UC San Diego
UC San Diego Previously Published Works

Title

Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria.

Permalink

https://escholarship.org/uc/item/1zd454pp

Journal

Science (New York, N.Y.), 378(6624)

ISSN

0036-8075

Authors

Balakrishnan, Rohan
Mori, Matteo
Segota, Igor
et al.

Publication Date

2022-12-01

DOI

10.1126/science.abk2066

Peer reviewed

Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria

Rohan Balakrishnan[1,†], **Matteo Mori**[1,†], **Igor Segota**[2], **Zhongge Zhang**[3], **Ruedi Aebersold**[4,5], **Christina Ludwig**[6], **Terence Hwa**[1,3,*]

[1]Department of Physics, University of California at San Diego, La Jolla, California 92093-0374.

[2]Departments of Medicine and Pharmacology, University of California at San Diego, La Jolla, California 92093.

[3]Section of Molecular Biology, Division of Biological Sciences, University of California at San Diego, La Jolla, California 92093.

[4]Faculty of Science, University of Zurich, Zurich, Switzerland.

[5]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Switzerland.

[6]Bavarian Center for Biomolecular Mass Spectrometry (BayBioMS), Technical University of Munich (TUM), Freising, Germany.

Abstract

Protein concentrations are set by a complex interplay between gene-specific regulatory processes and systemic factors, including cell volume and shared gene expression machineries. Elucidating this interplay is crucial for discerning and designing gene regulatory systems. We quantitatively characterized gene-specific and systemic factors affecting transcription and translation genome-wide for *E. coli* across many conditions. The results revealed two design principles that make regulation of gene expression insulated from concentrations of shared machineries: RNA polymerase activity is fine-tuned to match translational output, and translational characteristics are similar across the majority of mRNAs. Consequently, in bacteria protein concentration is set primarily at the promoter level. A simple mathematical formula relates promoter activities and protein concentrations across growth conditions, enabling quantitative inference of gene regulation from 'omics data.

ONE SENTENCE SUMMARY

Coordinated transcription-translation fluxes and predominance of transcriptional control lead to a quantitative formulation of the Central Dogma in bacteria.

Gene expression involves transcription of genes into mRNAs, followed by translation of mRNAs into proteins. Protein concentrations are in turn determined by the balance between protein synthesis and dilution (Fig. 1A, Fig. S1) for exponentially growing bacteria in which protein degradation is negligible(1, 2). Intuitively, doubling the transcription initiation rate of a specific gene by an activator should result in doubling the concentration of the corresponding mRNA and protein in the absence of post-transcriptional regulation(3–5). However, if the transcription initiation rate of every gene were doubled, say by a global activator, protein concentrations could not possibly double: First, the synthesis of proteins is constrained by the translational capacity of the ribosomes(6, 7). Second, in the well-characterized model bacterium *E. coli* the total number of proteins per cell volume does not vary much across different growth conditions (Fig. S2 and Ref. (8, 9)), making it impossible to change the concentrations of the majority of proteins in the same direction even if there is no constrain in gene expression capacity. Thus, the canonical single-gene picture of bacterial gene expression is not necessarily compatible with global constraints during broad changes in gene expression, as they occur, e.g., upon changes in nutrient conditions or exposure to antibiotics. These effects make it difficult to link the transcriptional and translational regulation to the concentrations of mRNAs and proteins.

As the expression of each gene is ultimately determined by the rates at which the respective mRNA and proteins are synthesized and diluted (Fig. S1), we designed a battery of experiments to determine these rates by measuring the absolute mRNA and protein concentrations and their fluxes, for *E. coli* growing exponentially under various conditions. Our findings establish characteristics of promoters and mRNAs that defied our expectations, and reveal design principles underlying *E. coli*'s gene regulation program which enable the cell to allocate its proteome in accordance to functional needs while complying with cellular constraints. We established a simple, quantitative relation which connects gene regulatory activities to mRNA and protein concentrations.

RESULTS

Translation initiation rates are similar across mRNAs and growth conditions

A proteomics workflow was developed to quantify the abundance of individual *E. coli* proteins(10) by combining the versatility of data-independent acquisition (DIA) mass spectrometry(11, 12) and the accuracy of ribosome profiling(13). We determined the protein number fractions $\psi_{p,i} \equiv [P_i]/[P]$ for >1900 proteins (labeled by i), with $[P] = \sum_i [P_i]$ being the total protein concentration, defined here as number of proteins per cell volume. RNA-sequencing was used to determine the mRNA number fractions $\psi_{m,i} \equiv [mR_i]/[mR]$ for the corresponding mRNAs, with $[mR] \equiv \sum_i [mR_i]$ being the total mRNA concentration; see SI Methods. In both cases, our data showed high reproducibility (Fig. S3A– C). The result for *E. coli* K-12 cells growing exponentially in glucose minimal medium is shown as a scatter plot of $\psi_{p,i}$ vs $\psi_{m,i}$ in Fig. 1B. We observed a strong correlation ($r = 0.80$) of mRNA and

protein abundances along the diagonal (red line) across a vast range of abundances (10^{-2} to 10^{-6}). The histogram of $\psi_{p,i}/\psi_{m,i}$ is peaked around 1, with 50% of the genes within 1.7-fold (Fig. S3E). We repeated the measurements for cells growing exponentially in three types of growth limiting conditions in minimal medium (carbon limitation, anabolic limitation, and translational inhibition (14, 15), with growth rates ranging between 0.3/h and 0.9/h. A similar number of gene products were detected in these conditions and the resulting scatter plots and histograms (Fig. S3F through S) looked very similar to those from bacteria grown in glucose minimal medium (Fig. S3DE). These results, summarized in Fig. 1C and 1D, indicate that the fractional abundances of mRNA and proteins are approximately the same, i.e.,

$$\psi_{m,i} \approx \psi_{p,i}, \tag{1}$$

for the vast majority of expressed genes in all growth conditions tested. The strong correlation between mRNA and protein fractions is also supported but less emphasized in several recent quantitative studies of *E. coli* protein expression(13, 16–18) (Fig. S3T–W).

To probe how changes in protein and mRNA fractions relate to each other across growth conditions, we generated additional proteomics and transcriptomics datasets for more conditions under each type of growth limitations (Fig. S4A), so that a smooth growth-rate dependence could be obtained individually for the mRNA and protein fractions. We extrapolated these data to compute the fold-change (FC) in the protein and mRNA fractions, FC($\psi_{p,i}$) and FC($\psi_{m,i}$) respectively, for each gene i (Fig. S4B). The fold-change was calculated between the "reference condition" (WT cells grown in glucose minimal medium) and one with ~3x slower growth, for each of the three types of growth limitation imposed. Their ratio, FC($\psi_{p,i}$)/FC($\psi_{m,i}$), was even more tightly distributed than $\psi_{p,i}/\psi_{m,i}$ for each type of growth limitation (compare Fig. 1E with Fig. 1D), indicating that the mRNA and protein fractions tightly co-varied for the majority of genes. The few exceptions which did not co-vary usually occurred in only one of the growth limitations, and mostly corresponded to known targets of post-transcriptional regulation (Fig. S4C–E, Table S5.

Total mRNA abundance matches translational capacity

From the steady state relation between concentrations of individual mRNAs and proteins (Fig. S1)

$$\alpha_{p,i}[mR_i] = \lambda[P_i] \tag{2}$$

where $a_{p,i}$ is the translation initiation rate of each mRNA mR$_i$ and λ denotes the growth rate, we can sum over contributions from all genes to obtain a relation between the flux of total protein synthesis and dilution,

$$\bar{\alpha}_p[mR] = \lambda[P], \tag{3}$$

with $\bar{\alpha}_p \equiv \sum_i \alpha_{p,i}\psi_{m,i}$ being the average translational initiation rate (over all mRNAs). Because the ratio of Eq. (2) and Eq. (3) gives

$$\alpha_{p,i}/\bar{\alpha}_p = \psi_{p,i}/\psi_{m,i}, \tag{4}$$

we see that Fig. 1D also provides the distribution of the relative translation initiation rates $\left(\alpha_{p,i}/\bar{\alpha}_p\right)$. The observed similarity between mRNA and protein fractions, Eq. (1), implies that the translation initiation rates $\alpha_{p,i}$ are similar for the majority of mRNAs for each growth condition. Thus, the average translational initiation rate $\bar{\alpha}_p$ can be taken as representative of the majority of mRNAs.

Because the total protein concentration $[P]$ changes less than 10% across nutrient-limited conditions (Fig. S2H), the total protein synthesis flux $\lambda[P]$ changes almost linearly with the growth rate λ. Consequently, Eq. (3) constrains the total mRNA concentration $[mR]$ and/or the average initiation rate $\bar{\alpha}_p$ to change with cellular growth. To understand how this constraint is accommodated, we quantified the total mRNA amount by combining hybridization of ^{3}H-uracil labeled RNA to genomic DNA with quantitative Northern blotting (Fig. S5, SI Methods). The result for carbon-limited growth (Fig. S5H), was then converted to cellular concentration, i.e., $[mR]$ (SI Note S1), and shown as the red symbols in Fig. 2A (left vertical axis). These data allowed us to use Eq. (3) to obtain the average translation initiation rate $\bar{\alpha}_p$. The approximately linear growth-rate dependence of $[mR]$ makes $\bar{\alpha}_p$ only weakly growth-rate dependent (Fig. S6A, left axis). The value of $\bar{\alpha}_p$ in turn allowed us to obtain the distribution of $\alpha_{p,i}$, the translational initiation rate of individual mRNAs, using Eq. (4) and the distributions of $\psi_{p,i}/\psi_{m,i}$ (Fig. 1D). The results for the reference and a slow, carbon-limiting growth condition exhibited weak dependence of $\alpha_{p,i}$ on both the mRNA species and growth condition (Fig. 2B and Fig. S6B).

Constancy of ribosome spacing across mRNA and nutrient conditions

To understand how the relation between the total mRNA concentration and the total protein synthesis flux (Eq. (3)) arises, we note that the total flux of peptide synthesis, is given by $\lambda[P]\,\bar{\ell}_P$, where $\bar{\ell}_P$ is the average length of a protein, ~250 aa across conditions (Fig. S2G). Because genome-wide translation elongation rates are narrowly distributed (Fig. S6C), this flux corresponds to the product of the concentration of actively translating ribosomes ($[Rb]_{act}$) and the speed of translational elongation (ε) as depicted in Fig. S6D, i.e.,

$$\lambda[P]\,\bar{\ell}_P = \varepsilon[Rb]_{act}. \tag{5}$$

Both $[Rb]_{act}$ and ε have been characterized for a broad range of nutrient conditions(7). The active ribosome concentration is plotted in Fig. 2A (right vertical axis). The data exhibit a congruence with the total mRNA concentration, revealing a coordination of mRNA abundance and the cellular translational capacity, with an average spacing between translating ribosomes close to 200 nt (Fig. 2C and Fig. S6E). This translates to an average of number of ribosomes per mRNA $\bar{r} \approx 4$ for typical mRNAs 750 nt long. The observed proportionality between total mRNA and active ribosome concentrations implies that the average translational initiation rates $\bar{\alpha}_p$ and elongation rates ε are proportional (Fig. S6A); this proportionality extends to individual mRNAs due to the similarity in protein and mRNA

fractions observed above. By comparing the concentrations of protein and mRNAs (Fig. S6FG), we confirmed that for most mRNA species the ribosome spacing is indeed clustered around 200 nt across conditions (Fig. 2D and E; Figs. S6HI). The data are bounded by ~40 nt per ribosome in accordance to the physical packing limit(19).

mRNA degradation is largely condition independent

We investigated the mechanism behind the observed proportionality between the concentrations of total mRNA and active ribosomes (Fig. 2A). mRNA concentration is set by the balance between its synthesis and degradation (Fig. S1). We performed kinetic experiments to determine mRNA degradation rates δ_i genome-wide in the reference and the slowest carbon-limiting condition by inhibiting transcription initiation with rifampicin and quantifying the relative mRNA concentrations at short time intervals by RNA-seq (Fig. S7A–C, SI Methods, Table S3). As an example, we measured time courses of changes in the relative mRNA concentrations of genes of the *nuo* operon in the two growth conditions (Fig. 3A–C). The time course can be described as a delayed exponential decay, with the lag time reflecting the time needed for the RNA polymerase to reach the gene from the transcription start (Fig. S7D), and the decay rate attributed to the turnover of that mRNA. This analysis yielded degradation rates for ~2700 mRNAs (SI Methods, Table S6). Genome-wide, mRNA degradation rates were strongly correlated in the two growth conditions (Fig. S7E). The average degradation rate was very similar (Fig. S7F, vertical dashed lines), even after weighting by mRNA abundances (Fig. S7G). In particular, the fold-change in δ_i is sharply peaked, with 90% of genes in the range 0.50 to 1.57 (Fig. 3D). Overall, the data indicated a lack of dependence of degradation rates on either the mRNA species or the growth condition for most mRNAs, as observed in other studies(20, 21). The cases where rates changed significantly could be associated to known post-transcriptional regulation (Fig. S7HI).

Total mRNA synthesis flux is adjusted to match translational capacity

From the concentration and degradation rates of each mRNA species, $[mR_i]$ and δ_i respectively, we can obtain the mRNA degradation flux, $\delta_i[mR_i]$, whose distributions are shown in Fig. 3E for the reference and slow growth conditions. A shift in the median of the two distributions is seen (vertical dashed lines), reflecting growth dependence of the total degradation flux, $\sum_i \delta_i[mR_i] = \bar{\delta}[mR]$, where $\bar{\delta} \equiv \sum_i \delta_i \psi_{m,i}$ is the average degradation rate across mRNAs. By the balance of mRNA synthesis and degradation in steady state growth (Fig. S1; see also SI Note S3), the total mRNA synthesis flux J_{mR} can be expressed as

$$J_{mR} \equiv \sum_i \alpha_{m,i}[G_i] = \bar{\delta}[mR] \,. \tag{6}$$

Because the average degradation rate ($\bar{\delta}$) is affected little by growth conditions (Fig. S7G), Eq. (6) predicts that the observed growth dependence of the total mRNA concentration $[mR]$ (Fig. 2A) is caused primarily by changes in mRNA synthesis flux, J_{mR}.

We tested this prediction by directly measuring the total mRNA synthesis flux J_{mR} across the range of carbon-limited growth conditions, by pulse-labelling cultures with ^{3}H-uracil and hybridizing the labelled RNA to genomic DNA over short time intervals (Fig. S8).

These data showed a strong growth-rate dependence Fig. 3F (red symbols, left vertical axis), closely matching the observed growth dependence of the total mRNA concentration (reproduced as black symbols in Fig. 3F, right vertical axis). Note that the total mRNA fluxes inferred from the degradation rates, $\bar{\delta}[mR]$ (orange crosses) were within 20% of the directly measured synthesis fluxes, showcasing the consistency of these two very different measurement approaches. Putting together, the results in Figs. 2 and 3 show that the global constraint Eq. (3) is enforced primarily by matching total mRNA synthesis flux J_{mR} with translational capacity (Fig. 3G).

mRNA synthesis flux and transcriptional regulation

The synthesis flux of each mRNA species is given molecularly by the product of the transcription initiation rate per gene, $a_{m,i}$, and the "gene concentration", $[G_i]$ (Fig. 1A and S1). The growth-rate dependence of gene concentration is in turn given by the product of number of chromosome replication origins (Ori) per cell volume, $[Ori]$, and the "gene dose" relative to the Ori, $g_i \equiv [G_i]/[Ori]$. Thus, the total mRNA synthesis flux can be expressed as

$$J_{mR} = [Ori] \cdot \sum_i \alpha_{m,i} g_i. \tag{7}$$

The relative gene dose g_i across growth rates can be obtained from the chromosomal position of the gene(22, 23) and the chromosome replication time (Fig. S9AB). Further including a weak growth-rate dependence of the Ori concentration(24) (Fig. S9C–E), we obtain negative growth-rate dependences for the concentration of genes $[G_i] = [Ori] \cdot g_i$ at all chromosomal positions (Fig. 4A). It is then clear from Eq. (7) that the strong positive growth-rate dependence seen for the total mRNA synthesis flux J_{mR} (Fig. 3F) cannot be accounted for by opposite growth-rate dependences of gene concentrations and must involve systematic changes in the promoter activities $a_{m,i}$. This was seen more explicitly by computing the distributions of the promoter activity $a_{m,i}$, obtained for each gene using the known degradation fluxes $\delta_i[mR_i]$ and gene concentrations $[G_i]$ at steady state (Eq. S3 in Fig. S1; see also SI Note S3; data in Table S6). The results (Fig. 4B) showed a broad range of promoter activity, spanning 4 orders of magnitude, with the high-end (~0.3/s in reference condition) approaching the maximum of ~1/s given the transcriptional elongation speed of ~50 nt/s and a transcription elongation complex footprint of ~40 nt (25–27). A shift is seen between the distributions for reference (grey) and C-limited (red) growth conditions. In fact, for most genes, $a_{m,i}$ in C-limited conditions decreases to about 1/3 of its value in reference condition (Fig. 4C), a change that is comparable to the reduction in growth rate. Thus, the coordination of mRNA synthesis flux with the growth rate (Fig. 3F) is likely a result of genome-wide changes in transcription initiation between these conditions.

To look further into the determinants of transcription initiation, we used a canonical model of transcription regulation (Fig. 4D)(28, 29) where the transcription initiation rate $a_{m,i}$ for gene i is given by the product of the concentration of available RNA polymerases ($[RNAP]_{av}$) and the promoter on-rate k_i, i.e.,

$$\alpha_{m,i} = k_i \cdot [RNAP]_{av}, \tag{8}$$

where k_i captures the regulatory activities of all transcription factors acting on the promoter driving gene i(28, 29). Using this expression for $a_{m,i}$, the balance of mRNA synthesis and degradation (Eq. S3 in Fig. S1) can be written as

$$[\text{RNAP}]_{\text{av}}[Ori] \cdot k_i g_i = \delta_i \cdot [mR_i]. \tag{9}$$

Quantitative relations connect transcriptional regulation to gene expression

From Eq. (9), we can derive two fundamental relations connecting transcription regulation to gene expression (see also SI Note S4). Summing Eq. (9) over all genes, the balance of the total transcription flux becomes

$$[\text{RNAP}]_{\text{av}}[Ori] \cdot \mathscr{K} = \bar{\delta} \cdot [mR], \tag{10}$$

where $\mathscr{K} \equiv \sum_i k_i g_i$ describes the total on-rate for promoters across the genome and is a measure of the total regulatory activity on transcription (weighted by gene dose). Given the proportionality between active ribosome and total mRNA concentrations, $\bar{r} = [Rb]_{\text{act}}/[mR] \approx 4$ (Fig. 2A), Eq. (10) can be written as:

$$[\text{RNAP}]_{\text{av}}[Ori] \cdot \mathscr{K} = (\bar{\delta}/\bar{r}) \cdot [Rb]_{\text{act}}. \tag{11}$$

This relation represents a fundamental constraint between the overall transcription activity ($[RNAP]_{\text{av}}\mathscr{K}$), the DNA content (via $[Ori]$) and the translational activity of the cell.

Another important relation can be obtained by taking the ratio of Eqs. (9) and (10). Noting that the mRNA degradation rates are closely distributed around the average $\bar{\delta}$ and independent of growth conditions, i.e., $\delta_i \approx \bar{\delta}$ (Fig. S7FG), we obtain $k_i g_i/\mathscr{K} \approx [mR_i]/[mR] = \psi_{\text{m},i}$. This relation extends further to the fractional protein abundances $\psi_{\text{p},i} = [P_i]/[P]$ due to the established relation between protein and mRNA fractions (Eq. (1) and Fig. 1), leading to

$$\frac{k_i g_i}{\mathscr{K}} \approx \psi_{\text{m},i} \approx \psi_{\text{p},i}. \tag{12}$$

This expression relates the (gene-dose weighted) regulatory activity on specific promoters ($k_i g_i$) to the mRNA and protein levels as determined by transcriptomics ($\psi_{\text{m},i}$) and proteomics ($\psi_{\text{p},i}$). Importantly, $\psi_{\text{p},i} = [P_i]/[P]$ gives approximately the cellular protein concentration $[P_i]$ since the total protein concentration varies only mildly with the growth rate, on average $[P] \approx 4.5 \cdot 10^6/\mu\text{m}^3$ (Fig. S2H). Thus, Eq. (12) quantitatively connects regulatory activities at the promoter level (k_i) to cellular protein concentrations $[P_i]$, without explicit reference to the macroscopic machineries of gene expression. Eqs. (11) and (12) are the central quantitative results of this study. We suggest Eq. (12) be viewed as a quantitative statement of the Central Dogma of bacterial gene expression, with Eq. (11) describing a system-level constraint on transcription and translation. In the following, we separately explore some consequences of these two central relations.

Global coupling in gene expression

According to Eq. (12), the mRNA and protein fractions corresponding to a given gene i are dependent not only on the regulatory activity on that gene, $k_i g_i$, but also on the total regulatory activity, $\mathscr{K} \equiv \sum_j k_j g_j$. The latter dependence couples the expression of all genes in the cell, as illustrated in Fig. 4E. This dependence is explicitly seen when comparing fold-changes in gene expression across two different conditions:

$$FC([P_i]) = \mathrm{FC}(\psi_{p,i}) = \mathrm{FC}(k_i g_i)/\mathrm{FC}(\mathscr{K}) . \tag{13}$$

In different growth conditions where the promoter on-rates k_i of many genes are affected, we generally expect the total rate $\mathscr{K}$ to vary, i.e. $\mathrm{FC}(\mathscr{K}) \neq 1$. Consequently, changes in the regulatory activity of a gene are generally expected to be different from the changes in the fractional abundances of the corresponding mRNA and protein. In fact, the latter might change even if the corresponding regulatory activity $k_i g_i$ is unchanged, due to the overall change in regulatory activity $\mathscr{K}$; illustrated in Fig. 4E.

To determine how the total regulatory activity $\mathscr{K}$ may change across growth conditions, we return to the spectrum of carbon-limited growth conditions. The growth-rate dependence of $\mathscr{K}$ can be deduced by applying the relation (13) to "constitutively expressed" (i.e., unregulated) genes, for which k_i is constant. For this purpose, we inserted constitutively expressed *lacZ* genes at various locations x on the chromosome, with known gene doses $g(x)$ (Fig. 4A). The total rate $\mathscr{K}$ can then be obtained by measuring the concentration of LacZ, $[LacZ(x)]$, for different growth rates, as

$$\mathscr{K} \propto \frac{g(x)}{[LacZ(x)]} . \tag{14}$$

The data in carbon-limited conditions (Fig. 4F) show that $[LacZ(x)]$ generally decreased at faster growth, with a steeper trend when *lacZ* was inserted near *ter* (blue) compared to near *ori* (orange). Upon calculating $\mathscr{K}$ using Eq. (14), the data collapsed on the same positive growth-rate dependence (Fig. 4G). Although this set of experiments established the relative changes in $\mathscr{K}$ across conditions, the absolute scale of $\mathscr{K}$ can be determined from Eq. (10) using the measured $[Ori]$ (Fig. 4A) and the measured $J_{\mathrm{mR.}}$ (Fig. 3F) for $\bar{\delta}[mR]$ (Eq. 5). As discussed in SI Note S5, the abundance of available RNAP, $[RNAP]_{\mathrm{av}}$ can be estimated in reference condition to be approximately $1300/\mu m^3$, leading to $\mathscr{K} \sim 1.27 \mu m^3/\mathrm{min}$ in reference condition.

Identifying promoter on-rates

Knowledge of the magnitude of $\mathscr{K}$, together with the mRNA abundances and degradation rates, allowed us to compute the promoter on-rates k_i for each gene i across growth conditions (see SI Note S4). The results for ~2500 genes (Fig. 5A) displayed a broad distribution across more than 3 orders of magnitude. Since k_i were computed by combining several different genome-wide measurements (SI Note 4), we sought to validate them by examining the on-rates of genes belonging to the same operon, to see if the variation in the

k_i of these co-transcribed genes are are indeed small as intuitively expected. Reassuringly, the coefficients of variation in k_i within operons were not only much smaller than the values obtained by randomly extracting k_i from the different operons, but also significantly smaller than the variation in the concentrations of proteins associated to the same regulatory units (Fig. 5B, Fig. S10A).

The complete set of gene expression rates generated in this work, including the promoter on-rates, mRNA degradation rates and translation initiation rates (Table S6) allowed us to investigate at the genome scale the dominance of different factors controlling protein concentrations in *E. coli*, as well as their changes across conditions. By plotting the promoter on-rates k_i against the protein concentrations in reference condition (Fig. 5C, top), we see that the large range of protein concentrations can be largely attributed to differences in the promoter on-rates, as opposed to other factors such as mRNA degradation rate, translation initiation rate, or the gene dose (other panels in Fig. 5C), in agreement with the simple scenario expressed by Eq. (12).

Proteins present at low concentrations tend to have lower translation initiation rates $a_{p,i}$ and larger mRNA degradation rates δ_i compared to those at high concentrations (middle panels in Fig. 5C; Fig. S10B); both effects tend to reduce the average number of proteins produced during the mRNA's lifetime(3) (Fig. S10C–F). When focusing on co-transcribed genes, we were also able to identify post-transcriptional effects (Fig. S11AB) and evidence of premature transcriptional termination (Fig. S11C). Still, these effects are rare and of small magnitude compared to the vast range of promoter on-rates.

Unraveling the innate and regulatory effects on gene expression

Fold changes in protein abundance across conditions showed an almost perfect correlation with promoter on-rates (Fig. 5D, top), whereas the effect of post-transcriptional regulation or gene copy number was negligible (Fig. 5D, other panels; Fig. S10G). Thus, protein concentrations appear to be adjusted across conditions by modulating the promoter on-rate, as described by Eq. (13). Yet, the typical changes in promoter on-rates were very small (Fig. 5D) even though the full range of k_i spans more than three orders of magnitude (Fig. 5C; see also Fig. S10HI). Indeed, based on proteomics data collected for a wide range of growth conditions from Mori et al(10) (e.g., Fig. S12A–E), we find that that protein abundances for 2/3 of the genes vary by less than 5-fold (Fig. S12F) even though the absolute abundances vary over three orders of magnitude. Particularly small changes are observed for proteins engaged in housekeeping activities (S12G–I; see also Fig. S11D–F) or those encoded by essential genes (Fig. S12F and J). Taken together, we conclude that innate promoter sequences determine the typical concentrations of most proteins, with transcriptional regulation providing fine adjustments depending on conditions. Some exceptions involve proteins needed in specific growth conditions (e.g. TCA proteins in aerobic conditions), whose genes are under strict transcriptional control (Fig. S12K–O).

Overall, promoter on-rates for genes involved in protein synthesis increased with growth rate, and largely accounted for the increase in the total promoter on-rate (Fig. 5E). Changes in k_i's for ribosomal proteins and elongation factors were similar as expected because they are largely co-transcribed (Fig. 5F, triangles; Fig. S13A–D). Notably, they

also matched the changes for the rRNA promoters (Fig. 5F, crosses). Thus, despite the known post-transcriptional regulation acting on ribosomal proteins(30), the coordination between r-proteins and rRNAs is largely implemented by matching their promoter on-rates. In contrast, the k_i's of other translation-affiliated proteins (e.g., initiation factors, tRNA synthetases) in different regulatory units present a variety of dependencies on the growth rate (Fig. S13E–H). Analogously, we found a variety of behaviors for the k_i of genes known to be controlled by cAMP-CRP and typically expressed in carbon-poor conditions(15) (Fig. S14). The strong increases in the protein levels of many catabolic proteins in C-limitation are due to a combination of faster promoter on-rates and reduced total regulatory activity $\mathscr{K}$ at slow growth, highlighting the passive effects described in Fig. 4E.

Control of global mRNA synthesis by the anti-sigma factor Rsd

The constraint between transcription and translation, Eq. (11), must be somehow satisfied despite the observed changes in active ribosomes (Fig. 2A), DNA concentration (Fig. 4A) and total regulatory activity (Fig. 4G). The combination $[Ori] \cdot \mathscr{K}$ has only a moderate dependence on growth rate (Fig. 6A). Instead, the concentration of available RNAP, estimated as the ratio of the mRNA synthesis flux (J_{mR}) and $[Ori] \cdot \mathscr{K}$ based on Eq. (6) and (10), exhibited a stronger growth rate dependence (left axis in Fig. 6B), approximately matching that of the concentration of active ribosomes (right axis).

A simple mechanism to change the availability of RNA polymerases is to change the abundance of the transcription machinery itself. However, our quantitative proteomics data showed that the cellular concentrations of RNA polymerase components, including the house-keeping factor σ^{70} (encoded by the gene *rpoD*), are all constant across the growth rate range studied (Fig. 6C, squares). We checked expression levels of the two known modulators of σ^7 function, 6S RNA(31, 32) and the anti-σ^{70} protein Rsd(33–35). Although the concentration of 6S RNA is ~ 1/100 that of σ^{70} (grey triangles, Fig. 6C) and thus unlikely to affect the global transcription flux in these conditions, the Rsd concentration rose to that of σ^{70} as growth rate is reduced (Fig. 6C, red triangles). Thus, Rsd could be a regulator of global transcription by sequestering σ^{70} during exponential growth (Fig. 6D), even though it is thought to have its primary role in stationary phase(33, 36). We tested this scenario by characterizing the total mRNA synthesis flux in a *rsd* strain. mRNA synthesis flux became nearly independent of growth rate (Fig. 6E, filled red circles), exceeding that of the wild type strain (open red circles), especially at slow growth when Rsd is highly expressed. Without *rsd*, the synthesis flux was no longer matched to the translational capacity (compare filled symbols), in contrast to the tight matching observed in wild type strain (open symbols). Concomitantly, the *rsd* strain exhibited a growth defect that was proportional to the amount of Rsd expression in wild-type cells in slow growth conditions where Rsd is expressed (Fig. 6F). Given the approximate constancy of mRNA turnover across growth conditions for wild type cells (Fig. S7E–G), we propose that Rsd may have a central role in controlling total mRNA concentrations (Figs. 2A and 3F).

DISCUSSION

We used comprehensive transcriptomic and proteomic studies, complemented by quantitative measurement of total mRNA abundance and transcription flux, to determine the absolute mRNA and protein abundances, the mRNA degradation rates as well as the promoter on-rates, for 1500+ genes in *E. coli* for many growth conditions during steady-state growth (Tables S3–5). The results revealed two simple rules on promoter and mRNA characteristics which profoundly shape how *E. coli* responds to environmental changes while coping with global constraints: (*i*) promoter on-rates span over 3 orders of magnitudes across genes, but vary much less (at most ~5-fold) across conditions for most genes. Thus, each gene is expressed within an innate abundance range across conditions, e.g., with ribosomal genes belonging to the most abundant and DNA replication proteins, belonging to one of the least abundant classes. (*ii*) mRNA characteristics, including translation initiation rate and mRNA degradation rate, vary little (<2-fold for half of the genes) across genes and conditions. The translation initiation rates are sufficiently rapid to maintain a high density of ribosomes on the mRNA (5 ribosomes per kb, Fig. 2A and Fig. S6E), resulting in high protein production despite short mRNA half-lives. The rules governing promoter and mRNA characteristics deduced here dictate, to a large extent, *E. coli*'s strategy to implement gene regulation while complying with the constraints on total protein concentration and a limited translation capacity. This can be cast into two design principles of gene expression as summarized below.

Global coordination between transcription and translation

The concentration of translating ribosomes changes strongly with the bacterial growth rate (Fig. 2A). Because of the constant density of translating ribosomes on most mRNA, the total mRNA concentration must also scale similarly. As the total mRNA pool is specified by the total mRNA synthesis rate (given the constant mRNA turnover rate across conditions), total mRNA synthesis is balanced with translating ribosomes concentration, a crucial condition captured by Eq. (11). We refer to this balance as the *principle of transcription-translation coordination*. We showed that *E. coli* implements this coordination across nutrient conditions primarily by adjusting the available RNAP concentration via the anti-σ^{78} factor Rsd (Fig. 6D–F).

If this coordination is broken, as in the case of *rsd* mutant (Fig. 6E), then Rule *ii* cannot hold as long as the constraints on translation capacity and protein density hold. An oversupply of mRNA with respect to ribosomes is expected to decrease the rate of translation initiation (due to competition for limited ribosomes) and increase the rate of mRNA degradation (due to reduced protection of mRNA by elongation ribosomes against RNase activity(37, 38). Aside from the futile cycle involving the synthesis and degradation of mRNAs and affecting growth (Fig 6F), breaking Rule *ii* would complicate the otherwise simple relation between transcriptional regulation and protein concentrations of the wild type system.

The predominant role of transcriptional control in setting protein concentrations

The similarity of mRNA characteristics (Rule *ii*) together with the vast disparity of promoter characteristics (Rule *i*) across genes in a given condition implies that protein abundances are predominantly set by the promoter characteristics, specifically, the promoter on-rates. Furthermore, as mRNA characteristics remain similar across different growth conditions (Rule *ii*), changes in protein concentrations across conditions must arise primarily from changes in the promoter on-rates, i.e., via transcriptional regulation (Fig. 5D). We refer to this strong effect of transcription on gene expression as the *principle of transcriptional predominance*.

The strong mRNA-protein correlations are at odds with early studies on bacterial gene expression(39–45). These differences may originate in part from technological advances in transcriptomics and proteomics over the years(10, 45, 46). Further, given the very different turnover rates of proteins and mRNAs, accurate mRNA-protein comparison requires ensuring steady-state conditions, which was a crucial component of our experimental design (SI Methods). We note that strong mRNA-protein correlation similar to what we report here was also contained in a number of recent datasets(13, 16–18); see Fig. S3T–W. However, such comparisons were not articulated as the main messages in these studies, as their focus was on the variability in translational characteristics among mRNAs. While we also see such differences (Fig. S4C–E), our data (as well as those from Ref. 14–17) show that such variabilities do not represent the general behavior of most genes.

Although setting protein concentrations transcriptionally appears simple, a quantitative relation between promoter on-rates and protein concentrations is complicated by the fact that the total protein concentration does not vary by much, despite strong changes in total transcriptional activity. Because protein concentrations do not depend on the absolute flux of the corresponding mRNAs, but rather on the fraction of the total transcription flux, a hypothetical doubling of all promoter on-rates, as alluded to in Introduction, will have no effect on mRNA and protein concentrations. Furthermore, the relations between promoter on-rates and protein concentrations are insulated from growth-dependent differences in shared machinery such as RNAP and ribosomes.

Global transcriptional coupling and its consequences

Because the protein output of a given promoter depends on the total regulatory activity $\mathscr{K}$ in Eq. (12), non-intuitive relations between promoters and protein concentration can arise when $\mathscr{K}$ changes across conditions. The latter is likely whenever there is a substantial change in growth conditions, e.g., due to changes in the on-rates of ribosomal genes (Fig. S12H, 5H). Hence, it is generally incorrect to infer regulatory activities directly from changes in mRNA and protein concentrations (Fig. 4E). This effect of global coupling which reflects the passive regulatory effect would hardly affect the result of most classical studies, which typically involved large changes in the output of a few individual promoters. However, for the majority of genes whose expression change moderately across conditions (e.g. <2-fold for 60% of the genes under C-limited growth; see Fig. 5D), the effect of global coupling would be substantial.

Our work provides a quantitative framework to distinguish gene-specific regulatory effects from global interactions in gene expression studies. Knowing the promoter on-rates k_i for individual genes facilitates a direct, promoter-centric view of regulation across conditions at the genome-scale (Fig. 5). This improves upon previous estimates of "promoter activity" based on protein synthesis fluxes(47), which mix systemic effects, such as RNA polymerase availability, with gene-specific regulatory effects (SI Note 4). These results would thus be of use in deciphering the behaviors of endogenous genetic circuits, and in guiding the design of synthetic circuits in variable growth conditions(48–50).

The results described here are specific to bacteria. Eukaryotic gene expression involves complex post-transcriptional regulation, including protein secretion(51) and degradation via ubiquination and autophagy(52). Global constraints are less understood there, in particular the extent to which protein density may vary across conditions. Indeed, even quantifying the cell volume may be difficult as large portions within a cell may be occupied by sub-cellular compartments (e.g., vacuoles) that do not contribute to the cytosol. Nonetheless, our study provides a framework to quantitatively explore gene expression in such complex systems.

Materials and Methods

Experimental methods for cellular growth, RNA sequencing, and quantification of mRNA abundance, synthesis fluxes and degradation rates, as well as numerical and statistical methods, are reported in the Supplementary Material.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

FUNDING

DATA AND MATERIAL AVAILABILITY

The proteomic mass spectrometry files as well as data analysis files have been published previously(10) and are accessible through the ProteomeXchange Consortium via the PRIDE partner repository: http://www.ebi.ac.uk/pride/archive/projects/PXD014948. The used *E. coli* spectral library for DIA/SWATH data analysis is available via SWATHAtlas: http://www.peptideatlas.org/PASS/PASS01421. The genome-wide parameters of gene expression generated in this study, including transcription and translation initiation rates, mRNA

degradation rates, promoter on-rates, as well as protein, mRNA and gene concentrations for reference condition (glucose minimal medium) and slow glucose-limited growth are reported in Table S6. Additional analysis, including modelling and details on the calculation of the parameters, are included in Notes S1 to S5. Raw RNA sequencing reads are uploaded to Gene Expression Omnibus (GEO accession number GSE205717). Strains generated in this study can be requested from the lead contact Terence Hwa (hwa.ucsd.edu).

References and notes

1. Goldberg AL, St John AC, Intracellular protein degradation in mammalian and bacterial cells: Part 2. Annu. Rev. Biochem. 45, 747–803 (1976). [PubMed: 786161]

2. Nath K, Koch AL, Protein degradation in Escherichia coli. II. Strain differences in the degradation of protein and nucleic acid resulting from starvation. J. Biol. Chem. 246, 6956–6967 (1971). [PubMed: 4942328]

3. Paulsson J, Models of stochastic gene expression. Phys. Life Rev. 2, 157–175 (2005).

4. Klumpp S, Zhang Z, Hwa T, Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. Cell. 139, 1366–1375 (2009). [PubMed: 20064380]

5. Lin J, Amir A, Homeostasis of protein and mRNA concentrations in growing cells. Nat. Commun. 9 (2018), doi:10.1038/s41467-018-06714-z.

6. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T, Interdependence of Cell Growth. Science (80-.). 330, 1099–1102 (2010).

7. Dai X, Zhu M, Warren M, Balakrishnan R, Patsalo V, Okano H, Williamson JR, Fredrick K, Wang Y-P, Hwa T, Frederick K, Wang Y-P, Hwa T, Reduction of translating ribosomes enables Escherichia coli to mantain elongation rates during slow growth. Nat. Microbiol. 2, 16231 (2016). [PubMed: 27941827]

8. Milo R, What is the total number of protein molecules per cell volume? A call to rethink some published values. BioEssays. 35, 1050–1055 (2013). [PubMed: 24114984]

9. E. R. O, Yuki K, Sven van T., Robust surface-to-mass coupling and turgor-dependent cell width determine bacterial dry-mass density. Proc. Natl. Acad. Sci. 118, e2021416118 (2021). [PubMed: 34341116]

10. Mori M, Zhang Z, Esfahani AB, Lalanne J, Collins BC, Schmidt A, Schubert OT, Lee D, Li GW, Hwa T, Ludwig C, From coarse to fine : The absolute Escherichia coli proteome under diverse growth conditions. Mol. Syst. Biol. 17, e9536 (2021). [PubMed: 34032011]

11. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R, Mol. Cell. Proteomics, in press, doi:10.1074/mcp.O111.016717.

12. Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R, Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. Mol. Syst. Biol. 14, e8126 (2018). [PubMed: 30104418]

13. Li G-W, Burkhardt D, Gross C, Weissman JS, Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell. 157, 624–35 (2014). [PubMed: 24766808]

14. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang Y-P, Lenz P, Yan D, Hwa T, Coordination of bacterial proteome with metabolism by cyclic AMP signalling (2013), doi:10.1038/nature12446.

15. Hui S, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T, Williamson JR, Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. Mol Syst Biol. 11, 784 (2015). [PubMed: 25678603]

16. Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A, Xie XS, "Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells," (available at http://science.sciencemag.org/).

17. Del Campo C, Bartholomäus A, Fedyunin I, Ignatova Z, Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. PLOS Genet | DOI (2015), doi:10.1371/journal.pgen.1005613.

18. Choe D, Lee JH, Yoo M, Hwang S, Sung BH, Cho S, Palsson B, Kim SC, Cho B-K, Adaptive laboratory evolution of a genome-reduced Escherichia coli, doi:10.1038/s41467-019-08888-6.

19. Mohammad F, Green R, Buskirk AR, A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution, doi:10.7554/eLife.42591.001.

20. Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S, Cohen SN, "Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays," (available at www.pnas.orgcgidoi10.1073pnas.112318199).

21. Coffman RL, Norris TE, Koch AL, Chain elongation rate of messenger and polypeptides in slowly growing Escherichia coli. J. Mol. Biol. 60, 1–11 (1971). [PubMed: 4937191]

22. Bremer H, Churchward G, An examination of the Cooper-Helmstetter theory of DNA replication in bacteria and its underlying assumptions. J. Theor. Biol. 69, 645–654 (1977). [PubMed: 607026]

23. Cooper C, Stephen and Helmstetter, Chromosome Replication and the Division Cycle of Escherichia coli B/r. J. Mol. Biol. 31, 519–540 (1968). [PubMed: 4866337]

24. Zheng H, Bai Y, Jiang M, Tokuyasu TA, Huang X, Zhong F, Wu Y, Fu X, Kleckner N, Hwa T, Liu C, General quantitative relations linking cell growth and the cell cycle in Escherichia coli. Nat. Microbiol. 5, 995–1001 (2020). [PubMed: 32424336]

25. Krummel B, Chamberlin MJ, Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. J. Mol. Biol. 225, 239–250 (1992). [PubMed: 1593619]

26. Ring BZ, Yarnell WS, Roberts JW, Function of E coli RNA polymerase σ factor σ70 in promoter-proximal pausing. Cell. 86, 485–493 (1996). [PubMed: 8756730]

27. Zhu M, Mori M, Hwa T, Dai X, Disruption of transcription-translation coordination in Escherichia coli leads to premature transcriptional termination. Nat. Microbiol 4 (2019), doi:10.1038/s41564-019-0543-1.

28. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R, Transcriptional regulation by the numbers: Models. Curr. Opin. Genet. Dev. 15, 116–124 (2005). [PubMed: 15797194]

29. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Kuhlman T, Phillips R, Transcriptional regulation by the numbers: Applications. Curr. Opin. Genet. Dev. 15, 125–135 (2005). [PubMed: 15797195]

30. Yates JL, Arfsten AE, Nomura M, In vitro expression of Escherichia coli ribosomal protein genes: Autogenous inhibition of translation. Proc. Natl. Acad. Sci. U. S. A. 77, 1837–1841 (1980). [PubMed: 6445562]

31. Wassarman KM, Storz G, 6S RNA regulates E. coli RNA polymerase activity. Cell. 101, 613–623 (2000). [PubMed: 10892648]

32. Wassarman KM, 6S RNA, A Global Regulator of Transcription, doi:10.1128/microbiolspec.RWR-0019-2018.

33. Jishage M, Ishihama A, "A stationary phase protein in Escherichia coli with binding activity to the major subunit of RNA polymerase (transcriptase controlstationary phase controlenzyme activity controlRNA polymerase storage)" (1998), (available at www.pnas.org.).

34. Gruber TM, Gross CA, Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. Annu. Rev. Microbiol. 57, 441–466 (2003). [PubMed: 14527287]

35. Lal A, Krishna S, Sai A, Seshasayee N, Regulation of Global Transcription in Escherichia coli by Rsd and 6S RNA (2018), doi:10.1534/g3.118.200265.

36. Piper SE, Mitchell JE, Lee DJ, Busby SJW, A global view of Escherichia coli Rsd protein and its interactions. Mol. Biosyst. 5, 1943–1947 (2009). [PubMed: 19763331]

37. Deana A, Belasco JG, Lost in translation: The influence of ribosomes on bacterial mRNA decay. Genes Dev. 19, 2526–2533 (2005). [PubMed: 16264189]

38. Hui MP, Foley PL, Belasco JG, Messenger RNA Degradation in Bacterial Cells. 7, 19 (2014).

39. Nie L, Wu G, Zhang W, Correlation of mRNA expression and protein abundance affected by multiple sequence features related to translational efficiency in Desulfovibrio vulgaris: a quantitative analysis. Genetics. 174, 2229–43 (2006). [PubMed: 17028312]

40. Lu P, Vogel C, Wang R, Yao X, Marcotte EM, Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat. Biotechnol. 25, 117–124 (2007). [PubMed: 17187058]

41. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C, Global signatures of protein and mRNA expression levels. Mol. Biosyst 5 (2009), doi:10.1039/b908315d.

42. Laurent JM, Vogel C, Kwon T, Craig SA, Boutz DR, Huse HK, Nozue K, Walia H, Whiteley M, Ronald PC, Marcotte EM, Protein abundances are more conserved than mRNA abundances across diverse taxa. Proteomics. 10, 4209–12 (2010). [PubMed: 21089048]

43. Maier T, Schmidt A, Güell M, Kühner S, Gavin A-C, Aebersold R, Serrano L, Quantification of mRNA and protein and integration with protein turnover in a bacterium. Mol. Syst. Biol. 7, 511 (2011). [PubMed: 21772259]

44. Omasits U, Quebatte M, Stekhoven DJ, Fortes C, Roschitzki B, Robinson MD, Dehio C, Ahrens CH, Directed shotgun proteomics guided by saturated RNA-seq identifies a complete expressed prokaryotic proteome. Genome Res. 23, 1916–27 (2013). [PubMed: 23878158]

45. Liu Y, Beyer A, Aebersold R, On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell. 165 (2016), pp. 535–550. [PubMed: 27104977]

46. Ozsolak F, Milos PM, RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. 12, 87–98 (2011). [PubMed: 21191423]

47. Gerosa L, Kochanowski K, Heinemann M, Sauer U, Dissecting specific and global transcriptional regulation of bacterial gene expression. Mol. Syst. Biol. 9, 658 (2013). [PubMed: 23591774]

48. Brophy JAN, Voigt CA, Principles of genetic circuit design. Nat. Methods. 11, 508–520 (2014). [PubMed: 24781324]

49. Michalodimitrakis K, Isalan M, Engineering prokaryotic gene circuits. FEMS Microbiol. Rev. 33, 27–37 (2009). [PubMed: 19016883]

50. Nandagopal N, Elowitz MB, Synthetic biology: Integrated gene circuits. Science (80-.). 333, 1244–1248 (2011).

51. Sun Z, Brodsky JL, Protein quality control in the secretory pathway. J. Cell Biol. 218, 3171–3187 (2019). [PubMed: 31537714]

52. Dikic I, Proteasomal and autophagic degradation systems. Annu. Rev. Biochem. 86, 193–224 (2017). [PubMed: 28460188]

53. Karp PD, Ong WK, Paley S, Billington R, Caspi R, Fulcher C, Kothari A, Krummenacker M, Latendresse M, Midford PE, Subhraveti P, Gama-Castro S, Muñiz-Rascado L, Bonavides-Martinez C, Santos-Zavaleta A, Mackie A, Collado-Vides J, Keseler IM, Paulsen I, The EcoCyc Database. EcoSal Plus. 8, 1–19 (2018).

54. Westblade LF, Ilag LL, Powell AK, Kolb A, Robinson CV, Busby SJW, Studies of the Escherichia coli Rsd-σ70 complex. J. Mol. Biol. 335, 685–692 (2004). [PubMed: 14687566]

55. Soupene E, Van Heeswijk WC, Plumbridge J, Stewart V, Bertenthal D, Lee H, Prasad G, Paliy O, Charernnoppakul P, Kustu S, Physiological Studies of Escherichia coli Strain MG1655: Growth Defects and Apparent Cross-Regulation of Gene Expression. J. Bacteriol. 185, 5611–5626 (2003). [PubMed: 12949114]

56. Brown SD, Jun S, Complete Genome Sequence of Escherichia coli NCM3722. Genome Announc 3, e00879–15 (2015). [PubMed: 26251500]

57. Basan M, Zhu M, Dai X, Warren M, Sévin D, Wang Y-P, Hwa T, Inflating bacterial cells by increased protein synthesis. Mol Syst Biol. 11, 836 (2015). [PubMed: 26519362]

58. Datsenko KA, Wanner BL, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97, 6640–5 (2000). [PubMed: 10829079]

59. Kuhlman TE, Cox EC, Site-specific chromosomal integration of large synthetic constructs. Nucleic Acids Res. 38, e92 (2010). [PubMed: 20047970]

60. Kuhlman TE, Cox EC, Gene location and DNA density determine transcription factor distributions in Escherichia coli. Mol. Syst. Biol. 8, 610 (2012). [PubMed: 22968444]

61. Langmead B, Salzberg SL, Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9, 357–9 (2012). [PubMed: 22388286]

62. Anders S, Pyl PT, Huber W, Genome analysis HTSeq-a Python framework to work with high-throughput sequencing data. 31, 166–169 (2015).

63. Pigott GH, Midgleyt JEM, "Characterization of Rapidly Labelled Ribonucleic Acid in Escherichia coli by Deoxyribonucleic Acid-Ribonucleic Acid Hybridization" (1968), (available at http://portlandpress.com/biochemj/article-pdf/110/2/251/762136/bj1100251.pdf).

64. Gillespie D, Spiegelman S, A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. J. Mol. Biol. 12, 829–842 (1965). [PubMed: 4955314]

65. Friesen JD, Control of messenger RNA synthesis and decay in Escherichia coli. J. Mol. Biol. 20, 559–573 (1966). [PubMed: 5338988]

66. Chen H, Shiroguchi K, Ge H, Sunney Xie X, Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. Mol Syst Biol. 11, 781 (2015). [PubMed: 25583150]

67. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heinemann M, The quantitative and condition-dependent Escherichia coli proteome. Nat. Biotechnol. 34, 104–110 (2016). [PubMed: 26641532]

68. Woldringh CL, Binnerts JS, Mans A, Variation in Escherichia coli buoyant density measured in Percoll gradients. J. Bacteriol. 148, 58–63 (1981). [PubMed: 6270065]

69. Si F, Li D, Cox SE, Sauls JT, Azizi O, Sou C, Schwartz AB, Erickstad MJ, Jun Y, Li X, Jun S, Invariance of Initiation Mass and Predictability of Cell Size in Escherichia coli. Curr. Biol. 27, 1278–1287 (2017). [PubMed: 28416114]

70. Gottesman S, Proteases and their targets in Escherichia coli. Annu. Rev. Genet. 30, 465–506 (1996). [PubMed: 8982462]

71. Li GW, How do bacteria tune translation efficiency? Curr. Opin. Microbiol. 24, 66–71 (2015). [PubMed: 25636133]

72. Zarai Y, Margaliot M, Tuller T, On the Ribosomal Density that Maximizes Protein Translation Rate (2016), doi:10.1371/journal.pone.0166481.

73. Buskirk AR, Green R, Ribosome pausing, arrest and rescue in bacteria and eukaryotes, doi:10.1098/rstb.2016.0183.

74. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K, "Translation Elongation Factor EF-P Alleviates Ribosome Stalling at Polyproline Stretches," (available at http://science.sciencemag.org/).

75. Thanaraj TA, Argos P, Ribosome-mediated translational pause and protein domain organization. Protein Sci. 5, 1594–1612 (1996). [PubMed: 8844849]

76. Keiler KC, Biology of trans-translation. Annu. Rev. Microbiol. 62, 133–51 (2008). [PubMed: 18557701]

77. Keiler KC, Waller PRH, Sauer RT, Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science (80-.). 271, 990–993 (1996).

78. Kuhlman T, Zhang Z, Saier MH, Hwa T, Combinatorial transcriptional control of the lactose operon of Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 104, 6043–6048 (2007). [PubMed: 17376875]

79. Heimann JD, The extracytoplasmic function (ECF) sigma factors (2002), vol. 46.

80. Mooney RA, Darst SA, Landick R, Sigma and RNA polymerase: An on-again, off-again relationship? Mol. Cell. 20, 335–345 (2005). [PubMed: 16285916]

81. Mazumder A, Kapanidis AN, Recent Advances in Understanding σ70-Dependent Transcription Initiation Mechanisms. J. Mol. Biol. 431, 3947–3959 (2019). [PubMed: 31082441]

82. Mooney RA, Landick R, Tethering σ70 to RNA polymerase reveals high in vivo activity of σ factors and σ70-dependent pausing at promoter-distal locations. Genes Dev. 17, 2839–2851 (2003). [PubMed: 14630944]

83. Shimamoto N, Kamigochi T, Utiyama H, Release of the σ subunit of Escherichia coli DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. J. Biol. Chem. 261, 11859–11865 (1986). [PubMed: 2427513]

84. Mukhopadhyay J, Kapanidis AN, Mekler V, Kortkhonjia E, Ebright YW, Ebright RH, Translocation of σ70 with RNA polymerase during transcription: Fluorescence resonance energy transfer assay for movement relative to DNA. Cell. 106, 453–463 (2001). [PubMed: 11525731]

85. Kapanidis AN, Margeat E, Laurence TA, Doose S, Ho SO, Mukhopadhyay J, Kortkhonjia E, Mekler V, Ebright RH, Weiss S, Retention of transcription initiation factor σ70 in transcription elongation: Single-molecule analysis. Mol. Cell. 20, 347–356 (2005). [PubMed: 16285917]

86. Raffaelle M, Kanin EI, Vogt J, Burgess RR, Ansari AZ, Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. Mol. Cell. 20, 357–366 (2005). [PubMed: 16285918]

87. Sun Z, Yakhnin AV, FitzGerald PC, Mclntosh CE, Kashlev M, Nascent RNA sequencing identifies a widespread sigma70-dependent pausing regulated by Gre factors in bacteria. Nat. Commun. 12, 1–14 (2021). [PubMed: 33397941]

88. Harden TT, Wells CD, Friedman LJ, Landick R, Hochschild A, Kondev J, Gelles J, Bacterial RNA polymerase can retain σ70 throughout transcription. Proc. Natl. Acad. Sci. U. S. A. 113, 602–607 (2016). [PubMed: 26733675]

89. Zhu M, Mu H, Jia M, Deng L, Dai X, Control of ribosome synthesis in bacteria: the important role of rRNA chain elongation rate. Sci. China Life Sci. 64, 795–802 (2020). [PubMed: 32840734]

90. Mauri M, Klumpp S, A Model for Sigma Factor Competition in Bacterial Cells. PLoS Comput. Biol. 10, 29–34 (2014).

91. Cayley S, Lewis BA, Guttman HJ, Record MT, Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity: Implications for protein-DNA interactions in vivo. J. Mol. Biol. 222, 281–300 (1991). [PubMed: 1960728]

92. Kennell D, Titration of the gene sites on DNA by DNA-RNA hybridization. II. The Escherichia coli chromosome. J. Mol. Biol. 34, 85–103 (1968). [PubMed: 4938545]

93. Norris TE, Koch AL, Effect of growth rate on the relative rates of synthesis of messenger, ribosomal and transfer RNA in Escherichia coli. J. Mol. Biol. 64, 633–649 (1972). [PubMed: 4553856]

94. Sowa SW, Gelderman G, Leistra AN, Buvanendiran A, Lipp S, Pitaktong A, Vakulskas CA, Romeo T, Baldea M, Contreras LM, Integrative FourD omics approach profiles the target network of the carbon storage regulatory system. Nucleic Acids Res. 45, 1673–1686 (2017). [PubMed: 28126921]

95. Chen S, Zhang A, Blyn LB, Storz G, MicC, a second small-RNA regulator of omp protein expression in Escherichia coli. J. Bacteriol. 186, 6689–6697 (2004). [PubMed: 15466019]

96. Wang J, Rennie W, Liu C, Carmack CS, Prévost K, Caron MP, Massé E, Ding Y, Wade JT, Identification of bacterial sRNA regulatory targets using ribosome profiling. Nucleic Acids Res. 43, 10308–10320 (2015). [PubMed: 26546513]

97. De Lay N, Gottesman S, A complex network of small non-coding RNAs regulate motility in Escherichia coli. Mol. Microbiol. 86, 524–538 (2012). [PubMed: 22925049]

98. Thomason MK, Fontaine F, De Lay N, Storz G, A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in Escherichia coli. Mol. Microbiol. 84, 17–35 (2012). [PubMed: 22289118]

99. Baracchini E, Bremer H, Control of rRNA synthesis in Escherichia coli at increased rrn gene dosage: Role of guanosine tetraphosphate and ribosome feedback. J. Biol. Chem. 266, 11753–11760 (1991). [PubMed: 1711040]

100. Ryals J, Little R, Bremer H, Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. J. Bacteriol. 151, 1261–1268 (1982). [PubMed: 6179924]

101. Massé E, Gottesman S, A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 99, 4620–4625 (2002). [PubMed: 11917098]

102. Lalaouna D, Morissette A, Carrier MC, Massé E, DsrA regulatory RNA represses both hns and rbsD mRNAs through distinct mechanisms in Escherichia coli. Mol. Microbiol. 98, 357–369 (2015). [PubMed: 26175201]

103. Klein G, Raina S, Small regulatory bacterial RNAs regulating the envelope stress response. Biochem. Soc. Trans. 45, 417–425 (2017). [PubMed: 28408482]

104. Lalanne JB, Taggart JC, Guo MS, Herzel L, Schieler A, Li GW, Evolutionary Convergence of Pathway-Specific Enzyme Expression Stoichiometry. Cell. 173, 749–761.e38 (2018). [PubMed: 29606352]

105. Feucht BU, Saier MH, Fine control of adenylate cyclase by the phosphoenolpyruvate:sugar phosphotransferase systems in Escherichia coli and Salmonella typhimurium. J. Bacteriol. 141, 603–610 (1980). [PubMed: 6245052]

106. Busby S, Ebright RH, Transcription activation by catabolite activator protein (CAP). J. Mol. Biol. 293, 199–213 (1999). [PubMed: 10550204]

107. Kolb A, Busby S, Buc H, Gorges S, Adhya S, Transcriptional regulation by cAMP and its receptor protein. Annu. Rev. Biochem. 62, 749–795 (1993). [PubMed: 8394684]

108. Epstein W, Rothman Denes LB, Hesse J, Adenosine 3':5' cyclic monophosphate as mediator of catabolite repression in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 72, 2300–2304 (1975). [PubMed: 166384]

109. Park S-J, Tseng C-P, Gunsalus RP, Regulation of succinate dehydrogenase sdhCDAB operon expression in Escherichia coli in response to carbon supply and anaerobiosis: role of ArcA and Fnr. Mol. Microbiol. 15, 473–482 (1995). [PubMed: 7783618]

110. Lynch AS, Lin ECC, Transcriptional control mediated by the ArcA two-component response regulator protein of Escherichia coli: Characterization of DNA binding at target promoters. J. Bacteriol. 178, 6238–6249 (1996). [PubMed: 8892825]

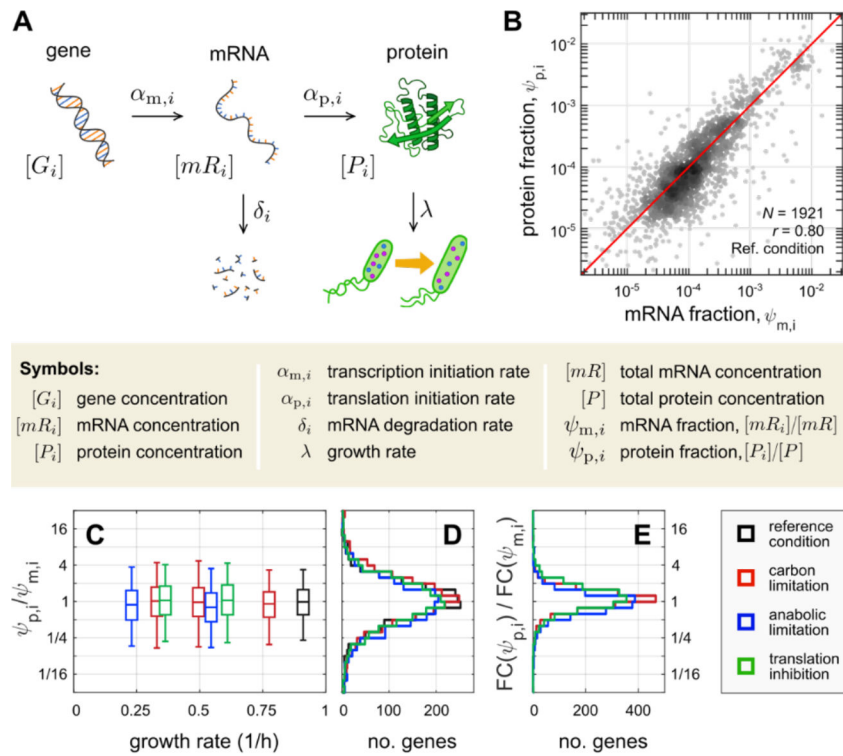**Figure 1. Genome-wide mRNA and protein comparison.**

(A) Schematic illustration of the basic processes determining mRNA and protein concentrations in exponentially growing bacteria. The rate of each process can potentially vary across both genes and conditions; the symbols used throughout the study are described alongside the respective cellular processes (see also Fig. S1).

(B) For *E. coli* K-12 strain NCM3722 growing exponentially in glucose minimal medium (reference condition, growth rate 0.91/h), the fractional number abundances of proteins ($\psi_{p,i}$, obtained from DIA/SWATH mass spectrometry(10) and of mRNAs ($\psi_{m,i}$, obtained from RNA-sequencing; see Methods) for each gene *i* are shown as scatter plot (number of genes and Pearson correlation coefficient in figure). The red line represents the diagonal, $\psi_{p,i} = \psi_{m,i}$.

(C) The ratios of protein and mRNA fractions, $\psi_{p,i}/\psi_{m,i}$ are distributed around 1 for exponentially growing cultures under all growth conditions studied (Fig. S3E–S). These include the reference condition (black), as well as conditions of reduced growth, achieved by limiting carbon catabolism (red), anabolism (blue), or inhibiting translation (green); see SI Methods. Boxes and the whiskers represent 50% and 90% of the genes, respectively; x-axis values give the corresponding growth rates. See Tables S1 and S2 for list of strains and conditions in this study, and Table S3–4 for transcriptomics and proteomics data.

(D) Distributions of the ratios $\psi_{p,i}/\psi_{m,i}$ obtained in reference condition and the slowest-growing of each of the three types of limitations; same color code as (C). The same plots also give the distributions of the relative translational initiation rate, $\alpha_{p,i}/\bar{\alpha}_p$; see text.

(E) The fold-changes in protein and mRNA fractions for each gene *i* between the reference condition and the slowest growth condition, FC($\psi_{p,i}$) and FC($\psi_{m,i}$), were computed as described in Fig. S4 for each one of the three growth limitations; the distribution of their

ratio $FC(\psi_{p,i})/FC(\psi_{m,i})$ is shown using the same color code as (C). The histograms are narrowly distributed around 1, with more than half of the genes within 35% from the median. See Table S5 for the fold changes in translation efficiency for each gene.
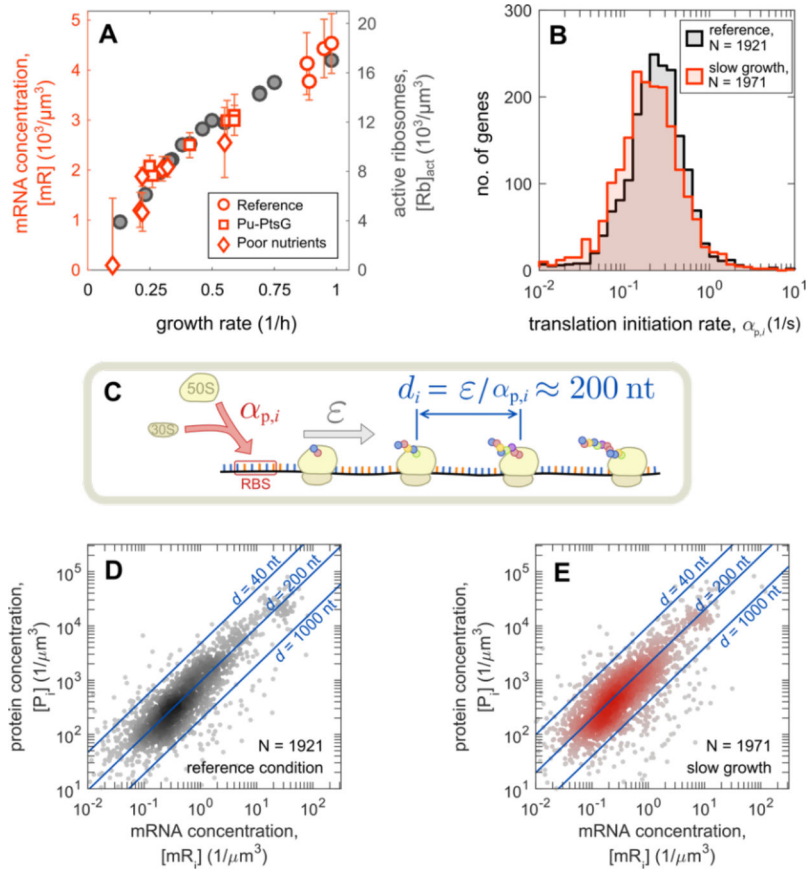
Figure 2. Coordination of mRNA and ribosome abundances.

(A) Left axis (red symbols): total concentration of mRNA is plotted against the growth rate. Total mRNA abundance and associated standard deviations based on 3 measurements obtained as described in Fig. S5 and Methods. The measurements were performed for a range of growth conditions, including reference, glucose uptake titration (Pu-*ptsG*, see Table S1) and a host of poor carbon sources. Right axis (grey symbols): concentration of active ribosomes in nutrient-limited conditions, converted from the data in Ref.(7) (reported per culture volume) using the total cellular volume shown in Fig. S2C–E.

(B) Translation initiation rates, $\alpha_{p,i} = \bar{\alpha}_p \cdot (\psi_{p,i}/\psi_{m,i})$, in reference (black) and carbon-limited (red) growth.

(C) The spacing between consecutive translating ribosomes on an mRNA is given by the ratio between the ribosome elongation rate (similar across mRNAs, Fig. S6 and Ref.(7)) and the translation initiation rate $a_{p,i}$, which is also narrowly distributed (see panel B). Our data give an average ribosome spacing of $d \approx 200$ nt; see Fig. S6D.

(D) Absolute mRNA and protein concentration for each gene in reference condition, computed by combining the fractional abundances $\psi_{m,i}$ and $\psi_{p,i}$ with total mRNA abundances (panel A), total protein abundances and cell volume (see Fig. S2 and Note S1). Blue lines indicate the corresponding values of inter-ribosome spacing d, calculated from the known elongation rates (~15.3 aa/s).

(E) Same as panel (D), but for slow growth in the most C-limiting condition (growth rate ~ 0.35/h, elongation rate ~12.4 aa/s (7)).
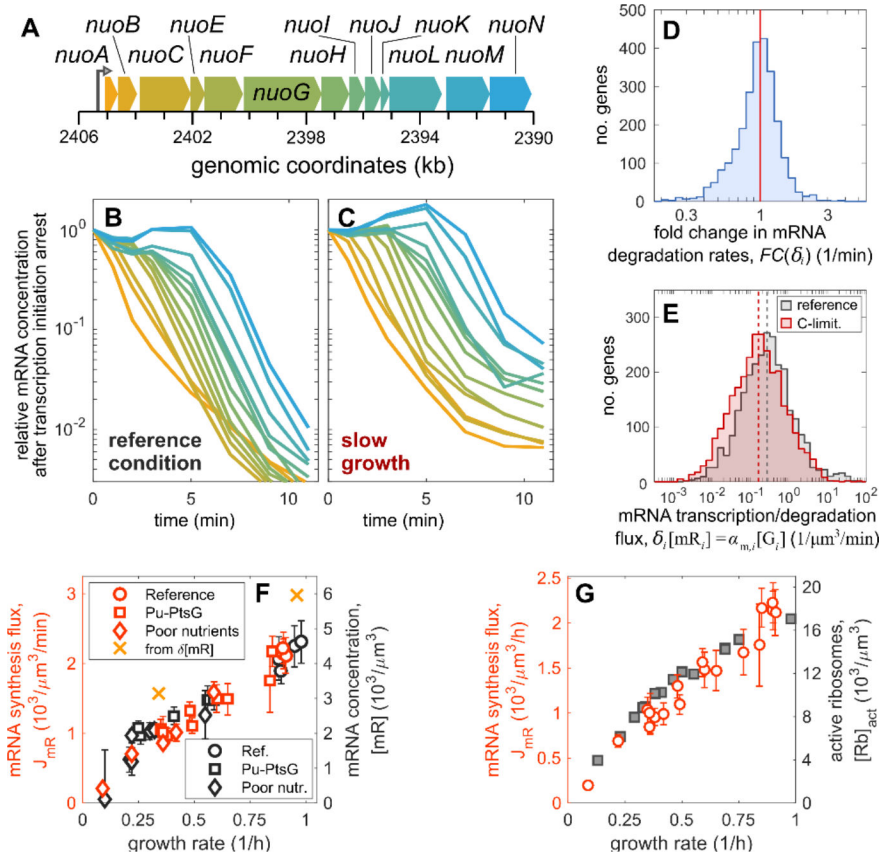
Figure 3. mRNA degradation and synthesis.

(A-C) Degradation of mRNA transcribed from the long *nuo* operon (A) in reference condition (B) and carbon-limited condition (C). The abundance of mRNA was measured by RNA-seq over the course of 11 minutes following the blockage of transcription initiation by rifampicin (SI Methods, Fig. S7). While the abundance of the mRNA of genes proximal to the promoter (*nuoA*, orange) drops immediately after rifampicin treatment (at time $t = 0$), a lag is observed for genes progressively more distant from the promoter (from orange to blue). The lag time corresponds to the time elapsed between the transcription of the proximal and distant genes by RNAPs which initiated transcription before the application of rifampicin (Fig. S7D).

(D) Histogram of fold-change of the mRNA degradation rates, $FC(\delta_i)$, between carbon limited medium and reference condition for $N = 2550$ genes. Half of the fold changes are within 25% from unity, and 90% of the fold changes are in the range 0.50 to 1.57, implying that the degradation rates for most mRNAs do not change significantly between the reference and carbon-limited growth conditions.

(E) Distribution of the mRNA degradation fluxes, $\delta_i[mR_i]$, computed from the mRNA concentration and degradation rates. These quantities should equate the mRNA synthesis fluxes, $\alpha_{m,i}[G_i]$, in steady state conditions. Dashed lines indicate the median fluxes, $0.194/\mu m^3/min$ in reference condition and $0.108/\mu m^3/min$ at slow growth.

(F) Left axis (red symbols): total mRNA synthesis flux $J_{mR} = \sum_i \alpha_{m,i}[G_i]$ (transcripts synthesized per cell volume per unit time), for a variety of growth conditions as indicated

(see Table S2 for growth conditions). The slope of radiolabel incorporated into mRNA over time was used to obtain the mRNA synthesis flux while the error bars represent the standard deviation from 6 measurements at different time-points following the label addition (Fig. S10). The orange crosses indicate the total mRNA synthesis flux obtained from summing $\delta_i \cdot [mR_i]$ using the data in (E). Right axis (black symbols): absolute mRNA abundances (same data as Fig. 2A).

(G) Left axis (red symbols): total RNA synthesis flux vs. growth rate (same data as in panel (F)). Right axis (grey symbols): concentration of active ribosomes (same data as Fig. 2A).
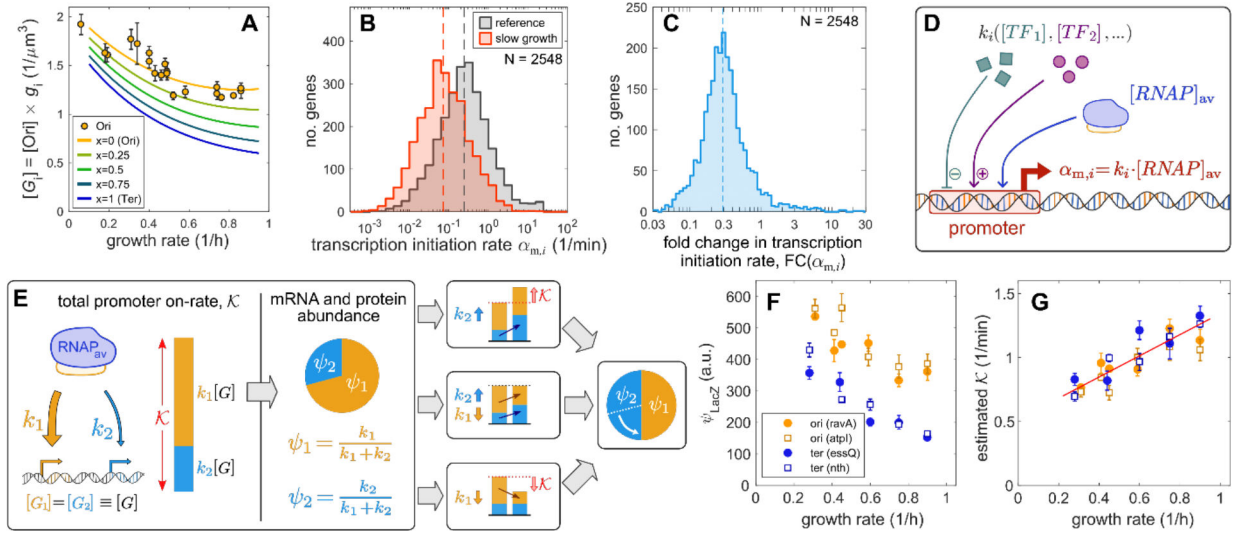
Figure 4. Quantitative relations between promoter on-rates and mRNA, protein abundances.
(**A**) Growth rate dependence of gene concentration $[G_i]$ at various distances x from the origin of replication Ori (solid lines). These are computed as the product of the Ori concentration $[Ori]$ (orange circles, shown in Fig. S9C with raw data and standard errors from Ref.(24)) and the gene dose $g_i = [G_i]/[Ori]$ (Fig. S9B); see Fig. S9 for details.
(**B**) Distribution of transcription initiation rates $\alpha_{m,i}$ in reference condition (black) and slow growth (red), computed using the available mRNA abundances and degradation rates (see SI Note S3 for details). Dashed lines indicate the median initiation rates in the two conditions (2.64/min for reference condition, 0.87/min for slow growth).
(**C**) Fold change of the transcription initiation rates $FC(\alpha_{m,i})$ between reference condition and slow growth. The data shows a generalized decrease of initiation rates, with a median reduction of 0.29 (dashed line) at slow growth ($\lambda = 0.3$/h) compared to the reference condition ($\lambda = 0.91$/h). (**D**) Illustration of a canonical model of transcriptional regulation(28, 29), with the transcription initiation rate for gene i, $\alpha_{m,i}$, depending on the promoter on-rate k_i, which is modulated by transcription factors (TF_1, TF_2, …), as well as on the cellular concentration of available RNA polymerases ($[RNAP]_{av}$), as described by Eq. (8).
(**E**) Cartoon illustrating the dependence of mRNA and protein abundances on the promoter on-rates, as described by Eq. (12). Consider two genes with promoter on-rates k_1 (orange) and k_2 (blue) and identical gene concentration $[G_1] = [G_2] \equiv [G]$; the corresponding mRNA and protein fractions ($\psi_{m,1} = \psi_{p,1} \equiv \psi_1$ and $\psi_{m,2} = \psi_{p,2} \equiv \psi_2$, respectively) depend on both promoter on-rates via the total regulatory activity $\mathcal{K} = (k_1 + k_2)[G]$ (in red). Three possible scenarios are illustrated. Top: If k_2 increases, while k_1 remains constant, then $\mathcal{K}$ increases, resulting in the reduction of protein and mRNA abundances for the orange gene despite it not being downregulated at the transcriptional level. Bottom: If only k_1 decreases while k_2 remains constant (bottom), then the proteins and mRNAs for the blue gene increase despite the lack of change at its promoter level. Middle: If $\mathcal{K}$ is unchanged (due to compensating changes in k_1 and k_2 in this case), then the changes in protein and mRNA fractions would reflect changes at the regulatory level.
(**F**) *E. coli* strains harboring constitutive expression of *lacZ* at various locations near *oriC* (orange) and near *terC* (blue; loci listed in the legend) were grown in carbon-limited

conditions (see Tables S1–S2 for strains and conditions). LacZ protein abundance per culture volume (OD·mL), obtained from the slopes of β-gal activity versus OD_{600} (Miller units), is shown; error bars indicate standard errors from 4 measurements (Methods).

(G) The relative change in the total regulatory activity $\mathscr{K}$ across growth rates was estimated from the relative change in LacZ abundance using the data in panel (F) and Eq. (14) in the text. To do so, the LacZ abundance per culture volume was converted to protein fraction by dividing by total protein mass per culture volume (Fig. S2F). The result shows a linear dependence of the total regulatory activity on the growth rate (red line). The absolute scale $\mathscr{K}$ was set for the reference condition using Eq. (10) with the values for the total mRNA synthesis flux J_{mR} obtained from Fig. 3F, the $oriC$ concentration from Fig. 4A, and the available RNAP concentration estimated as described in SI Note S5.
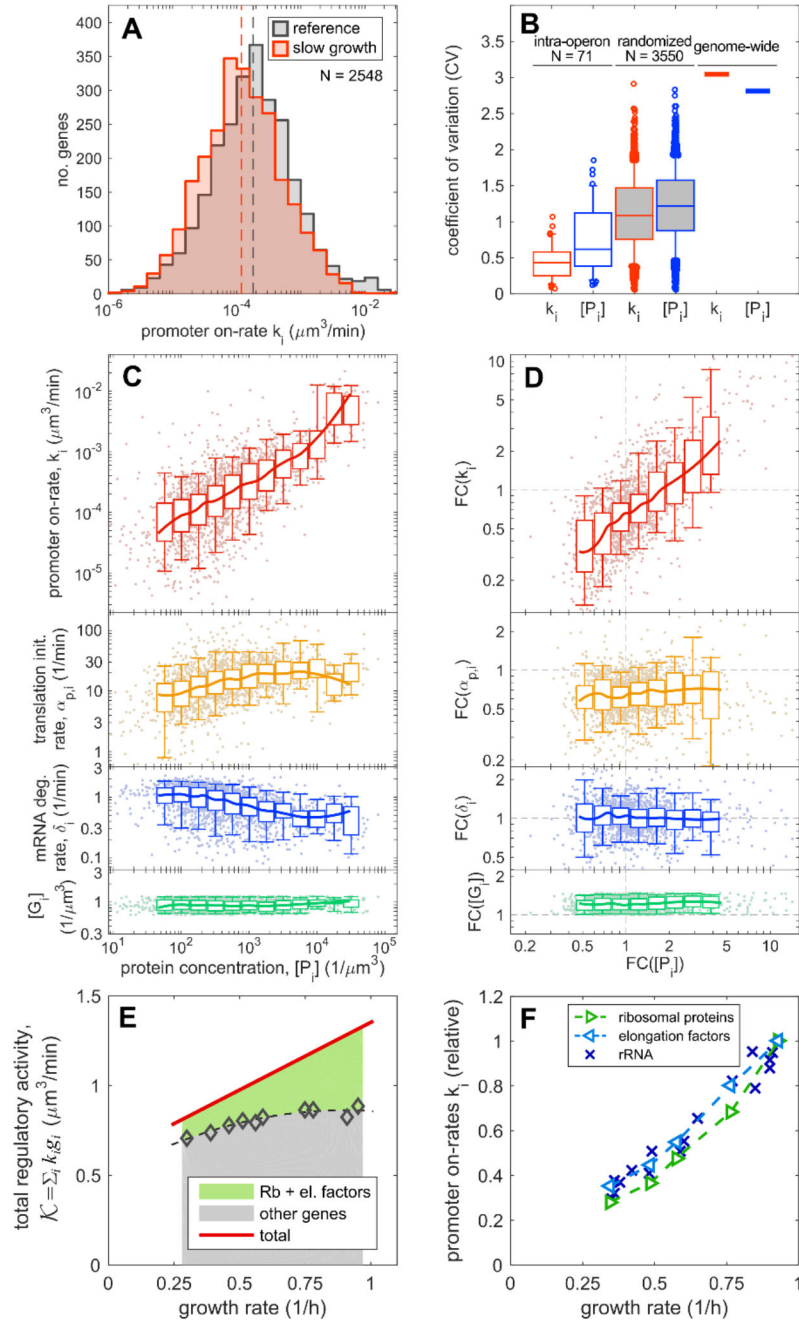
Figure 5. Gene expression is primarily determined by the promoter on-rates.

(A) Distribution of promoter on-rates k_i in the reference and slow growth condition, obtained from the distribution of the translation initiation rate and the concentrations of available RNAP, $k_i = \alpha_{m,i}/[RNAP]av$ (see Eq. (8)), as described in SI Note S4. The median promoter-on rate (vertical dashed lines) shifts from $1.63 \cdot 10^{-4}$ μm^3 in reference condition ($\lambda \sim 0.9$/h) to $1.07 \cdot 10^{-4}$ μm^3/min in slow growth ($\lambda \sim 0.3$/h). This change is much less than the ~3-fold change in both the growth rate and the median transcription initiation rates (Fig. 4BC).

(B) For 71 operons containing at least 3 genes as annotated in Ecocyc(53), we computed the coefficient of variation (CV) in the promoter on-rates k_i or in the protein concentrations $[P_i]$

for genes within each operon in reference condition. The average intra-operon CVs for the promoter on-rates are significantly smaller than that computed for the protein concentrations $[P_i]$ ($p < 7 \cdot 10^{-7}$, unpaired t-test); see also Fig. S10A. As a control, we randomly shuffled the genes across the operons 50 times, leading to sets of 3550 CVs (grey-filled boxes), and considered the CVs computed using all available genes (lines on the right). The CV for the promoter on-rates are also significantly smaller than all the other distributions ($p < 3 \cdot 10^{-35}$ when comparing to the randomized cases) and the genome-wide CVs. Boxes and whiskers indicate 50% and 90% intervals, respectively; median CVs are indicated by the central lines within the boxes.

(C) Promoter on-rates k_i, translation initiation rates $a_{\mathrm{p},i}$, mRNA degradation rates δ_i and gene concentrations $[G_i]$ are the four molecular parameters determining cellular concentration of a protein in a given growth condition (Fig. 1A, with the transcription initiation rate $a_{\mathrm{m},i}$ given by k_i via Eq. (8)). These four molecular parameters are plotted against the protein concentrations $[P_i]$ in reference condition, binned according to the observed protein concentrations. Boxes and whiskers indicate 50% and 90% central intervals for the binned data; the solid lines represent moving averages.

(D) Same as panel (C), but for the fold changes (FC) of each quantity across growth conditions (slow growth compared to reference). All molecular parameters and concentrations shown in panels A-D are listed in Table S6.

(E) The sum of promoter on-rates weighted by gene dose, $\mathscr{K} = \sum_i k_i g_i$ (red line; same as in Fig. 4G) is partitioned between the contribution from ribosomal proteins and translation elongation factors (green) and the rest of genes (grey area). Symbols indicate the partitioning obtained from the computed k_i across growth rates. The growth rate dependence of $\mathscr{K}$ largely stems from that of the promoter on-rates of the translational genes.

(F) Growth rate dependence of promoter on-rates summed over different groups of genes: ribosomal proteins, elongation factors (encoded by *fusA*, *tufAB* and *tsf*), and the rRNA operons. The activity of the rRNA operons was estimated from the synthesis flux of stable RNA (SI Methods and Fig. S14).
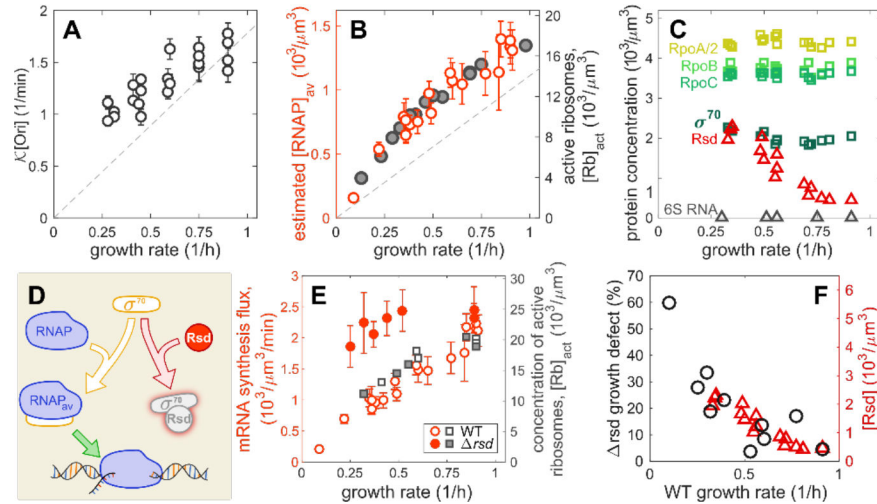
Figure 6. The role of the anti-sigma factor Rsd in global regulation of mRNA synthesis.
(A) Value of $\mathscr{K} \cdot [Ori]$ across growth rates, obtained from the values (data and standard errors) of the total regulatory activity $\mathscr{K}$ shown in Fig. 4G, multiplied by the interpolated values for $[Ori]$ at the same growth rates (Fig. 4A). For comparison, the dashed line shows direct proportionality to the growth rate.
(B) Concentration of available RNA polymerases (red symbols, left axis), estimated from the ratio between the measured mRNA synthesis flux (data and errors in Fig. 3G) and $\mathscr{K} \cdot [Ori]$ (using the interpolated curves in Fig. 4A and 4G). Note that this quantity shows a stronger dependence on the growth rate compared to $\mathscr{K} \cdot [Ori]$ in panel (A) and has the same growth-rate dependence as the concentration of active ribosomes (grey symbols, right axis).
(C) The concentrations of various components of the transcription machinery in carbon-limited conditions is plotted against the growth rate. Components of the core enzyme, RpoABC, and the major sigma factor σ^{70} are shown as squares. Known modulators of σ^{70}, Rsd and 6S RNA are shown as triangles. The protein concentrations are determined from mass spectrometry(10), while the concentration of 6S RNA is determined from RNA-sequencing and the concentration of total mRNA concentration (Fig. S4).
(D) Cartoon illustrating the control of RNA polymerase (RNAP) availability through the known σ^{70}-sequestration function of Rsd(33, 54).
(E) Comparison of mRNA synthesis fluxes between wild type (open symbols) and *rsd* strain (filled symbols). Left axis: total mRNA synthesis flux of *rsd* strain (red filled circles) and wild type (red open circles); standard errors are computed as in Fig. 3F. Right axis: concentration of active ribosomes computed from the measured total RNA for the two strains and the fraction of active ribosomes observed in carbon limited growth(7).
(F) The growth defect of *rsd* strain, defined as % reduction in growth rate compared to wild type cells in the same growth condition (black circles, left axis), is plotted against the growth rate of wild type cells for the range of carbon-limited growth conditions. The observed growth reduction matches Rsd expression of wild type cells in the same conditions (red triangles, right axis; same data as in panel C).