperature, and then left overnight at 4^o C. The excess dye was then removed using a Centrispin 20 column (Princeton Separations). After addition of glycerol and BSA to 30% and 1 mg/ml respectively, the samples were flash-frozen in liquid N_2, and aliquots were stored at -80^{o}C.

Reconstitution of doubly-labeled holoenzyme

Cy5-σ^{70}RNAP^{549} holoenzyme (see Appendix S4) was reconstituted by incubating $121\ n$M of RNAP^{549} and $280\ n$M of Cy5-σ^{70} for 30 min at 37^oC.

Colocalization Single-Molecule Spectroscopy (CoSMoS) Experiments

Single-molecule total internal reflection fluorescence microscopy was performed[41] at excitation wavelengths 532 and $633\ n$m, for observation of DNA^{Cy5} template (and/or Cy5-σ^{70}) and RNAP^{549}, respectively; focus was automatically maintained[42]. A stage heating device was used to keep the samples at 30^oC. Single-molecule observations were performed in glass flow chambers (volume $\sim 30\ \mu$l) passivated with a mPEG-SG2000:biotin-PEG-SVA5000 (Laysan Bio) $200:1$ w/w mixture as described in[43]. Neutravidin (#21125; Life Technologies) was introduced at $220\ n$M in KO buffer ($50\ m$M TrisOAc, $100\ m$M KOAc, $8\ m$M Mg(OAC)$_2$, $27\ m$M NH_4(OAc), $0.1\ m$g/ml bovine serum albumin (BSA) (#126615 EMB Chemicals), pH 8.0), incubated for 45 s, and washed out (this and all subsequent wash steps used two washes each of two chamber volumes of KO buffer). The chamber was then incubated with $\sim 1\ n$M Cy5-DNA (DNA$^{Cy5}_{Short}$ or DNA$^{Cy5}_{Long}$) in KO buffer for ~ 20 min and washed out with Imaging buffer (KO buffer supplemented with an O_2 scavenging system: 4.5 mg/ml glucose, 40 units/ml glucose oxidase, and $1,500$ units/ml catalase[43]).

For the experiments to measure the dwell time of RNAP^{549} on DNA at different concentrations of σ^{70}, $\sim 1\ n$M of RNAP^{549} was introduced into the chamber in Imaging buffer supplemented with $3.5\%\ w/v$ PEG $8,000$ and $1\ m$g/ml BSA for ~ 10 min. Image acquisition was performed by alternating 1 s exposures to 532 and $633\ n$m, at 450 and $200\ \mu$W respectively (all laser powers were measured incident to the micromirror optics), and the flow chamber was washed with Imaging buffer supplemented with $3.5\%\ w/v$ PEG $8,000$ and containing 0, $148\ n$M, $297\ n$M, $593\ n$M, $741\ n$M or $1.19\ \mu$M of σ^{70}. In the cases where the concentration of σ^{70} was lower than $1.19\ \mu$M, the appropriate amount of σ^{70} storage buffer was added in replacement so that all experiments were performed at the same solute concentrations: $47.0\ m$M TrisOAc, pH 8.0, $93.0\ m$M KOAc, $7.4\ m$M Mg(OAc)$_2$, $25\ m$M NH_4(OAc), $3\%\ w/v$ $8,000$ PEG, $0.2\ m$M Tris-Cl$^-$, $2.0\ m$M NaCl, $0.4\ \mu$M $ZnCl_2$, 20 μM $MgCl_2$, $4.5\ m$g/ml glucose, 40 units/ml gluclose oxidase, $1,500$ units/ml catalase, 0.6% glycerol, $2\ \mu$M EDTA, $0.1\ m$g/ml BSA, $2\ m$M DTT, $10\ n$M DTT-quenched Cy5.5 maleimide dye.

The experiments to measure dwell time of σ^{70}RNAP on DNA (see Appendix S4) were performed similarly to the ones described in the previous paragraph, with three differences. First, a photobleaching step was performed after DNA surface attachment. Cy5 photobleaching was induced by $633\ n$m excitation at ~ 1 mW in the presence of Imaging buffer without DTT. Second, instead of σ^{70}, $\sim 1\ n$M of Cy5-σ^{70}RNAP^{549} (which also contained an additional $1.5\ n$M Cy5-σ^{70}) was introduced into the chamber in Imaging buffer supplemented with $3.5\%\ w/v$ PEG $8,000$ and no subsequent wash was performed. The final composition of the solution was $47.5\ m$M of TrisOAc, pH 8.0, $95.1\ m$M KOAc, $7.4\ m$M Mg(OAc)$_2$, $25.7\ m$M NH_4(OAc), $3.3\%\ w/v$ $8,000$ PEG, $0.1\ m$g/ml BSA, $1\ m$M DTT, $0.1\ m$M Tris-Cl$^-$, 0.3% glycerol, $1\ m$M NaCl, $8\ \mu$M $MgCl_2$, $36\ n$M $ZnCl_2$, $1.8\ \mu$M

EDTA, $1.1\,n$M SNAP-RNAP, $2.5\,n$M unreacted SNAP-Surface 549, $2.5\,n$M Cy5-σ^{70}, $4.5\,m$g/ml glucose, 40 units/ml glucose oxidase, $1,500$ units/ml catalase. Third, image acquisition was performed by continuous exposure to 532 and $633\,n$m lasers, at 450 and $200\,\mu$W respectively, at an acquisition rate of 1 frame per second.

CoSMoS Data analysis

Analysis of CoSMoS video recordings was done using custom software and algorithms for mapping between wavelength channels, spatial drift correction, and detection of spot colocalization as described[41]. In each recording, we selected $\text{DNA}^{\text{Cy5}}_{\text{Long}}$ or $\text{DNA}^{\text{Cy5}}_{\text{Short}}$ fluorescence spots that co-localized with RNAP^{549} spots at $t = 0$. For the selected DNA molecules, we computed RNAP^{549} fluorescence intensity time records by summing the intensity over 3×3 pixel squares centered at DNA molecule locations in each recorded frame. Fluorescence intensity values were corrected for background fluorescence and non-uniform illumination across the microscope field of view, yielding normalized values for spot intensities[44]. This allowed us to directly compare integrated intensity values for different spots located throughout the field of view. The numbers of decreasing intensity steps in the resulting time traces were counted to assess the initial number of RNAP^{549} molecules present at each DNA molecule location (Fig. S2A). Records that showed more than a single RNAP^{549} molecule bound at $t = 0$ were excluded from subsequent analysis. The times of the first image with no spot at each DNA location were taken to be the dwell times of the RNAP^{549} molecules present at the beginning of the recording. Spots that persisted until the end of the recording were separately counted as censored dwell times equal to the recording duration.

Fits to RNAP-DNA complex dwell time distributions

The probability distribution of RNAP-DNA complex dwell times measured in each individual experiment was modeled as the sum of two exponential terms

$$f_1 = a_{\text{fast}} k_{\text{fast}} \exp\left(-k_{\text{fast}} t\right) + (1 - a_{\text{fast}}) k_{\text{slow}} \exp\left(-k_{\text{slow}} t\right) \tag{13}$$

or the sum of three exponential terms

$$\begin{aligned} f_2 = & a_{\text{fast}} k_{\text{fast}} \exp\left(-k_{\text{fast}} t\right) + a_{\text{inter}} k_{\text{inter}} \exp\left(-k_{\text{inter}} t\right) \\ & + (1 - a_{\text{inter}} - a_{\text{fast}}) k_{\text{slow}} \exp\left(-k_{\text{slow}} t\right) . \end{aligned} \tag{14}$$

For each distribution, lifetimes of RNAP^{549} binding events that terminated by disappearance of the fluorescent spot, and those that were censored by the end of the experiment, were jointly fit using the maximum likelihood algorithm by an approach analogous to the one used in[45]. Confidence intervals were calculated by bootstrapping[41].

Extraction of model parameters k_{b}, k_{off}, and $k_{\mathrm{off,s}}$

To get values for k_{b}, k_{off}, and $k_{\mathrm{off,s}}$, we globally fit dwell times collected at all concentrations of σ^{70} to the models in equation 13 (for $[\sigma^{70}] = 0$) or equation 14 (for $[\sigma^{70}] > 0$), where k_{fast} was explicitly constrained to depend on $[\sigma^{70}]$:

$$k_{\mathrm{fast}} = k_{\mathrm{off}} + \frac{k_{\mathrm{off,s}}\, k_{\mathrm{b}}\, [\sigma^{70}]}{k_{\mathrm{off,s}} + k_{\mathrm{b}}\, [\sigma^{70}].} \tag{15}$$

In this formulation, k_{inter} and k_{slow} are assumed to be independent of $[\sigma^{70}]$, and therefore were globally fit across distributions obtained for template $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$. For the experiment performed with template $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Short}}$, we fit independent values k'_{inter} and k'_{slow} for these parameters. Censored data were treated using the same approach as described above for the individual experiment fits to dwell time distributions.

Fit to holoenzyme-DNA dwell time distribution

To account for non-specific binding of holoenzyme to the glass flow chamber (see Appendix S4), we first randomly selected locations on the chamber surface that did not contain DNA molecules. We then fit the distribution of dwell times for Cy5-σ^{70}RNAP^{549} holoenzyme molecules bound at these non-DNA locations to a biexponential model

$$f_{\mathrm{nD}}(t) = \left[a_1^{\mathrm{nD}} k_1^{\mathrm{nD}} \exp\left(-k_1^{\mathrm{nD}} t\right) + (1 - a_1^{\mathrm{nD}}) k_2^{\mathrm{nD}} \exp\left(-k_2^{\mathrm{nD}} t\right)\right] / f_{\mathrm{nD}}^0, \tag{16}$$

with

$$f_{\mathrm{nD}}^0 = a_1^{\mathrm{nD}} \exp\left(-k_1^{\mathrm{nD}} t_{\mathrm{min}}\right) + (1 - a_1^{\mathrm{nD}}) \exp\left(-k_2^{\mathrm{nD}} t_{\mathrm{min}}\right), \tag{17}$$

where $f_{\mathrm{nD}}(t)$ is normalized so that it integrates to 1 over all dwell times t greater than the minimum detectable dwell time $t_{\mathrm{min}} = 0.2$ s.

By analogy to[41], the dwell time distribution for Cy5-σ^{70}RNAP^{549} at DNA locations was then fit to the background-corrected model

$$\begin{aligned} f_{\mathrm{D}}(t) =& (F_{\mathrm{D}} - F_{\mathrm{nD}}) \left[a_1^{\mathrm{D}} k_1^{\mathrm{D}} \exp\left(-k_1^{\mathrm{D}} t\right) + a_2^{\mathrm{D}} k_2^{\mathrm{D}} \exp\left(-k_2^{\mathrm{D}} t\right)\right. \\ &\left. + (1 - a_1^{\mathrm{D}} - a_2^{\mathrm{D}}) k_3^{\mathrm{D}} \exp\left(-k_3^{\mathrm{D}} t\right)\right] / f_{\mathrm{D}}^0 + F_{\mathrm{nD}} f_{\mathrm{nD}}^{\mathrm{ML}}(t), \end{aligned} \tag{18}$$

with

$$\begin{aligned} f_{\mathrm{D}}^0 =& a_1^{\mathrm{D}} \exp\left(-k_1^{\mathrm{D}} t_{\mathrm{min}}\right) + a_2^{\mathrm{D}} \exp\left(-k_2^{\mathrm{D}} t_{\mathrm{min}}\right) \\ &+ (1 - a_1^{\mathrm{D}} - a_2^{\mathrm{D}}) \exp\left(-k_3^{\mathrm{D}} t_{\mathrm{min}}\right), \end{aligned} \tag{19}$$

where F_{D} and F_{nD} represent the total binding frequency at DNA and non-DNA locations, respectively, and

$f_{\mathrm{nD}}^{\mathrm{ML}}$ stands for f_{nD} evaluated on the maximum likelihood estimators obtained by fitting the non-DNA locations data. As before, we jointly fit the censored and uncensored data using the maximum likelihood method[45].

Simulation of search times

In order to calculate the mean search time $\langle t_{\mathrm{f}} \rangle$ required for the RNAP to find the secondary promoter, we used numerical simulations of the mechanism depicted in Fig. 1. In particular, we used the Gillespie algorithm[46] to generate the stochastic trajectory of an RNAP molecule on a DNA molecule. In the simulation, the state of the system is characterized by the position of the RNAP on DNA and whether it is bound to σ^{70} or not. Initially, the RNAP is at position $x = 0$ (which corresponds to the position of the terminator from the primary transcription unit), and it is not bound to σ^{70}. We then draw a time t_1 at random from an exponential distribution with $p(t) = \lambda \exp(-\lambda t)$ with $\lambda = k_{\mathrm{b}}[\sigma^{70}] + k_{\mathrm{off}}$, and choose between two possible transitions: binding σ^{70} or dissociating from the DNA. Which of the transitions takes place is chosen at random according to their relative weights $\frac{k_{\mathrm{b}}[\sigma^{70}]}{k_{\mathrm{b}}[\sigma^{70}]+k_{\mathrm{off}}}$ and $\frac{k_{\mathrm{off}}}{k_{\mathrm{b}}[\sigma^{70}]+k_{\mathrm{off}}}$. If the RNAP dissociates from the DNA, its attempt to find the secondary promoter is considered unsuccessful, and the simulation starts over with a new trial. If the RNAP binds σ^{70}, the position on the DNA x_1 at which binding occurs is dependent on the amount of diffusion away from the primary terminator. x_1 is determined by drawing at random from a normal distribution with mean $\mu = 0$ and variance $\mathrm{var} = 2Dt_1$. The time t_2 that is required to diffuse from x_1 to the secondary promoter located at x_p is then drawn at random from the first passage time density

$$p(t_2) = \frac{|x_p - x_1|}{\sqrt{4\pi D t_2}} \exp\left(-\frac{(x_p - x_1)^2}{4Dt_2}\right).$$

An RNAP-σ^{70} complex dissociation time t_3 is drawn at random from an exponential distribution with $\lambda = k_{\mathrm{off,s}}$. If $t_2 < t_3$, the RNAP-σ complex is considered to have found the secondary promoter at time $t_{\mathrm{f}} = t_1 + t_2$. If $t_2 > t_3$, the attempt to find the secondary promoter was unsuccessful. The whole process is repeated multiple times, generating a distribution of search times.

Genome-wide analysis of terminator-promoter distances

Using the promoter and terminator annotations reported by Conway et al[29], we measured the distance of each operon-ending terminator to the nearest operon initial promoter in the *E. coli* genome.

ACKNOWLEDGMENTS

We thank Bob Landick, and Rachel Mooney, and for providing us with plasmids and proteins. We would like to thank members of the Landick, Kondev and Gelles labs for insightful discussion. We are grateful to Johnson Chung for help with microscopy, and with Liuyu Chen for help in plasmid preparation. This work was funded by grants from NIGMS (R01 GM081648 to J.G.), from NSF (DMR-1610737 to JK), and the Simons Foundation (to JK).

AUTHOR CONTRIBUTIONS

D.T., L.F., J.G. and J.K. designed the research; D.T., K.I. and A.C. performed the research. D.T. analyzed the data; D.T., J.G., and J.K. drafted the manuscript and all authors contributed to writing the final version.

AUTHOR COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

[1] A Feklístov, BD Sharon, SA Darst, CA Gross, Bacterial sigma factors: A historical, structural, and genomic perspective. *Annual Review of Microbiology* **68**, 357–376 (2014).

[2] RR Burgess, AA Travers, JJ Dunn, EK Bautz, Factor stimulating transcription by RNA polymerase. *Nature* **221**, 43–46 (1969).

[3] DF Browning, SJ Busby, Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology* **14**, 638–650 (2016).

[4] M Fischbach, CA Voigt, Prokaryotic gene clusters: A rich toolbox for synthetic biology. *Biotechnology Journal* **5**, 1277–1296 (2010).

[5] F Jacob, D Perrin, C Sanchez, J Monod, L'opéron: groupe de gènes à expresion coordonnée par un opérateur. *Comptes Rendus Academie des Sciences Paris* **250**, 514–520 (1960).

[6] H Zhang, Y Yin, V Olman, Y Xu, Genomic Arrangement of Regulons in Bacterial Genomes. *PLOS ONE* **7**, e29496 (2012).

[7] Y Yin, H Zhang, V Olman, Y Xu, Genomic arrangement of bacterial operons is constrained by biological pathways encoded in the genome. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6310–6315 (2010).

[8] M Zampieri, N Soranzo, D Bianchini, C Altafini, Origin of co-expression patterns in E. coli and S. cerevisiae emerging from reverse engineering algorithms. *PLoS ONE* **3**, 1–10 (2008).

[9] JO Korbel, LJ Jensen, CV Mering, P Bork, Analysis of genomic context: Prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nature Biotechnology* **22**, 911–917 (2004).

[10] L Pannier, E Merino, K Marchal, J Collado-Vides, Effect of genomic distance on coexpression of coregulated genes in E. coli. *PLoS ONE* **12**, 1–20 (2017).

[11] S Kruglyak, H Tang, Regulation of adjacent yeast genes. *Trends in Genetics* **16**, 109–111 (2000).

[12] EJ Williams, DJ Bowles, Coexpression of Neighboring Genes in the Genome of Arabidopsis thaliana. *Genome Research* **14**, 1060–1067 (2004).

[13] LF Liu, JC Wang, Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7024–7027 (1987).

[14] J Tan, L Shu, HY Wu, Activation of the leu-500 promoter by adjacent transcription. *Journal of Bacteriology* **176**, 1077–1086 (1994).

[15] KY Rhee, et al., Transcriptional coupling between the divergent promoters of a prototypic LysR-type regulatory system, the ilvYC operon of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 14294–14299 (1999).

[16] TT Harden, et al., Alternative transcription cycle for bacterial RNA polymerase. *Nature Communications* **11**, 448 (2020).

[17] W Kang, et al., Transcription reinitiation by recycling RNA polymerase that diffuses on DNA after releasing terminated RNA. *Nature Communications* **11**, 450 (2020).

[18] W Kang, S Hwang, JY Kang, C Kang, S Hohng, Hopping and flipping of rna polymerase on dna during recycling for reinitiation after intrinsic termination in bacterial transcription. *International Journal of Molecular Sciences* **22**, 1–13 (2021).

[19] L You, et al., Structural basis for intrinsic transcription termination. *Nature 2023* pp. 1–7 (2023).

[20] H Maeda, N Fujita, A Ishihama, Competition among seven Escherichia coli subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Research* **28**, 3497–3503 (2000).

[21] FYH Wu, LR Yarbrough, CW Wu, Conformational Transition of Escherichia coli RNA Polymerase Induced by the Interaction of σ Subunit with Core Enzyme. *Biochemistry* **15**, 3254–3258 (1976).

[22] S Redner, *A Guide to First-Passage Processes.* (Cambridge University Press), (2001).

[23] KM Arndt, MJ Chamberlin, Transcription termination in Escherichia coli. *Journal of Molecular Biology* **202**, 271–285 (1988).

[24] K Inlow, D Tenenbaum, L Friedman, J Kondev, J Gelles, Mechanism of Post-Termination Recycling of E. coli RNA Polymerase by the Swi2/Snf2 ATPase RapA (manuscript in preparation). (2023).

[25] V Vogt, Breaks in DNA stimulate transcription by core RNA polymerase. *Nature* **223**, 854–855 (1969).

[26] S Choubey, J Kondev, A Sanchez, Distribution of Initiation Times Reveals Mechanisms of Transcriptional Regulation in Single Cells. *Biophysical Journal* **114**, 2072–2082 (2018).

[27] I Golding, J Paulsson, SM Zawilski, EC Cox, Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036 (2005).

[28] IL Grigorova, NJ Phleger, VK Mutalik, CA Gross, Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5332–5337 (2006).

[29] T Conway, et al., Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *mBio* **5** (2014) e01442-14.

[30] PC Blainey, AM Van Oijen, A Banerjee, GL Verdine, XS Xie, A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5752–5757 (2006).

[31] P Hammar, et al., The lac repressor displays facilitated diffusion in living cells. *Science* **336**, 1595–1598 (2012).

[32] PC Blainey, et al., Nonspecifically bound proteins spin while diffusing along DNA. *Nature Structural and Molecular Biology* **16**, 1224–1229 (2009).

[33] M Stracy, et al., Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E4390–E4399 (2015).

[34] M Stracy, et al., Transient non-specific DNA binding dominates the target search of bacterial DNA-binding proteins. *Molecular cell* **81**, 1499–1514 (2021).

[35] G Fang, EP Rocha, A Danchin, Persistence drives gene clustering in bacterial genomes. *BMC Genomics* **9**, 4 (2008).

[36] P Nicolas, et al., Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. *Science* **335**, 1103–1106 (2012).

[37] KA Padgett, JA Sorge, Creating seamless junctions independent of restriction sites in PCR cloning. *Gene* **168**, 31–35 (1996).

[38] LE Tetone, et al., Dynamics of GreB-RNA polymerase interaction allow a proofreading accessory protein to patrol for transcription complexes needing rescue. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E1081–E1090 (2017).

[39] JP Mumm, LJ Friedman, J Gelles, Mechanism of upstream promoter element stimulation of transcription at a ribosomal RNA promoter determined by single-molecule imaging. *bioRxiv* (2020) doi: https://doi.org/10.1101/2020.02.17.953182.

[40] S Callaci, E Heyduk, T Heyduk, Conformational Changes of Escherichia coli RNA Polymerase 70 Factor Induced by Binding to the Core Enzyme. *Journal of Biological Chemistry* **273**, 32995–33001 (1998).

[41] LJ Friedman, J Gelles, Multi-wavelength single-molecule fluorescence analysis of transcription mechanisms. *Methods (San Diego, Calif.)* **86**, 27 (2015).

[42] DJ Crawford, AA Hoskins, LJ Friedman, J Gelles, MJ Moore, Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA (New York, N.Y.)* **14**, 170–179 (2008).

[43] LJ Friedman, J Gelles, Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell* **148**, 679–689 (2012).

[44] LD Jesús-Kim, et al., DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7. *eLife* **10**, 1–83 (2021).

[45] JP Bombardier, et al., Single-molecule visualization of a formin-capping protein 'decision complex' at the actin filament barbed end. *Nature communications* **6**, 8707 (2015).

[46] DT Gillespie, Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry* **81**, 2340–2361 (2002).

Supporting Information for

RNA polymerase sliding on DNA can couple the transcription of nearby bacterial operons

Debora Tenenbaum, Koe Inlow, Larry Friedman, Anthony Cai, Jeff Gelles, Jane Kondev

Corresponding authors: Jeff Gelles, Jane Kondev
E-mail: gelles@brandeis.edu, kondev@brandeis.edu

This PDF file includes:

Supporting Information Text

Appendix S1: Approximation of coupling distance $d_c^{\approx}$ for the case $[\sigma^{70}] \approx [\sigma^{70}]_c$.

$d_c^{\approx}$ is defined in the main text as the distance for which the coupling efficiency E decays by a factor of e: $E(d_c^{\approx}) = E(0)e^{-1}$, which in this case converts to

$$k_b[\sigma^{70}] \exp\left(-\sqrt{\frac{k_{\text{off,s}}}{D}} d_c^{\approx}\right)[d_c^{\approx}(\sqrt{k_{\text{off,s}}D})^{-1} + k_{\text{off,s}}^{-1}] = k_b[\sigma^{70}]/(ek_{\text{off,s}}) \ . \quad [1]$$

Defining $x = \sqrt{\frac{k_{\text{off,s}}}{D}} d_c^{\approx}$, this reduces to the equation,

$$e^{x-1} = x + 1, \quad [2]$$

whose solution is given by

$$x = -\text{W}(-e^{-2}) - 1, \quad [3]$$

where $\text{W}(z)$ is the product log function. The only real positive solution to eqn. 3 is given by

$$x = -\text{W}_{-1}(-\text{e}^{-2}) - 1 \approx 2.15, \quad [4]$$

which gives $d_c^{\approx} \approx 2.15\sqrt{\frac{D}{k_{\text{off,s}}}}$.

Appendix S2: General approximation of coupling distance.

Values for model parameters D, $k_{\text{b}}[\sigma^{70}]$, k_{off}, and $k_{\text{off,s}}$ were randomly drawn from log-uniform distributions in the ranges:
$D : (10^3 - 10^6)$ bp^{2}s^{-1}
$k_{\text{b}}[\sigma^{70}] : (10^{-3} - 10^2)$ s^{-1}
$k_{\text{off}} : (10^{-4} - 1)$ s^{-1}
$k_{\text{off,s}} : (10^{-3} - 10^2)$ s^{-1}.

The characteristic coupling distance d_c (i.e., the distance for which the coupling efficiency decreases by a factor of e) was numerically calculated from Eq. 8, and plotted against the approximated value $\sqrt{D(k_b[\sigma^{70}] + k_{\text{off}})^{-1}} + \sqrt{D(k_{\text{off,s}})^{-1}}$ (Fig. S1), confirming that the coupling distance can be approximated by

$$d_c \approx \sqrt{D(k_b[\sigma^{70}] + k_{\text{off}})^{-1}} + \sqrt{D(k_{\text{off,s}})^{-1}} \quad [5]$$

for all parameter values tested.

Appendix S3: Selection of DNA templates that colocalized with a single RNAP^{549} molecule at $t = 0$.

Because of the length of the DNA$^{\text{Cy5}}$ templates employed (3,033 bp for DNA$^{\text{Cy5}}_{\text{Long}}$ and 586 bp for DNA$^{\text{Cy5}}_{\text{Short}}$), multiple RNAP^{549} could be bound to the same DNA molecule simultaneously. We restricted the analysis to those DNA$^{\text{Cy5}}_{\text{Long}}$ or DNA$^{\text{Cy5}}_{\text{Short}}$ spots that colocalized with a single RNAP^{549} molecule at $t = 0$.

To quantify the number of RNAP^{549} molecules bound to each DNA, We used the integrated fluorescence time records. In the cases where the RNAP^{549} spots present at $t = 0$ disappeared during the duration of the experiment, the number of decreasing intensity steps was counted. Fig. S2A shows examples for one and two RNAP^{549} molecules present initially. DNA templates containing more than two RNAP^{549} molecules bound were rare.

For the cases in which the RNAP^{549} spots present at $t = 0$ did not disappear during the experiment, the absolute integrated intensity value was used to estimate the number of RNAP^{549} molecules present (Fig. S2B).

Appendix S4: Experimental evidence for and interpretation of the intermediate and slow dwell time distribution components

In addition to the σ^{70}-dependent component characterized by the dissociation rate k_{fast}, we observed components in the RNAP-DNA dwell time distributions that did not change systematically with σ^{70} concentration. We investigated the origin of these σ^{70}-independent components as follows:

Our preparation of RNAP^{549} showed low levels of σ^{70}, indicating contamination of the core RNAP with σ^{70}RNAP holoenzyme (Fig. S5). This raised the possibility that RNAP dwell time components with the σ^{70}-independent rate constants k_{inter} and/or k_{slow} are due to σ^{70}RNAP. To test this hypothesis,we measured the dwell times of σ^{70}RNAP on DNA and compared them to dwell time components measured for the RNAP-σ^{70}-DNA (RNAP-DNA complex that had bound σ^{70} from solution). Specifically, we performed a control experiment with doubly-labeled σ^{70}RNAP holoenzyme (Cy5-σ^{70}RNAP^{549}) on template DNA$^{\text{Cy5}}_{\text{Long}}$ (Fig. S6A) and scored only dwell times at DNA locations where there was simultaneous colocalization of RNAP^{549} and Cy5-σ^{70} spots. This criterion excluded from our measurements dwell times of core RNAP molecules that did not have

bound σ^{70}. When the Cy5-σ^{70} fluorescent spot disappeared before the corresponding RNAP^{549} spot (e.g., Fig. S6B, right), the dwell time of the latter was measured.

In this experiment, roughly $10-15\%$ of $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$ spots showed colocalization with more than one RNAP^{549} molecule bound simultaneously, often accompanied by Cy5-σ^{70}, at some intervals during the recording. To simplify the interpretation of the data, we removed those records from the analysis, retaining only records that showed at most one core RNAP molecule at a time (e.g., Fig. S6B).

A fit of the resulting dwell time distribution (Fig. S6C) to a three-exponential background-corrected model (Eq. 18) (1) yielded the rate constants shown in Table S2. The fastest rate k_1^{D} is approximately consistent with time constants previously observed for σ^{70}RNAP on nonspecific DNA (2). k_2^{D} and k_3^{D} are approximately k_{inter} and k_{slow}, respectively (Figs. S6D and E). This is consistent with the hypothesis that k_{inter} and k_{slow} are due to the minor σ^{70} contamination of the SNAP-RNAP sample used, and do not represent σ^{70}-induced dissociation of RNAP from DNA.

The relative amount of the component with rate constant k_{slow}, $a_{\mathrm{slow}} = 1 - a_{\mathrm{fast}} - a_{\mathrm{inter}}$, was roughly constant across experiments that used the same DNA preparation, even when $[\sigma^{70}]$ was different. Unexpectedly, a_{slow} was different between experiments using different preparations of the same DNA, and was almost absent in one of them (Fig. S7A). We propose that the events from this longest time component are due to stable binding of RNAP^{549} or σ^{70}RNAP^{549} to some kind of imperfections in a fraction of the DNA molecules, such as nicks or gaps, which may be a consequence of the method used to prepare the DNAs and therefore will in general differ in abundance for different DNA preparations. In contrast, the fraction of the other (i.e., non-slow) dwells that are in the intermediate component, $a_{\mathrm{inter}}/(a_{\mathrm{inter}} + a_{\mathrm{fast}})$, was roughly constant for two different $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$ preparations and for the two different DNA sequences, $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Long}}$ and $\mathrm{DNA}^{\mathrm{Cy5}}_{\mathrm{Short}}$ (Fig. S7B). A possible explanation for the binding events with characteristic dissociation rate k_{inter} is that they arise from holoenzyme binding to rare tight binding sequences (3–5) that may be present in both of the DNA circles tested. More experiments will be needed to test this hypothesis.

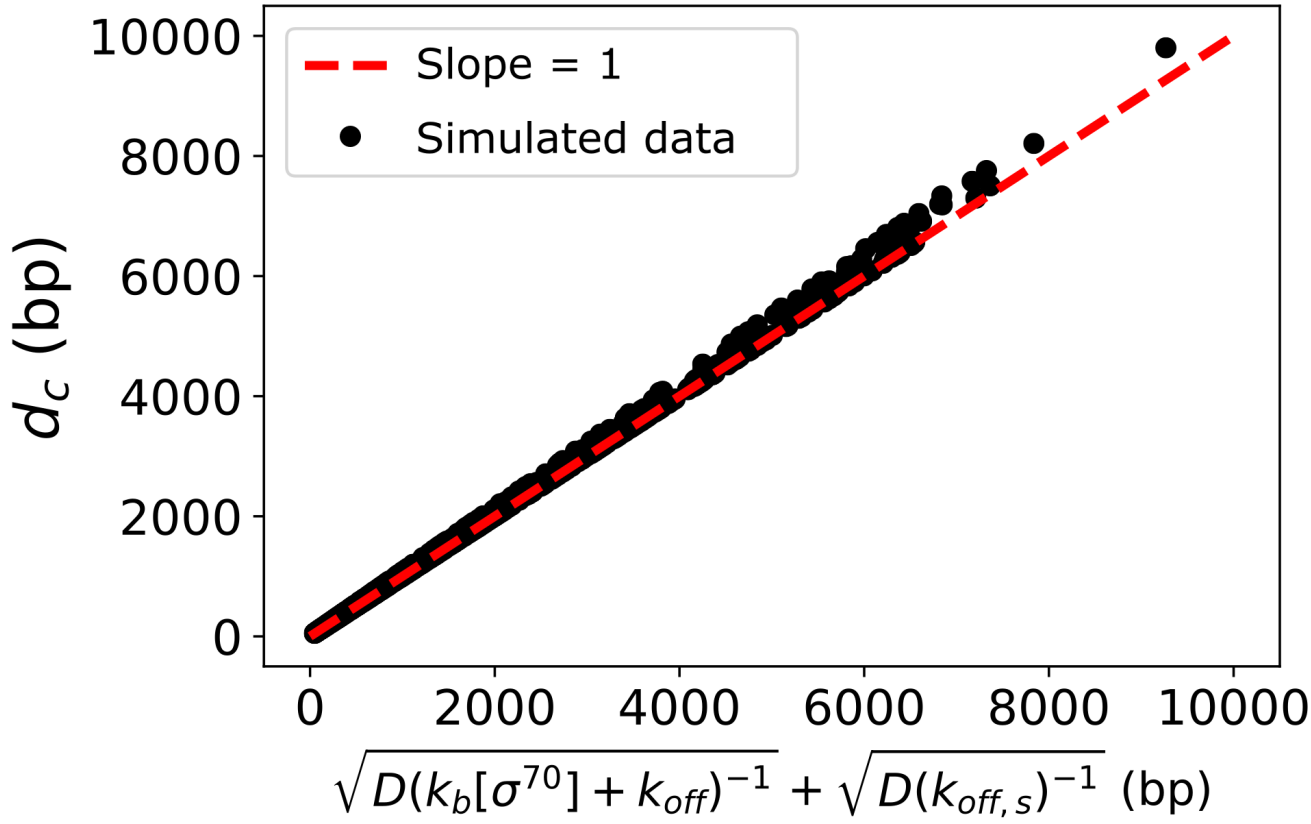

Fig. S1. Approximation of coupling distance d_c. Coupling distance was numerically calculated from Equation 8 as the value of d over which the efficiency E decreases by a factor of e, using parameter values chosen at random (black dots, see Appendix S2). That exact calculation agrees well with our approximate expression for d_c (red line), confirming that d_c can be approximated by $\sqrt{D(k_b[\sigma^{70}]+k_{\text{off}})^{-1}} + \sqrt{D(k_{\text{off,s}})^{-1}}$.

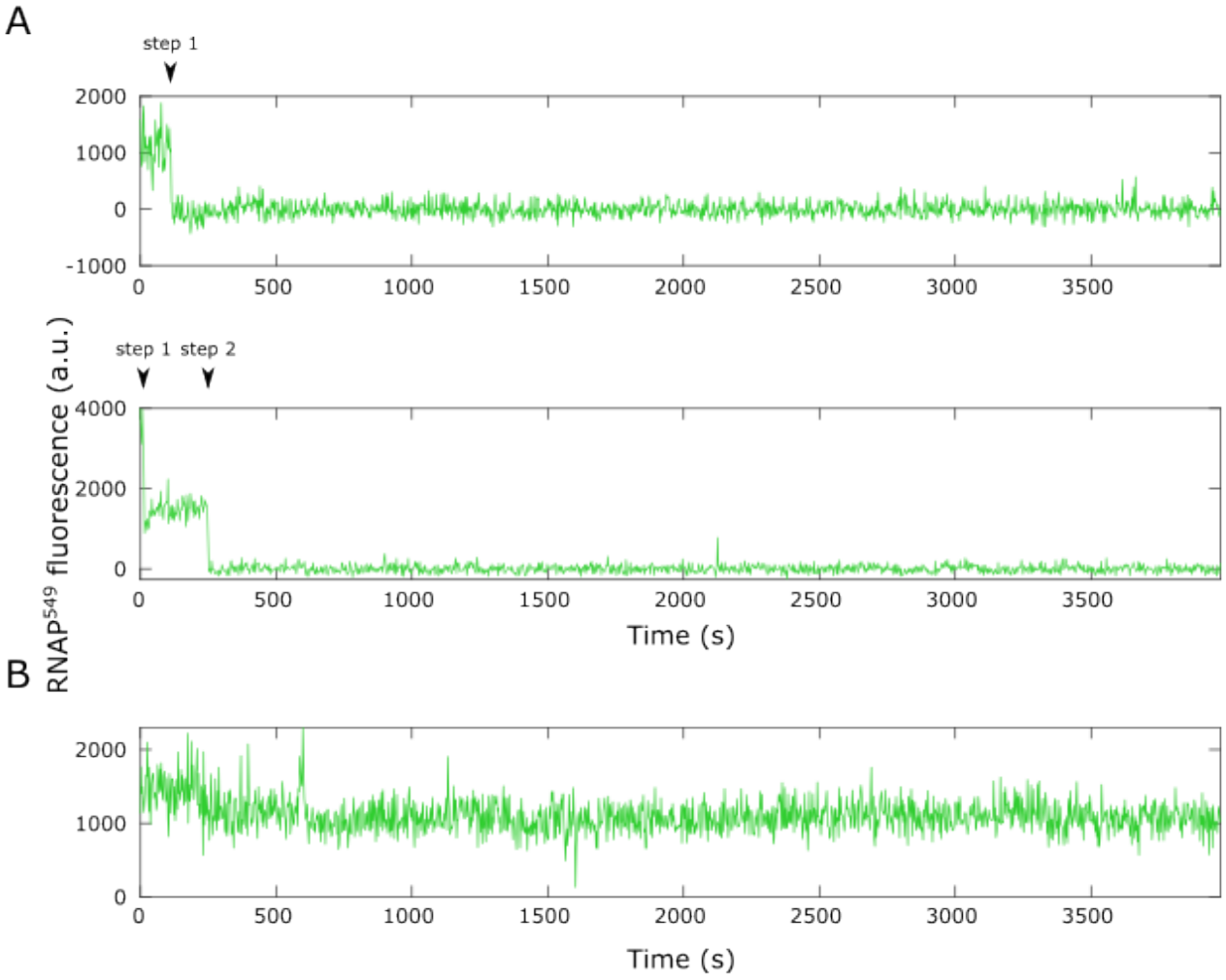

Fig. S2. Counting the number of RNAP^{549} molecules bound to an individual DNA molecule at $t = 0$. A. Records showing one (top) and two (bottom) RNAP^{549} molecules that dissociate or photobleach during the experiment. **B**. Record showing a single RNAP^{549} molecule that lasts throughout the experiment, as indicated by the roughly constant normalized fluorecesce intensity of $\sim 1,000$. In all records, zero fluorescence corresponds to the diffuse background fluorescence detected in the absence of a DNA-colocalized RNAP^{549} molecule.

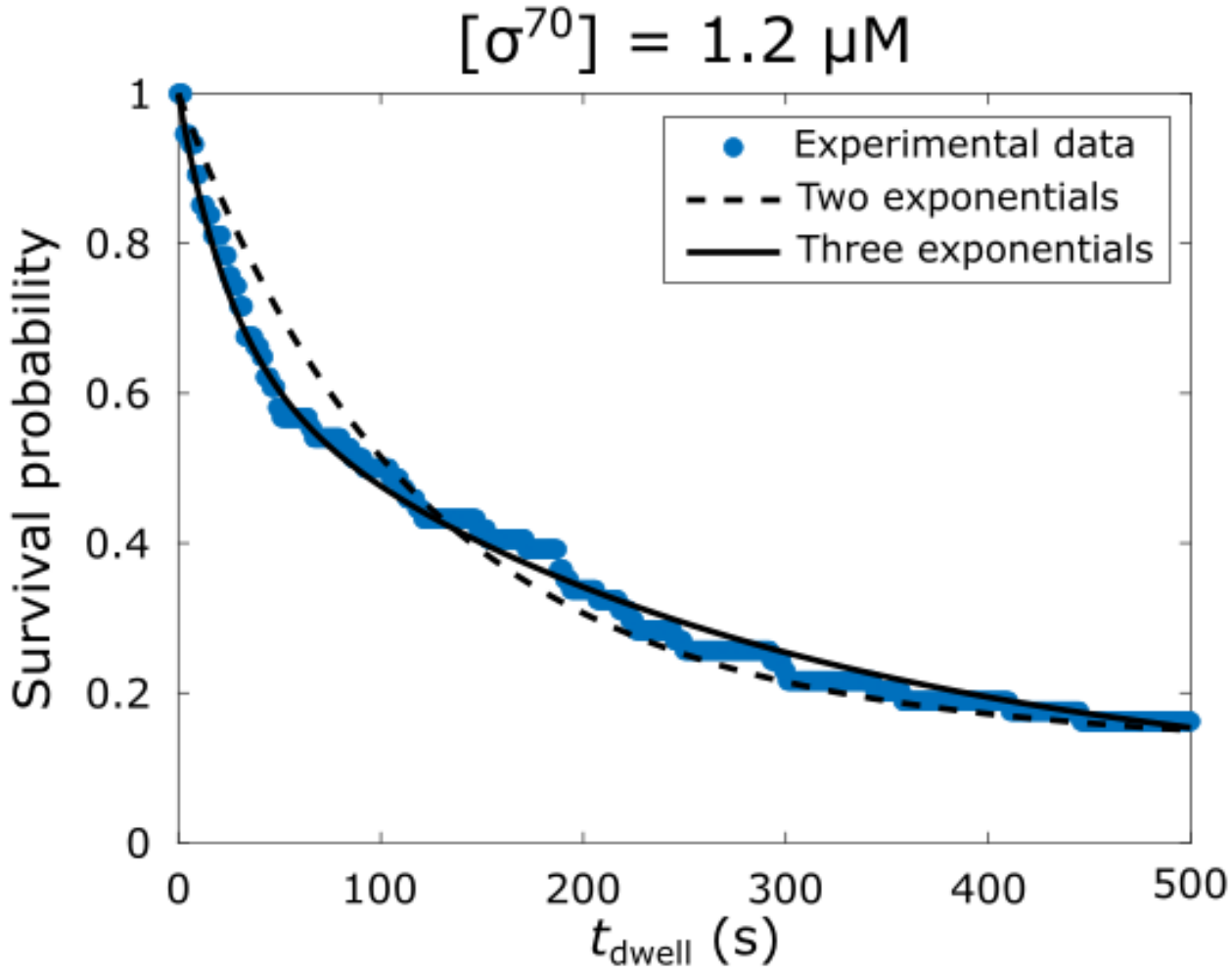

Fig. S3. Comparison between two-exponential and three-exponential fits to an example RNAP^{549} − DNA$^{Cy5}_{Long}$ dwell time distribution. The distribution obtained at 1.19 μM σ^{70} (blue) was fit to a sum of two exponentials model (Eq. 13 with $a_{\text{fast}} = 8.2(6.9-9.2)\times10^{-1}$, $k_{\text{fast}} = 8.7(6.0-14.2)\times10^{-3}$ s^{-1}, $k_{\text{slow}} = 4.7(1.2-10.0)\times10^{-4}$ s^{-1}; dashed line) and a sum of three exponentials model (Eq. 14 with $a_{\text{fast}} = 3.4(2.4-4.9)\times10^{-1}$, $k_{\text{fast}} = 3.9(2.4-5.4)\times10^{-2}$ s^{-1}, $a_{\text{inter}} = 5.8(4.4-6.7)\times10^{-1}$, $k_{\text{inter}} = 3.9(2.5-5.2)\times10^{-3}$ s^{-1}, $k_{\text{slow}} = 1.6(0-3.5)\times10^{-4}$ s^{-1}; solid line). The values are presented with 68% C.I.s.

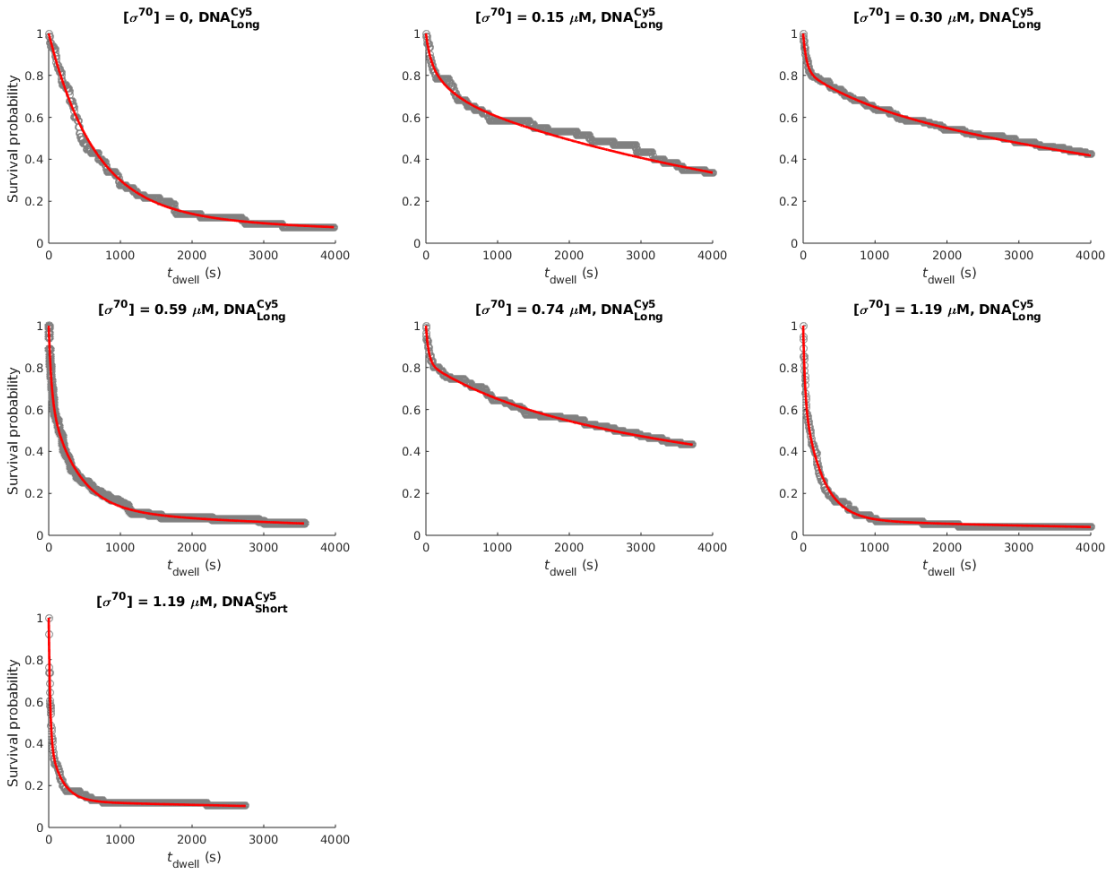

Fig. S4. RNAP^{549} – DNA dwell time distributions and local fits with two different DNA constructs in the presence of different concentrations of σ^{70}. Each distribution was fit to the sum of two exponentials (for $[\sigma^{70}] = 0$) or the sum of three exponentials (for $[\sigma^{70}] > 0$). The fit parameters are reported in Table 1.

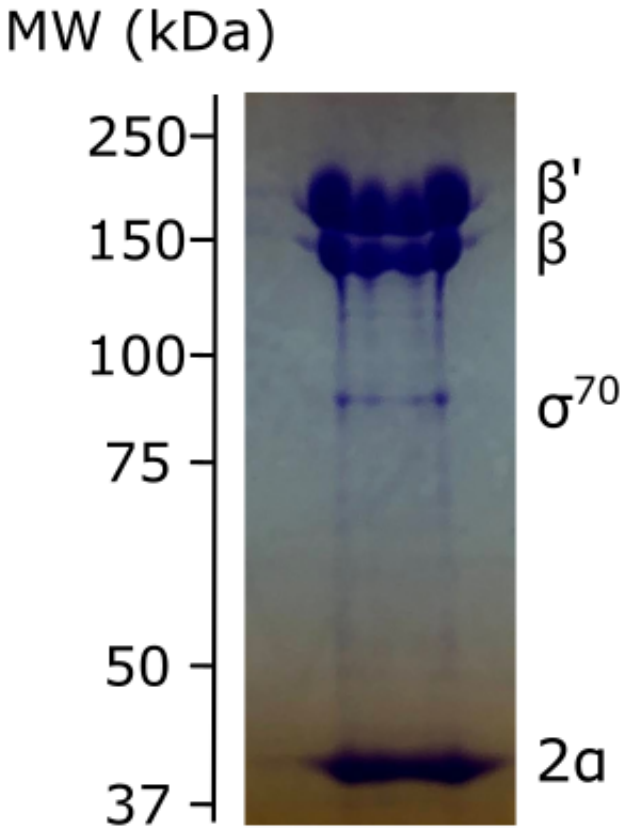

Fig. S5. Extent of σ^{70} contamination in the RNAP-SNAP preparation used in this study. Densitometry of the Coomassie-Blue stained SDS-PAGE gel indicates the presence of σ^{70} at ~ 7.5 mole percent of core.

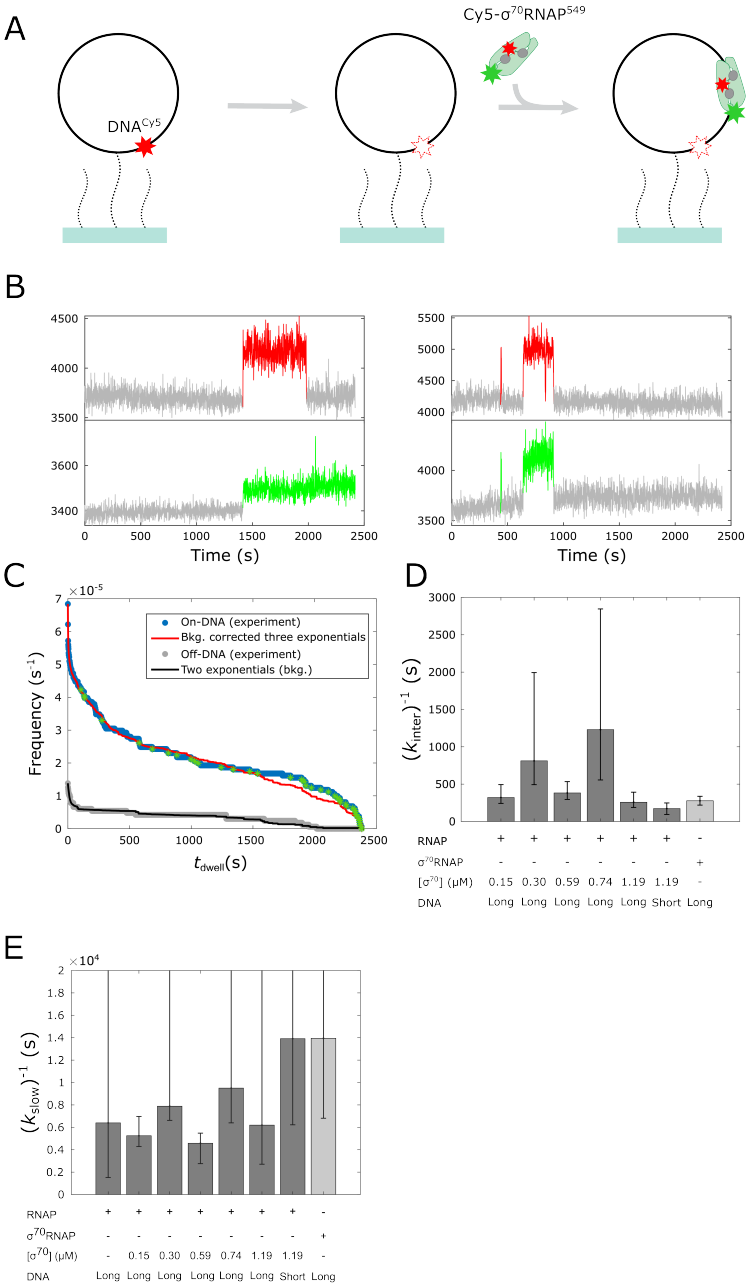

Fig. S6. Control experiment with doubly-labeled σ^{70}RNAP holoenzyme. See Appendix S4. **A.** Experiment design. Fluorescently labeled, promoterless circular DNA molecules (DNA^{Cy5}_{Long}; black circles with red stars) were tethered to the surface of a glass flow chamber (blue) through polyethylene glycol linkers (dotted black curves). After recording the position of the DNA molecules, the fluorescent dye was photobleached (white star). Doubly-labeled RNAP holoenzyme Cy5-σ^{70}RNAP^{549} was added and two-color fluorescence time records were acquired. **B.** Examples of two-color Cy5-σ^{70}RNAP^{549} fluorescence time records recorded at the locations of two individual DNA molecules. Time intervals in which the fluorescently labeled protein subunits were detected are indicated (color). **C.** Cumulative frequency distributions for Cy5-σ^{70}RNAP^{549} molecules binding to DNA^{Cy5}_{Long} (blue) or binding to non-DNA positions (gray) with a dwell time greater than or equal to the specified value. Stars indicate censored dwell times that last until the end of the experiment. Figure also shows a two-exponential distribution function (black) fit to the non-DNA data and a three-exponential background-corrected distribution function (red) fit to the DNA data (1). Distribution function plots were numerically simulated to include the effects of censoring. The fit parameters are summarized in Table S2. **D.** The characteristic durations (with 68% C.I.s) of intermediate-length binding events from experiments with unlabeled σ^{70} (dark gray; Fig. 3A; k_{inter} in table 1) is consistent with the characteristic duration of intermediate-length specific binding events of Cy5-σ^{70}RNAP^{549} holoenzyme (light gray; see panel C; k_2 in Table S2). **E.** Same as D, but for long-duration time constants. While C.I. lower bounds are well determined, most of the C.I. upper bounds are poorly defined because the reciprocal rate constants exceed the durations of the experiments, which range from $2,749$ s to $4,570$ s.

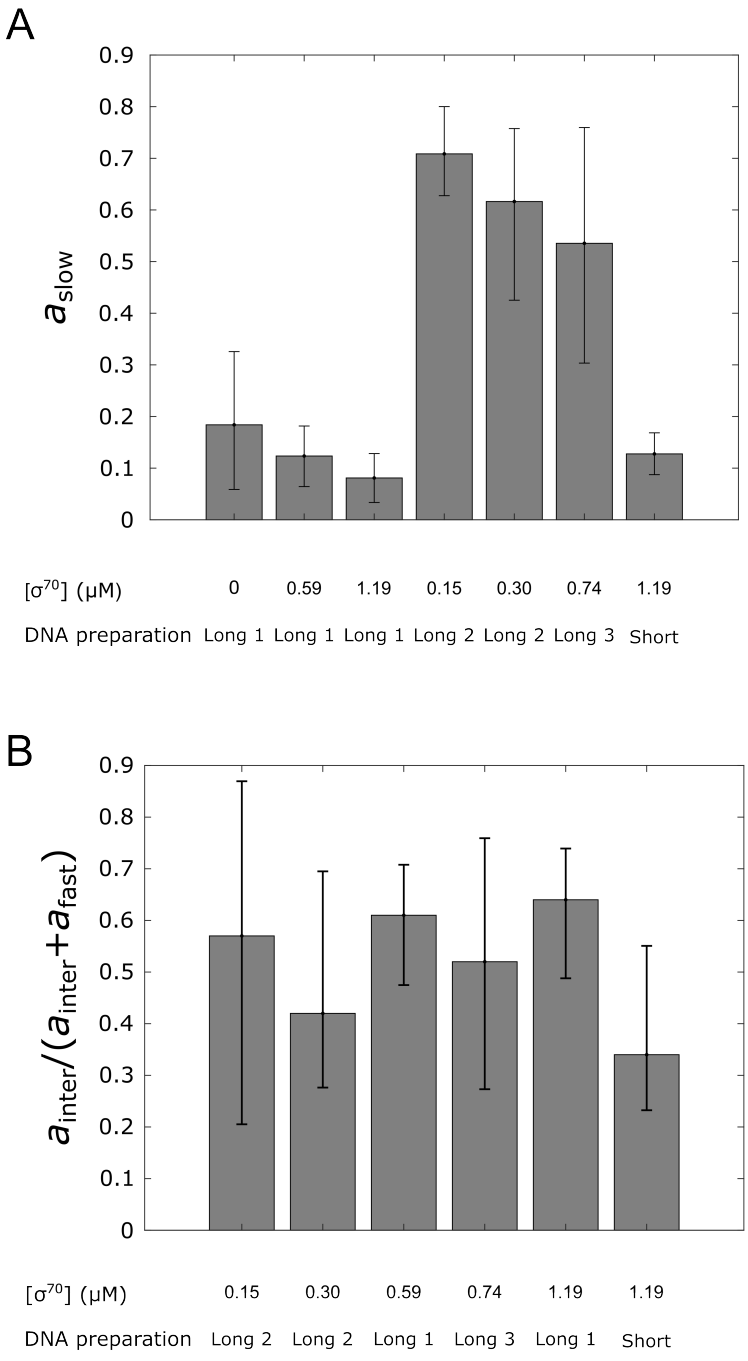

Fig. S7. Amplitudes of the slow and intermediate dwell time components in experiments with different DNA preparations and different $[\sigma^{70}]$. **A.** Relative amplitude of the slow dissociation component a_{slow} for experiments performed using three different preparations of circular DNA DNA^{Cy5}_{Long}, all built following the same protocol, and one preparation of DNA^{Cy5}_{Short}. The value of $a_{slow} = 1 - a_{inter} - a_{fast}$ (Eq. 14) depends on the DNA templates preparation used, supporting the hypothesis that the longest binding events are due to stable binding to imperfections in some of the DNA molecules. **B.** Fraction of the non-slow dwells that are in the intermediate component, as given by $a_{inter}/(a_{inter} + a_{fast})$ in the same experiments as in (A). This quantity is roughly constant across experiments on all DNA constructs and preparations.

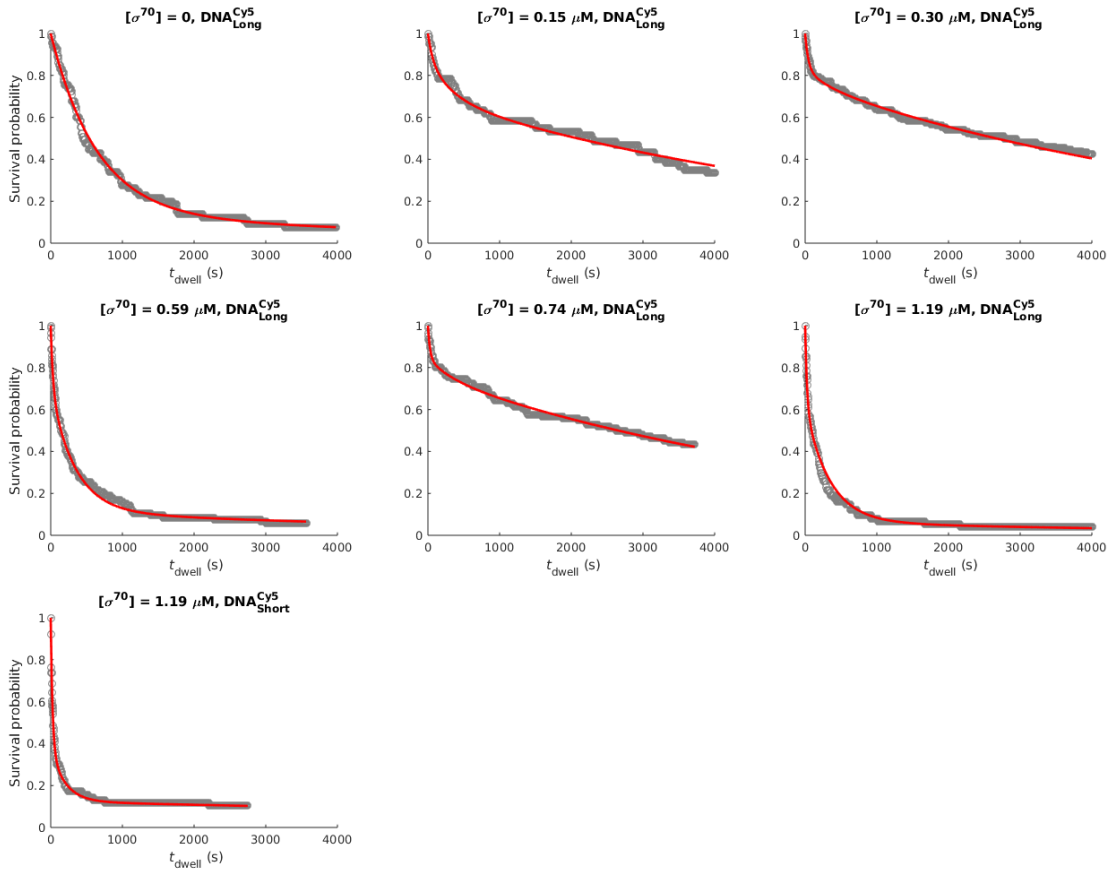

Fig. S8. RNAP^{549} – DNA dwell time distributions and global fits with two different DNA constructs in the presence of different concentrations of σ^{70}. Displayed fits (red) are to the global model in which k_{inter} and k_{slow} are fixed across experiments, and k_{fast} is constrained to a hyperbolic dependence on σ^{70} concentration. This contrasts with Fig. S4 in which we used independent local parameter values for each dataset. Parameters for the fits shown here are summarized in Table 2 and Table S3. Experimental data (gray) are the same as in Fig. S4.

Table S1. Fraction of DNA and non-DNA locations that had one RNAP^{549} molecule bound at $t = 0$.

	Fraction of locations	
$[\sigma^{70}](\mu\text{M})$	DNA	non-DNA
0	18%	1.8%
0.15	15%	0.5%
0.30	31%	0.9%
0.59	17%	1.6%
0.74	15%	0.7%
1.19	25%	1.3%

The non-DNA binding was measured by observing the binding of RNAP^{549} molecules at random positions in the microscope field of view that contained no visible DNA$^{\text{Cy5}}_{\text{Long}}$ molecules.

Table S2. Fit parameters for distribution of Cy5-σ^{70}RNAP^{549} dwell times on DNA$^{\text{Cy5}}_{\text{Long}}$.

Non-DNA					
F_{nD} $(\times 10^{-5}\ \text{s}^{-1})$	a_1^{nD} $(\times 10^{-1})$	k_1^{nD} $(\times 10^{-2}\ \text{s}^{-1})$	k_2^{nD} $(\times 10^{-4}\ \text{s}^{-1})$		
1.4	$5.2\ (4.2-6.1)$	$6.0\ (5.0-9.0)$	$2.0\ (1.0-4.0)$		
DNA					
F_{D} $(\times 10^{-5}\ \text{s}^{-1})$	a_1^{D} $(\times 10^{-1})$	k_1^{D} $(\times 10^{-1}\ \text{s}^{-1})$	$a_2^{\text{D}}(\times 10^{-1})$	k_2^{D} $(\times 10^{-3}\ \text{s}^{-1})$	k_3^{D} $(\times 10^{-5}\ \text{s}^{-1})$
6.8	$2.4(1.9-3.0)$	$5.1\ (3.8-6.4)$	$3.8(3.0-4.4)$	$3.6\ (3.0-4.6)$	$7.2\ (0-14.7)$

Non-DNA data represents nonspecific binding of Cy5-σ^{70}RNAP^{549} to the surface of the glass flow chamber, and was taken into account in fitting the DNA data. The models used for the fits and their parameters are defined in Eqs. 16 and 18. Values are presented with their 68% C.I.s.

Table S3. Fit parameters for survival lifetime probability distributions obtained by globally fixing k_{inter} and k_{slow} and considering k_{fast} dependence on $[\sigma^{70}]$.

$[\sigma^{70}]$ (μM)	N	a_{fast} ($\times 10^{-1}$)	a_{inter} ($\times 10^{-1}$)	k_{inter} ($\times 10^{-3}$ s^{-1})	k_{slow} ($\times 10^{-4}$ s^{-1})
0	65	$8.6(7.9-9.2)$	-	$2.9(2.5-3.5)$	$1.6(1.4-1.8)$
0.15	60	$1.2(0.4-2.0)$	$1.8(0.8-2.9)$	"	"
0.30	139	$1.7(1.3-2.1)$	$0.7(0.1-1.2)$	"	"
0.59	106	$3.0(2.3-3.8)$	$5.8(5.1-6.5)$	"	"
0.74	106	$1.6(1.1-2.0)$	$0.8(0.3-1.4)$	"	"
1.19	74	$3.8(2.9-4.7)$	$5.5(4.7-6.5)$	"	"
1.19*	74	$6.2(5.1-7.0)$	$2.6(1.8-3.6)$	$5.1(3.7-7.8)$	$0.7(0-1.3)$

($*$) Experiment performed with a 586 bp long circular template made from a different DNA sequence ($\text{DNA}^{\text{Cy5}}_{\text{Short}}$). N is the number of DNA sites with co-localized RNAP that were analyzed. The values are presented with 68% C.I.s

References

1. LJ Friedman, J Gelles, Multi-wavelength single-molecule fluorescence analysis of transcription mechanisms. *Methods (San Diego, Calif.)* **86**, 27 (2015).
2. TT Harden, et al., Alternative transcription cycle for bacterial RNA polymerase. *Nature Communications* **11**, 448 (2020).
3. P Melançon, MT Record, RR Burgess, Nitrocellulose Filter Binding Studies of the Interactions of Escherichia coli RNA Polymerase Holoenzyme with Deoxyribonucleic Acid Restriction Fragments: Evidence for Multiple Classes of Nonpromoter Interactions, Some of Which Display Promoter-like Proper. *Biochemistry* **21**, 4318–4331 (1982).
4. AM Huerta, J Collado-Vides, Sigma70 promoters in Escherichia coli: specific transcription in dense regions of overlapping promoter-like signals. *Journal molecular biology* **333**, 261–278 (2003).
5. M Kawano, G Storz, BS Rao, JL Rosner, RG Martin, Detection of low-level promoter activity within open reading frame sequences of Escherichia coli. *Nucleic Acids Research* **33**, 6268 (2005).