

Fundamental relationship between operon organization and gene expression

Han N. Lim¹, Yeong Lee, and Razika Hussein

Department of Integrative Biology, University of California, Berkeley, CA 94720

Edited by Richard M. Losick, Harvard University, Cambridge, MA, and approved May 19, 2011 (received for review April 8, 2011)

Half a century has passed since the discovery of operons (groups of genes that are transcribed together as a single mRNA). Despite the importance of operons in bacterial gene networks, the relationship between their organization and gene expression remains poorly understood. Here we show using synthetic operons in *Escherichia coli* that the expression of a given gene increases with the length of the operon and as its position moves farther from the end of the operon. These findings can be explained by a common mechanism; increasing the distance from the start of a gene to the end of the operon (termed the “transcription distance”) provides more time for translation to occur during transcription, resulting in increased expression. We confirmed experimentally that the increased expression is indeed due to increased translation. Furthermore our analysis indicates the translation initiation rate for an mRNA is six-fold greater during transcription than after its release, which amplifies the impact of the transcription distance on gene expression. As a result of these mechanisms, gene expression increases by ~40% for each 1,000 nucleotides of transcription distance. In summary, we demonstrate that a fundamental relationship exists between gene expression and the number, length, and order of the genes in an operon. This relationship has important implications for understanding the functional basis of genome organization and practical applications for synthetic biology.

systems biology | mathematical model | polycistronic

Operons are a central feature of bacterial gene regulation (1). Each operon consists of a group of adjacent genes that are cotranscribed as a single mRNA. It is estimated that 50% of the genes in *Escherichia coli* are transcribed at least some of the time as part of an operon (2). The organization of genes in operons can alter gene expression when specific regulatory mechanisms are present such as translational coupling (3), polarity (4), and/or feedback (5, 6). However, it is unclear whether operon organization has any impact on gene expression in the absence of these specific mechanisms.

In this study we used synthetic operons to systematically examine whether the number, length, or order of the genes in an operon modulates gene expression. These synthetic operons lack the mRNA-specific, regulatory mechanisms commonly found in native operons. We found that gene expression increases with the distance between the start of a gene and the end of the operon (“transcription distance”) due to increased translation. These findings reveal that operon organization can modulate levels and patterns of gene expression via a general mechanism.

Results

Gene Expression Increases with Operon Length. The relationship between gene expression and operon length was examined in four sets of operons constructed with the genes encoding the cyan (*cfp*), yellow (*yfp*), and/or mCherry (*rfp*) fluorescent proteins (Fig. 1 A–D and Fig. S1). The operons were constitutively transcribed from the pLlacO-1 promoter in the absence of the Lac repressor. The translation of the first gene in each operon was controlled by the highly efficient T710RBS7 ribosome binding sequence (RBS) and the other genes were translated from the moderately efficient

st7 RBS (7) (Fig. S2). RFP was not used as a reporter because fluorescence varied with cell density and growth rate.

In the four sets of operons, *cfp* or *yfp* was placed in the first position (closest to the promoter) and operon length was increased by adding downstream genes. The mean expression of *cfp* and *yfp* in these operons was plotted as a function of the distance from the start codon of the reporter gene to the end of the operon (the transcription distance, λ) (Fig. 1 A–D). The transcription distance varies with operon length as well as with gene position in an operon (the latter is examined below). We found that gene expression increased linearly with the transcription distance in all four sets of operons.

To gather further support for the linear relationship between expression and transcription distance, two additional panels of operons were constructed with varying sized fragments of the *lacZ* gene [with codons 11–41 deleted (Δ M15), resulting in an inactive β -galactosidase protein (8)] downstream of *cfp*. In these panels we also altered the RBS for *cfp* (resulting in a large decrease in CFP production) to determine whether the translational efficiency has an impact on the observed linear relationship (Fig. 1 E and F). The length of *lacZ* was extended from a fixed point at the 5' end of the gene (“lacZ1”, Fig. 1E) or it was extended from the 3' end of the gene (“lacZ2”, Fig. 1F). In both panels a significant linear relationship exists between gene expression and transcription distance (Fig. 1 E and F).

A quadratic regression was performed on the data from the lacZ1 ($R^2 = 0.36$, $df = 39$, $P = 0.0002$) and lacZ2 panels ($R^2 = 0.51$, $df = 14$, $P = 0.007$). Although these fits were significant, they were not significantly better than the linear regression ($P = 0.83$ and 0.62 , respectively). Therefore, there is no evidence to indicate loss of the linear relationship between gene expression and transcription distance at longer distances.

A substantial amount of variation was observed in the fluorescence values at each transcription distance in the experiments (Fig. 1 A–F). Most of this variation reflects differences between cultures of the same strain measured on the same or different days. This variation was substantially reduced in operons that were integrated into the chromosome (see below), which suggests that it primarily arises from fluctuations in plasmid numbers (9).

Expression of a Gene Increases As Its Position Moves Proximally Within an Operon. Within an operon, the transcription distance of a gene increases as its position moves away from the end of the operon. Therefore, we would expect the expression of a gene at the first position to be higher than that of an identical gene at the second position, which should be higher than that of an identical gene at the third position (assuming the same RBS at the different positions). To test this prediction, two and three gene operons were

Author contributions: H.N.L. designed research; H.N.L., Y.L., and R.H. performed research; H.N.L. analyzed data; and H.N.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: hanlim@berkeley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1105692108/-DCSupplemental.

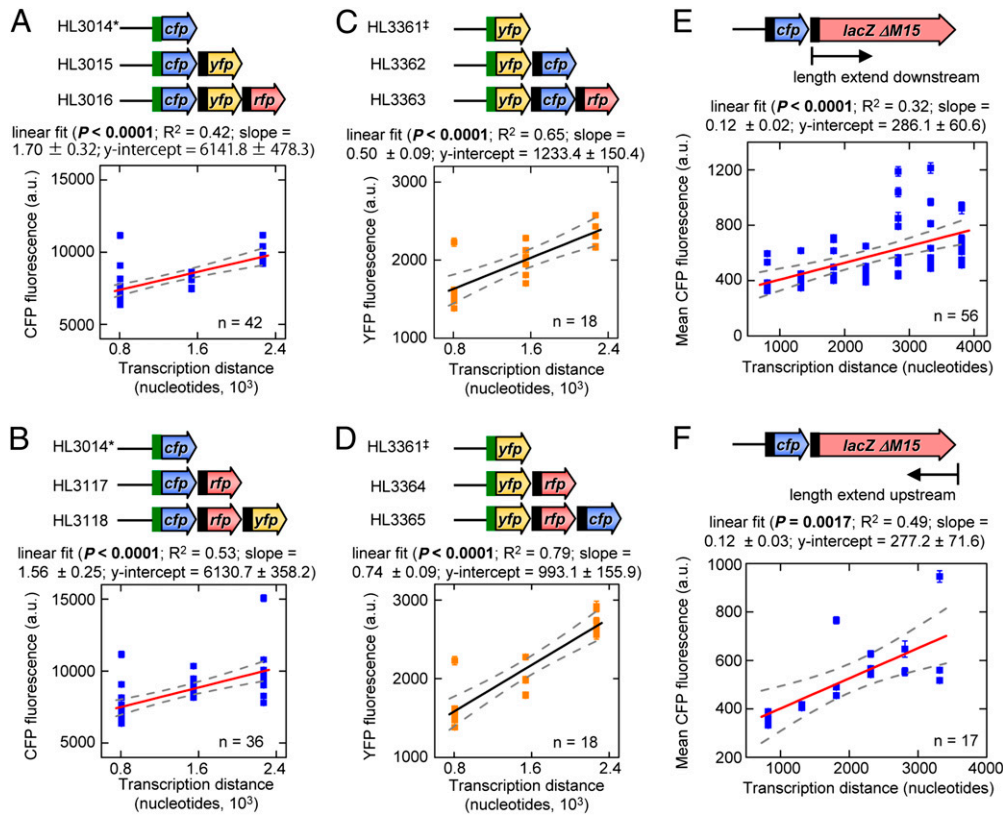


Fig. 1. Linear relationship between transcription distance and gene expression in operons of different length. Each data point is the mean fluorescence of the first gene in each operon \pm SEM from a separate culture. In the schematics of the operons, the green box indicates T710RBS7 and the black box indicates st7. (A and B) Two sets of operons with *cfp* in the first position of the operon. *, the monocistronic CFP fluorescence values are the same in both A and B. (C and D) Plots are the same as above except the two sets of operons have *yfp* in the first position. †, the monocistronic YFP fluorescence values are the same in both C and D. (E and F) Gene expression at different transcription distances in the lacZ1 (E) and lacZ2 (F) panels.

constructed with *cfp* and *yfp* in the first, second, or third position. All genes in the operons had the same RBS (st7).

The expression of CFP and YFP was compared in pairs of two and three gene operons that were identical except the positions of *cfp* and *yfp* were swapped (Fig. 2 A and B). As predicted, expression was always greater for the gene farthest from the end of the operon (i.e., the more proximal gene). The presence of *rfp* in the first position substantially increased the expression of all of the genes in an operon via a mechanism that is unclear (*rfp* also had an effect on gene expression that depended on whether it was in the second or the third position). Nonetheless, operons with *rfp* in the first position still showed significantly higher expression in the more proximal gene. This result indicates that sequence differences that alter absolute levels of transcription and translation do not alter the relationship between the transcription distance and gene expression. The CFP and YFP fluorescence values at each position were plotted as a function of transcription distance (excluding operons with *rfp* in the first position because the higher fluorescence would disproportionately affect the regression). This analysis showed a significant linear relationship between gene expression and transcription distance *within* an operon (Fig. 2 C and D).

General Model of Operon Translation. We created a general model to understand the relationship between gene expression and transcription distance. A detailed derivation of the model, with accompanying notes on the assumptions and simplifications, is in *SI Materials and Methods*. In the model, translation was separated into three phases; “transcriptional translation” (translation during mRNA transcription), “in-transit translation” (translation initi-

ated during transcription and completed after mRNA release), and “posttranscriptional translation” (translation initiated and completed after mRNA release) (Fig. 3A).

Transcriptional translation can commence once the start codon of a gene is transcribed and continues until the RNA polymerase encounters the terminator and releases the mRNA. The amount of time available for transcriptional translation is therefore determined by the transcription distance (λ) divided by the transcription rate (ρ) minus the lag time to create the first protein. The lag time is calculated by dividing the gene length in nucleotides (L) by the number of nucleotides per codon (μ) and the translation rate (σ_1 ; units of codons per second). Multiplying this time period by the rate of protein production (β_1 ; units are proteins per mRNA per second) gives the amount of transcriptional translation per mRNA, which is

$$\beta_1 \left(\frac{\lambda}{\rho} - \frac{L}{\mu\sigma_1} \right). \quad [1]$$

By definition $L \leq \lambda$ and it has been shown that $\mu\sigma_1 \sim \rho$ (10, 11). Both β_1 and β_2 (defined below) depend on the rate of translation initiation (ω_1 and ω_2 , respectively) and the fractions of these initiations that result in a complete protein (θ_1 and θ_2 , respectively) (*SI Materials and Methods*).

In-transit translation is typically determined by the number of ribosomes spanning the length of the gene before the mRNA is released (*SI Materials and Methods*). This number can be calculated by dividing the gene length (measured in codons) by the average spacing between each ribosome. The spacing is determined by the translation rate during transcription (σ_1) divided

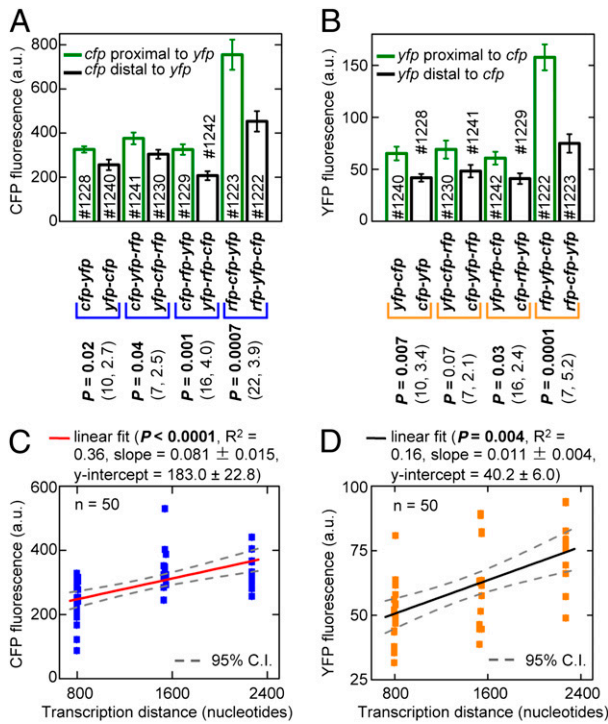


Fig. 2. Expression of a gene depends on its position within an operon. Error bars are \pm SEM. # indicates the HL strain number. (A and B) Mean CFP and YFP fluorescence in operon pairs. In each pair, the mean was compared using the two-tailed *t* test with a significance threshold of 0.05. The values in parentheses are the degrees of freedom and the *t* value, respectively. (C and D) Mean CFP and YFP fluorescence as a function of the transcription distance within operons. Each data point is from a separate culture.

by the average time between each *successful* translation initiation event ($1/\beta_1$). Therefore, the amount of in-transit translation per mRNA is

$$\frac{\beta_1 L}{\mu\sigma_1} \quad [2]$$

Posttranscriptional translation occurs after the mRNA is released until it is degraded. Therefore, the time available is determined by the mRNA lifetime (δ) minus the lag time to create the first protein. The lag time is the gene length in codons (L/μ) divided by the translation rate after mRNA release (σ_2 ; units are codons per second). Therefore, the amount of posttranscriptional translation per mRNA is

$$\beta_2 \left(\delta - \frac{L}{\mu\sigma_2} \right), \quad [3]$$

where β_2 is the protein production rate for free mRNA.

The total protein per mRNA is the sum of the protein produced by transcriptional, in-transit, and posttranscriptional translation. To obtain the total amount of protein in the cell at steady state, which is experimentally measured, the total protein per mRNA must be multiplied by the number of mRNAs transcribed per second (m), which is governed by the promoter's strength and divided by the protein degradation rate constant (γ in units of s^{-1}). That is,

$$\text{total amount of protein for a given gene in the cell} = \frac{m}{\gamma} \left[\beta_1 \left(\frac{\lambda}{\rho} - \frac{L}{\mu\sigma_1} \right) + \frac{\beta_1 L}{\mu\sigma_1} + \beta_2 \left(\delta - \frac{L}{\mu\sigma_2} \right) \right]. \quad [4]$$

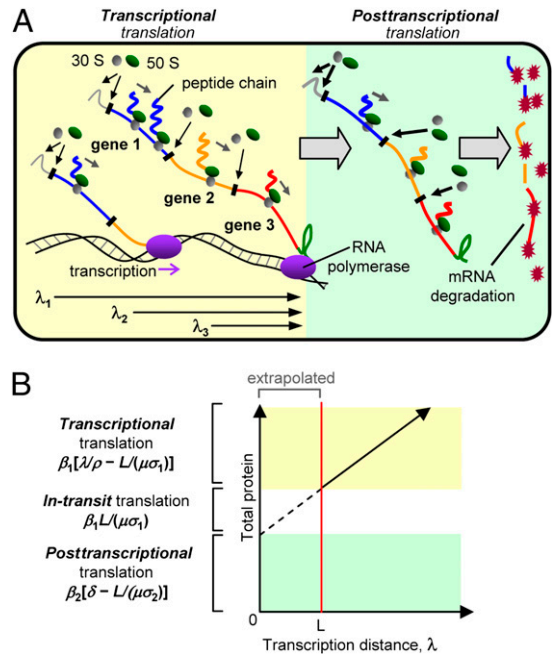


Fig. 3. Model of operon transcription and translation. (A) Translation during transcription (transcriptional translation) and following mRNA release (posttranscriptional translation). λ_1 , λ_2 , and λ_3 are the transcription distances for genes 1, 2, and 3, respectively. (B) The contributions to total protein in the cell. The estimated in-transit translation assumes the translation and transcription rates in units of codons per second are approximately equal.

The values for transcriptional, in-transit, and posttranscriptional translation can be obtained from plots of gene expression as a function of transcription distance (Fig. 3B).

Normalizing the total protein by the amount of protein from posttranscriptional translation gives

$$\frac{\beta_1}{\beta_2\rho(\delta - L/\mu\sigma_2)}\lambda + 1 = \varepsilon\lambda + 1. \quad [5]$$

The advantage of this normalization is that the slope, termed the translation coefficient (ε), is independent of mRNA production, protein degradation, and the fluorescent reporter. Therefore, the translation coefficient can be compared across different datasets. It can also be used to determine the ratio of transcriptional and posttranscriptional protein production by

$$\frac{\beta_1}{\beta_2} = \varepsilon\rho\left(\delta - \frac{L}{\mu\sigma_2}\right). \quad [6]$$

In summary, the model provides a simple explanation for the linear increase in gene expression with the transcription distance. It occurs because the transcription distance determines the length of time an mRNA is attached to the DNA and therefore the amount of transcriptional translation.

Operon Length and Gene Position Have Quantitatively Similar Effects on Gene Expression.

We examined the translation coefficient (ε) across the different experiments by normalizing the fluorescence values (i.e., total protein) by the extrapolated fluorescence values at $\lambda = 0$ (i.e., posttranscriptional translation). The translation coefficients were as follows (data source in parentheses): 2.77×10^{-4} (Fig. 1A), 2.54×10^{-4} (Fig. 1B), 4.05×10^{-4} (Fig. 1C), 7.45×10^{-4} (Fig. 1D), 4.19×10^{-4} (Fig. 1E), 4.33×10^{-4} (Fig. 1F), 4.43×10^{-4} (Fig. 2C), and 2.74×10^{-4} (Fig. 2D). The translation

coefficients from the different datasets are remarkably similar given the error in the fits; that the measurements are *in vivo*; and that manipulating the order, length, and sequence of the genes will have had confounding effects on expression (e.g., via altered mRNA folding). The differences most likely reflect errors in the data fits but they may also indicate true variation in the constants (β_1 , β_2 , δ , ρ , and σ_2).

The average translation coefficient was $4.06 \times 10^{-4} \pm 0.60 \times 10^{-4}$ (range = 7.45×10^{-4} – 2.54×10^{-4} and SD = 1.58×10^{-4}). That is, gene expression increases by 40% for each 1,000 nucleotides of transcription distance. The differences in the coding sequence, 5'-UTR, and RBS between the datasets do not appear to have had a major effect on the translation coefficient.

Dynamics of Gene Expression Within an Operon. The dynamics of expression were examined to confirm that gene expression increases with the transcription distance due to increased protein production (as predicted by the model) rather than decreased protein degradation or a delay in the expression of distal genes (12, 13). Gene expression was measured in two operons (5' *cfp-rfp-yfp* 3' and 5' *yfp-rfp-cfp* 3') after induction of transcription with isopropyl- β -D-thiogalactopyranoside (IPTG) (Fig. 4 A–C). We found that CFP and YFP expression at the first position in the operons was higher than at the third position for each time point. The protein concentration at each time point was normalized by the final concentration to determine their relative expression (R). R is independent of protein production and depends only on the protein degradation rate and any delay in the induction (*SI Materials and Methods*). When R was calculated, it was the same for positions 1 and 3 at each time point (Fig. 4 D and E). Therefore, protein degradation and any delay after induction are the same at both positions, indicating the differences in expression must be due to greater protein production in the more proximal position of the operons.

Transcription Distance Primarily Modulates Translation. Protein production may increase for genes with longer transcription dis-

tances due to increased translation from each mRNA or a higher mRNA concentration. To distinguish between these possibilities, gene expression and the mRNA concentration were compared between a monocistronic *cfp* gene and the *cfp-lacZ* operon (Fig. 5A). We found that CFP expression was 1.5-fold (± 0.1) greater for *cfp-lacZ* compared with *cfp* (Fig. 5B) but there was no significant difference in their mRNA concentrations as measured by quantitative RT-PCR (the 5' end of *cfp* was amplified in both mRNAs; Fig. 5C). Therefore, the increased CFP expression with the longer transcription distance is due to greater translation from each mRNA.

We also performed Northern blots to measure the mRNA concentrations and their decay (Fig. 5 D–G). The probe was located at the 5' end of *cfp* (a probe using the entire *cfp* sequence gave similar results; Fig. S3). The Northern blots showed the full-length *cfp-lacZ* mRNA transcript had slightly lower abundance than full-length *cfp* mRNA (Fig. 5 E and F). Again there was no increase in the *cfp-lacZ* mRNA concentration that would account for the increased CFP expression. It was further demonstrated that the *cfp* and *cfp-lacZ* mRNAs have similar half-lives; both are ~ 4 min (dividing the half-life by \log_2 yields the mRNA lifetime = 352 s; Fig. 5G). It should be noted that because the mRNA lifetime is calculated from the degradation rate, it does not include the time taken to transcribe the mRNA.

The mRNA lifetime ($\delta = 352 \pm 34$ s), the translation coefficient ($\epsilon = 4.06 \times 10^{-4} \pm 0.60 \times 10^{-4}$), gene length ($L = 720$ nucleotides), and the transcription rate measured under similar conditions ($\rho = 42 \pm 2$ nucleotides/s) (14) were substituted into Eq. 6. Furthermore, we assume that $\mu\sigma_2 \sim \rho$. This calculation yielded a β_1/β_2 ratio of 5.7 ± 1.2 . That is, the protein production rate is approximately six times greater during mRNA transcription than after its release.

Chromosomal Gene Expression Increases with Transcription Distance.

We tested whether the transcription distance also modulates the expression of genes in the chromosome. Strains were created with *cfp* and *yfp* at separate locations (*intS* and *galK*) (9) or at the same location as an operon (Fig. 6A). Unfortunately, we were unable to insert operons with three fluorescent genes into the chromosome. We found that expression was greater at the first position of an operon compared with a monocistronic gene with a shorter transcription distance (Fig. 6B). In contrast, the last gene in an operon and a monocistronic gene, which have similar transcription distances, have approximately the same expression levels.

Discussion

In this study we found that gene expression increases linearly with the distance from the start of a gene to the end of the operon (transcription distance). This relationship was observed in multiple sets of operons of different lengths, at different gene positions in multiple operons, with multiple coding sequences, and with different 5'-UTR sequences. Furthermore, the relative increase in translation per nucleotide of transcription distance (i.e., the translation coefficient) was similar across the different experiments and operons. Together these findings provide compelling support for a relationship between gene expression and the transcription distance and they indicate a common mechanism.

We proposed a general model that shows that genes with longer transcription distances have increased expression because they have a longer period for translation during transcription. There are three points in support of the model. First, it correctly predicts from first principles that the relationship between gene expression and the transcription distance is linear. Second, increasing the transcription distance was found to increase translation as predicted. Third, it provides a single explanation for why varying operon length and varying gene position have the same effect on gene expression (i.e., the same translation coefficient).

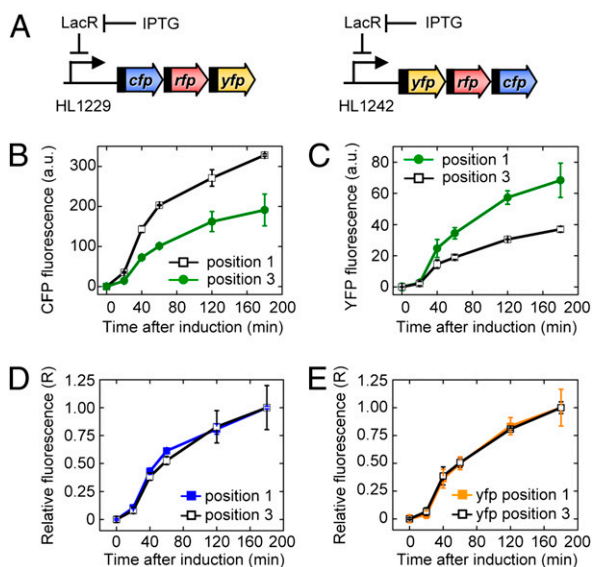


Fig. 4. The effect of gene position on the dynamics of expression. Data are the mean fluorescence \pm SEM. (A) The experimental system used to measure the dynamics of gene expression at different positions within the operon following the induction of transcription with IPTG. All genes had the st7 RBS. (B and C) Mean CFP and YFP fluorescence in the first and third positions of a three-gene operon. (D and E) Relative CFP and YFP fluorescence in the first and third positions of the operon (see main text).

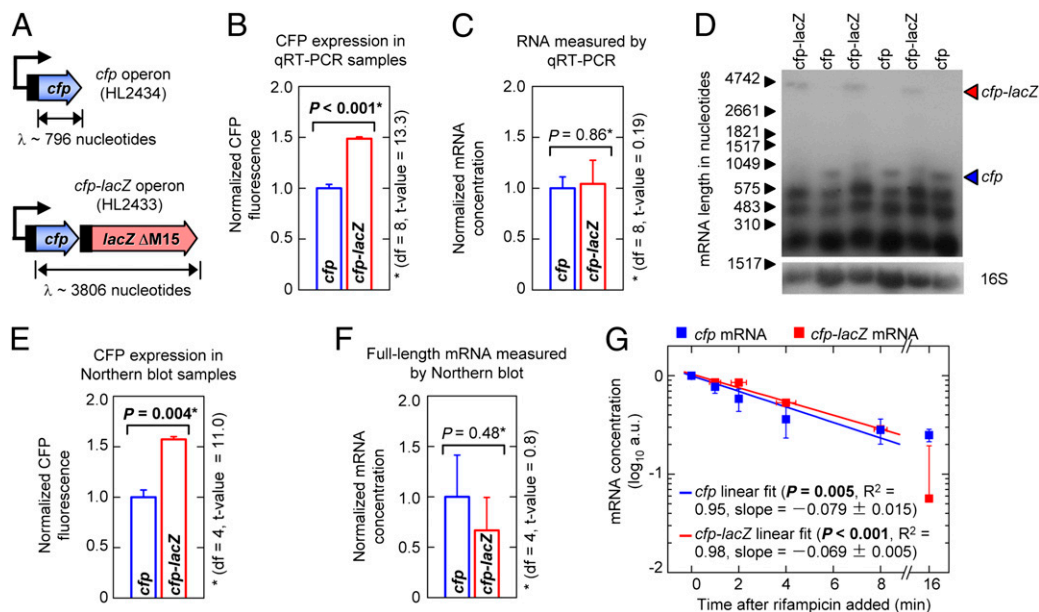


Fig. 5. Increasing transcription distance increases translation. Error bars are \pm SEM. *, the means were compared using a two-tailed t test with a significance threshold of 0.05. (A) Schematics of the *cfp* and *cfp-lacZ* operons. (B) Normalized CFP fluorescence in the *cfp* and *cfp-lacZ* operons. The difference in the expression of the short and long operons was slightly less than previously observed; this difference may be due to the higher cell density required for RNA extractions. (C) The relative concentration of *cfp* and *cfp-lacZ* mRNA normalized to the 16S RNA control as determined by quantitative RT-PCR. (D) Northern blots showing full-length *cfp* and *cfp-lacZ* mRNA. The contrast of the whole image was altered solely to enable visualization of the full-length *cfp* and *cfp-lacZ* mRNA transcripts; it played no role in the analysis. (E) CFP expression for the samples shown in D. (F) Relative concentration of full-length *cfp* and *cfp-lacZ* mRNA for the samples shown in D. (G) Decay of full-length *cfp* and *cfp-lacZ* mRNA measured by Northern blots following treatment with rifampicin ($t = 0$). The fits did not include the final time points due to the inaccuracy of the measurements at such low concentrations.

The most intriguing prediction of the model is that the production rate of proteins during transcription (β_1) is sixfold greater than after mRNA release (β_2). This result means that increasing the transcription time by 24 s (resulting from a 1-kb increase in transcription distance) has an equivalent effect on expression to increasing the mRNA lifetime by 144 s. The difference in transcriptional and posttranscriptional protein production rates could be due to local differences in ribosome concentrations and/or due to a different mRNA structure during transcription (15). In support of the former, ribosomes have been shown to be preferentially located at sites of active transcription in *Bacillus subtilis* (16) and there is mounting evidence that spatial localization is important for bacterial translation (17, 18). Both mechanisms would increase translation initiation.

The transcription distance had a measurable impact on expression in a wide variety of operons and there is evidence to suggest an association between gene expression and transcription distance in some native operons (19, 20). Therefore, although the transcription distance has only a moderate effect on gene expression, its role should not be ignored. In addition, altering gene expression by varying the transcription distance is fundamentally different from changing the RBS (21) or the transcription rate (22, 23); consequently it may have unique effects on gene noise and for coordinating expression from multiple genes. Furthermore, varying the transcription distance has a predictable effect on gene expression and this could be exploited to tune patterns and levels of gene expression in synthetic and native operons (24). In particular, it could be used to optimize gene order in operons to increase the output of a pathway (25) and to generate specific stoichiometries for protein complexes.

In conclusion, we show that operon organization can modulate levels and patterns of gene expression. It was known that the proximal genes in an operon can influence the expression of distal genes (e.g., polar mutations). Here we demonstrated that the converse also occurs; distal genes can regulate proximal genes in

an operon via their effect on the transcription distance. These findings also provide an example of how synthetic biology can help deconstruct complex biological processes. In this case, synthetic operons enabled the effect of operon organization on gene expression to be decoupled from the regulatory mechanisms that exist in native operons, thereby making it easier to identify. The next and more difficult task will be to investigate the role of the transcription distance in modulating the expression of non-fluorescent genes in native operons.

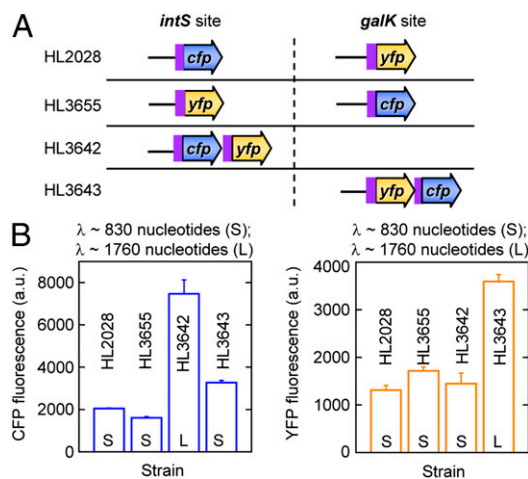


Fig. 6. Chromosomal gene expression increases with the transcription distance. (A) The arrangement of the genes at the *intS* and *galK* sites in the chromosome. The pLacO-1 promoter is used for all genes. The 5'-UTR, RBS, and the first 11 codons of the T7 10 gene (purple box) are fused to *cfp* and *yfp*. (B) The mean CFP and YFP fluorescence for each strain. Strains were measured in triplicate and error bars indicate the SEM.

Materials and Methods

Plasmids and Strains. Details of the strains and plasmids are in Fig. S1 A and B, Tables S1–S3, and *SI Materials and Methods*.

Measurements of Gene Expression. For steady-state and dynamics experiments, cells during early exponential growth ($OD_{600} = 0.1$) were harvested and placed on ice followed by centrifugation at $16,100 \times g$ for 1 min to concentrate the cells. The cells were resuspended in LB media and gene expression was measured by fluorescence microscopy using a TE2000E microscope (Nikon) with an X-cite 120PC lamp (Exfo) and 100 \times objective with phase 3 contrast. Images were captured with a Pixus 1024 pixel CCD camera (Princeton Instruments). Image capture and analysis were automated with Metamorph 7.0 (Molecular Devices). Only objects ± 2 SD from the mean were included in the analysis to eliminate cell debris and rare, exceptionally bright fluorescent cells. The mean fluorescence was determined by subtracting the background autofluorescence in cells without the plasmid. Further details are in *SI Materials and Methods*.

Quantitative RT-PCR Measurements. Total RNA was extracted from five separate exponentially growing cell cultures using TRIzol (Invitrogen). The RNA

was treated with DNase I (New England Biolabs) and cDNA was synthesized using the iScript select cDNA synthesis kit with random primers (Bio-Rad). Quantitative PCR was performed with oligonucleotides located at the 5' end of *cfp* or at the 3' end of *rrsB* (control for extraction and cDNA synthesis). Full details are in *SI Materials and Methods*.

Northern Blots. For each strain, the total RNA was extracted from three separate cultures at $OD_{600} = 0.2$ – 0.5 . Northern blots were performed using standard protocols and the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche). The membranes were visualized on radiographic film and quantified using Quantity One Analysis software (Bio-Rad). Further details are in *SI Materials and Methods*. The mRNA lifetime was determined by a linear fit to a semilog plot of the relative amount of mRNA as a function of time. The slope divided by $\log_{10}e$ is the decay constant (k) and $1/k$ is the mRNA lifetime.

ACKNOWLEDGMENTS. We thank Lusha Liang for constructing the chromosomal operons, Dena Block for constructing some of the plasmids, and David Adamson for helpful discussions and comments on the model and the manuscript. This work was supported by a University of California (Berkeley) Power Excellence Award (W5646).

- Rocha EP (2008) The organization of the bacterial genome. *Annu Rev Genet* 42: 211–233.
- Okuda S, et al. (2007) Characterization of relationships between transcriptional units and operon structures in *Bacillus subtilis* and *Escherichia coli*. *BMC Genomics* 8:48.
- Oppenheim DS, Yanofsky C (1980) Translational coupling during expression of the tryptophan operon of *Escherichia coli*. *Genetics* 95:785–795.
- Wek RC, Sameshima JH, Hatfield GW (1987) Rho-dependent transcriptional polarity in the *ilvGMEDA* operon of wild-type *Escherichia coli* K12. *J Biol Chem* 262:15256–15261.
- Mattheakis LC, Nomura M (1988) Feedback regulation of the *spc* operon in *Escherichia coli*: Translational coupling and mRNA processing. *J Bacteriol* 170:4484–4492.
- Yamada M, Saier MH, Jr (1988) Positive and negative regulators for glucitol (*gut*) operon expression in *Escherichia coli*. *J Mol Biol* 203:569–583.
- Vellanoweth RL, Rabinowitz JC (1992) The influence of ribosome-binding-site elements on translational efficiency in *Bacillus subtilis* and *Escherichia coli* in vivo. *Mol Microbiol* 6:1105–1114.
- Prentki P (1992) Nucleotide sequence of the classical *lacZ* deletion delta M15. *Gene* 122:231–232.
- Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186.
- Lacroute F, Stent GS (1968) Peptide chain growth of γ -galactosidase in *Escherichia coli*. *J Mol Biol* 35:165–173.
- Young R, Bremer H (1976) Polypeptide-chain-elongation rate in *Escherichia coli* B/r as a function of growth rate. *Biochem J* 160:185–194.
- Berberich MA, Kovach JS, Goldberger RF (1967) Chain initiation in a polycistronic message: Sequential versus simultaneous derepression of the enzymes for histidine biosynthesis in *Salmonella typhimurium*. *Proc Natl Acad Sci USA* 57:1857–1864.
- Alpers DH, Tomkins GM (1966) Sequential transcription of the genes of the lactose operon and its regulation by protein synthesis. *J Biol Chem* 241:4434–4443.
- Gotta SL, Miller OL, Jr., French SL (1991) rRNA transcription rate in *Escherichia coli*. *J Bacteriol* 173:6647–6649.
- de Smit MH, van Duin J (1990) Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc Natl Acad Sci USA* 87:7668–7672.
- Mascarenhas J, Weber MH, Graumann PL (2001) Specific polar localization of ribosomes in *Bacillus subtilis* depends on active transcription. *EMBO Rep* 2:685–689.
- Montero Llopis P, et al. (2010) Spatial organization of the flow of genetic information in bacteria. *Nature* 466:77–81.
- Nevo-Dinur K, Nussbaum-Shochat A, Ben-Yehuda S, Amster-Choder O (2011) Translation-independent localization of mRNA in *E. coli*. *Science* 331:1081–1084.
- Kovács K, Hurst LD, Papp B (2009) Stochasticity in protein levels drives colinearity of gene order in metabolic operons of *Escherichia coli*. *PLoS Biol* 7:e1000115.
- Nishizaki T, Tsuge K, Itaya M, Doi N, Yanagawa H (2007) Metabolic engineering of carotenoid biosynthesis in *Escherichia coli* by ordered gene assembly in *Bacillus subtilis*. *Appl Environ Microbiol* 73:1355–1361.
- Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27:946–950.
- Murphy KF, Balási G, Collins JJ (2007) Combinatorial promoter design for engineering noisy gene expression. *Proc Natl Acad Sci USA* 104:12726–12731.
- Cox RS, 3rd, Surette MG, Elowitz MB (2007) Programming gene expression with combinatorial promoters. *Mol Syst Biol* 3:145.
- Khalil AS, Collins JJ (2010) Synthetic biology: Applications come of age. *Nat Rev Genet* 11:367–379.
- Klipp E, Heinrich R, Holzhütter HG (2002) Prediction of temporal gene expression. Metabolic optimization by re-distribution of enzyme activities. *Eur J Biochem* 269: 5406–5413.