Supporting Information

Lim et al. 10.1073/pnas.1105692108

SI Materials and Methods

Construction of the Plasmids and Strains. Full details of the strains and plasmids are provided in Tables S1 and S2 and Fig. S1 *A* and *B*. The oligonucleotide sequences used to construct the plasmids are provided in Table S3. The *lacZ* sequence (with the ΔM15 mutation) was PCR amplified from pLEX/LacZ (Invitrogen). The mCFP, mYFP, and mCherry sequences were PCR amplified from plasmids provided by R. Tsien (University of California at San Diego, La Jolla, CA). The fluorescent proteins are monomeric (1, 2) and thus the fluorescence signal should be proportional to their concentration. There was no significant "bleed-through" between the fluorescence channels except from RFP to CFP but the magnitude was small (Fig. S1).

The st7 (3) and T710RBS7 ribosome binding sequences were created synthetically. The latter is a fusion of part of the noncoding sequence of the T7 10 gene to st7 (5′ cgccagcgcacaacgg-tttccctctagaaataattttgtttaact**ttaaggaggaaaaaaa**ATG 3′, with the start codon in uppercase and st7 sequence in boldface type). The T710RBS7 sequence was created because expression of the fluorescent proteins from the unmodified RBS and noncoding sequence of the T7 10 gene was very high and toxic to the cell when the gene was located on a medium-copy plasmid. (Note: it was not toxic when a single copy was inserted into the chromosome). The T1 terminator that was located at the end of each operon, the ColE origin, and the antibiotic resistance genes were obtained from the pZ system (4).

The chromosomal strains were constructed by inserting the promoter, the fluorescent gene, and the terminator via the lambda Red method at the *intS* and *galK* sites (5, 6). To create the two gene operons we first inserted one gene and then integrated the second gene immediately downstream. We then removed the promoter and terminator sequences in between the two genes to create a two-gene operon. The full 5′-UTR sequence from T7 10, which was used to regulate the translation of fluorescent reporters inserted into the chromosome, was obtained from the pET-11a plasmid (Stratagene).

The lacZ1 panel strains are HL2434, HL2763, HL2764, HL2765, HL2766, HL2767, and HL2433. The lacZ2 panel strains are HL2434, HL3251, HL3252, HL3253, HL3254, and HL3255.

Further Details on the Measurement of Gene Expression. Single colonies were grown overnight in LB media with 50 μg/mL kanamycin and diluted 1/25,000 in fresh LB media with antibiotic (± IPTG) and then grown for 4 h at 37 °C at 200 rpm to the early exponential growth phase (OD_{600} = 0.1). Dynamic experiments were performed in duplicate in the same manner except overnight cultures were diluted 1/1,500 in LB media with 50 μg/mL kanamycin and grown for 3 h. Several different dilutions of cells were then made (1/5–1/625) with fresh LB media and 50 μg/mL kanamycin and 1 mM IPTG so that there were flasks of different cell densities at the specified time points. We selected the flask with cells in early to midexponential growth at each time point for measurement.

For each sample, images of ∼500 cells were captured across 5–10 fields of view per slide. The measurement error without biological variation was 5.6 ± 2.5% (SEM). The error was determined from five separate measurements of CFP and YFP expression (with background autofluorescence subtracted) from a single flask for two different exponentially growing strains (HL3015 and HL1240).

The sets of excitation filter/dichroic mirror/emission filter used for fluorescence microscopy were 436 ± 10 nm/455 nm/480 ± 20 nm (CFP), 500 ± 10 nm/515 nm/535 ± 15 nm (YFP), and 575 ± 25 nm/610 nm/640 ± 25 nm (mCherry).

mRNA Measurements by Quantitative RT-PCR. Total RNA was extracted from five exponentially growing cell cultures using TRIzol (Invitrogen). The RNA was treated with DNase I (New England Biolabs) to remove contaminating DNA and then cDNA was synthesized using the iScript select cDNA synthesis kit and random primers (Bio-Rad). The concentration of cDNA was quantified by PCR using iQ SYBR Green Supermix (Bio-Rad) with the iQ5 Real-Time PCR detection system (Bio-Rad). A −RT sample was also prepared using an identical amount of DNase I-treated RNA without reverse transcriptase. The −RT sample was amplified in parallel with the cDNA sample to measure the concentration of any residual contaminating DNA.

The *cfp* and *cfp-lacZ* mRNAs were amplified with oligonucleotides located at the 5′ end of *cfp* (cyfp18F and cyfp122R). The stable 16S ribosomal RNA (encoded by the *rrsB* gene) was also measured to normalize for differences in RNA extraction, loading, and the efficiency of cDNA synthesis between samples. The amount of mRNA was calculated by

$$\frac{E_{CYFP}{}^{CT(\text{cDNA})} - E_{CYFP}{}^{CT(-\text{RT})}}{E_{rrsB}{}^{CT(\text{cDNA})} - E_{rrsB}{}^{CT(-\text{RT})}},$$

where E_{CYFP} and E_{rrsB} are the PCR efficiencies for the *cfp* and *rrsB* oligonucleotide pairs, respectively. The PCR efficiencies were determined from three samples at five different dilutions. The *CT* values were determined automatically by the iQ5 software.

Northern Blots. The steady-state mRNA concentration was measured in three separate cultures for each strain at OD_{600} = 0.2–0.5. Total RNA was extracted using the RNeasy mini kit and treated with RNase-free, DNase I (Qiagen). A 1:1 or 1.5:1 volume of RNA and loading buffer (250 μL 100% formamide, 83 μL 37% formaldehyde, 50 μL 10× Mops, 50 μL 100% glycerol, 10 μL 2.5% bromophenol blue, and 57 μL H_2O) plus a 1:4 ratio of RNA Molecular Weight Marker I DIG labeled (Roche) and loading buffer were electrophoresed on 1% agarose gels with 2% formaldehyde in 1× Mops buffer. The RNA was transferred to Hybond N+ membranes (GE Healthcare Life Sciences), using standard upward capillary transfer. The *cfp* and *rrsB* PCR products (which were the same as used for quantitative RT-PCR) and full-length *cfp* without the RBS were used as probes. The probes were labeled with DIG-11-dUTP by random-primed DNA labeling. The membranes were fixed with UV light (5 min exposure at 302 nm in the Bio-Rad Gel Doc XR imaging system) and then hybridized, washed, and detected according to the protocol described in the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche). The membranes were visualized on radiographic film and then digital images were captured by transillumination of the film using the Gel Doc XR imaging system. Band intensity was quantified on nonsaturated exposures with an algorithm based on the area and intensity of the bands, using Quantity One Analysis software (Bio-Rad).

To measure mRNA decay, rifampicin (Fisher Bioreagents) was added to exponentially growing cultures to a final concentration of 560 μM to terminate transcription. Cells were harvested at the specified times and 0.25 vol of stop solution was added (95% ethanol and 5% phenol). The extraction of total RNA and the measurement of the mRNA concentration by Northern blotting were performed as described above.

General Model for the Translation of Operons. We created a model from first principles that examines the potential effect of the transcription distance on gene expression. The model takes into account that bacterial translation occurs both during transcription ("transcriptional translation") and after mRNA release ("post-transcriptional translation"). There is also a period of transition when ribosomes bind during mRNA transcription but do not complete translation until after the mRNA has been released. Translation during this transition phase is referred to as "in-transit translation" because the ribosomes are in transit at the time of mRNA release. In all translation phases, the mRNA is simultaneously translated by multiple ribosomes moving at approximately the same rate (7, 8).

Transcriptional translation can commence once the coding sequence of a gene begins to be transcribed and continues until the RNA polymerase encounters a terminator causing mRNA release. Therefore, the maximum time available for transcriptional translation is proportional to the distance between a gene's start codon and the end of the operon (i.e., the transcription distance, λ, measured in nucleotides) divided by ρ, the average transcription rate in nucleotides per second. A complete protein cannot be generated until the full length of the gene has been transcribed and the ribosome reaches the end of the gene (L is the gene length in nucleotides). Therefore, there is a lag time before the first complete protein is made that is determined by the length of the gene measured in codons (i.e., L/μ, where $\mu = 3$ nt/codon) divided by the translation rate σ_1 in units of codons per second. That is, the time available to produce complete proteins from a gene during mRNA transcription is the total time minus the lag time,

$$\frac{\lambda}{\rho} - \frac{L}{\mu\sigma_1}, \qquad \textbf{[S1]}$$

where by definition $L \leq \lambda$. The first ribosome cannot progress faster than the RNA polymerase is creating the mRNA; therefore assuming the translation rate and transcription rate are constant, $\mu\sigma_1 \leq \rho$. Experimentally, the evidence indicates that $\mu\sigma_1 \sim \rho$ (9, 10), which is used below to simplify the equations. The time between the transcription of the RBS and the binding of the first ribosome is assumed to be very short and therefore was ignored.

If the translation rate during transcription is constant, then the number of proteins generated after the lag time ($L/\mu\sigma_1$) will simply be determined by the rate at which ribosomes are successfully loaded onto the mRNA at the RBS (i.e., the translation initiation rate), that is, after the lag time translation termination occurs at the same rate as translation initiation. The translation initiation rate during transcription in units of events per mRNA per second, ω_1, is determined by the strength of the RBS. Some of the ribosomes that initiate translation may dissociate during translation; therefore a term, θ_1 (units: proteins per event) is included in the model to account for the fraction remaining ($0 \leq \theta_1 \leq 1$). Therefore, the amount of transcriptional translation per mRNA is

$$\omega_1\theta_1\left(\frac{\lambda}{\rho} - \frac{L}{\mu\sigma_1}\right). \qquad \textbf{[S2]}$$

If $\lambda/\rho \leq L/(\mu\sigma_1)$, there will be no transcriptional translation and all of the translation that is initiated during transcription will contribute to in-transit translation.

In-transit translation is the amount of protein produced by ribosomes that bound during transcription but did not complete the synthesis of the full-length protein until after the mRNA was released. If the mRNA is released before the *first* full-length protein is created [i.e., $\lambda/\rho \leq L/(\mu\sigma_1)$], then the number of partially completed proteins is simply determined by the number of ribosomes that have bound (which is the product of the translation initiation rate and the fraction of ribosomes that do not dissociate) during the amount of time determined by the transcription distance. Therefore, the amount of in-transit translation per mRNA is

$$\omega_1\,\theta_1\,\frac{\lambda}{\rho}, \text{ where } \lambda \geq L. \qquad \textbf{[S3]}$$

When λ/ρ is equal to or exceeds the lag time [$L/(\mu\sigma_1)$], then ribosomes can span the entire gene length (see below). If $\lambda/\rho < L/(\mu\sigma_1)$, then only fraction f of the gene can be covered by ribosomes. f is the ratio of the transcription time (λ/ρ) to the lag time $L/(\mu\sigma_1)$. Consequently, Eq. S3 can be rewritten as

$$\frac{\omega_1\theta_1 L f}{\mu\sigma_1}, \text{ where } 0 \leq f \leq 1. \qquad \textbf{[S4]}$$

If mRNA release occurs after the *first* full-length protein is created, then there is enough time for ribosomes to span the full length of the gene sequence in the mRNA. The number of ribosomes (and therefore the number of partially translated proteins that will complete translation after mRNA release) is determined by the length of the gene (in codons) and the average spacing between each ribosome on an mRNA. The spacing is determined by the translation rate (σ_1) divided by the time between successful translation initiation events [$1/(\omega_1\theta_1)$]. Therefore, when the transcription time is equal to or exceeds the lag time, the amount of in-transit translation per mRNA is equal to

$$\frac{L}{\mu(\sigma_1/\omega_1\theta_1)} = \frac{\omega_1\theta_1 L}{\mu\sigma_1}. \qquad \textbf{[S5]}$$

Note that Eq. S5 is the same as Eq. S4 when $f = 1$, that is, when the transcription distance time is equal to or exceeds the time lag. As stated above, $f = 1$ in most cases because typically $\rho \sim \mu\sigma_1$.

Posttranscriptional translation follows the release of an mRNA until it is degraded. Therefore, the maximum amount of time available for posttranscriptional translation is determined by δ, the mean mRNA lifetime measured in seconds. The mRNA lifetime does not depend on its length as we and others have shown (11). The key components of the RNA degradosome appear to be located at the cell membrane (12–14) away from the site of transcription (15); therefore mRNA degradation during transcription was not incorporated into the model. Once mRNA degradation has been initiated, it is fast (16) compared with the mRNA lifetime.

A lag time also exists for posttranscriptional translation. The lag time between the binding of the first ribosome following mRNA release and the release of the first full-length protein is determined by the length of the gene in codons [i.e., L/μ] divided by the translation rate during posttranscriptional translation (σ_2). As above with transcriptional translation, the time taken to bind the first ribosome following mRNA release is assumed to be very short. Therefore, the time available for posttranscriptional translation is the mRNA lifetime minus the lag time, which is equal to

$$\delta - \frac{L}{\mu\sigma_2}. \qquad \textbf{[S6]}$$

The amount of posttranscriptional translation per mRNA will therefore be

$$\omega_2\theta_2\left(\delta - \frac{L}{\mu\sigma_2}\right), \qquad \textbf{[S7]}$$

where ω_2 is the translation initiation rate during posttranscriptional translation in units of events per mRNA per second and θ_2 is the fraction of ribosomes not dissociating during trans-

lation of the free mRNA, resulting in translation termination and the release a complete protein ($0 \le \theta_2 \le 1$; units are proteins per event). If the mRNA lifetime is equal to or less than the lag time (i.e., $\delta \le L/\mu\sigma_2$), then no full-length protein will be generated during posttranscriptional translation.

The total protein per mRNA is the sum of the protein produced by transcriptional, in-transit, and posttranscriptional translation, respectively, which is equal to

$$\omega_1\theta_1\left(\frac{\lambda}{\rho}-\frac{L}{\mu\sigma_1}\right) + \frac{\omega_1\theta_1 Lf}{\mu\sigma_1} + \omega_2\theta_2\left(\delta-\frac{L}{\mu\sigma_2}\right). \quad \text{[S8]}$$

Experimentally, we measure the total amount of protein in the cell at steady state rather than the amount of protein per mRNA. The total protein in the cell is equal to

$$\frac{m[\omega_1\theta_1(\lambda/\rho - L/\mu\sigma_1) + \omega_1\theta_1 Lf/\mu\sigma_1 + \omega_2\theta_2(\delta - L/\mu\sigma_2)]}{\gamma}, \quad \text{[S9]}$$

where m is the number of mRNAs produced per second (which is determined by the promoter's strength) and γ is the degradation rate of the protein in seconds. In our experiments m and γ are constant.

It has been established in several studies that proteins are translated at a rate that is equivalent to transcription (i.e., $\rho \sim \mu\sigma_1$) (9, 10). Therefore, the lag time is generally about the same amount of time as the transcription of a gene; consequently $f \sim 1$. The translation initiation rates and the fraction of ribosomes that do not dissociate during translation can be grouped, resulting in $\beta_1 = \omega_1\theta_1$ and $\beta_2 = \omega_2\theta_2$. β_1 and β_2 are the number of proteins produced per mRNA per second during transcriptional translation and posttranscriptional translation. With these assumptions and simplifications Eq. **S9** can be rewritten as

$$\text{total protein per cell} = \frac{m[\beta_1(\lambda/\rho) + \beta_2(\delta - L/\mu\sigma_2)]}{\gamma}. \quad \text{[S10]}$$

The numerator in this expression corresponds to ϕ in Eq. **S14**. The equation can be simplified to

$$\text{total protein per cell} = b\lambda + c, \quad \text{[S11]}$$

where $b = (m\beta_1)/(\rho\gamma)$, and $c = [m\beta_2(\delta - L/(\mu\sigma_2)]/\gamma$. This form of the expression clearly shows the total amount of protein increases linearly with the transcription distance.

As stated in the main text, the different contributions of transcriptional, in-transit, and posttranscriptional translation can be obtained from a plot of total protein as a function of transcription distance (Fig. 3*B*). The CFP and YFP fluorescence levels in each cell are assumed to be directly proportional to the total amount of protein per cell. The extrapolated amount of protein at $\lambda = 0$ is the amount generated by posttranscriptional translation. The difference in the amount of protein at $\lambda = L$ and $\lambda = 0$ is the amount of protein due to in-transit translation (with the assumption that $\rho \sim \mu\sigma_1$). The amount of transcriptional translation is the amount of protein at transcription distances exceeding the gene length, again assuming that $\rho \sim \mu\sigma_1$.

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

$$\frac{\beta_1}{\beta_2\rho(\delta - L/\mu\sigma_2)}\lambda + 1 = \varepsilon\lambda + 1. \quad \text{[S12]}$$

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

$$\frac{\beta_1}{\beta_2} = \varepsilon\rho\left(\delta - \frac{L}{\mu\sigma_2}\right), \quad \text{[S13]}$$

which is derived from Eq. **S12**. The ratio of β_1 and β_2 will reflect differences in ribosome recruitment, translation initiation, and ribosome dissociation. The values for β_1 and β_2 will depend on the strength of the RBS but their ratio would be expected to remain constant. That is, altering the RBS may increase or decrease the stability of ribosomes bound to the mRNA or the proportion of the bound ribosomes that actually progress to translation initiation. If the RBS has the same proportional effect on both transcriptional and posttranscriptional translation then the β_1/β_2 ratio and consequently the translation coefficient will be independent of RBS efficiency (as was observed).

Modeling the Dynamics of Gene Expression Within an Operon. The position of a gene within an operon determines its transcription distance and therefore its average protein production rate (protein per cell per unit time). The protein production rate also determines the dynamics of expression. The standard dynamic equation for gene expression is

$$\frac{dP}{dt} = \phi - \gamma P, \quad \text{[S14]}$$

(17), where ϕ is the production rate of protein P and γ is its degradation rate. This equation assumes the mRNA concentration is at steady state because mRNAs typically have shorter lifetimes than proteins. This assumption holds for our experiments; the mean mRNA lifetime was 352 s (*Results*) and the mean protein (CFP and YFP) lifetime that is determined by the cell growth rate (17) is ~2,256 ± 180 s.

Integrating the above equation yields the total protein concentration as a function of time,

$$P(t) = \frac{\phi}{\gamma}(1 - e^{-\gamma t}), \quad \text{[S15]}$$

with the initial amount being zero. Therefore, a higher production rate ϕ due to a longer transcription distance results in higher expression at any time point after induction. This prediction was confirmed experimentally by measuring expression at the first and third positions in two operons (5′ *cfp-rfp-yfp* 3′ and 5′ *yfp-rfp-cfp* 3′) after induction of transcription with IPTG.

The higher expression of the first gene at any time point after induction could possibly be due to a difference in the degradation rate of the protein or a time delay between the induction of a gene at positions 1 and 3 instead of being due to a difference in protein production. To exclude the former two mechanisms we normalized the protein concentration at each time point $P(t)$ by the amount at the final time point $P(T)$ to eliminate the prefactor ϕ/γ. This relative expression R at gene position j, R_j, is equal to

$$\frac{1 - e^{-\gamma_j(\tau_j + t)}}{1 - e^{-\gamma_j(\tau_j + T)}}. \quad \text{[S16]}$$

R_j is independent of protein production and depends only on γ_j and any time delay τ_j that might exist. Therefore, if the degradation rates for proteins produced at positions 1 and 3 are the same ($\gamma_1 = \gamma_3$) and there is no difference in the delay in the induction at position 3 compared with position 1 ($\tau_1 = \tau_3$), then R at positions 1 and 3 (R_1 and R_3, respectively) will have an identical function. That is, R_1 and R_3 will have the same values at all time points; analysis of the data showed that this was indeed the case (Fig. 4 *D* and *E*). Therefore, the observed difference in the ex-

pression levels of genes at positions 1 and 3 during induction is not due to a time delay or different protein degradation rates; therefore it must be due to differences in protein production.

Although we have shown that the observed difference in the expression levels of genes at positions 1 and 3 is not due to a time delay, it is still possible that a time delay exists that is the same for the genes at both positions. Furthermore, it is possible that there is a difference in the time delay at positions 1 and 3 that is too small to be detected with the fluorescent reporters (and also too small to explain the difference in expression at different gene positions) due to their slow maturation times (1, 18) and long lifetimes (19).

1. Shaner NC, et al. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* 22: 1567–1572.
2. Zacharias DA, Violin JD, Newton AC, Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296:913–916.
3. 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.
4. Lutz R, Bujard H (1997) 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* 25:1203–1210.
5. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640–6645.
6. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186.
7. Slayter H, Kiho Y, Hall C, Rich A (1968) An electron microscopic study of large bacterial polyribosomes. *J Cell Biol* 37:583–590.
8. Miller OL, Jr., Hamkalo BA, Thomas CA, Jr (1970) Visualization of bacterial genes in action. *Science* 169:392–395.
9. Lacroute F, Stent GS (1968) Peptide chain growth of -galactosidase in Escherichia coli. *J Mol Biol* 35:165–173.
10. 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.
11. Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, Cohen SN (2002) Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci USA* 99:9697–9702.
12. Khemici V, Poljak L, Luisi BF, Carpousis AJ (2008) The RNase E of Escherichia coli is a membrane-binding protein. *Mol Microbiol* 70:799–813.
13. Lewis PJ, Thaker SD, Errington J (2000) Compartmentalization of transcription and translation in Bacillus subtilis. *EMBO J* 19:710–718.
14. Liou GG, Jane WN, Cohen SN, Lin NS, Lin-Chao S (2001) RNA degradosomes exist in vivo in Escherichia coli as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc Natl Acad Sci USA* 98:63–68.
15. Pogliano J (2002) Dynamic cellular location of bacterial plasmids. *Curr Opin Microbiol* 5:586–590.
16. Cannistraro VJ, Kennell D (1994) The processive reaction mechanism of ribonuclease II. *J Mol Biol* 243:930–943.
17. Alon U (2007) *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall/CRC, London).
18. Heim R, Cubitt AB, Tsien RY (1995) Improved green fluorescence. *Nature* 373:663–664.
19. Andersen JB, et al. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* 64:2240–2246.
20. Peters JE, Thate TE, Craig NL (2003) Definition of the Escherichia coli MC4100 genome by use of a DNA array. *J Bacteriol* 185:2017–2021.
21. Franchini AG, Egli T (2006) Global gene expression in Escherichia coli K-12 during short-term and long-term adaptation to glucose-limited continuous culture conditions. *Microbiology* 152:2111–2127.

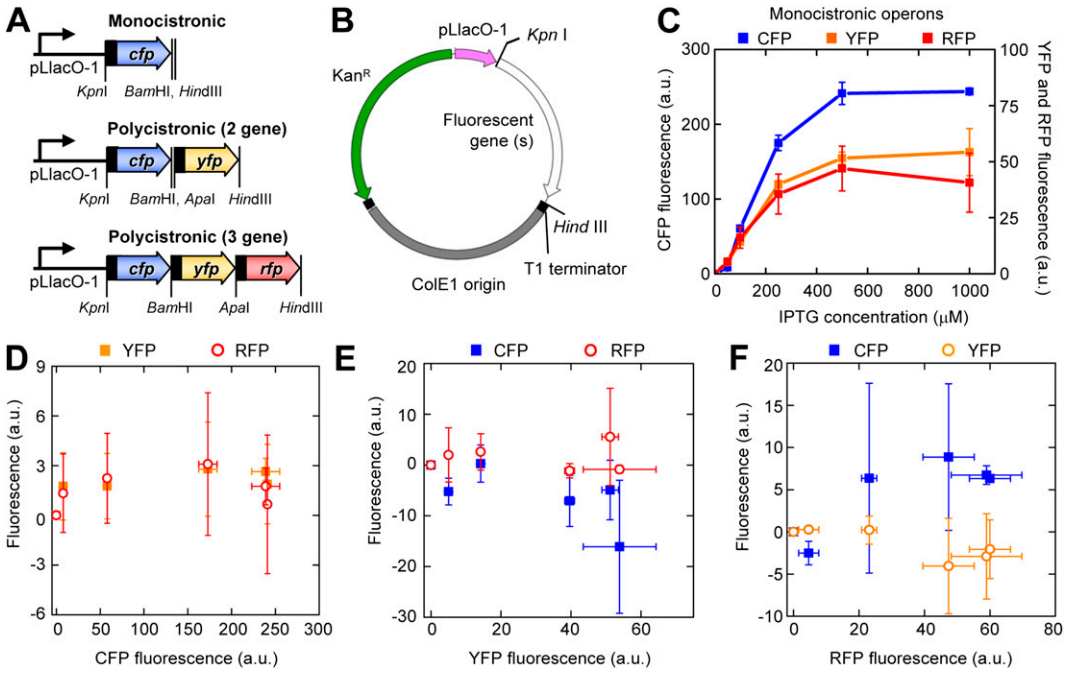

Fig. S1. The construction of synthetic operons and measurements of fluorescence "bleed-through". Data are presented as the mean fluorescence ± SEM. The mean fluorescence values were corrected for autofluorescence by subtracting the values measured at 0 mM IPTG (uninduced). Linear regression was performed using Origin 7.5 and the *P* value was determined using the two-tailed *t* test with the null hypothesis that the slope is 0. (*A*) The *cfp*, *yfp*, and *rfp* genes were assembled in different combinations to create monocistronic and polycistronic operons with various lengths and gene orders, using the restriction enzyme sites shown. Black boxes indicate the st7 ribosome binding site. (*B*) The organization of the plasmid carrying the operons showing the KpnI and HindIII restriction enzyme sites where the operons were inserted. (*C*) The mean CFP, YFP, and RFP fluorescence produced by the monocistronic *cfp*, *yfp*, and *rfp* genes, respectively, at different IPTG concentrations. Transcription was regulated by the pLlacO-1 promoter in a strain carrying the Lac repressor. (*D*) The amount of RFP and YFP fluorescence measured at different levels of monocistronic CFP fluorescence. There was no statistically significant linear correlation between CFP fluorescence and YFP (R^2 = 0.38, P = 0.19) or RFP fluorescence (R^2 = 0.06, P = 0.65). (*E*) The amount of CFP and RFP fluorescence measured at different levels of monocistronic YFP fluorescence. There was no statistically significant correlation between YFP fluorescence and CFP (R^2 = 0.50, P = 0.11) or RFP (R^2 = 0.03, P = 0.73) fluorescence. (*F*) The amount of CFP and YFP fluorescence at different levels of monocistronic RFP expression. A statistically significant correlation was identified between RFP fluorescence and CFP (R^2 = 0.91, P = 0.002, slope = 0.33 ± 0.05, y-intercept = 1.92 ± 1.48 a.u.) but the magnitude compared with the CFP expression that was generally observed was small. There was no statistically significant correlation between RFP fluorescence and YFP fluorescence (R^2 = 0.32, P = 0.24).

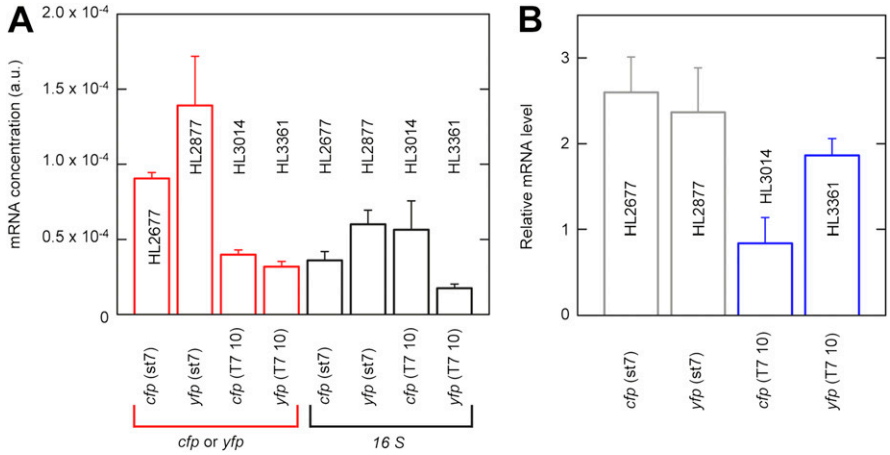

Fig. S2. Comparison of mRNA concentrations with the modified T7 10 sequence and the RBS st7 sequence. Data are presented as the mean fluorescence ± SEM. (*A*) The mRNA concentrations as measured by quantitative RT-PCR for strains with the 5′-UTR, the RBS, and part of the coding sequence of the T7 10 gene and with the st7 sequence. The red bars show the *cfp* or *yfp* mRNA concentrations and the black bars indicate the levels of the 16S loading control. (*B*) The normalized mRNA levels (*cfp* or *yfp* mRNA concentration divided by the 16S concentration).

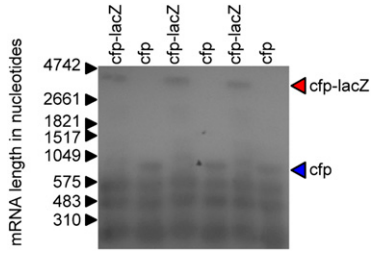

Fig. S3. Northern blot using the complete *cfp* sequence as a probe. The blot shown in Fig. 5*D* was stripped and hybridized with a probe generated using the full length of the *cfp* sequence. The contrast of the whole image was altered to aid the visualization of the full-length *cfp* and *cfp-lacZ* mRNA transcripts in the absence of backlight.

Table S1. Plasmids

Plasmids	Genes in operon (5′–3′)*	Construction
pHL265	Plasmid backbone	KanR, colE origin, pLlacO-1 promoter derived from the pZ plasmid system†
pHL274	*rfp*	pHL265 *backbone* plus *rfp* PCR product
pHL275	*cfp*	pHL265 *backbone* plus *cfp* PCR product
pHL276	*yfp*	pHL265 *backbone* plus *yfp* PCR product
pHL288	*cfp-yfp*	pHL275 plus *yfp* PCR product
pHL289	*yfp-rfp*	pHL276 plus *rfp* PCR product
pHL285	*rfp-cfp*	pHL274 plus *cfp* PCR product
pHL286	*rfp-yfp*	pHL274 plus *yfp* PCR product
pHL290	*yfp-cfp*	pHL276 plus *cfp* PCR product
pHL308	*rfp-yfp-cfp*	pHL285 plus *yfp* PCR product
pHL309	*rfp-cfp-yfp*	pHL286 plus *cfp* PCR product
pHL310	*cfp-yfp-rfp*	pHL287 plus *yfp* PCR product
pHL311	*cfp-rfp-yfp*	pHL288 plus *rfp* PCR product
pHL312	*yfp-cfp-rfp*	pHL289 plus *cfp* PCR product
pHL313	*yfp-rfp-cfp*	pHL290 plus *rfp* PCR product
pHL471	{KanR, pLlacO-1::T710::*cfp*::T1 terminator} template for integration into the chromosome	KanR, pLlacO-1:T710::*cfp*::T1 terminator and ColE. KanR, ColE and T1 terminator from pZ system†
pHL538	{KanR with FRT sites, pLlacO-1::T710::*cfp*} template for integration into the chromosome	KanR with FRT sites‡, pLlacO-1:T710::*cfp*::T1 terminator and ColE. ColE and T1 terminator from pZ system†
pHL582	{KanR with FRT sites, pLlacO-1::T710::*yfp*} template for integration into the chromosome	KanR with FRT sites‡, pLlacO-1:T710::*yfp*::T1 terminator and ColE. ColE and T1 terminator from pZ system†
pHL717	*cfp-lacZ* (1–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–2,982 bp) PCR product
pHL961	*cfp-lacZ* (1–498 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–498 bp) PCR product
pHL962	*cfp-lacZ* (1–1,002 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–1,002 bp) PCR product
pHL963	*cfp-lacZ* (1–1,500 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–1,500 bp) PCR product
pHL964	*cfp-lacZ* (1–2,001 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–2,001 bp) PCR product
pHL965	*cfp-lacZ* (1–2,502 bp)	pHL288 (minus *yfp*) plus *lacZ (*1–2,502 bp) PCR product
pHL1049	T710RBS7*cfp*	pHL275 (minus *cfp*) plus T710RBS7*cfp* PCR product
pHL1050	T710RBS7*cfp-yfp*	pHL288 (minus *cfp*) plus T710RBS7*cfp* PCR product
pHL1051	T710RBS7*cfp-yfp-rfp*	pHL310 (minus *cfp*) plus T710RBS7*cfp* PCR product
pHL1062	*cfp-lacZ* (499–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ* (499–2,982 bp) PCR product
pHL1063	*cfp-lacZ* (1,003–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ* (1,003–2,982 bp) PCR product
pHL1064	*cfp-lacZ* (1,501–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ* (1,501–2,982 bp) PCR product
pHL1065	*cfp-lacZ* (2,002–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ* (2,002–2,982 bp) PCR product
pHL1066	*cfp-lacZ* (2,503–2,982 bp)	pHL288 (minus *yfp*) plus *lacZ* (2,503–2,982 bp) PCR product
pHL1106	T710RBS7*cfp-rfp*	pHL1051 plus *rfp* PCR product
pHL1107	T710RBS7*cfp-rfp-yfp*	pHL1051 plus *rfp-yfp* from pHL311
pHL1144	T710RBS7*yfp*	pHL1049 (minus *cfp*) plus T710RBS7*yfp* PCR product
pHL1167	{CamR with FRT sites + T1 terminator + pLlacO-1::T710RBS::*yfp*::*aspA* terminator} template for integration into the chromosome	CamR with FRT sites§ + T1 terminator + pLlacO-1::T710RBS::*cfp*::*aspA* terminator¶; ColE and T1 terminator from pZ system†
pHL1168	{CamR with FRT sites + T1 terminator + pLlacO-1::T710RBS::*cfp*::*aspA* terminator} template for integration into the chromosome	CamR with FRT sites§ + T1 terminator + pLlacO-1::T710RBS::*yfp*::*aspA* terminator¶; ColE and T1 terminator from pZ system†
pHL1173	T710RBS7*yfp-cfp*	pHL288 (minus *yfp*) plus T710RBS7*yfp* from pHL1144
pHL1174	T710RBS7*yfp-cfp-rfp*	pHL312 (minus *cfp*) plus T710RBS7*yfp* from pHL1144
pHL1181	{KanR, pLlacO-1::T710::*yfp*::T1 terminator} template for integration into the chromosome	KanR, pLlacO-1:T710::*yfp*::T1 terminator and ColE; KanR, ColE, and T1 terminator from pZ system†
pHL1183	T710RBS7*yfp-rfp*	pHL287 (minus *cfp*) plus T710RBS7*yfp* from pHL1144
pHL1184	T710RBS7*yfp-rfp-cfp*	pHL313 (minus *yfp*) plus T710RBS7*yfp* from pHL1144

*Unless otherwise stated all genes have the RBS st7 sequence.
†See ref. 4.
‡PCR amplified from pKD13 (5).
§PCR amplified from pkD3 (5).
¶AspA term amplified from pLEX (Invitrogen).

Table S2. Strains

Strain	Description	Source/reference
MC4100	*E. coli* K-12 derivative	(20)
HL716	MG1655 + *lacIq* at *intS* site	This study
HL1219	HL716 + pHL274	This study
HL1222	HL716 + pHL308	This study
HL1223	HL716 + pHL309	This study
HL1226	HL716 + pHL275	This study
HL1228	HL716 + pHL288	This study
HL1229	HL716 + pHL311	This study
HL1230	HL716 + pHL312	This study
HL1238	HL716 + pHL276	This study
HL1239	HL716 + pHL289	This study
HL1240	HL716 + pHL290	This study
HL1241	HL716 + pHL310	This study
HL1242	HL716 + pHL313	This study
HL2028	MG1655, Δ*lacI*, pLlacO-1::T710::*cfp* {PCR amplified from pHL538} at *intS*, pLlacO-1::T710::*yfp* {PCR amplified from pHL582} at *galK*	This study
HL2433	MC4100 + pHL717	This study
HL2434	MC4100 + pHL275	This study
HL2677	MC4100 + pHL275	This study
HL2763	MC4100 + pHL961	This study
HL2764	MC4100 + pHL962	This study
HL2765	MC4100 + pHL963	This study
HL2766	MC4100 + pHL964	This study
HL2767	MC4100 + pHL965	This study
HL2877	MC4100 + pHL276	This study
HL3014	MC4100 + pHL1049	This study
HL3015	MC4100 + pHL1050	This study
HL3016	MC4100 + pHL1051	This study
HL3117	MC4100 + pHL1106	This study
HL3118	MC4100 + pHL1107	This study
HL3251	MC4100 + pHL1062	This study
HL3252	MC4100 + pHL1063	This study
HL3253	MC4100 + pHL1064	This study
HL3254	MC4100 + pHL1065	This study
HL3255	MC4100 + pHL1066	This study
HL3361	MC4100 + pHL1144	This study
HL3362	MC4100 + pHL1173	This study
HL3363	MC4100 + pHL1174	This study
HL3364	MC4100 + pHL1183	This study
HL3365	MC4100 + pHL1184	This study
HL3355	MG1655 + *KanR* + pLlacO::T710::*cfp*::T1 term {PCR amplified from pHL471} at the *gtrAB* site in reverse orientation	This study
HL3368	MG1655 + *KanR* + pLlacO::T710::*yfp*::T1 term {PCR amplified from pHL1181} at the *galM* site in reverse orientation	This study
HL3515	HL3355 + T1 term + pLlacO::T710::*yfp*::*aspA* term {PCR amplified from pHL1167} at the *intS* site immediately downstream of pLlacO::T710::*cfp*::T1 term	This study
HL3546	HL3368 + T1 term + pLlacO::T710::*cfp*::*aspA* term { PCR amplified from pHL1168} at the *galK* site immediately downstream of pLlacO::T710::*yfp*::T1 term	This study
HL3642	HL3515, Δ T1 term, Δ *lacI*	This study
HL3643	HL3546, Δ T1 term, Δ *lacI*	This study
HL3655	MG1655, Δ*lacI*, pLlacO-1::T710::*cfp* {PCR amplified from pHL538} at *galK*, pLlacO-1::T710::*yfp* {PCR amplified from pHL582} at *intS*	This study

term, terminator.

Table S3. Oligonucleotides

Oligonucleotide	Description	Sequence (5′–3′)
pLacOAatSalF	Amplify pLlacO-1 with AatII and SalI	catgacgtctcgaactgcaggtcgactgatccataaatgtgagcggataacattgacatt
pLacKpnIR2	Amplify pLlacO-1 with KpnI	cttaggtaccagtatcttgttatccgctcacaatgtcaatgttatccgctcacatttatg
RCYFPRBSAatIIKpnIF	Amplify *cfp*, *yfp*, and *rfp* with st7 RBS and AatII and KpnI	cgcgacgtccataggtacctaaggaggaaaaaaaatggtgagcaagggcgaggag
RCYFPRBSBamApaF	Amplify *cfp*, *yfp*, and *rfp* with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatggtgagcaagggcgaggag
RCYFPRBSBamF	Amplify *cfp*, *yfp*, and *rfp* with st7 RBS and BamHI	cgcggatcccatacggccctaaggaggaaaaaaaatggtgagcaagggcgaggag
RCYHindBamR	Amplify *cfp*, *yfp*, and *rfp* with BamHI and HindIII	cgcaagcttcataggatccttacttgtacagctcgtccatgcc
RCYApaIR	Amplify *cfp*, *yfp*, and *rfp* with ApaI	catgggcccttacttgtacagctcgtccatgcc
RCYFPNoRBSEcoRF	Amplify *yfp* for insertion into pHL1049 (minus *cfp*). Also amplify full-length *cfp* probe for the Northern blot.	catgaattcatggtgagcaagggcgaggag
CYFP18F	For RT-PCR and Northern blot probes	ggagctgttcaccggggtggtgc
CYFP122R	For RT-PCR and Northern blot probes	gatgaacttcagggtcagcttgc
LacZRBSBamApaF	Amplify *lacZ* with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatgaccatgattacggattcactg
LacZBamHindR	Amplify *lacZ* with BamHI and HindIII	cgcggatcccataaagcttttatttttgacaccagaccaactgg
LacZ498M15BamHindR	Amplify *lacZ* to nucleotide 498 (lacZΔM15) with BamHI and HindIII	cgcggatcccataaagcttttaactgccgtcactccagcgcag
LacZ1002M15BamHindR	Amplify *lacZ* to nucleotide 1,002 (lacZΔM15) with BamHI and HindIII	cgcggatcccataaagcttttaatgcagaggatgatgctcgtgacg
LacZ1500M15BamHindR	Amplify *lacZ* to nucleotide 1,500 (lacZΔM15) with BamHI and HindIII	cgcggatcccataaagcttttactctccaggtagcgaaagccattt
LacZ2001M15BamHindR	Amplify *lacZ* to nucleotide 2,001 (lacZΔM15) with BamHI and HindIII	cgcggatcccataaagcttttagagccagagttgcccggcgctctc
LacZ2502M15BamHindR	Amplify *lacZ* to nucleotide 2,502 (lacZΔM15) with BamHI and HindIII	cgcggatcccataaagcttttattgaccactaccatcaatccggta
LacZ499RBSBamApaF	Amplify *lacZ* from nucleotide 499 (lacZΔM15) with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatgctggaagatcaggatatgtgg
LacZ1003RBSBamApaF	Amplify *lacZ* from nucleotide 1,003 (lacZΔM15) with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatgcaggtcatggatgagcagacg
LacZ1501RBSBamApaF	Amplify *lacZ* from nucleotide 1,501 (lacZΔM15) with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatgcgcccgctgatcctttgcgaa
LacZ2002RBSBamApaF	Amplify *lacZ* from nucleotide 2002 (lacZΔM15) with st7 RBS and *BamH*I and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatggtacgcgtagtgcaaccgaac
LacZ2503RBSBamApaF	Amplify *lacZ* from nucleotide 2,503 (lacZΔM15) with st7 RBS and BamHI and ApaI	cgcggatcccatagggccctaaggaggaaaaaaaatggcgattaccgttgatgttgaa
pkD1FIntCF	Used to PCR amplify and integrate pLlacO-1::T710::*cfp* or pLlacO-1::T710::*yfp* at *intS* with pHL538 and pHL582 as template	atagttgttaaggtcgctcactccaccttctcatcaagccagtccgcccagtgtaggctggagctgcttc
CYFPIntCR	Use with pkD1FIntCF	ccgtagatttacagttcgtcatggttcgcttcagatcgttgacagccgcattacttgtacagctcgtccatgcc
pkD1FGalKF	Used to PCR amplify and integrate pLlacO-1::T710::*cfp* or pLlacO-1::T710::*yfp* at *galK* with pHL538 and pHL582 as template	ttcatattgttcagcgacagcttgctgtacggcaggcaccagctcttccggtgtaggctggagctgcttc
CYFPGalKR	Use with pkD1FGalKF	gtttgcgcgcagtcagcgatatccattttcgcgaatccggagtgtaagaattacttgtacagctcgtccatgcc
gtrBFterm0ext	Used to PCR amplify and insert pLlacO-1::T710::*cfp*::T1 term at *gtr*AB position using pHL471 as template	tcatttttttgactctcttgatgatgtatttcgggcgtttttttggtttcaaatcaacaggagtccaagcgagctctcg
IntCFRCCYFPR	Use with gtrBFterm0ext	tgggcggactggcttgatgagaaggtggagtgagcgaccttaacaactatttacttgtacagctcgtccatgcc

Table S3. Cont.

Oligonucleotide	Description	Sequence (5′–3′)
GalMFterm0ext	Used to PCR amplify and insert pLlacO-1::T710::*yfp*::T1 term at *galM* position using pHL1181 as template	ctggtgatttgaacaatatgagataaagccctcatgacgagggcgtaacaatcaacaggagtccaagc gagctctcg
GalKFRCCYFPR	Use with GalMFterm0ext	cggaagagctggtgcctgccgtacagcaagctgtcgctgaacaatatgaattacttgtacagctcgtccatgcc
CYFPendpkD1F	Used to PCR amplify and insert {CamR with FRT s + T1 term + pLlacO-1::T710RBS::*cfp*::AspTerm} or {CamR with FRT s + T1 term + pLlacO-1::T710RBS::*yfp*::AspTerm} from pHL1167 or pHL1168 at *intS* and *galK* sites	tcgtgaccgccgccgggatcactcacggcatggacgagctgtacaagtaagtgtaggctggagctgcttc
IntCColER	Use with CYFPendpkD1F	ccgtagatttacagttcgtcatggttcgcttcagatcgttgacagccgcaagctgataccgctcgccgca gccgaacg
GalKColER	Use with CYFPendpkD1F	gtttgcgcgcagtcagcgatatccattttcgcgaatccggagtgtaagaaagctgataccgctcgccgca gccgaacg
rrnB1361F*	Used to PCR amplify for *rrsB* RT-PCR and Northern blot	gaatgccacggtgaatacgtt
rrnB1475R	Used to PCR amplify for *rrsB* RT-PCR and Northern blot	cacaaagtggtaagcgccct

*Sequence from ref. 21.