

Transcription attenuation in synthetic promoters in tandem formation

Vatsala Chauhan¹, Ines S. C. Baptista¹, Rahul Jagadeesan¹, Suchintak Dash¹, and Andre S. Ribeiro^{1,*}

¹Faculty of Medicine and Health Technology, Tampere University, 33520, Finland.

*Correspondence: andre.sanchesribeiro@tuni.fi

ABSTRACT

Closely spaced promoters are ubiquitous in prokaryotic and eukaryotic genomes. How their structure and dynamics relate remains unclear, particularly for tandem formations. To study their transcriptional interference, we engineered two pairs and one trio of synthetic promoters in non-overlapping, tandem formation, in single-copy plasmids. From *in vivo* measurements in *E. coli* cells, we found that promoters in tandem formation have attenuated transcription rates. The attenuation strength can be widely fine-tuned by the promoters' positioning, natural regulatory mechanisms, and other factors, including the antibiotic rifampicin, which hampers RNAP promoter escape. From this, and supported by *in silico* models, we concluded that the attenuation emerges from premature terminations generated by collisions between RNAPs elongating from upstream promoters and RNAPs occupying downstream promoters. Moreover, we found that these collisions can cause one or both RNAPs to fall-off. The broad spectrum of possible, externally regulated, attenuation strengths in synthetic tandem promoters should make these structures valuable internal regulators of future synthetic circuits.

Keywords: Tandem promoter formations, transcription attenuation, synthetic genetic circuits.

INTRODUCTION

Closely spaced promoters in convergent, divergent, and tandem geometries are widely present in living organisms, including *Escherichia coli*¹⁻⁷. They are known to have high conservation levels^{2,8}, but it remains unclear what are their selective advantages.

Provided that the elongation regions of two genes overlap, RNAPs starting from their promoters can be expected to interact. Studies reported⁹ that promoters in convergent formation have weakened RNA production, due to collisions between the RNAPs elongating from one TSS and RNAPs bound to the other. Meanwhile, in promoters in tandem formation, RNAPs bound to one promoter can block (at least transiently) the RNAPs elongating from the other promoter¹⁰. Moreover, natural genes controlled by tandem promoters that overall, i.e. whose transcription start sites (TSSs) are closer than ~35 bps, have significantly weaker expression rates (on average) than more distanced ones¹¹. Likely, if overlapping, when one RNAP occupies one TSS, it occludes the other TSS from RNAP binding.

It should be possible to use the interactivity between RNAPs of promoters closely spaced, in tandem formation, as a means to engineer synthetic genes with fine-tuned dynamics. Models have explored the potential dynamics of promoters in tandem formation as a function of properties such as the component promoters' strength^{11,12}. However, empirical validation is largely lacking and the potential influence of other parameters (e.g., transcription factor regulation) remains largely unexplored. Also, empirical data is lacking on the outcome of collisions between elongating RNAPs and RNAPs bound to promoters, and how it could be regulated (e.g., by tuning promoter escape rates).

Here, we studied interference between promoters in tandem formation. For this, we engineered synthetic constructs of promoters in tandem formation using, as building blocks, three genetically modified natural promoters: $P_{LacO3O1}$, P_{tetA} , and P_{BAD} ^{13–15}. The internal composition of our synthetic constructs is illustrated in Figure 1. In all constructs, the promoter(s) control the expression of an mCherry protein to track their transcription dynamics. We used the constructs to study interference as a function of the regulation state of the component promoters, and when subject to an antibiotic that directly interferes with transcription initiation¹⁶.

From Figure 1, the fluorescence probes of the tandem constructs allow tracking the overall dynamics of the component promoters (Figures 1A₄ to 1A₆), but not the dynamics of each component promoter independently. For that, we used individual promoter constructs (Figures 1A₁ to 1A₃). From here onwards, we referred to the constructs carrying promoters in tandem formation as “tandem promoters” (e.g., tandem promoters “ $P_{tetA}^U-P_{LacO3O1}^D$ ”). Meanwhile, we referred to the constructs carrying a single promoter as “individual promoters” (e.g., “individual promoter P_{tetA} ”).

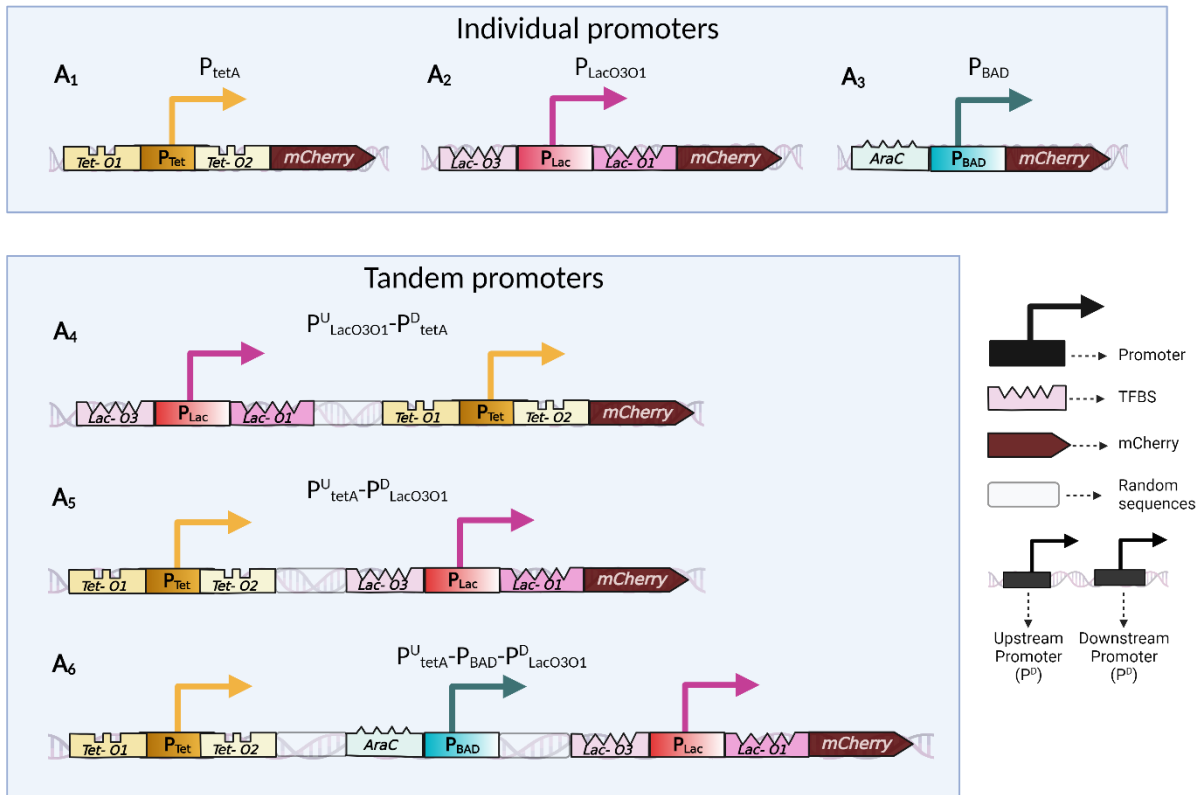


Figure 1: Schematic representation of the synthetic promoters in tandem formation along with their transcription factor binding sites. Bioparts A₁-A₆ were each inserted into single-copy pBAC plasmids, respectively. Each biopart is followed by an mCherry coding region. Bioparts A₁, A₂, and A₃ are the individual promoter constructs. Bioparts A₄, A₅, and A₆ are the constructs with (pairs or trios of) tandem promoters. The component promoters in individual formation are P_{tetA}, P_{LacO3O1}, and P_{BAD}, respectively. The two dual synthetic promoters in tandem formation are P^U_{LacO3O1}-P^D_{tetA} and P^U_{tetA}-P^D_{LacO3O1}, respectively, (U and D stand for upstream and downstream respectively). Finally, the trio of promoters in tandem formation is P^U_{tetA}-P_{BAD}-P^D_{LacO3O1}.

RESULTS

The promoters used to build the tandem promoters have been profusely studied. P_{LacO3O1} was engineered from the natural Lac promoter of the lactose operon in *E. coli*¹⁷, by removing the operator O₂^{18,19}. This decreases by 2- to 3-fold the repression strength of the wild-type tetrameric Lac repressor (LacI)¹⁸. The homotetrameric Lac repressor protein represses by binding the DNA operator sequences²⁰. The binding forces DNA loops^{21,22} that make the promoter less accessible²³. Contrarily, IPTG, a structural analog of the natural inducer Lactose²⁴, indirectly induces P_{LacO3O1} by binding to LacI, which hampers LacI's ability to bind to the promoter²⁰.

The second promoter used, P_{tetA}, was extracted in its natural from the Tet operon in *E. coli*, where it controls the expression of the tetA gene²⁵. The Tet operon is involved in tetracycline resistance and can self-repress due to carrying a second gene, tetR, coding for TetR²⁶. This protein binds to

the operator sites of the promoters of *tetA* and *tetR* and prevents their transcription²⁷. Meanwhile, tetracycline binds to TetR hampering its ability to bind to the DNA, thus enabling the expression of *tetA* and *tetR*²⁸. aTc, an analog of tetracycline, can induce P_{tetA} by the same process²⁹.

The final promoter used, P_{BAD} , was also extracted in its natural form. Originally, it is a component of the L-arabinose operon of *E. coli*³⁰. It can be repressed by dimers of AraC, which can form a DNA loop that blocks transcription³¹. Arabinose can induce P_{BAD} by binding to the dimeric AraC. This binding changes the conformation of AraC, which breaks the DNA loop, allowing RNAP to bind to P_{BAD} ³².

A recent work used similar promoters to study the dynamics of genes in convergent, divergent, and tandem formations³³. However, unlike in our constructs, the elongation regions were separated by Rho independent hairpin loops. Thus, the RNAPs transcribing from one promoter should not collide with RNAPs transcribing from the other promoter. Among other, they used the constructs to study the influence of supercoiling buildup on closely spaced genes.

Next, we describe the assembly of the genetic constructs. Afterwards, we study their dynamics and show that, in tandem formation, their strength is reduced due to collisions between RNAPs leading to premature transcription terminations. We also show that this phenomenon can be fine-tuned by inducers of the regulatory mechanisms of the component promoters and by a transcription-targeting antibiotic. Moreover, we use a stochastic model to show that the phenomena identified suffices to explain the observed dynamics of the promoters in tandem formation. In the end, we discuss how our constructs may become valuable components of future synthetic genetic circuits.

Assembly of the synthetic, non-overlapping tandem promoters in single-copy plasmids

We designed the tandem constructs using Snapgene (GSL Biotech) and assembled them at Integrated DNA Technology, Iowa, U.S.A. Next, we introduced them into single-copy plasmids (pBAC)³⁴. The original sequences of $P_{LacO3O1}$, P_{tetA} , and P_{BAD} (Methods Section “Bacterial strains, growth conditions, induction, and antibiotic”) were not altered. We distanced the transcription start sites (TSSs) of the pairs of promoters in tandem formation by 150 bps (Figures 1A₄ and 1A₅), while the trio of promoters were separated by 200 bps, each (Figure 1A₆). For that, we introduced random sequences between the promoters generated using [Random.org](https://random.org). Snapgene confirmed the absence of sequences known to lead to the formation of secondary RNA structures (e.g., hairpin loops) as well as of sequences coding for translational products.

In addition to the tandem promoters (“biopart”), each plasmid has a sequence coding for a fluorescent protein, mCherry³⁵ (plasmid accession number KX264176.1^{34,35}), immediately downstream of the biopart. The DNA code includes a strong RBS, to ensure strong translation

rates to robustly track promoter activity. Moreover, the maturation time of mCherry (FPbase ID: ZERB6) is 15 minutes³⁶, to closely report RNA production rates. We also produced three control strains, each carrying one individual promoter (P_{LacO3O1} , P_{tetA} , and P_{BAD} , respectively) (Figures 1A₁-A₃).

Since all constructs have the same RNA degradation rate and mCherry translation and degradation rates, differences in fluorescence intensities between strains should be caused mostly by differences in overall transcription rates alone. To best ensure this, small differences in growth rates between strains were accounted for (next section). Finally, since some tetracycline derivatives exhibit fluorescence³⁷, we tested if the inducer aTc could interfere with the measurements. However, within the range of concentrations used, aTc did not significantly affect cell fluorescence (Figure S1).

Cell morphology and physiology. Variability in single cell expression levels.

We searched for potential physiological and morphological differences between strains that could affect the constructs' expression. First, we monitored cell growth rates of all strains in all measurement conditions studied throughout the manuscript (Figure 2A). Strains carrying the plasmids have a slightly slower growth rate than the wild type, WT, strain (8% weaker). For comparison, rifampicin (5 $\mu\text{g}/\text{mL}$) reduced growth by much more (41%). Given the existence of differences, from here onwards, we accounted for them by normalizing mean population expression levels obtained by spectrophotometry using data on cell populations sizes using data from O.D.₆₀₀ measurements (Methods Section "Spectrophotometry").

Next, we observed cell sizes using pulse width from flow cytometry as a proxy³⁸. The pulse width differs little between strains and conditions (Figure 2B). In agreement, we could not find a significant correlation between mean single-cell protein expression levels (E) and mean single-cell pulse widths (Figure 2C). Thus, gene expression intensities obtained by flow cytometry were not normalized by cell sizes.

Finally, we searched for correlations (in flow cytometry data) between mean single-cell expression levels and cell-to-cell variability (as measured by the squared coefficient of variation CV^2). The high R^2 value of a curve (equation 1) accounting for a noise floor " n "³⁹ (Figure 2D) suggests that the constructs follow the natural genome-wide behavior^{40,41}.

$$CV^2 = \frac{C}{\varepsilon} + n \tag{1}$$

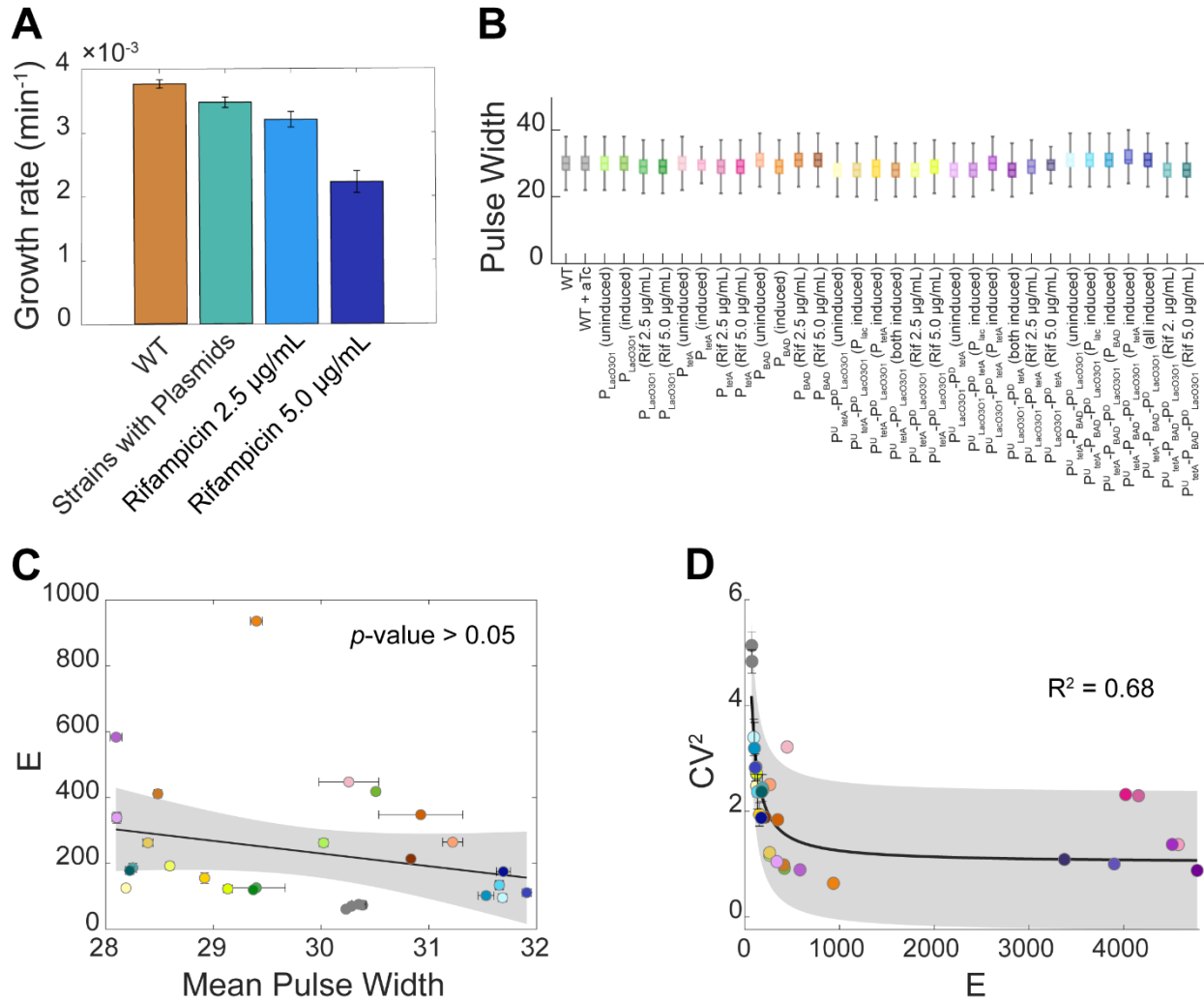


Figure 2: (A) Average growth rates (O.D.₆₀₀ measurements by spectrophotometry). **(B)** Boxplots of the distributions of single-cell pulse width, used as a proxy for cell size. The outliers (values higher or lower than 1.5·IQR, where IQR is the interquartile range) are not shown. **(C)** Scatter plot of the mean single-cell fluorescence, E , and the mean pulse width. Also shown are the best-fitting line and corresponding p -value, after excluding outliers (not shown) (Methods section “Fitting and statistical analysis”). The correlation is not significant at the 5% significance level, even when including outliers (Figure S2). **(D)** Scatter plot between the CV^2 and the mean single-cell fluorescence intensities (E). Also shown is the best-fitting curve and the corresponding R^2 (Methods section “Fitting and statistical analysis”). The shaded area represents the 95% confidence interval. The equation and coefficient values of the fitted curve are shown in Supplementary Table S4. Finally, the error bars in **(A)**, **(C)**, **(D)** correspond to the standard error of the mean of 3 biological replicates.

Promoters in tandem formation have attenuated overall expression rates

Next, we studied the dynamics of the dual synthetic tandem promoters $P_{\text{LacO301}}^{\text{U}}-P_{\text{tetA}}^{\text{D}}$ and $P_{\text{tetA}}^{\text{U}}-P_{\text{LacO301}}^{\text{D}}$ from single-cell fluorescence data under three induction schemes: (i) upstream promoter induced; (ii) downstream promoter induced; and (iii) both promoters induced. For comparison, we also studied the individual promoters P_{LacO301} and P_{tetA} , when and not induced.

The mean, standard deviation, CV^2 , skewness and kurtosis of the distributions of single-cell expression levels are shown in Supplementary File 1.

From the data, first, we quantified the attenuation in the expression of the promoters, when in tandem formation, when compared with the component promoters. Let T stand for the tandem promoters, U stand for the individual upstream promoter, and D stand for the individual downstream promoter. We define “absolute attenuation”, α_A , in the tandem promoters as:

$$\alpha_A = E_U + E_D - E_T \quad (2)$$

From Figure 3, in both tandem constructs under all induction schemes, $\alpha_A > 0$. Thus, in general, placing these promoters in tandem formation reduces overall protein expression rates.

Meanwhile, for $P_{tetA}^U - P_{LacO3O1}^D$ alone: $E_T < E_U$ and $E_T < E_D$ (Figures 3A₁-3A₃). I.e., placing the weaker promoter downstream heavily reduced the RNA production rate of the upstream promoter, causing $P_{tetA}^U - P_{LacO3O1}^D$ expression to be weaker than the downstream promoter alone.

Conversely, for $P_{LacO3O1}^U - P_{tetA}^D$, in one condition we find that: $E_T > E_U$ and $E_T < E_D$ (Figure 3B₂). Interestingly, we also observe that when $P_{LacO3O1}$ is induced: $E_T > E_U$ and $E_T > E_D$ (Figures 3B₁ and 3B₃). I.e., the tandem promoters $P_{LacO3O1}^U - P_{tetA}^D$ express more strongly than the individual downstream promoter.

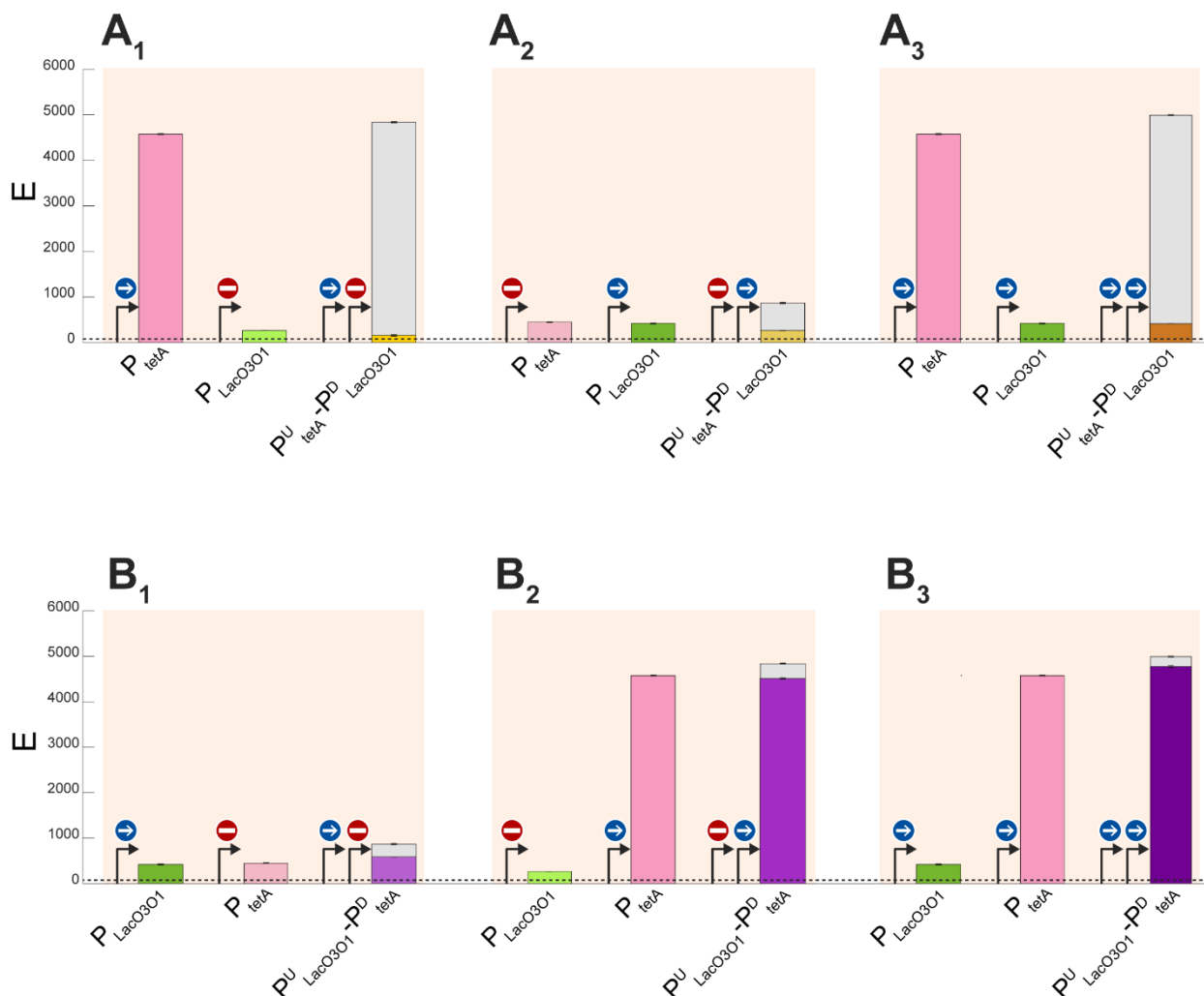


Figure 3: Mean expression intensity (E) of the tandem and the individual promoters, under various induction schemes. The blue sign with a straight white arrow stands for “fully induced”, while the red “Do Not Enter” sign stands for “fully repressed” (**A₁-A₃**) P_{tetA} , $P_{LacO3O1}$, and $P_{tetA}^{U}-P_{LacO3O1}^{D}$, respectively. (**B₁-B₃**) P_{tetA} , $P_{LacO3O1}$, and $P_{LacO3O1}^{D}-P_{tetA}^{U}$, respectively. In all plots, the height of the gray bars equals the absolute attenuation, α_A , defined in equation (2). I.e., they correspond to the “loss” in mean expression intensity of two promoters, when placing them in tandem formation. The error bars correspond to the standard error of the mean of 3 biological replicates. The dashed horizontal line near 0 marks the autofluorescence intensity of WT cells. The expression levels of the individual promoters, P_{tetA} and $P_{LacO3O1}$, are shown more than once to facilitate comparing the tandem constructs with the component promoters in individual formation.

From the above, we interpret the positive absolute attenuation in both constructs in all conditions (Figures 3A₁-3A₃ and 3B₁-3B₃) as evidence that many RNAPs elongating from the upstream promoter are prematurely terminated (fall-off). When the downstream promoter is repressed, the fall-offs are likely due to the repressors’ binding (Figure 4C). Conversely, when the downstream promoter is active, the fall-offs are likely due to collisions between the elongating

RNAP from the upstream promoter with an RNAP bound to the downstream promoter (Figure 4D).

Moreover, we interpret the existence of a tandem construct with weaker activity than either of its component promoters (Figures 3A₁-3A₃) as evidence that several collisions cause fall-offs of both the RNAP bound to the downstream promoter as well as the elongating RNAP (Figure 4D₃). If this never occurred, the tandem promoters should be at least as strong as the downstream promoter alone.

Finally, we interpret the existence of a tandem construct with stronger activity than the individual downstream promoter (Figures 3B₁ and 3B₃) as evidence that, in some tandem promoters, neither the repressors (Figure 3B₁), nor RNAPs initiating transcription (Figure 3B₃) at the downstream promoter, can block all RNAPs elongating from the upstream promoter (Figures 4C₂ and 4D₂).

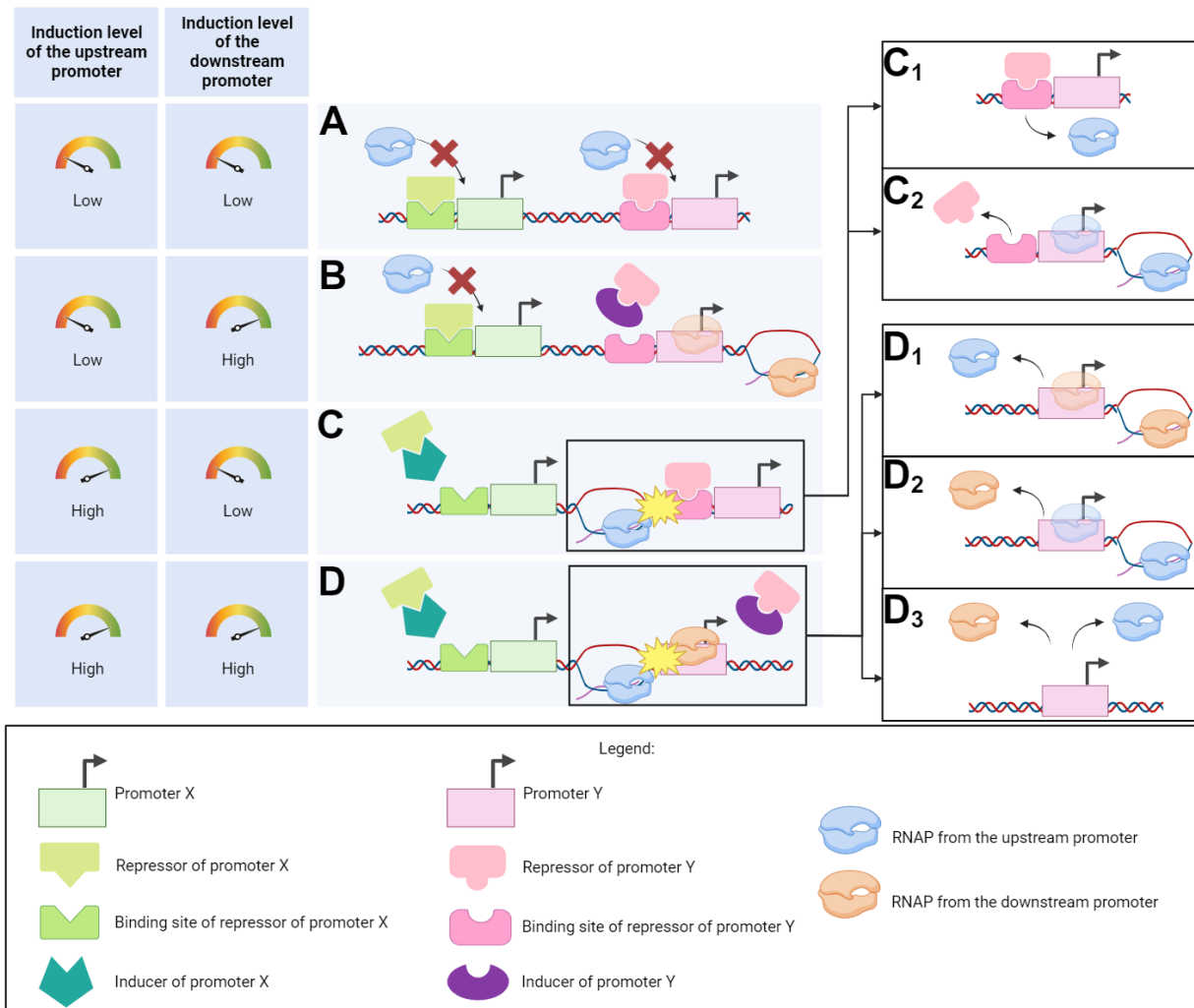


Figure 4: Illustration of transcription interference events generating attenuation in tandem promoters under different induction schemes. **(A)** Weak-to-no induction of both the upstream and the downstream promoters. **(B)** Weak-to-no induction of the upstream promoter along with high induction of the downstream promoter. **(C)** High induction of the upstream promoter and weak-to-no induction of the downstream promoter. **(D)** High induction of both the upstream and downstream promoters. In **(C)** collisions between RNAPs elongating from the upstream promoter and the repressor at the downstream promoter are likely to occur. Those collisions can cause **(C₁)** fall-off of the elongating RNAP, or **(C₂)** fall-off of the repressor bound to the operator site. Similarly, in **(D)** the RNAP elongating from the upstream promoter can collide with the RNAP occupying the downstream promoter. This causes **(D₁)** the elongating RNAP to fall-off, **(D₂)** the RNAP at the downstream promoter to fall-off, or **(D₃)** both RNAPs to fall-off.

Attenuation in synthetic tandem constructs can be fine-tuned by induction of the component promoters

We investigated if the attenuation can be fine-tuned, i.e., is sensitive to external regulation of the activity of the component promoters. We tested in $P_{tetA}^U - P_{LacO3O}^D$, since it exhibited stronger attenuation. To compare attenuations as a function of inductions strength, we defined relative attenuation, α_R , as:

$$\alpha_R = \frac{E_U + E_D}{E_T} \quad (3)$$

We started with a fully induced $P_{LacO3O1}^D$ and then gradually induced P_{tetA}^U . Visibly, $P_{tetA}^U - P_{LacO3O1}^D$ is less induced than P_{tetA} alone (Figure 5A₁), and α_R increases with induction of the upstream promoter (Figure 5A₂). This can be explained by an increase in the fraction of RNAPs from the upstream promoter that fail to complete elongation. Such fall-offs could be explained by the transient occupation of the (active) downstream promoter by RNAPs initiating transcription.

Next, we started with a fully induced P_{tetA}^U and gradually induced the downstream promoter, $P_{LacO3O1}^D$. Again, the individual promoter (P_{tetA}) is more strongly induced than the tandem construct, implying the existence of attenuation (Figure 5B₁). Nevertheless, contrary to before, α_R decreased with induction (Figure 5B₂). This is expected since repressors bound to the DNA are expected to be more efficient in blocking RNAPs elongating from the upstream promoter than RNAPs occupying the downstream promoter. Consequently, as the repression mechanism of the downstream promoter is inactivated, more RNAPs from the upstream promoter can complete elongation.

Given all of the above, we conclude that α_R can be fine-tuned by the regulatory mechanisms of the component promoters in diverse ways.

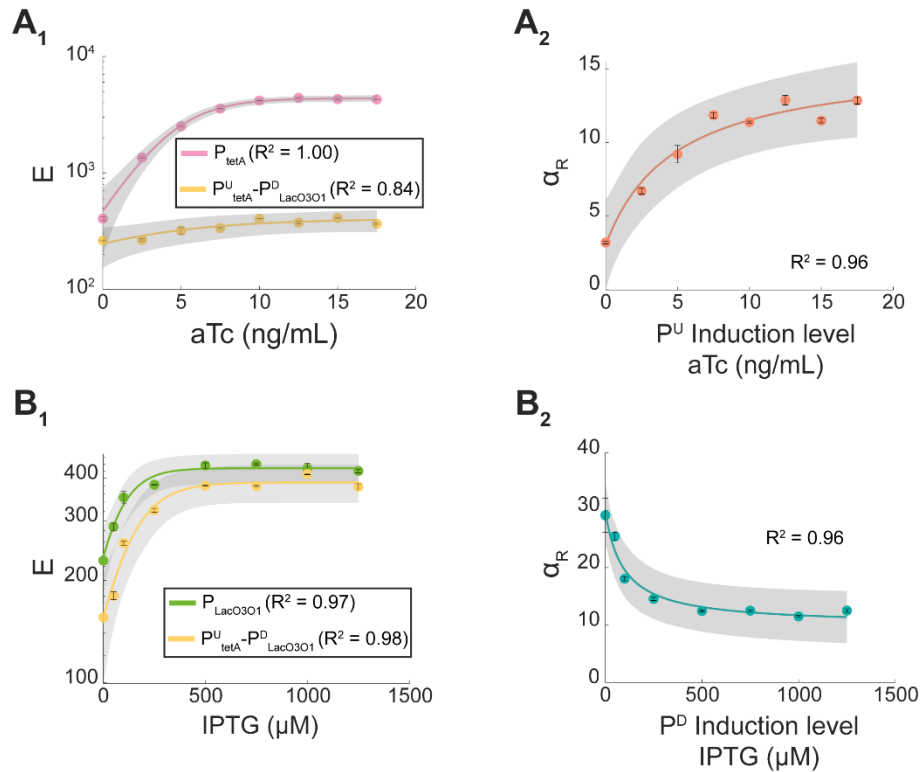


Figure 5: (A₁) Mean single-cell expression level (E) of P_{tetA} and of $P_{tetA}^U-P_{LacO3O1}^D$, when inducing P_{tetA} . In all conditions, $P_{LacO3O1}^D$ is fully induced. (A₂) Attenuation of the tandem promoters with increasing induction of the upstream promoter. (B₁) Mean single-cell expression level (E) of $P_{LacO3O1}$ and of $P_{tetA}^U-P_{LacO3O1}^D$, when inducing $P_{LacO3O1}$. In all conditions, P_{tetA}^U is fully induced. (B₂) Attenuation of the tandem promoters with increasing induction of the downstream promoter. In all plots, the error bars correspond to the standard error mean of 3 biological replicates. Also shown are best fitting curves and the corresponding R^2 (Methods section “Fitting and statistical analysis”). The shaded areas represent the 95% confidence intervals. The equation and coefficient values of the fitted curve are shown in Supplementary Table S4. Note that the vertical axes of (B₁) and (A₁) are in logarithmic scale.

Adding another promoter in tandem formation enhances attenuation

We next introduced an additional promoter, P_{BAD} , in between two tandem promoters ($P_{tetA}^U-P_{LacO3O1}^D$). From this resulted the trio of tandem promoters: $P_{tetA}^U-P_{BAD}-P_{LacO3O1}^D$. To study its attenuation, we defined, where M stands for “Middle”:

$$\alpha_A = E_U + E_M + E_D - E_T \quad (4)$$

Expectedly, $P_{tetA}^U-P_{BAD}-P_{LacO3O1}^D$ activity is also weaker than the sum of individual promoter activities. I.e., in all induction schemes (Figures 6A₁-A₄), $\alpha_A > 0$.

Conversely, unlike the dual tandem promoters, the E_T of the trio of tandem promoters is smaller than E_U , E_M , and E_D , alone, in all induction schemes (Figure 6A): $E_T < E_U, E_M, E_D$. This is evidence for the occurrence of many collisions between RNAPs elongating from the upstream promoters

with repressors (causing the fall-off of the RNAP) and with RNAPs occupying the downstream promoters (causing the fall-offs of both colliding RNAPs).

The strongest α_R (Equation 5) still occurs when the most upstream promoter is the only one induced (Figure 6B), suggesting that, as expected, collisions between elongating RNAPs and repressors bound to the DNA usually cause the collision of the RNAP. Interestingly, α_R is also high when all promoters are induced. The simplest explanation is that numerous RNAPs elongating from the upstream promoters are prematurely terminated by collisions with RNAPs occupying the downstream promoters, and that several collisions cause both colliding RNAPs to fall-off.

$$\alpha_R = \frac{E_U + E_M + E_D}{E_T} \quad (5)$$

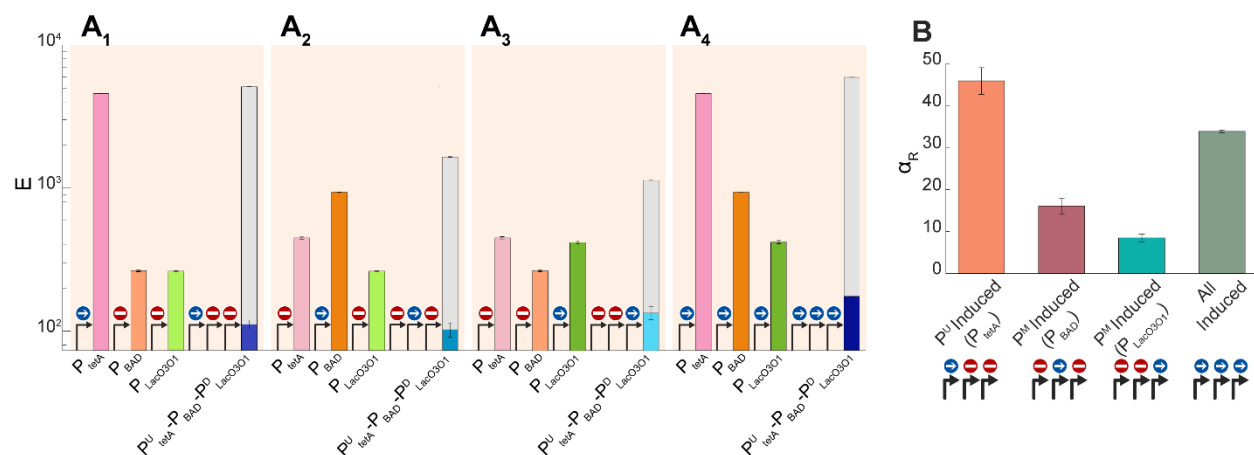


Figure 6: (A₁-A₄) Mean expression intensities (E , in logarithmic scale) of the trio tandem promoters (P_{tetA}^U - P_{BAD} - $P_{LacO103}^D$) and of each individual promoter. We tested inducing only the most upstream promoter (A₁), inducing only the middle promoter (A₂), inducing only the most downstream promoter (A₃), and inducing all promoters at the same time (A₄). The gray bars measure the absolute attenuation, α_A , defined in equation (4). **(B)** Relative attenuation, α_R (Equation 5), of the trio of tandem promoters under various induction schemes. In all plots, the error bars correspond to the standard error of the mean of 3 biological replicates.

Attenuation is enhanced when hampering promoter escape by adding rifampicin

Above, we argued that several RNAPs are prematurely terminated by collisions of RNAPs elongating from upstream promoters with RNAPs sitting at downstream promoters. If this holds true, increasing the fraction of time spent by RNAPs occupying a downstream promoter should increase α_R .

To test this, we subjected cells to rifampicin (Methods section “Fitting and statistical analysis”). This antibiotic binds to the β sub-unit of RNAP⁴². When bound, the RNAP cannot escape beyond 2-3 nucleotides away from the TSS^{43,44}. Thus, rifampicin should not only reduce the activity of

each component promoter, but also increase α_R , due to increasing the occupancy time of RNAPs at the downstream promoter (illustrated in Figure 7A). Other direct effects on gene expression are not expected since, e.g., rifampicin does not affect stable transcription elongation complexes^{45,46}.

As expected, rifampicin reduced the activity of both P_{tetA} as well as $P_{LacO3O1}$ in individual formations (Figures 7B₁ and 7B₂, respectively). $P_{LacO3O1}$ was (relatively) more affected than P_{tetA} , for unknown reasons (also visible in spectrophotometry data in Figure S3). Also, for unknown reasons, P_{tetA} activity reduction was approximately linear with rifampicin concentration, while $P_{LacO3O1}$ had a sharp initial decrease in expression, but further increasing rifampicin concentration did have significant additional effects.

Most importantly, we also observed that rifampicin gradually increased α_R in both tandem constructs (Figures 7C₁ and 7C₂), as hypothesized. Noteworthy, the higher sensitivity of $P_{LacO3O1}$ to rifampicin explains why α_R of $P_{tetA}^U-P_{LacO3O1}^D$ changed the most (Figures 7C₁ and 7C₂).

These results exemplify how the attenuation in tandem promoters can be controlled by external regulation and, with that, produce different (yet predictable) gene expression dynamics.

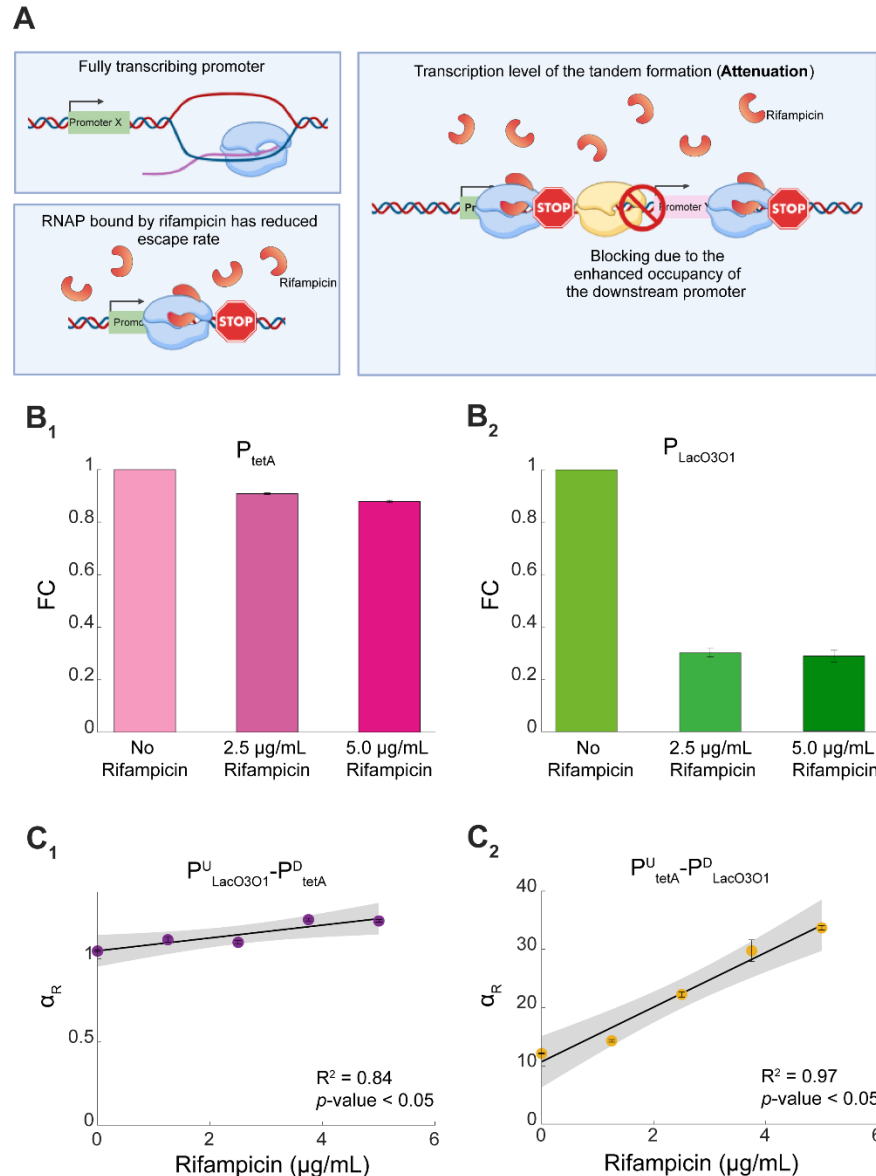


Figure 7: Effects of rifampicin on the attenuation in tandem promoters. **(A)** Illustration of how rifampicin reduces the activity of an individual promoter, as well as of promoters in tandem formation. The latter should be further affected by enhanced attenuation. **(B₁-B₂)** Fold-change in mean expression levels (FC) of P_{tetA} and $P_{LacO3O1}$ due to rifampicin. **(C₁-C₂)** Relative attenuation (α_R) of $P_{tetA}^U-P_{LacO3O1}^D$ and $P_{LacO3O1}^U-P_{tetA}^D$, when subject to rifampicin. The error bars correspond to the standard error of the mean of 3 biological replicates. Shown are the best-fitting lines and their p -value and R^2 (Methods section “Fitting and Statistical Analysis”). The shaded areas represent the 95% confidence interval. The equation and coefficient values of the fitted lines are shown in Supplementary Table S4.

Model of transcription attenuation of promoters in tandem formation due to RNAP collisions and consequent fall-offs can explain the empirical data

Above, we observed that the tandem promoters always exhibit attenuation. In $P_{tetA}^U-P_{LacO3O1}^D$, the attenuation was strong, in general, making expression levels lower than in the individual downstream promoter. Instead, in $P_{LacO3O1}^U-P_{tetA}^D$, the attenuation was weak, allowing, in some conditions, almost as strong expression as the sum of expressions of the individual component promoters. In both tandem promoters, attenuation was present when the downstream promoter was activated, as well as when it was repressed, in which case it was stronger.

In this section, we propose a general stochastic model for tandem promoters. Shortly, the model includes a two-step transcription initiation process at each promoter, to account for promoter occupancy by RNAPs. Repressors can also occlude promoters, while inducers inactivate repressors. Also, RNAs degrade. Finally, we model collisions between RNAPs, as well as between RNAPs and repressors occluding promoters. Critically, we assume that the collisions between RNAPs can cause fall-offs of one or both RNAPs. Meanwhile, collisions between RNAPs and repressors always cause the RNAPs to fall-off. A complete model description is provided in Supplementary Section S1. Using this model, we show that various degrees of attenuation observed above can be explained by fall-offs emerging from RNAP-RNAP and RNAP-repressor collisions.

First, we tested whether increasing the time that the downstream promoter is occupied by initiating RNAPs increases attenuation (due to increased rate of collisions and subsequent fall-offs). For this, we decreased the rate of RNAP escape from the downstream promoter (k_{esc}^D) (similar to the effects of increasing rifampicin, Figures 7C₁-C₂). This increased the mean amount of time that the downstream promoter is occupied by RNAPs (Figure 8A). This, in turn, increased linearly the relative attenuation (Figure 8B), as in Figures 7C₁-C₂. Also, in agreement with the empirical data (Supplementary file 1), it decreased E_D (Figure 8C) and E_T (Figure 8D).

Next, we tested increasing the frequency f_2 with which RNAP-RNAP collisions cause both RNAPs to fall-off (instead of only one RNAP falling-off). We observed, first, a small decrease in the fraction of time that the downstream promoter is occupied by RNAPs (Figure 8A), as expected from the increased fall-offs of the initiating RNAPs. In agreement, we observe increased α_R with increased f_2 (Figure 8B). Meanwhile, the quantitative relationships between E_D and E_T with the average RNAP escape rates (Figure 8C-8D) are not influenced (since both change accordingly).

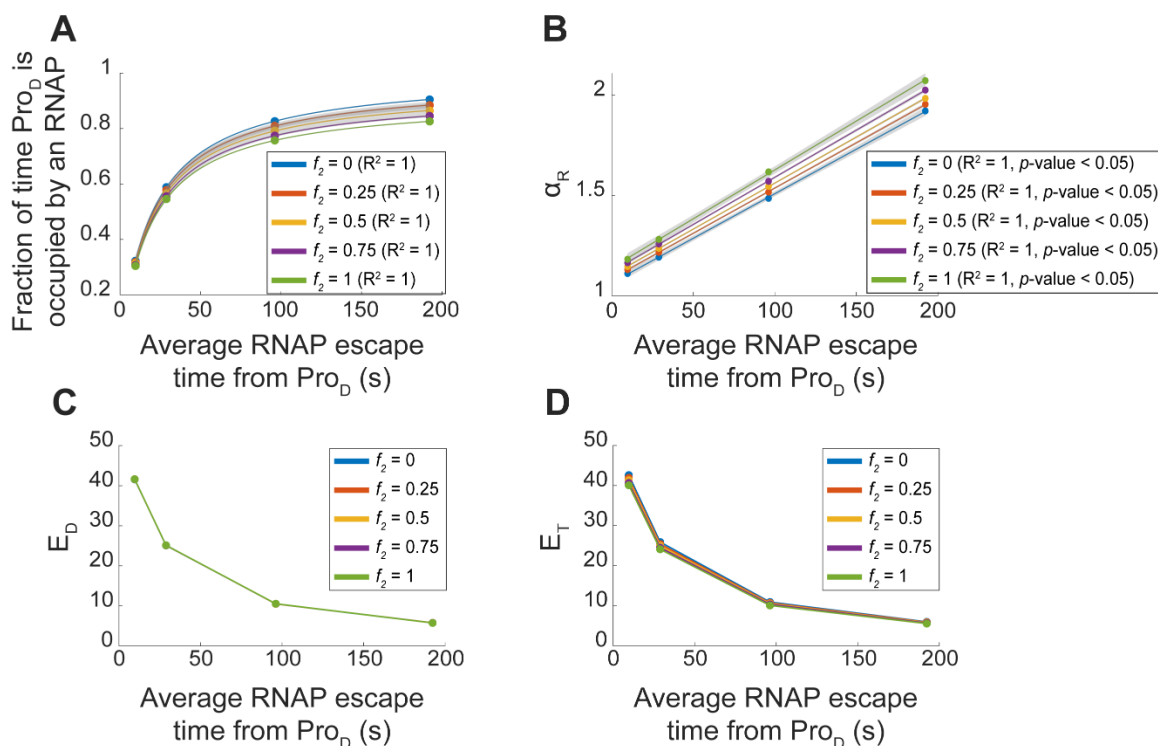


Figure 8: Estimations using the stochastic models of the effects of increasing the average time for RNAPs to escape the downstream promoter (Pro_D), by decreasing k_{esc}^D , as well as of changing the frequency with which RNAP collisions cause both RNAPs to fall-off, instead of only one falling-off. **(A)** Fraction of time that the downstream promoter is occupied by an RNAP as a function of the inverse of the RNAP escape rate. Note that the latter is affected also by collisions between RNAPs. **(B)** Relative attenuation (α_R) in tandem promoters as a function of the average RNAP escape times from downstream promoters. **(C-D)** Expression levels of the individual downstream promoter (E_D) and of the tandem promoters (E_T) as a function of the average RNAP escape times from the downstream promoter. In all plots, each colored line represents a different relative frequency (f_2) of both RNAPs falling-off upon colliding (as opposed to only one RNAP falling-off). (A) and (B) also show the best fitting functions along with their R^2 values (and p -values in case of linear fits) (Methods section "Fitting and statistical analysis"). The shaded areas represent the 95% confidence intervals. The equation and coefficient values of the fitted lines are shown in Supplementary Table S4.

We also observed that, for higher frequencies of double RNAP fall-offs, E_T can become slightly weaker than E_D (Supplementary Figure S4A). We explored this by tuning the binding and escape rates of RNAPs to Pro_U and Pro_D . From Supplementary Figure S4B, there is a wide range of parameter values (Supplementary Table S3) for which E_T can be considerably weaker than E_D (in agreement with the empirical data in Figures 3A₁-3A₃).

Finally, we tested with the model if the induction-repression mechanisms can increase and decrease α_R with increasing induction, depending on whether it is the upstream or the downstream promoter that are initially repressed. Supplementary Figure S5 supports that this is possible.

Overall, we conclude that the model can explain all behaviors observed in the two tandem constructs.

DISCUSSION

Above, we found that premature transcription terminations of RNAPs elongating from upstream promoters and of RNAPs occupying downstream promoters can cause significant attenuation of tandem constructs. The attenuation can differ widely with the active transcription initiation dynamics of the component promoters, allowing for the expression of the tandem constructs to range from weaker than either component promoter up to similar (but never higher) than the sum of the expression dynamics of the two promoters. Moreover, the attenuation can be further externally tuned using the natural regulatory mechanisms of each promoter and/or using antibiotics targeting, e.g., transcription. This tunability should allow tandem formations to execute fine-tuned expression dynamics.

Tandem constructs dynamics were largely predictable by a simple model that only requires knowledge of the dynamics of the component promoters in individual formation, along with the relative rate of collisions that lead to two (instead of one) RNAP fall-offs. This rate should differ with the binding affinity of RNAPs or repressors to the downstream promoter and can be empirically estimated.

This model-based predictability is of significance since it remains an important and difficult challenge in synthetic biology. For example, we are yet to efficiently predict how random DNA sequences express, even in bacteria, albeit significant recent successes^{47–54}.

On the contrary, knowledge of the dynamics of individual, natural promoters (e.g. in *E. coli*) is relatively easy to acquire and the data is rapidly increasing^{41,55–58}. Using these increasing libraries, along with the model proposed, it should be feasible to engineer novel tandem promoters with desired, diverse dynamics. These novel constructs could then be used, e.g., as building blocks for future, more complex circuits.

As an example, genetic switches are circuits composed of two genes repressing one another. Because of this, they are expected to have two possible states: either one gene is “ON” and the other one is “OFF”, or vice-versa. One of the most promising applications of these circuits is as components of information processing circuits (e.g., for storing information). The main difficulty in achieving this is in tuning a switch to be both sufficiently sensitive to changes supposed to control their state, while also being robust in maintaining the state, if no relevant changes occur in the cell. This requires specific expression levels of both genes of the switch⁵⁹, and depend on promoter occupancy times by RNAP and repressors, repressor-activator binding and unbinding rates, and other parameters⁶⁰. Finding natural promoters with the desired features or changing their sequence to achieve them is presently complex. Meanwhile, our results suggest that

starting with a large library (such as^{41,57}), and then use the model to find combinations of two promoters in specific states that achieve a desired dynamics is a more promising approach.

As a side note, while not observed here, we do not exclude the possibility that a few promoters in tandem formation can be stronger than the sum of the two component promoters independently. For example, this could potentially be possible if elongating RNAPs from the upstream promoter would dislodge repressors at the downstream promoter, while not falling-off themselves. Weak repressor-DNA binding affinities could make this possible. However, we do not expect it to be common in natural genomes. Nevertheless, removing repressor binding sites (or altering their sequences) could facilitate it in synthetic circuits.

In the future, we plan to explore broadly how to tune the attenuation in synthetic as well as in natural promoters in tandem formation. In addition to repression mechanisms and antibiotics hampering RNAP promoter escape, it may be possible to use supercoiling regulation for this, as it affects RNAP-promoter binding^{61,62}. In that case, the location of the tandem promoters in the DNA (e.g. if in highly or weakly expressing topological domains^{63,64}) could already suffice to influence the attenuation. Other influential variables are the durations of open complex formation and promoter escape, as they influence RNAP occupancy times of the promoter⁶⁵. Changing both could even allow changing attenuation without changing transcription rates¹¹.

Overall, our findings, including the model proposed to predict attenuation levels, should allow establishing a novel pipeline for engineering promoters (in tandem formation) whose overall kinetics can be fined-tuned within a relative wide dynamic state space. This pipeline should contribute to the development of novel components for complex synthetic genetic circuits with predictable behaviors, making their assembly faster and cheaper. Subsequently, the novel complex circuits could contribute to the engineering of bacterial strains whose metabolic tasks have minimal resource consumption, so as to improve the efficiency of bioindustrial processes.

METHODS

Bacterial strains, growth conditions, induction, and antibiotics

We used the *E. coli* strain DH5 α -PRO (identical to DH5 α Z1, here named “wild type”, WT)^{13,14,66}. This strain produces the necessary regulatory proteins (LacI, AraC, and TetR) that tightly regulate each of our promoter constructs¹³.

First, chemically competent (CC) *E. coli* DH5 α -PRO cells were prepared for plasmid transformations. For each strain, 10 ng of the plasmid DNA was mixed with 50 μ L DH5 α -PRO CC (1:10 ratio), and the mixture was incubated on ice for 30 min. Next, the mixture was kept at 42 °C in a water bath for 1 minute. Finally, 800 μ L of LB medium was added to the mixture, which was kept at 37 °C under aeration at 250 RPM for 1 hour.

From each mixture, 200 μL was plated using the spread plate method on fresh LB agar plates (2%), prepared by supplementing with antibiotics (34 $\mu\text{g}/\text{mL}$ chloramphenicol). Finally, the plates were kept overnight at 37 $^{\circ}\text{C}$. The next day, three colonies were picked from each plate and inoculated in fresh LB medium, supplemented with antibiotics (34 $\mu\text{g}/\text{mL}$ chloramphenicol). Afterwards, the cells were incubated at 30 $^{\circ}\text{C}$ overnight with shaking at 250 RPM. The resulting cells were diluted into fresh M9 media (0.03 O.D.₆₀₀) and were supplemented with 0.4% glycerol, amino acids, and vitamin solutions^{11,14}.

The control condition was M9 medium at 30 $^{\circ}\text{C}$. Cells were induced and given antibiotic treatment at the mid-exponential phase (O.D.₆₀₀ \approx 0.25-0.3). Flow cytometry data was collected 180 minutes after induction. The spectrophotometry time series started immediately after induction and continued for 700 min.

The LacO₃O₁ promoter was induced with 1 mM IPTG, the TetA promoter was induced with 15 ng/mL aTc, and the BAD promoter was induced with 0.1% arabinose (induction curves in Figure S6)^{14,66,67}. The promoters are referred to as P_{LacO₃O₁}, P_{tetA}, and P_{BAD}, respectively, for simplicity. In some experiments, cells were subjected to rifampicin (2.5 $\mu\text{g}/\text{mL}$, 1.25 $\mu\text{g}/\text{mL}$, 3.75 $\mu\text{g}/\text{mL}$ and 5.0 $\mu\text{g}/\text{mL}$)⁶⁸. Rifampicin (on its own) does not influence cell fluorescence (Figure S7).

Spectrophotometry

We performed spectrophotometry to measure optical density at 600 nm (O.D.₆₀₀), which we used to monitor cell growth. We used a Biotek Synergy HTX Multi-Mode Reader. We also measured mean cell fluorescence. We used the excitation and emission wavelengths for mCherry (575/15 nm and 620/15 nm, respectively) with a gain of 50. Time series data was captured every 20 minutes. We normalized the fluorescence by the O.D.₆₀₀ to estimate the average single-cell fluorescence. For this, O.D.₆₀₀ curves were time shifted so that they reached 5% of their maximal O.D.₆₀₀ at the same time point., as in⁵⁷.

Flow cytometry

We performed flow cytometry using an ACEA NovoCyte Flow Cytometer controlled by the software Novo Express V1.50. Cells were diluted 1:10000 into 1 ml of PBS, vortexed for 10 sec. For each condition, we performed 3 biological replicates, acquiring 50000 events for each replicate. We detected mCherry using the PE-Texas Red channel with excitation 561 nm and emission 615/20. We collected events at a flow rate of 14 $\mu\text{L}/\text{minute}$ (core diameter of 7.7 μM). We have set a minimum detection threshold in FSC-H at 5000 to remove small particle interference. We also discarded the 1% of the highest PE-Texas Red-H intensities detected, to remove outliers.

Fittings and statistical analysis

All best-fitting lines (e.g., Figure 2C) were obtained by linear regression (MATLAB). Fits of other functions (Figure 2D and Figure 8A) were obtained by the nonlinear least squares method (MATLAB). Their goodness of fit was estimated from R^2 values.

Meanwhile, to determine if there is a significant correlation between any pair of variables, y and x , we used t -statistic with the null hypothesis that the slope of the best-fitting line (of a scatter plot between the variables) is not different from zero. For p -values smaller than 0.05, we reject the null hypothesis.

To identify outlier data points (in Figure 2C) we used an iterative procedure to discard outliers in linear correlations⁴⁰. Data points (x,y) were classified as outliers when their vertical distance from the best linear fit is larger than the standard deviation of the distribution of y values (Figure S2). When applied, the process always converged in 1 iteration.

Finally, the associated standard error of the mean for calculations with fluorescence was calculated using an error propagation method. In general, if $z = f(x, y)$, then the error of z is

$$SEM_z = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 \cdot (SEM_x)^2 + \left(\frac{\partial z}{\partial y}\right)^2 \cdot (SEM_y)^2}$$
, where $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are the partial derivatives of z with respect to x and y , respectively.

Stochastic simulations

We used stochastic simulations to simulate models used to estimate how changes in the dynamics of the promoters (e.g., RNAP occupancy time, repressor occupancy time, and frequency of dual RNAP fall-offs compared to one RNAP fall-off) influence the degree of attenuation in tandem promoter formations. The model is described in Supplementary Section S1. Simulations were performed using SGNS2⁶⁹, a stochastic gene networks simulator whose dynamics follow the Stochastic Simulation Algorithm⁷⁰. The time length of each simulation was set to 5×10^5 seconds (with a sampling interval of 100 seconds). This time length sufficed to reach quasi-equilibrium in RNA numbers. The average behavior of each condition was characterized from 100 independent runs, as it always sufficed to obtain consistent results. Finally, for the various conditions, we assumed the parameter values in Supplementary Table S1 and the initial amounts of reactants in Supplementary Table S2.

COMPETING INTEREST

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We thank Jane and Aatos Erkko Foundation [10-10524-38 to A.S.R.]; Sigrid Jusélius Foundation [230181 to A.S.R.]; Suomalainen Tiedeakatemia (to S.D. and to R.J.); EDUFI Fellowship [TM-21-11655 to R.J.]; Finnish Cultural Foundation [00222452 and 00232457 to I.S.C.B.]; Tampere University Graduate Program (to V.C.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors also acknowledge the Tampere Flow Cytometry Facility for their service. Finally, Figures 1, 4, and 7A were created with BioRender.com.

AUTHOR CONTRIBUTIONS

A.S.R. and V.C. conceived the study. A.S.R. supervised the study. V.C. designed the constructs and planned and executed most measurements. I.S.C.B. executed most data analysis. S.D. assisted in the design of the constructs and measurements. V.C. and R.J. contributed to the data analysis. A.S.R., R.J., and I.S.C.B. developed the model. I.S.C.B. implemented the model. A.S.R., V.C., and I.S.C.B. drafted all documents, which were revised by all co-authors.

REFERENCES

- (1) Adachi, N.; Lieber, M. R. Bidirectional Gene Organization: A Common Architectural Feature of the Human Genome. *Cell* **2002**, *109* (7), 807–809. [https://doi.org/10.1016/s0092-8674\(02\)00758-4](https://doi.org/10.1016/s0092-8674(02)00758-4).
- (2) Trinklein, N. D.; Aldred, S. F.; Hartman, S. J.; Schroeder, D. I.; Otilar, R. P.; Myers, R. M. An Abundance of Bidirectional Promoters in the Human Genome. *Genome Res.* **2004**, *14* (1), 62–66. <https://doi.org/10.1101/gr.1982804>.
- (3) Herbert, M.; Kolb, A.; Buc, H. Overlapping Promoters and Their Control in Escherichia Coli: The Gal Case. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83* (9), 2807–2811. <https://doi.org/10.1073/pnas.83.9.2807>.
- (4) Bendtsen, K. M.; Erdosy, J.; Csiszovszki, Z.; Svenningsen, S. L.; Sneppen, K.; Krishna, S.; Semsey, S. Direct and Indirect Effects in the Regulation of Overlapping Promoters. *Nucleic Acids Res.* **2011**, *39* (16), 6879–6885. <https://doi.org/10.1093/nar/gkr390>.
- (5) Beck, C. F.; Warren, R. A. Divergent Promoters, a Common Form of Gene Organization. *Microbiol. Rev.* **1988**, *52* (3), 318–326. <https://doi.org/10.1128/mr.52.3.318-326.1988>.
- (6) Prescott, E. M.; Proudfoot, N. J. Transcriptional Collision between Convergent Genes in Budding Yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (13), 8796–8801. <https://doi.org/10.1073/pnas.132270899>.
- (7) Mackie, G. A.; Parsons, G. D. Tandem Promoters in the Gene for Ribosomal Protein S20. *J. Biol. Chem.* **1983**, *258* (12), 7840–7846.
- (8) Korbel, J. O.; Jensen, L. J.; von Mering, C.; Bork, P. Analysis of Genomic Context: Prediction of Functional Associations from Conserved Bidirectionally Transcribed Gene Pairs. *Nat. Biotechnol.* **2004**, *22* (7), 911–917. <https://doi.org/10.1038/nbt988>.
- (9) Callen, B. P.; Shearwin, K. E.; Egan, J. B. Transcriptional Interference between Convergent Promoters Caused by Elongation over the Promoter. *Mol. Cell* **2004**, *14* (5), 647–656. <https://doi.org/10.1016/j.molcel.2004.05.010>.

- (10) Horowitz, H.; Platt, T. Regulation of Transcription from Tandem and Convergent Promoters. *Nucleic Acids Res.* **1982**, *10* (18), 5447–5465. <https://doi.org/10.1093/nar/10.18.5447>.
- (11) Chauhan, V.; Bahrudeen, M. N. M.; Palma, C. S. D.; Baptista, I. S. C.; Almeida, B. L. B.; Dash, S.; Kandavalli, V.; Ribeiro, A. S. Analytical Kinetic Model of Native Tandem Promoters in *E. Coli*. *PLoS Comput. Biol.* **2022**, *18* (1), e1009824. <https://doi.org/10.1371/journal.pcbi.1009824>.
- (12) Sneppen, K.; Dodd, I. B.; Shearwin, K. E.; Palmer, A. C.; Schubert, R. A.; Callen, B. P.; Egan, J. B. A Mathematical Model for Transcriptional Interference by RNA Polymerase Traffic in *Escherichia Coli*. *J. Mol. Biol.* **2005**, *346* (2), 399–409. <https://doi.org/10.1016/j.jmb.2004.11.075>.
- (13) Lutz, R.; Bujard, H. Independent and Tight Regulation of Transcriptional Units in *Escherichia Coli* via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements. *Nucleic Acids Res.* **1997**, *25* (6), 1203–1210. <https://doi.org/10.1093/nar/25.6.1203>.
- (14) Oliveira, S. M. D.; Häkkinen, A.; Lloyd-Price, J.; Tran, H.; Kandavalli, V.; Ribeiro, A. S. Temperature-Dependent Model of Multi-Step Transcription Initiation in *Escherichia Coli* Based on Live Single-Cell Measurements. *PLoS Comput. Biol.* **2016**, *12* (10), e1005174. <https://doi.org/10.1371/journal.pcbi.1005174>.
- (15) Muthukrishnan, A. B.; Kandhavelu, M.; Lloyd-Price, J.; Kudasov, F.; Chowdhury, S.; Yli-Harja, O.; Ribeiro, A. S. Dynamics of Transcription Driven by the TetA Promoter, One Event at a Time, in Live *Escherichia Coli* Cells. *Nucleic Acids Res.* **2012**, *40* (17), 8472–8483. <https://doi.org/10.1093/nar/gks583>.
- (16) Campbell, E. A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S. A. Structural Mechanism for Rifampicin Inhibition of Bacterial Rna Polymerase. *Cell* **2001**, *104* (6), 901–912. [https://doi.org/10.1016/s0092-8674\(01\)00286-0](https://doi.org/10.1016/s0092-8674(01)00286-0).
- (17) Jacob, F.; Monod, J. Genetic Regulatory Mechanisms in the Synthesis of Proteins. *J. Mol. Biol.* **1961**, *3*, 318–356. [https://doi.org/10.1016/s0022-2836\(61\)80072-7](https://doi.org/10.1016/s0022-2836(61)80072-7).
- (18) Oehler, S.; Eismann, E. R.; Krämer, H.; Müller-Hill, B. The Three Operators of the Lac Operon Cooperate in Repression. *EMBO J.* **1990**, *9* (4), 973–979. <https://doi.org/10.1002/j.1460-2075.1990.tb08199.x>.
- (19) Oehler, S.; Amouyal, M.; Kolkhof, P.; von Wilcken-Bergmann, B.; Müller-Hill, B. Quality and Position of the Three Lac Operators of *E. Coli* Define Efficiency of Repression. *EMBO J.* **1994**, *13* (14), 3348–3355. <https://doi.org/10.1002/j.1460-2075.1994.tb06637.x>.
- (20) Gilbert, W.; Müller-Hill, B. Isolation of the Lac Repressor. *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *56* (6), 1891–1898. <https://doi.org/10.1073/pnas.56.6.1891>.
- (21) Swigon, D.; Coleman, B. D.; Olson, W. K. Modeling the Lac Repressor-Operator Assembly: The Influence of DNA Looping on Lac Repressor Conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (26), 9879–9884. <https://doi.org/10.1073/pnas.0603557103>.
- (22) Becker, N. A.; Peters, J. P.; Maher, L. J., 3rd; Lionberger, T. A. Mechanism of Promoter Repression by Lac Repressor-DNA Loops. *Nucleic Acids Res.* **2013**, *41* (1), 156–166. <https://doi.org/10.1093/nar/gks1011>.
- (23) Krämer, H.; Niemöller, M.; Amouyal, M.; Revet, B.; von Wilcken-Bergmann, B.; Müller-Hill, B. Lac Repressor Forms Loops with Linear DNA Carrying Two Suitably Spaced Lac Operators. *EMBO J.* **1987**, *6* (5), 1481–1491. <https://doi.org/10.1002/j.1460-2075.1987.tb02390.x>.

- (24) Lewis, M.; Chang, G.; Horton, N. C.; Kercher, M. A.; Pace, H. C.; Schumacher, M. A.; Brennan, R. G.; Lu, P. Crystal Structure of the Lactose Operon Repressor and Its Complexes with DNA and Inducer. *Science* **1996**, *271* (5253), 1247–1254. <https://doi.org/10.1126/science.271.5253.1247>.
- (25) Hillen, W.; Schollmeier, K.; Gatz, C. Control of Expression of the Tn10-Encoded Tetracycline Resistance Operon. II. Interaction of RNA Polymerase and TET Repressor with the Tet Operon Regulatory Region. *J. Mol. Biol.* **1984**, *172* (2), 185–201. [https://doi.org/10.1016/s0022-2836\(84\)80037-6](https://doi.org/10.1016/s0022-2836(84)80037-6).
- (26) Beck C F; Mutzel R; Barbé J; Müller W. A Multifunctional Gene (TetR) Controls Tn10-Encoded Tetracycline Resistance. *J. Bacteriol.* **1982**, *150* (2), 633–642. <https://doi.org/10.1128/jb.150.2.633-642.1982>.
- (27) Hillen, W.; Berens, C. Mechanisms Underlying Expression of Tn10 Encoded Tetracycline Resistance. *Annu. Rev. Microbiol.* **1994**, *48*, 345–369. <https://doi.org/10.1146/annurev.mi.48.100194.002021>.
- (28) Orth, P.; Schnappinger, D.; Hillen, W.; Saenger, W.; Hinrichs, W. Structural Basis of Gene Regulation by the Tetracycline Inducible Tet Repressor-Operator System. *Nat. Struct. Biol.* **2000**, *7* (3), 215–219. <https://doi.org/10.1038/73324>.
- (29) Murphy, K. F.; Adams, R. M.; Wang, X.; Balázsi, G.; Collins, J. J. Tuning and Controlling Gene Expression Noise in Synthetic Gene Networks. *Nucleic Acids Res.* **2010**, *38* (8), 2712–2726. <https://doi.org/10.1093/nar/gkq091>.
- (30) Schleif, R. Regulation of the L-Arabinose Operon of Escherichia Coli. *Trends Genet.* **2000**, *16* (12), 559–565. [https://doi.org/10.1016/s0168-9525\(00\)02153-3](https://doi.org/10.1016/s0168-9525(00)02153-3).
- (31) Seabold, R. R.; Schleif, R. F. Apo-AraC Actively Seeks to Loop. *J. Mol. Biol.* **1998**, *278* (3), 529–538. <https://doi.org/10.1006/jmbi.1998.1713>.
- (32) Smith, B. R.; Schleif, R. Nucleotide Sequence of the L-Arabinose Regulatory Region of Escherichia Coli K12. *J. Biol. Chem.* **1978**, *253* (19), 6931–6933.
- (33) Yeung, E.; Dy, A. J.; Martin, K. B.; Ng, A. H.; Del Vecchio, D.; Beck, J. L.; Collins, J. J.; Murray, R. M. Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Syst* **2017**, *5* (1), 11–24.e12. <https://doi.org/10.1016/j.cels.2017.06.001>.
- (34) Lee, J. W.; Gyorgy, A.; Cameron, D. E.; Pyenson, N.; Choi, K. R.; Way, J. C.; Silver, P. A.; Del Vecchio, D.; Collins, J. J. Creating Single-Copy Genetic Circuits. *Mol. Cell* **2016**, *63* (2), 329–336. <https://doi.org/10.1016/j.molcel.2016.06.006>.
- (35) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from *Discosoma* Sp. Red Fluorescent Protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567–1572. <https://doi.org/10.1038/nbt1037>.
- (36) Lambert, T. J. FPbase: A Community-Editable Fluorescent Protein Database. *Nat. Methods* **2019**, *16* (4), 277–278. <https://doi.org/10.1038/s41592-019-0352-8>.
- (37) Pautke, C.; Vogt, S.; Kreutzer, K.; Haczek, C.; Wexel, G.; Kolk, A.; Imhoff, A. B.; Zitzelsberger, H.; Milz, S.; Tischer, T. Characterization of Eight Different Tetracyclines: Advances in Fluorescence Bone Labeling. *J. Anat.* **2010**, *217* (1), 76–82. <https://doi.org/10.1111/j.1469-7580.2010.01237.x>.

- (38) Cunningham, A. Analysis of Microalgae and Cyanobacteria by Flow Cytometry. In *Flow Cytometry in Microbiology*; Lloyd, D., Ed.; Springer London: London, 1993; pp 131–142. https://doi.org/10.1007/978-1-4471-2017-9_10.
- (39) Hausser, J.; Mayo, A.; Keren, L.; Alon, U. Central Dogma Rates and the Trade-off between Precision and Economy in Gene Expression. *Nat. Commun.* **2019**, *10* (1), 1–15. <https://doi.org/10.1038/s41467-018-07391-8>.
- (40) Bar-Even, A.; Paulsson, J.; Maheshri, N.; Carmi, M.; O’Shea, E.; Pilpel, Y.; Barkai, N. Noise in Protein Expression Scales with Natural Protein Abundance. *Nat. Genet.* **2006**, *38* (6), 636–643. <https://doi.org/10.1038/ng1807>.
- (41) Taniguchi, Y.; Choi, P. J.; Li, G.-W.; Chen, H.; Babu, M.; Hearn, J.; Emili, A.; Xie, X. S. Quantifying E. Coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **2010**, *329* (5991), 533–538. <https://doi.org/10.1126/science.1188308>.
- (42) Hartkoorn, R. C.; Sala, C.; Magnet, S. J.; Chen, J. M.; Pojer, F.; Cole, S. T. Sigma Factor F Does Not Prevent Rifampin Inhibition of RNA Polymerase or Cause Rifampin Tolerance in Mycobacterium Tuberculosis. *J. Bacteriol.* **2010**, *192* (20), 5472–5479. <https://doi.org/10.1128/JB.00687-10>.
- (43) Campbell, E. A.; Pavlova, O.; Zenkin, N.; Leon, F.; Irschik, H.; Jansen, R.; Severinov, K.; Darst, S. A. Structural, Functional, and Genetic Analysis of Sorangicin Inhibition of Bacterial RNA Polymerase. *EMBO J.* **2005**, *24* (4), 674–682. <https://doi.org/10.1038/sj.emboj.7600499>.
- (44) McClure, W. R.; Cech, C. L. On the Mechanism of Rifampicin Inhibition of RNA Synthesis. *J. Biol. Chem.* **1978**, *253* (24), 8949–8956. [https://doi.org/10.1016/S0021-9258\(17\)34269-2](https://doi.org/10.1016/S0021-9258(17)34269-2).
- (45) Sippel, A. E.; Hartmann, G. R. Rifampicin Resistance of RNA Polymerase in the Binary Complex with DNA. *Eur. J. Biochem.* **1970**, *16* (1), 152–157. <https://doi.org/10.1111/j.1432-1033.1970.tb01066.x>.
- (46) Sippel, A.; Hartmann, G. Mode of Action of Rifamycin on the RNA Polymerase Reaction. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **1968**, *157* (1), 218–219. [https://doi.org/10.1016/0005-2787\(68\)90286-4](https://doi.org/10.1016/0005-2787(68)90286-4).
- (47) Avsec, Ž.; Agarwal, V.; Visentin, D.; Ledsam, J. R.; Grabska-Barwinska, A.; Taylor, K. R.; Assael, Y.; Jumper, J.; Kohli, P.; Kelley, D. R. Effective Gene Expression Prediction from Sequence by Integrating Long-Range Interactions. *Nat. Methods* **2021**, *18* (10), 1196–1203. <https://doi.org/10.1038/s41592-021-01252-x>.
- (48) Zhou, J.; Theesfeld, C. L.; Yao, K.; Chen, K. M.; Wong, A. K.; Troyanskaya, O. G. Deep Learning Sequence-Based Ab Initio Prediction of Variant Effects on Expression and Disease Risk. *Nat. Genet.* **2018**, *50* (8), 1171–1179. <https://doi.org/10.1038/s41588-018-0160-6>.
- (49) Kelley, D. R. Cross-Species Regulatory Sequence Activity Prediction. *PLoS Comput. Biol.* **2020**, *16* (7), e1008050. <https://doi.org/10.1371/journal.pcbi.1008050>.
- (50) Kelley, D. R.; Reshef, Y. A.; Bileschi, M.; Belanger, D.; McLean, C. Y.; Snoek, J. Sequential Regulatory Activity Prediction across Chromosomes with Convolutional Neural Networks. *Genome Res.* **2018**, *28* (5), 739–750. <https://doi.org/10.1101/gr.227819.117>.
- (51) Agarwal, V.; Shendure, J. Predicting mRNA Abundance Directly from Genomic Sequence Using Deep Convolutional Neural Networks. *Cell Rep.* **2020**, *31* (7), 107663. <https://doi.org/10.1016/j.celrep.2020.107663>.

- (52) Einav, T.; Phillips, R. How the Avidity of Polymerase Binding to the $-35/-10$ Promoter Sites Affects Gene Expression. *Proceedings of the National Academy of Sciences* **2019**, *116* (27), 13340 LP – 13345. <https://doi.org/10.1073/pnas.1905615116>.
- (53) Urtecho, G.; Tripp, A. D.; Insigne, K. D.; Kim, H.; Kosuri, S. Systematic Dissection of Sequence Elements Controlling $\sigma 70$ Promoters Using a Genomically Encoded Multiplexed Reporter Assay in *Escherichia Coli*. *Biochemistry* **2019**, *58* (11), 1539–1551. <https://doi.org/10.1021/acs.biochem.7b01069>.
- (54) Brewster, R. C.; Jones, D. L.; Phillips, R. Tuning Promoter Strength through RNA Polymerase Binding Site Design in *Escherichia Coli*. *PLoS Comput. Biol.* **2012**, *8* (12), e1002811. <https://doi.org/10.1371/journal.pcbi.1002811>.
- (55) Tierrafría, V. H.; Rioualen, C.; Salgado, H.; Lara, P.; Gama-Castro, S.; Lally, P.; Gómez-Romero, L.; Peña-Loredo, P.; López-Almazo, A. G.; Alarcón-Carranza, G.; Betancourt-Figueroa, F.; Alquicira-Hernández, S.; Polanco-Morelos, J. E.; García-Sotelo, J.; Gaytan-Nuñez, E.; Méndez-Cruz, C.-F.; Muñiz, L. J.; Bonavides-Martínez, C.; Moreno-Hagelsieb, G.; Galagan, J. E.; Wade, J. T.; Collado-Vides, J. RegulonDB 11.0: Comprehensive High-Throughput Datasets on Transcriptional Regulation in *Escherichia Coli* K-12. *Microb Genom* **2022**, *8* (5). <https://doi.org/10.1099/mgen.0.000833>.
- (56) Keseler, I. M.; Gama-Castro, S.; Mackie, A.; Billington, R.; Bonavides-Martínez, C.; Caspi, R.; Kothari, A.; Krummenacker, M.; Midford, P. E.; Muñiz-Rascado, L.; Ong, W. K.; Paley, S.; Santos-Zavaleta, A.; Subhraveti, P.; Tierrafría, V. H.; Wolfe, A. J.; Collado-Vides, J.; Paulsen, I. T.; Karp, P. D. The EcoCyc Database in 2021. *Front. Microbiol.* **2021**, *12*, 711077. <https://doi.org/10.3389/fmicb.2021.711077>.
- (57) Zaslaver, A.; Bren, A.; Ronen, M.; Itzkovitz, S.; Kikoin, I.; Shavit, S.; Liebermeister, W.; Surette, M. G.; Alon, U. A Comprehensive Library of Fluorescent Transcriptional Reporters for *Escherichia Coli*. *Nat. Methods* **2006**, *3* (8), 623–628. <https://doi.org/10.1038/nmeth895>.
- (58) Parisutham, V.; Chhabra, S.; Ali, M. Z.; Brewster, R. C. Tunable Transcription Factor Library for Robust Quantification of Regulatory Properties in *Escherichia Coli*. *Mol. Syst. Biol.* **2022**, *18* (6), e10843. <https://doi.org/10.15252/msb.202110843>.
- (59) Ribeiro, A. S. Dynamics and Evolution of Stochastic Bistable Gene Networks with Sensing in Fluctuating Environments. *Phys. Rev. E* **2008**, *78* (6), 1–9. <https://doi.org/10.1103/PhysRevE.78.061902>.
- (60) Potapov, I.; Lloyd-Price, J.; Yli-Harja, O.; Ribeiro, A. S. Dynamics of a Genetic Toggle Switch at the Nucleotide and Codon Levels. *Phys. Rev. E* **2011**, *84* (3), 1–9. <https://doi.org/10.1103/PhysRevE.84.031903>.
- (61) Chong, S.; Chen, C.; Ge, H.; Xie, X. S. Mechanism of Transcriptional Bursting in Bacteria. *Cell* **2014**, *158* (2), 314–326. <https://doi.org/10.1016/j.cell.2014.05.038>.
- (62) Ma, J.; Wang, M. D. DNA Supercoiling during Transcription. *Biophys. Rev.* **2016**, *8* (Suppl 1), 75–87. <https://doi.org/10.1007/s12551-016-0215-9>.
- (63) Postow, L.; Hardy, C. D.; Arsuaga, J.; Cozzarelli, N. R. Topological Domain Structure of the *Escherichia Coli* Chromosome. *Genes Dev.* **2004**, *18* (14), 1766–1779. <https://doi.org/10.1101/gad.1207504>.
- (64) Willenbrock, H.; Ussery, D. W. Chromatin Architecture and Gene Expression in *Escherichia Coli*. *Genome Biol.* **2004**, *5* (12), 252. <https://doi.org/10.1186/gb-2004-5-12-252>.

- (65) McClure, W. R. Mechanism and Control of Transcription Prokaryotes, *Annu. Rev. Biochem* **1985**, *54*, 171–204.
- (66) Startceva, S.; Kandavalli, V. K.; Visa, A.; Ribeiro, A. S. Regulation of Asymmetries in the Kinetics and Protein Numbers of Bacterial Gene Expression. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **2019**, *1862* (2), 119–128. <https://doi.org/10.1016/j.bbagr.2018.12.005>.
- (67) Kandavalli, V. K.; Tran, H.; Ribeiro, A. S. Effects of σ Factor Competition Are Promoter Initiation Kinetics Dependent. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2016**, *1859* (10), 1281–1288. <https://doi.org/10.1016/j.bbagr.2016.07.011>.
- (68) Weinstein, Z. B.; Zaman, M. H. Evolution of Rifampin Resistance in Escherichia Coli and Mycobacterium Smegmatis Due to Substandard Drugs. *Antimicrob. Agents Chemother.* **2019**, *63* (1). <https://doi.org/10.1128/AAC.01243-18>.
- (69) Lloyd-Price, J.; Gupta, A.; Ribeiro, A. S. SGNS2: A Compartmentalized Stochastic Chemical Kinetics Simulator for Dynamic Cell Populations. *Bioinformatics* **2012**, *28* (22), 3004–3005. <https://doi.org/10.1093/bioinformatics/bts556>.
- (70) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81* (25), 2340–2361. <https://doi.org/10.1021/j100540a008>.