

The protein translation machinery is expressed for maximal efficiency in *Escherichia coli*

Xiao-Pan Hu ¹, Hugo Dourado ¹, Peter Schubert ¹ & Martin J. Lercher ¹ 

Protein synthesis is the most expensive process in fast-growing bacteria. Experimentally observed growth rate dependencies of the translation machinery form the basis of powerful phenomenological growth laws; however, a quantitative theory on the basis of biochemical and biophysical constraints is lacking. Here, we show that the growth rate-dependence of the concentrations of ribosomes, tRNAs, mRNA, and elongation factors observed in *Escherichia coli* can be predicted accurately from a minimization of cellular costs in a mechanistic model of protein translation. The model is constrained only by the physicochemical properties of the molecules and has no adjustable parameters. The costs of individual components (made of protein and RNA parts) can be approximated through molecular masses, which correlate strongly with alternative cost measures such as the molecules' carbon content or the requirement of energy or enzymes for their biosynthesis. Analogous cost minimization approaches may facilitate similar quantitative insights also for other cellular subsystems.

¹Institute for Computer Science and Department of Biology, Heinrich Heine University, D-40225 Düsseldorf, Germany. ✉email: martin.lercher@hhu.de

Protein translation is central to the self-replication of biological cells. While the workings of its individual components are well understood, the translation apparatus is a complex machine with many degrees of freedom, where the same rate of protein production could be achieved with very different relative abundances of its components. Although a large body of quantitative experimental data on these abundances in *E. coli* across different growth conditions is available, it is still unclear according to which organizing principle(s)—if any—they are set by the cell. Given that translation is the energetically most expensive process in fast growing *E. coli* cells, accounting for up to 50% of the proteome¹ and 2/3 of cellular ATP consumption², it is likely that natural selection acted to optimize the efficiency of translation. But what exactly is efficiency in the evolutionary context?

In the late 1950s, it was hypothesized that ribosomes operate at a constant, maximal rate^{3,4}, consistent with the observed linear dependence of ribosome concentration on growth rate^{3,5–7}. This hypothesis was later proven untenable, as the activity of ribosomes was observed to increase with growth rate⁸. Klumpp et al.⁹ suggested that optimal translational efficiency corresponds to the parsimonious usage of translation-associated proteins, most notably ribosomal proteins, elongation factor Tu, and tRNA synthetases. While these authors were able to fit a coarse-grained phenomenological model to the data, their suggested evolutionary objective could also not explain the observed growth rate dependencies quantitatively (see Supplementary Note 1 for a

discussion of ref. ⁹ and of the phenomenological model of bacterial growth it is based on^{5,10}). Thus, it is currently unclear to what extent translation has indeed been optimized by natural selection, and—if such optimization indeed occurred—whether its action can be expressed in terms of a simple objective function.

In principle, these questions would best be addressed in the context of a whole-cell model of balanced growth that combines mechanistic descriptions of metabolism and protein production. However, while such models have been described conceptually^{11,12}, kinetic parameterizations are unavailable for a majority of the relevant enzymatic reactions¹³, preventing a truly mechanistic description that combines metabolism and protein translation. Thus, we here focus on protein production alone, taking the experimentally observed output of translation (proteome production rate and composition in a given growth condition), the corresponding input (charged tRNAs), and the kinetics of individual translation reactions as given. We then use this mechanistic description of translation to find the combination of the concentrations of mRNA, ribosomes, elongation factors, and tRNAs that results in minimal cellular costs in the given condition. Thus, our estimate of the optimal efficiency of the translation machinery is not based on the maximization of ribosome activity, but on the minimization of the combined cost of the complete translation machinery at an observed protein production output.

We base our cost definition on the experimental observation that cellular dry mass per cell volume is approximately constant

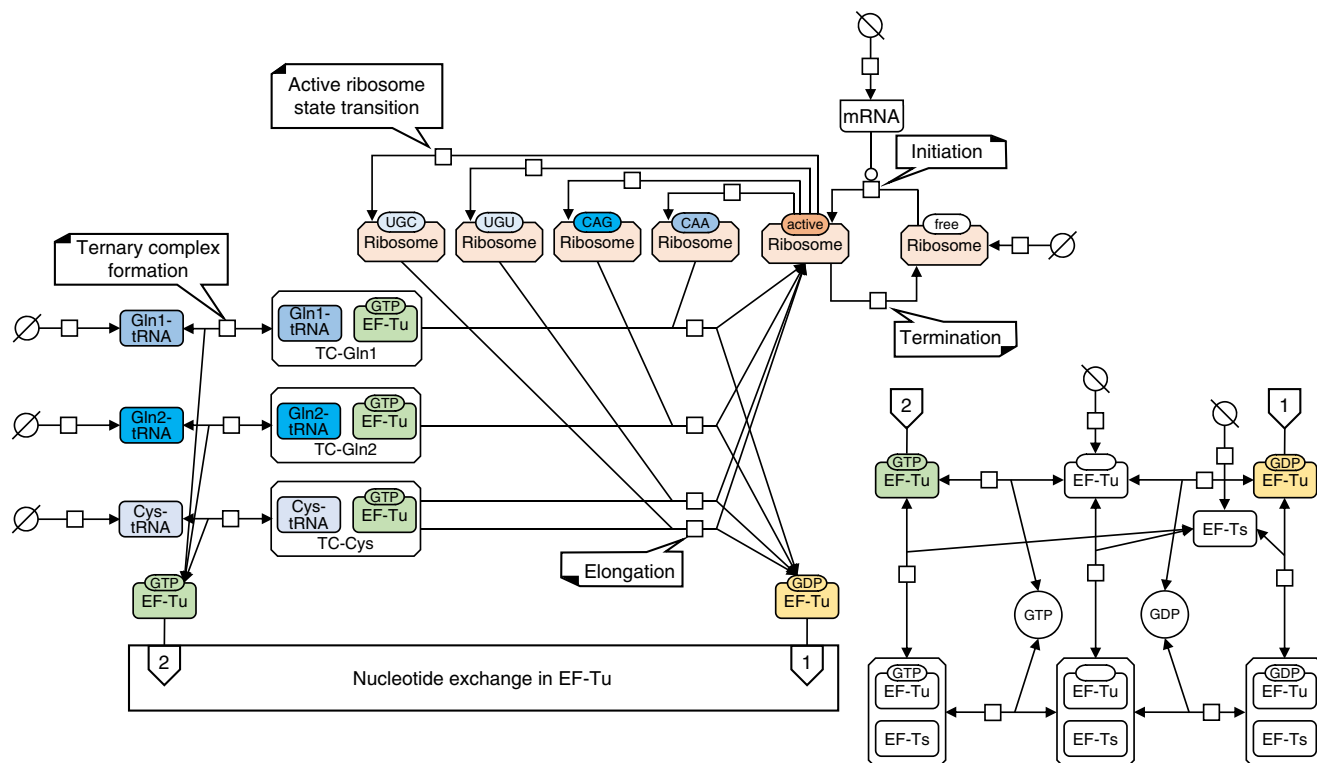


Fig. 1 Schematic overview of the translation model. A reduced pathway for elongation with amino acids cysteine (aminoacyl-tRNA Cys) and glutamine (aa-tRNA Gln1, Gln2) is represented in Systems Biology Graphical Notation. Initiation: free (unbound) ribosome gets converted to active ribosome, modulated by mRNA. Termination: active ribosome converts back to free ribosome at a rate fixed by the desired protein production rate. Active ribosome state transition: active ribosome instantaneously binds to codons (61 codons in full model, 4 here) at the fractions set by the specified proteome composition. Ternary complex formation: charged tRNAs (40 aa-tRNAs in full model, 3 here), replenished from a pool, combine with EF-Tu*GTP to form ternary complexes (40 TCs in the full model). Kinetic parameters of these reversible processes depend on the aa-tRNA. Elongation: labeled ribosome binds with the cognate TC to elongate the protein with the respective amino acid. The ribosome returns to its active state and EF-Tu*GDP is released. Other products of this reaction, such as deacylated tRNA, are not modeled. Nucleotide exchange (see right panel): EF-Tu*GDP is reactivated to EF-Tu*GTP in a sequence of steps modeled by reversible mass action kinetics. GTP and GDP pools are modeled with fixed concentrations. The nucleotide exchange is supported by EF-Ts, and the main flux is carried through the complexes formed by EF-Tu with EF-Ts.

across environments and growth rates in *E. coli*¹⁴, as is the total mass concentration in the cytosol¹⁵. If the cell allocates more of this limited mass concentration budget to one particular process, less is available to other processes. The upper bound for the cytosolic mass concentration, beyond which diffusion becomes inefficient, is a fundamental constraint on cellular growth^{16–18}, and we thus use the cytosolic mass concentration of a particular molecule type as an approximation to its cost. Theoretical models of cellular growth that account for all major biochemical and biophysical constraints indicate that the limit on cellular dry mass indeed represents a dominant constraint on bacterial growth rates¹².

We hypothesize that to maximize the *E. coli* growth rate in a given environment, natural selection minimizes the total mass concentration of translation components utilized to achieve the required protein production rate. A corresponding optimality principle has been used to understand the relationship between the concentrations of enzymes and their substrates¹⁹.

We find that a theoretical minimization of the combined cellular costs of the translation machinery components indeed leads to accurate predictions for their abundances, the resulting elongation rate, and the RNA/protein ratio. In addition to molecular masses, we also examine four alternative cost measures for cellular components that have been explored in the literature: (i) their protein content^{9,20}; (ii) their carbon content²¹; and (iii) the energy^{22,23} or (iv) the amount of catalysts^{5,24} required for their production. We find that these alternative cost measures are strongly correlated for the studied components of the translation machinery and lead to very similar predictions for their abundances; the only cost measure that leads to substantially different predictions is the protein content, which does not assign any cost to tRNA and mRNA molecules.

Results

Cost minimization in a mechanistic model of translation. To test our hypothesis, we constructed a translation model consisting of 274 biochemical reactions, including 119 reactions with non-linear kinetics. Figure 1 shows the modeled reactions for a subset of the 61 codons and the 40 species of charged tRNAs; for details see Methods, Supplementary Table 1, and Supplementary Data 1. This mechanistic model accounts for the concentrations of mRNA, the ribosome, the different charged tRNAs, and the elongation factors Ts (EF-Ts) and Tu (EF-Tu). We fully parameterized the model with molecular masses and kinetic constants measured experimentally^{25–27}; the only exceptions are the translation initiation parameters, which were previously estimated from gene expression data²⁵, and the ribosomal Michaelis constant for the ternary complexes, which was previously estimated based on the diffusion limit⁹. The model is based purely on biochemical and biophysical considerations; it contains no adjustable parameters, nor does it include any explicit growth-rate dependencies.

For *E. coli* growing under different experimental conditions, we used measured growth rates and protein concentrations²⁸ to determine the required translation rate and the proportions of the different amino acids incorporated into the elongating proteins. At this required protein production rate, we minimized the combined cost of the translation machinery in our model, treating the concentrations of all components as free variables; the values of individual reaction fluxes result deterministically from these concentrations according to the respective rate laws (Methods). As the modeled kinetic rate laws are non-linear, all optimizations were performed numerically. In repeated optimization runs with two different solvers, we never found alternative optima, indicating that the optimization problem may be convex.

The results shown in the main text and figures are based on the assumption that costs are proportional to molecular masses; results based on other cost functions are shown in Supplementary Figs. 5 and 7.

Predicted concentrations agree with observations. We first compared our predictions to experimental data for exponentially growing *E. coli* in different conditions^{8,28–31}. Figure 2 shows the results for growth in a glucose-limited chemostat at growth rate $\mu = 0.35 \text{ h}^{-1}$; for other conditions, see Supplementary Fig. 1. The mechanistic model accurately predicts the absolute concentrations of ribosomes, EF-Tu, EF-Ts, mRNA, and total tRNA in each condition. Predictions for individual tRNA concentrations are less accurate but are still mostly within a 2-fold error (Fig. 2, Supplementary Fig. 1); the discrepancies may be due to the simplifying assumption of the same ribosomal Michaelis constant K_m for all tRNA species⁹.

We next tested if this systems-level view on the total cost of translation explains the observed growth rate-dependencies of the expression of translation machinery components^{8,28,29,31,32}, of the elongation rate³², and of the RNA/protein ratio^{5,32}, considering experimental data across 20 diverse conditions (14 minimal media, including 3 stress conditions; 4 chemostats; and 2 rich media)²⁸. The predicted concentrations of ribosomes, EF-Tu, and EF-Ts increase with growth rate in line with experimental observations (Fig. 3).

Predicted absolute abundances of EF-Tu (Fig. 3a), EF-Ts (Fig. 3b), and mRNA (Supplementary Fig. 2a) account quantitatively for the experimental data^{8,28–31}, with average deviations (geometric mean fold-error) GMFE $\leq 27\%$ for the elongation factors and GMFE = 6% for mRNA. At low growth rates, experimentally observed concentrations of EF-Tu (Fig. 3a) and tRNA (Supplementary Fig. 2b) are higher than predicted. The model only includes charged (aminoacyl-) tRNA concentrations, and it is likely that the unknown fraction of uncharged tRNA explains at least part of this deviation. Overall, the largest deviations between observed concentrations and predictions are seen in the two non-minimal conditions, which also exhibit the fastest growth ($\mu > 1 \text{ h}^{-1}$). A recent analytical study of balanced cellular growth indicates that these deviations may result from the influence of an increased growth-related dilution of cofactors and other intermediate metabolites, a phenomenon not included in our simulations¹².

Active and de-activated ribosome fractions. At low growth rates ($\mu < 0.3 \text{ h}^{-1}$; Fig. 3c), observed ribosome concentrations exceed those predicted from cost minimization, a deviation consistent with a substantial reserve of deactivated ribosomes at low growth rates³². Such deactivated ribosomes may provide fitness benefits in changing environments^{33,34}, but cannot be maximally efficient in a constant environment and thus cannot be predicted by our optimization strategy. To allow a meaningful comparison between predictions and experiment, we thus estimated the experimental concentration of ribosomes actively involved in elongation (Methods). Cost minimization predicts these experimental estimates with high accuracy across the full range of assayed growth rates; observed values deviate from predictions on average by GMFE = 14% (Fig. 3d).

The remaining, non-active ribosome fraction comprises two parts: the deactivated ribosome reserve currently unavailable for translation³², and free, potentially active ribosomes not currently bound to mRNA (see Supplementary Note 2 for the nomenclature on ribosome states). As our model quantifies the abundance of both active and free ribosomes, their subtraction from observed total ribosome concentrations provides an

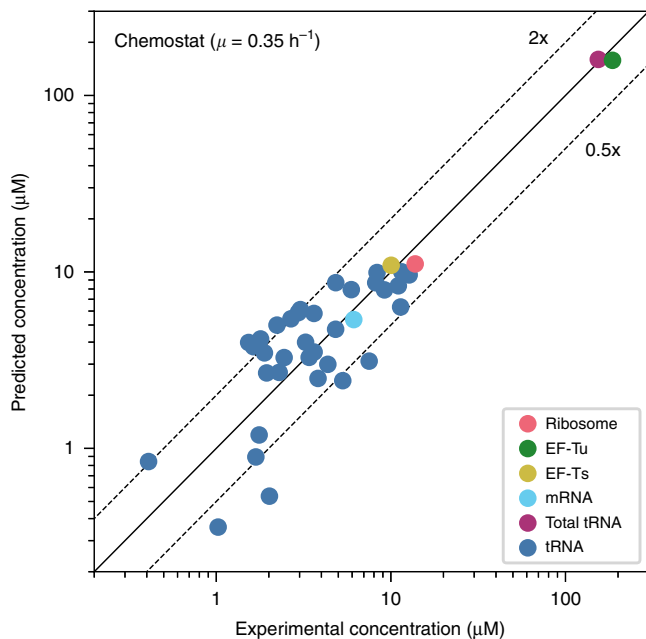


Fig. 2 Predicted vs. observed concentrations in a glucose-limited chemostat. Growth rate $\mu = 0.35 \text{ h}^{-1}$ (for other conditions, see Supplementary Fig. 1). The solid line shows the expected identity, whereas the upper and lower dashed lines show prediction errors of 2x and 0.5x, respectively. Predictions for ribosome, EF-Tu, EF-Ts, mRNA, and total tRNA are highly accurate, with Pearson's $R^2 = 0.99$ and geometric mean fold-error GMFE = 1.13, i.e., predictions based purely on a physico-chemical model and the assumption of cost minimization are on average 13% off. Predictions for individual tRNA species are somewhat less accurate, GMFE = 1.64. Experimentally determined concentrations of the ribosome (averaged over all ribosomal proteins), EF-Tu, and EF-Ts are from ref. ²⁸. mRNA³⁰ and tRNA²⁹ concentrations are interpolated values based on growth rates. Source data are provided as a Source Data file.

estimate of the deactivated ribosome reserve as a function of the growth rate (Fig. 4). While this reserve accounts for less than 20% of total ribosomes at fast to moderate growth, it reaches almost 50% at the lowest growth rate assayed in ref. ²⁸.

RNA/protein ratio and elongation rate. A linear correlation between the RNA/protein ratio and growth rate was discovered in the 1950s^{3,6,7,35} and forms the basis of phenomenological bacterial growth laws^{5,9,32}. Relating the predicted total RNA (ribosomal RNA + tRNA + mRNA) with measured protein concentrations²⁸ indeed results in a near-linear relationship, accurately matching observed values at high to intermediate growth rates ($\mu > 0.3 \text{ h}^{-1}$; Fig. 5a). At lower growth rates, model predictions are slightly too low, likely because of the deactivated ribosome reserve³² (Fig. 4). At low growth rates ($\mu = 0.12 \text{ h}^{-1}$), predictions of RNA and proteins allocated to an optimally efficient translation machinery (including deactivated ribosomes) account for 13% of total dry mass, rising almost linearly to 49% at high growth rates ($\mu = 1.9 \text{ h}^{-1}$; Supplementary Fig. 3).

The concentrations of the individual components of the translation machinery determine the average translation elongation rate (ribosomal activity), defined as the total cellular translation rate divided by the total active ribosome content³⁴. The predicted elongation rates closely match the experimental data³² over a broad range of growth rates (Fig. 5b).

Cost minimization predicts response to antibiotics. The expression of *E. coli*'s translation machinery reacts strongly to the

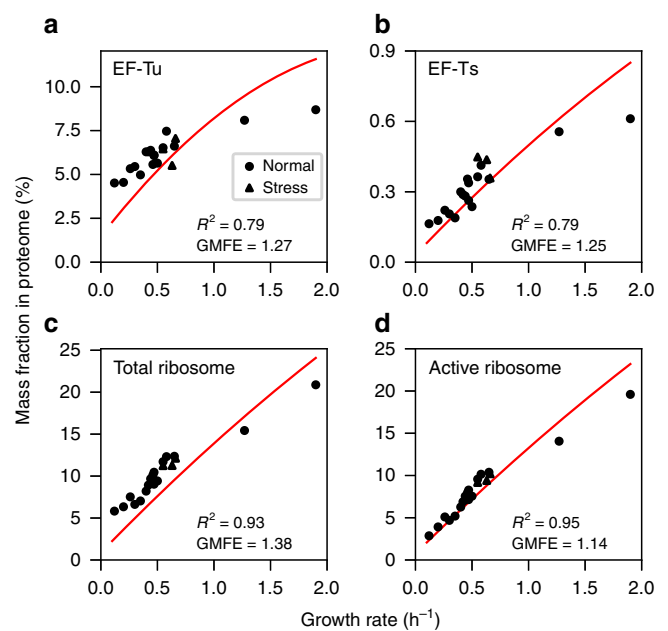


Fig. 3 Growth rate dependence of predicted (red lines) and observed concentrations. **a** EF-Tu, $R^2 = 0.79$, GMFE = 1.27. **b** EF-Ts, $R^2 = 0.79$, GMFE = 1.25. **c** Total ribosome concentration (arithmetic means across ribosomal proteins). **d** Actively elongating ribosomes, estimated from data in **c** according to ref. ³² (see Methods). Circles indicate normal conditions, triangles indicate stress conditions. Source data are provided as a Source Data file.

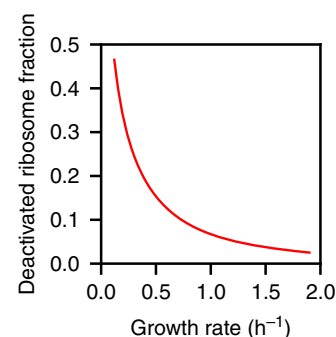


Fig. 4 Estimated fraction of deactivated ribosomes. The deactivated fraction reaches almost 50% for the lowest growth rate assayed in ref. ²⁸ and drops rapidly towards zero at higher growth rates. Source data are provided as a Source Data file.

exposure to antibiotics that inhibit the ribosome, such as chloramphenicol^{5,32,36}. The details of these changes can also be understood from our hypothesis of cost minimization. The concentrations of ribosomes and EF-Tu, the RNA/protein ratio, and the elongation rate of active ribosomes increase under chloramphenicol stress (Supplementary Fig. 4); these changes partially compensate for the reduced fraction of active ribosomes. The concentration of EF-Ts instead decreases with increasing chloramphenicol concentration (Supplementary Fig. 4c). EF-Ts contributes to translation by converting EF-Tu-GDP to EF-Tu-GTP, which then forms a ternary complex with charged tRNA. Under chloramphenicol stress, fewer ternary complexes are turned over, and hence less EF-Ts is needed.

Alternative cost measures lead to similar results. The biological fitness of *E. coli* cells depends on many factors, and hence any

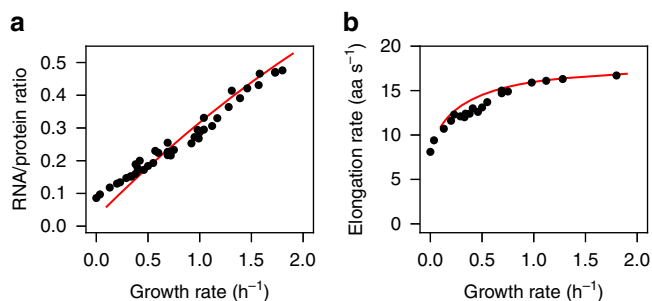


Fig. 5 Growth rate dependences of total RNA/protein ratio and ribosome activity. **a** Predicted total RNA concentration (mRNA + tRNA + rRNA) relative to observed total protein concentration at different cellular growth rates (red line) compared to experimental observations^{5,32}; $R^2 = 0.97$, GMFE = 1.12. **b** Predicted (red line) and experimentally determined³² elongation rates of actively translating ribosomes (ribosome activities); $R^2 = 0.93$, GMFE = 1.06. At the lowest assayed growth rates, non-growth-related translation—which is not included in the model—may become comparable to growth-related translation; at these growth rates, the numerical optimization of our model did not converge ($\mu < 0.1 \text{ h}^{-1}$), and thus the red lines are not extended into this region. Source data are provided as a Source Data file.

simple assignment of fitness costs to molecules can only be approximate. The results presented so far are based on the assumption that costs are proportional to molecular masses. To test if alternative cost measures lead to consistent results, we repeated our calculations using four distinct costs that have been employed in the literature.

Across most conditions, we obtained very similar predictions for the concentrations of translation machinery components when our model assigned molecular costs based on the carbon content of the molecules²¹, on the amount of energy^{22,23} spent on their production (ATP cost), or on the total investment into macromolecular catalysts²⁴ required for their production (synthesis cost) (Supplementary Figs. 5 and 6; we estimated ATP costs based on refs. ^{37,38}, and calculated synthesis costs using flux balance analysis with molecular crowding^{39–41}, see Methods). All components whose concentrations we predict consist of protein, RNA, or both, and all costs examined are approximately proportional to the lengths of RNA and protein molecules. Thus, the relative costs of all components are essentially a function of the RNA/protein cost ratio r , i.e., the cost of RNA per nucleotide divided by the cost of protein per amino acid. We assume that the cost of RNA per nucleotide is identical for tRNA and rRNA; the corresponding cost ratio is broadly similar between molecular masses ($r = 3.0$), carbon content ($r = 2.0$), ATP cost ($r = 1.6$), and synthesis costs ($r = 1.7–2.1$) across minimal growth conditions ($\mu < 1 \text{ h}^{-1}$; Supplementary Fig. 7). In contrast, assuming that costs are proportional to only the protein content of the molecular assemblies^{9,20} results in an RNA/protein cost ratio of zero. Predictions based on protein costs hence overestimate mRNA and tRNA concentrations (which cost nothing), resulting in corresponding underestimates of EF-Ts and especially EF-Tu concentrations (Supplementary Figs. 5, 6).

We note that in rich medium ($\mu = 1.9 \text{ h}^{-1}$), the RNA/protein cost ratio for synthesis costs is much lower than across minimal media, falling to $r = 0.21$ for tRNA and rRNA (Supplementary Fig. 7). This results in an overprediction of the observed tRNA concentration^{8,29,31} by a factor of almost 2 (Supplementary Fig. 6). Moreover, while the predicted tRNA concentration is also almost twice the predicted EF-Tu concentration, experimental estimates for tRNA and EF-Tu are very similar^{8,28,29,31,42}. We conclude that if the translation machinery has been optimized for

efficiency at high growth rates by natural selection, the synthesis cost of its components is unlikely to have been central to this optimization.

In sum, cost minimization in a mechanistic bottom-up model of optimal translation efficiency, fully parameterized with known kinetic constants and molecular masses, accounts quantitatively for the concentrations of all molecule species involved without any adjustable parameters. The optimal concentrations of different components change differentially with growth rate, explaining the observed scaling of *E. coli*'s translation machinery composition, RNA composition, and elongation rate. At least for the translation machinery part of the cellular economy, whose components consist largely of protein and RNA, approximate cost measures appear to be sufficient: several alternative cost measures provided predictions very similar to those based on molecular masses, emphasizing the generality of our findings.

We conclude that *E. coli*'s translation machinery works close to optimal cost efficiency. Accordingly, our findings are consistent with the hypothesis that natural selection has minimized a cost function similar to those examined here. Our results further support the idea that phenomenological growth laws of proteome composition^{5,9,32,36} may have their root in the costs associated with the non-protein molecules involved in particular processes, and that their explicit inclusion in systems biology models of cellular growth^{9,25,43,44} may eventually allow these models to abandon any reliance on phenomenological parameters.

Methods

Experimental concentrations of ribosomes, EF-Tu, and EF-Ts. We used molar concentrations (μM) in the model; thus, all experimental data were converted to molar concentrations. We first calculated the total protein density during exponential growth on a glucose minimal medium at growth rate $\mu = 0.58 \text{ h}^{-1}$. In this condition, the total protein mass per cell is 280 fg (Supplementary Note 3 in ref. ²⁸), and cell volume is 1.90 fL (the cell volume 2.84 fL modified by a factor of 0.67 according to Supplementary Note 3 in ref. ²⁸). Accordingly, the protein mass density on glucose is $\rho_{p,\text{glc}} = 147.15 \text{ g L}^{-1}$.

We then fitted a second-order polynomial function $\phi(\mu)$ to the fraction of total protein in dry mass provided in ref. ¹ across different growth rates μ . With $\phi(\mu)$, $\rho_{p,\text{glc}} = 147.15 \text{ g L}^{-1}$ at $\mu = 0.58 \text{ h}^{-1}$, and the observed constant dry mass density of *E. coli* across growth conditions^{14,15}, we obtained the condition-specific total protein concentration, ρ_p , for all other growth conditions based on the respective observed growth rates.

For a given growth rate, μ' , the total protein concentration $\rho_{p,\mu = \mu'}$ is given by

$$\rho_{p,\mu = \mu'} = \frac{\phi(\mu = \mu') \cdot \rho_{p,\text{glc}}}{\phi(\mu = 0.58)} \quad (1)$$

With the measured fraction of each protein ($f_{p,i}$) in the proteome²⁸, the molar concentration of each protein was then calculated as

$$c_i = \frac{\rho_p \cdot f_{p,i}}{MW_{p,i}}, \quad (2)$$

where $MW_{p,i}$ is the molecular weight of protein i in $\text{g}\cdot\text{mole}^{-1}$. The ribosome concentration was calculated as the arithmetic mean of the molar concentrations of all ribosomal proteins.

Experimental concentration of active ribosome. Active ribosomes are defined here as ribosomes engaged in peptide elongation. Dai et al. estimated the fraction of active ribosomes, f_{active} , in *E. coli* at different growth rates³². We fitted a Michaelis-Menten type equation to their data, resulting in $f_{\text{active}} = \mu/(0.124 + \mu)$. For each total ribosome concentration c_{ribosome} in ref. ²⁸, we then estimated the corresponding active ribosome concentration as $c_{\text{active-ribosome}} = f_{\text{active}} \cdot c_{\text{ribosome}}$.

Experimental concentrations of GTP and GDP. GTP and GDP concentrations are from ref. ⁴⁵. We chose the data for growth on glucose for all simulations (See Supplementary Note 3 for details).

Experimental concentration of mRNA. mRNA concentration was calculated from the data given in ref. ³⁰ as the ratio of mRNA copy number per cell and cell volume. To estimate the mRNA concentrations at the growth rates shown in Fig. 2 and Supplementary Fig. 1, we fitted a second order polynomial to the mRNA concentration as a function of growth rate; we then read off the values at the required growth rates. mRNA concentrations were assayed only at growth rates between

0.11 h⁻¹ and 0.49 h⁻¹, and we did not attempt to extrapolate values beyond this range.

Experimental concentration of tRNA. We collected three independent datasets of tRNA concentrations. Dataset 1²⁹ contains tRNA concentration for each individual tRNA, whereas both dataset 2⁸ and dataset 3³¹ contain only total tRNA concentrations. In each of these experiments, tRNA abundance was measured as the ratio of tRNA to ribosomal RNA (rRNA). We scaled these values to absolute tRNA concentrations assuming that the rRNA concentration corresponds to the ribosome concentration estimated from the proteomics data (see the subsection “Experimental concentrations of ribosomes, EF-Tu, and EF-Ts”). To estimate the tRNA concentrations at the growth rates shown in Fig. 2 and Supplementary Fig. 1, we used the same fitting procedure as for mRNA.

Concentrations of individual tRNAs and relationship to model. Our model differentiates tRNAs by their anticodons (see below for details). Thus, 40 tRNAs were used to represent all elongator tRNAs. The tRNAs modeled in this work are listed in Supplementary Table 1 together with their common names used in dataset 1²⁹ and their gene IDs.

In the experiments by Dong et al. (dataset 1)²⁹, tRNAs were classified into 41 distinct sets based on two-dimensional polyacrylamide gel electrophoresis. We combined two tRNA sets corresponding to different tRNA weights if they have the same anticodon (i.e., the pairs of Val2A + Val2B, Thr1 + Thr3, and Tyr1 + Tyr2). The experimenters could not distinguish between the tRNAs Gly1 and Gly2, as these have very similar molecular weights and isoelectric point; the same was true for Ile1 and Ile2. We estimated the individual concentrations of these four tRNAs based on the ratios 3:2 between Gly1 and Gly2 and 20:1 between Ile1 and Ile2 observed by Ikemura et al.⁴⁶.

To estimate the tRNA concentrations at the growth rates shown in Fig. 2 and Supplementary Fig. 1, we fitted the concentration of each tRNA in dataset 1 (measured for growth rates ranging from 0.28 h⁻¹ to 1.73 h⁻¹) to a second order polynomial of growth rate and extended this function to the required range, 0.12 h⁻¹ to 1.9 h⁻¹.

Combination of tRNAs that are predicted to be non-expressed. The predicted concentrations of 6 tRNAs are 0 μM; these are highlighted in red in Supplementary Table 2. This result is a straightforward consequence of the model structure. The relationship between codons and tRNAs is not one-to-one in *E. coli*. Consider a given codon (codon1) that has more than one cognate tRNA, say, tRNA1 and tRNA2. If tRNA2 is also the cognate tRNA of another codon (codon2), the predicted concentration of tRNA1 will be zero: for the same “price” (the same contribution to the limited total mass concentration), tRNA2 can service two codons, while tRNA1 can service only one. For example, codon GGG has two cognate tRNAs, gly1 and gly2; gly2 is also the cognate tRNA of codon GGA. Thus, both gly1 and gly2 can translate GGG, but gly2 can translate GGA, too, and is thus more valuable to the cell if we assume that both tRNAs are processed equally efficiently, as done in the model. Therefore, the predicted optimal concentration of gly1 will be zero (note that this might not occur in models that consider different ribosomal k_{cat} or K_m values for the two tRNAs). To compare our predictions to the experimental data²⁹, we combined tRNAs with predicted zero concentration with their co-functioning tRNAs in both the predictions and the experimental data. The resulting six combined tRNA pairs are: GLY1 + Gly2; Leu1 + Leu3; Leu4 + Leu5; Pro1 + Pro3; Ser2 + Ser1; Thr2 + Thr4. In all reported figures, the total number of tRNAs shown is thus 34.

Concentrations of amino acids and total protein. We first calculated a typical molecular protein mass for each condition. MW_p is the weighted average over the molecular masses of all proteins assayed by Schmidt et al., where the weights are the corresponding proteome fractions in that condition. The molar concentration of protein, $c_{protein}$, is given by

$$c_{protein} = \frac{\rho_p}{MW_p} \quad (3)$$

where ρ_p is the mass concentration of total protein at the given condition (Eq. 1).

The concentration of amino acids encoded by *codon-i* is given by:

$$c_{codon-i} = f_{codon-i} \cdot L_{protein} \cdot c_{protein} \quad (4)$$

where $L_{protein}$ is the abundance-weighted average protein length at the given condition; $f_{codon-i}$ is the frequency of codon *i* in the genome, where each gene is weighted by its relative abundance in the proteome.

Note that for an amino acid AA_j encoded by multiple synonymous codons, $c_{codon-i}$ is not the total concentration of AA_j in cellular proteins, but only of the fraction encoded by *codon-i*; the total concentration of AA_j is obtained by summing the $c_{codon-i}$ values for all synonymous codons for AA_j .

For simulations under chloramphenicol stress, $c_{protein}$ and $c_{codon-i}$ are not available. We approximated their values by the corresponding concentrations for growth on glucose in the absence of the antibiotic.

Mass fraction of translation machinery in total dry weight. The dry mass fraction of the translation machinery shown in Supplementary Fig. 3 includes ribosome, mRNA, charged tRNAs, EF-Tu, and EF-Ts; it does not include GDP, GTP, free tRNA, tRNA-synthetases, and elongation factor G (FusA). We converted from protein fractions to mass fractions of total dry weight using the relationship between total protein mass and dry mass discussed in the subsection “Experimental concentrations of ribosomes, EF-Tu, and EF-Ts”.

Experimental estimate: The mass fraction of the translation machinery in total dry weight is the sum of two parts: (1) protein and (2) RNA.

(1) We calculated the mass fraction of translational proteins (including ribosomal protein, EF-Tu, and EF-Ts) in dry weight from the proteomics data in ref. 28.

(2) We fitted the reported total RNA/protein ratio in refs. 5,32 to a second order polynomial of growth rate. We then used this fitted function to calculate the RNA/protein ratio at the growth rates assayed by Schmidt et al.²⁸, and multiplied this ratio with the dry mass fraction of protein.

Theoretical prediction: The predicted dry mass fraction of the translation machinery was the ratio of the total mass concentration of the translation machinery (including free ribosome, active ribosome, EF-Tu, EF-Ts, charged tRNA, and mRNA) to the total dry mass density. The total dry mass density was estimated as $\rho_{p,glc}/\phi(0.58 \text{ h}^{-1}) = 147.15 \text{ g L}^{-1}/0.631 = 233.30 \text{ g L}^{-1}$. For the prediction including de-activated ribosome concentrations (dashed line in Supplementary Fig. 3), we added estimates of de-activated ribosome concentrations according to Fig. 4 of the main text.

Molecular weights. Molecular weights of ribosome, charged tRNAs (aa-tRNAs), EF-Tu, EF-Ts, and the ternary complexes (TC, EF-Tu-GTP-aa-tRNA) were calculated from their sequences. The stoichiometry of ribosomal proteins and RNAs in the ribosome was obtained from the EcoCyc database⁴⁷; the stoichiometry of all components is 1 except for RplL, for which it is 4.

We used an average mRNA to represent the total mRNA. The molecular weight of an average mRNA (MW_{mRNA}) is the sum of two parts: (1) the molecular weight of the coding sequence (CDS) of mRNA ($MW_{mRNA-CDS}$), which was calculated from protein-expression-weighted mRNA length and nucleotide composition of *E. coli* protein-coding sequences in each growth condition²⁸; (2) the weight of the untranslated region (UTR) of mRNA ($MW_{mRNA-UTR}$), which was calculated from the average nucleotide composition of the genome and the typical length of UTR. The length of UTR was assumed to be 85 nt, a typical length of the untranslated region in *E. coli*⁴⁸. Thus,

$$MW_{mRNA} = MW_{mRNA-CDS} + MW_{mRNA-UTR} \quad (5)$$

In the simulations of translation under antibiotic stress, the molecular weight of chloramphenicol was set to 0.

Alternative costs of translation machinery components. The hypothesis underlying our analysis is that the components of the translation machinery are expressed to minimize the total cost of translation at a given protein production rate. For the calculations underlying the figures of the main text, we assumed that molecular costs are proportional to molecular mass concentrations. To test alternative cost measures proposed in the literature, we estimated the costs of each translation machinery component in terms of (1) its carbon content²¹ (carbon cost); (2) the total number of high-energy phosphate bonds required for its production (ATP cost); (3) the total enzyme mass required for its production (synthesis cost)²⁴; and (4) its protein content^{9,20} (protein content). Some of these cost measures are condition-dependent. The estimated costs are provided in Supplementary Data 2; the RNA/protein cost ratios are compared in Supplementary Fig. 7 and listed in Supplementary Table 3. All costs were estimated per component (i.e., per molecule or per macromolecular complex). Note that all components whose costs are considered in the model consist of protein, RNA, or both.

Carbon cost as an alternative cost measure. The carbon cost of a component is its total number of carbon atoms.

ATP cost as an alternative cost measure. The ATP cost of a component is the number of high-energy phosphate bonds (denoted ~P) that were invested into its production. The ATP costs includes (1) the ATP invested into the synthesis of the precursors (nucleoside triphosphates or amino acids) and (2) the ATP cost of polymerization during RNA transcription or protein translation.

The ATP cost of amino acid synthesis: For cells growing on minimal carbon media, the ATP cost of amino acid production was obtained from ref. 37. Since the ATP costs of a given amino acid are very similar across minimal media with different carbon sources³⁷, we used the ATP costs for amino acid synthesis on glucose for all minimal media considered. For *E. coli* growing on glycerol + amino acids²⁸, the ATP production cost for amino acids was assumed to be zero.

The polymerization cost of protein: The polymerization cost of protein is 4 ATP per amino acid: two ATP for tRNA charging, 1 ATP for EF-Tu in elongation, and 1 ATP for elongation factor G (EF-G) in elongation.

The ATP cost of NTP synthesis: For cells growing on minimal carbon media, the de novo synthesis cost of NTP was obtained from ref. 38. The glycerol + amino acids medium used in the proteomics study²⁸ also contains adenine and uracil; here, we assumed that the synthesis of ATP and GTP starts from adenine and that the synthesis of UTP and CTP starts from uracil. PRPP (5-phospho- α -D-ribose 1-diphosphate), whose production consumes 29 ATP³⁷, was considered to be the donor of ribose to the synthesis of NTPs. Finally, we estimated the total energy (\sim P) costs in the glycerol + amino acids medium for ATP, UTP, CTP, and GTP as 31, 31, 29, and 32, respectively. We did not attempt to estimate the ATP cost of NTP synthesis in the LB condition, as it is not clear to what extent NTPs are taken up from the medium.

The polymerization cost of RNA: The polymerization cost of RNA is 0 ATP per base, as no high-energy phosphate bonds beyond those of the polymerized NTPs are required.

The degradation cost of mRNA: The degradation rates of tRNA and rRNA are much lower than their production rates, and hence we did not account for their degradation. In contrast, mRNA is degraded much more quickly than tRNA and rRNA, and we thus considered the influence of degradation on the mRNA polymerization cost. At steady state, all degraded mRNA (in the form of nucleoside mono-phosphates, NMPs) is assumed to be recycled to re-transcribe mRNA. These recycled NMPs require two \sim P to form NTPs, so the mRNA recycling cost ($\text{cost}_{\text{NTP-deg}}$) is 2 \sim P per NTP. At a given mRNA concentration $c_{\text{mRNA-NTP}}$ (in units of NTPs built into mRNA), the production of mRNA must offset the combination of mRNA degradation and mRNA dilution by cellular growth at rate μ . The rate of ATP consumption for this production is thus given by

$$v_{\text{cost-ATP-mRNA}} = \mu \cdot \text{cost}_{\text{NTP}} \cdot c_{\text{mRNA-NTP}} + k_{\text{deg}} \cdot \text{cost}_{\text{NTP-deg}} \cdot c_{\text{mRNA-NTP}} \quad (6)$$

where cost_{NTP} is the synthesis cost of NTPs (estimated above, “The ATP cost of NTP synthesis”) and k_{deg} is the mRNA degradation rate constant. k_{deg} is calculated from mRNA half-life (t_{half}), $k_{\text{deg}} = \ln(2)/t_{\text{half}}$ with $t_{\text{half}} = 5$ min for all growth conditions⁴⁹. To obtain the ATP cost per NTP in mRNA, we must divide this rate by $\mu \cdot c_{\text{mRNA-NTP}}$:

$$\text{cost}_{\text{NTP-mRNA}} = \frac{v_{\text{cost-ATP-mRNA}}}{\mu \cdot c_{\text{mRNA-NTP}}} = \text{cost}_{\text{NTP}} + \frac{k_{\text{deg}}}{\mu} \cdot \text{cost}_{\text{NTP-deg}} = \text{cost}_{\text{NTP}} + \frac{2k_{\text{deg}}}{\mu} \quad (7)$$

Thus, mRNA degradation dominates the ATP cost of mRNA at very low growth rates, but becomes insignificant at growth rates higher than the mRNA degradation rate.

Synthesis cost as an alternative cost measure. The synthesis cost is the total macromolecular dry mass, which includes transporters, enzymes, RNA polymerase, and ribosome, that is needed to synthesize each component of the translation machinery. To estimate the synthesis cost of each component, we first estimated the macromolecular dry mass that is needed to synthesize one millimole of amino acid (cost_{AA}) and one millimole of nucleotide (cost_{nuc}). Based on these estimates, the synthesis cost of a protein is $\text{cost}_{\text{protein}} = L_{\text{protein}} \cdot \text{cost}_{\text{AA}}$, where L_{protein} is the protein length in amino acids, and the synthesis cost of an RNA molecule is $\text{cost}_{\text{RNA}} = L_{\text{RNA}} \cdot \text{cost}_{\text{nuc}}$, where L_{RNA} is the length of the RNA molecule in nucleotides.

We calculated cost_{AA} and cost_{nuc} using cFBA^{40,41}, which is an implementation of the MOMENT³⁹ algorithm for flux balance analysis with molecular crowding, featuring an improved treatment of co-functional enzymes. Briefly, cFBA assigns each enzyme a constant catalytic rate (k_{cat}) and molecular weight, and then finds the flux distribution that maximizes biomass production while not exceeding a threshold on the total enzyme mass.

Synthesis cost of protein: We first added a protein synthesis reaction to the iML1515 model implemented in cFBA. The stoichiometric coefficient of each amino acid consumed in this reaction was set to its proportion in the biomass reaction of the iML1515 model. In *E. coli*, there are approx. 9 TCs per ribosome at high growth rates^{29,42}; we thus designated the ribosome plus 9 ternary complexes (TCs) as the “enzyme” of the protein synthesis reaction. We parameterized this “enzyme” with $k_{\text{cat}} = 22 \text{ s}^{-1}$ and molecular weight = 2933.241 kD¹. We then set the objective function to the rate of protein production, v_{AA} , instead of the biomass production rate v_{bio} . We simulated different growth conditions by only allowing the model to import nutrients available in the respective medium, setting the lower bound of the corresponding exchange reactions to $-1000 \text{ (mmol} \cdot \text{g}_{\text{DW}}^{-1} \cdot \text{h}^{-1})$. For cells growing on LB medium, the lower bound of all exchange reactions was set to -1000 , i.e., all metabolites for which there is a transporter in the model can be taken up. We maximized the protein synthesis rate v_{AA} , given a limit on the dry mass fraction of macromolecules involved in metabolism (a “budget”) of $C = 0.27$. As the lower bound chosen for the exchange rates was very high, the optimizations were constrained by C , i.e., the optimal v_{AA} value is the maximal rate of protein production with this macromolecular budget. Accordingly, the synthesis cost of amino acids is $\text{cost}_{\text{AA}} = 0.27/v_{\text{AA}}$, expressed as the dry mass fraction required to produce $1 \text{ mmol} \cdot \text{g}_{\text{DW}}^{-1} \cdot \text{h}^{-1}$ of amino acids. The computed protein synthesis costs for all 20 conditions considered are shown in Supplementary Table 4.

Synthesis cost of stable RNA: To estimate the cost per nucleotide of stable RNA production, cost_{nuc} , we implemented an analogous algorithm. We first added an RNA synthesis reaction to the iML1515 model⁵⁰ and set this reaction as the objective function. The stoichiometry of NTPs consumed in this reaction was set to the corresponding fractions of NTPs in the biomass reaction of the iML1515

model. We designated the RNA polymerase (molecular weight = 389.11 kD) as the enzyme catalyzing this reaction. For stable RNA (tRNA and rRNA) synthesis, the turnover rate of the RNA polymerase (RNA-P)¹ is $k_{\text{RNAP-sRNA}} = 85 \text{ s}^{-1}$. We maximized the flux of the RNA synthesis reaction, v_{sRNA} , constrained by the macromolecular budget $C = 0.27$. The synthesis cost of stable RNA per nucleotide was then calculated as $\text{cost}_{\text{nuc-sRNA}} = 0.27/v_{\text{sRNA}}$, expressed as the dry mass fraction required to produce $1 \text{ mmol} \cdot \text{g}_{\text{DW}}^{-1} \cdot \text{h}^{-1}$ of NTP in stable RNA. The computed stable RNA synthesis costs for all 20 conditions considered are listed in Supplementary Table 4.

Synthesis cost of mRNA: In *E. coli*, mRNA transcription is slower than stable RNA transcription. The turnover rate of RNA-P for mRNA is $k_{\text{RNAP-mRNA}} = 66 \text{ s}^{-1}$ (i.e., the fastest rate of mRNA transcription has been observed to be about three times the maximal translation rate)⁵¹. As for the calculation of the ATP cost of mRNA, we need to account for mRNA degradation at rate k_{deg} (which is much faster than the degradation of stable RNA). We again assumed that nucleotides from degraded mRNA are re-used by RNA-P to synthesize mRNA. As before, we note that at a given mRNA concentration c_{mRNA} , the production of mRNA must offset the combination of mRNA degradation and mRNA dilution by cellular growth at rate μ . The rate of mRNA production by the RNA-P is thus

$$v_{\text{mRNA}} = \mu \cdot c_{\text{mRNA}} + k_{\text{deg}} \cdot c_{\text{mRNA}} \quad (8)$$

To obtain the concentration of RNA-P necessary to catalyze this rate, we need to divide this expression by $k_{\text{RNAP-mRNA}}$:

$$c_{\text{RNAP-mRNA}} = \frac{v_{\text{mRNA}}}{\mu} = c_{\text{mRNA}} \left(1 + \frac{k_{\text{deg}}}{\mu} \right) \quad (9)$$

Thus, the concentration of RNA-P required for mRNA production is larger by a factor $(1 + k_{\text{deg}}/\mu)$ when accounting for mRNA degradation than it would be otherwise. To account for mRNA degradation, we thus set the effective turnover number of RNA-P to $k_{\text{eff-mRNA}} = k_{\text{RNAP-mRNA}} / (1 + k_{\text{deg}}/\mu)$. We maximized the flux of the mRNA synthesis reaction, v_{mRNA} , constrained by the macromolecular budget $C = 0.27$. The synthesis cost of mRNA per nucleotide was then calculated as $\text{cost}_{\text{nuc-mRNA}} = 0.27/v_{\text{mRNA}}$, expressed as the dry mass fraction required to produce $1 \text{ mmol} \cdot \text{g}_{\text{DW}}^{-1} \cdot \text{h}^{-1}$ of NTP in mRNA. The computed costs for all 20 conditions considered are provided in Supplementary Table 4.

Protein content as an alternative cost measure. The protein content was considered as the only relevant cost of the translation machinery in two previous models^{9,20}. These models accurately predicted the ribosomal protein fraction in total protein, whereas they substantially underestimated the EF-Tu proteome fraction compared to measurements data^{9,20}.

Following these earlier works, we calculated the protein cost of a component as the number of amino acid residues it incorporates. Since mRNA and tRNA have zero protein content, this definition assigns no costs to their expression, potentially leading to a prediction of infinite concentrations. To avoid such pathological predictions, we set the protein content costs of mRNA and charged tRNAs to small, arbitrary values in the model (mRNA: 10 amino acid residues; tRNA: 5 amino acid residues).

Model overview. The mechanistic translation model encompasses the processes of translation initiation, elongation, termination, nucleotide exchange in EF-Tu, and ternary complex (TC) formation. Figure 1 of the main text illustrates the modeled reaction; for better readability, the figure shows only a subset of codon/charged tRNA combinations. In total, the model includes 274 reactions.

Translation initiation. During initiation, mRNA converts free ribosomes to active ribosomes:



Translation initiation consists of multiple elementary reactions⁵². However, a recent study found that at steady state, the kinetics of initiation effectively follow Michaelis–Menten kinetics, with mRNA in the enzyme position (with concentration c_{mRNA}) and free (unbound) ribosomes in the substrate position (with concentration $c_{\text{ribo-free}}$)⁵³,

$$v_{\text{il-init}} = k_{\text{cat-mRNA}} \cdot c_{\text{mRNA}} \cdot \frac{c_{\text{ribo-free}}}{K_{\text{M-ribo}} + c_{\text{ribo-free}}} \quad (10)$$

with $k_{\text{cat-mRNA}} = 1.33 \text{ s}^{-1}$ and $K_{\text{M-ribo}} = 8.5 \mu\text{M}$ from ref. 25.

We assume that the turnover number of mRNA for ribosome binding ($k_{\text{cat-mRNA}}$) is growth rate-independent, and hence that the observed growth rate-dependent activity is due to changes in the concentration of free ribosomes available for initiation⁵⁴. Thus, we used the maximal reported mRNA activity as an estimate of $k_{\text{cat-mRNA}}$.

Ternary complex formation and EF-Tu nucleotide exchange. Ternary complex formation and nucleotide exchange in EF-Tu are the processes by which the translation machinery recycles its substrates, the ternary complexes, for elongation.

TC formation: The binding of EF-Tu-GTP to charged-tRNA (aa-tRNA) forms the ternary complex (ternary complex formation in Fig. 1), which is the substrate of translation elongation.

Nucleotide exchange in EF-Tu. EF-Tu-GDP is released after the formation of a new peptide bond. Elongation factor Ts (EF-Ts) binds to EF-Tu-GDP and induces the exchange of GDP for GTP (right panel in Fig. 1).

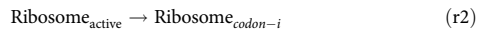
The individual steps of these two processes are modeled with mass action kinetics^{26,27}.

In the implementation of the model, we divide each reversible reaction into two irreversible reactions. The TC formation reaction is a set of reactions that include the binding of 40 aa-tRNAs to Tu-GTP, and its rate constants depend on which amino acid is involved²⁷ (for the parameter values see Supplementary Data 1).

Elongation. Elongation is a very complex process^{55,56}. For simplicity, we model elongation as a single reaction with an active ribosome as the enzyme and TC as the substrate, which was proposed by Klumpp et al.⁹. In this single reaction model, TC is discriminated by the anticodon and all TCs are treated with the same activity. Michaelis–Menten kinetics are used to describe the reaction rate⁹. There are 40 anticodons in total for all elongator tRNAs in *E. coli*; accordingly, our model uses 40 tRNAs to represent all tRNAs.

At steady state, the total translation rate (per cytosolic volume) of each codon remains constant. We do not model the translation of a whole protein. Instead, we decompose protein synthesis into the translation of 61 codons (see also “Modeling” below).

Active ribosome state: We distinguish active ribosomes according to their binding codons (codon in ribosome A site). Thus, there are 61 types of active ribosome in the model, distinguished in the model by subscripts indicating the codon currently presented by the ribosome:



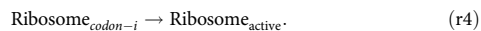
here, *codon-*i** is one of the 61 codons and $\text{Ribosome}_{\text{codon-}i}$ is the active ribosome that binds *codon-*i**. This reaction is constrained by mass balance, but is considered to be instantaneous. We assume that when an active ribosome binds with a specific codon, it only translates the codon’s cognate tRNA.

In the model, there are 61 codons and 40 tRNAs, and the relation between tRNA and codon is not one-to-one. Based on the number of cognate tRNAs, we partition the 61 codons into 2 classes: class 1 codons have one cognate tRNA, whereas class 2 codons have two cognate tRNAs. The lists of class 1 and class 2 codons are provided in Supplementary Table 5.

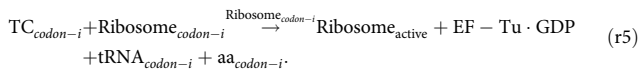
For class 1 codons ($n = 51$), the elongation reaction is:



where $\text{TC}_{\text{codon-}i}$ is the cognate TC of *codon-*i**, $\text{tRNA}_{\text{codon-}i}$ is the released free tRNA, and $\text{aa}_{\text{codon-}i}$ symbolizes the amino acid that was just appended to the growing peptide. The $\text{tRNA}_{\text{codon-}i}$ and $\text{aa}_{\text{codon-}i}$ are included here for completeness, but are not included explicitly in the optimized model, as they do not influence the results once appropriate exchange reactions have been added. Simultaneously, the $\text{Ribosome}_{\text{codon-}i}$ is converted to $\text{Ribosome}_{\text{active}}$, which is ready to participate in the next round of elongation,



For simplicity, we combine r3 and r4 into the new reaction r5, with $\text{Ribosome}_{\text{codon-}i}$ as a substrate and $\text{Ribosome}_{\text{active}}$ as a product (the same for r6 and r7),

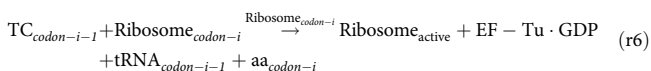


The translation rate of *codon-*i** is described by Michaelis–Menten kinetics,

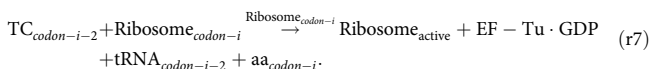
$$v_{\text{tl-codon-}i} = c_{\text{ribo-codon-}i} \cdot k_{\text{cat-ribo}} \cdot \frac{c_{\text{TC-codon-}i}}{c_{\text{TC-codon-}i} + c_{\text{TC-codon-}i-2} + K_{\text{M-TC}}}, \quad (\text{11})$$

where $c_{\text{ribo-codon-}i}$ is the concentration of ribosomes that present *codon-*i** ($\text{Ribosome}_{\text{codon-}i}$), $c_{\text{TC-codon-}i}$ is the concentration of cognate TC of *codon-*i** ($\text{TC}_{\text{codon-}i}$), $k_{\text{cat-ribo}} = 22 \text{ s}^{-1}$, and $K_{\text{M-TC}} = 3 \mu\text{M}$ (parameters from ref. ⁹).

For class 2 codons ($n = 10$), the active ribosome can translate two TCs and thus there are two reactions:



and



The translation rate of *codon-*i** is the sum of these two reactions:

$$v_{\text{tl-codon-}i} = v_1 + v_2, \quad (\text{12})$$

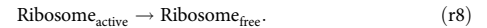
with

$$v_1 = c_{\text{ribo-codon-}i} \cdot k_{\text{cat-ribo}} \cdot \frac{c_{\text{TC-codon-}i-1}}{(c_{\text{TC-codon-}i-1} + c_{\text{TC-codon-}i-2}) + K_{\text{M-TC}}} \quad (\text{13})$$

and

$$v_2 = c_{\text{ribo-codon-}i} \cdot k_{\text{cat-ribo}} \cdot \frac{c_{\text{TC-codon-}i-2}}{(c_{\text{TC-codon-}i-1} + c_{\text{TC-codon-}i-2}) + K_{\text{M-TC}}}. \quad (\text{14})$$

Termination. In termination, an active ribosome is converted to a free ribosome:



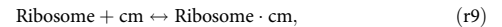
The termination rate is equal to the protein synthesis rate at steady state:

$$v_{\text{term}} = v_{\text{protein-syn}} = \mu \cdot c_{\text{protein}}, \quad (\text{15})$$

where c_{protein} (Eq. 3) is the absolute concentration of protein measured experimentally.

Modeling exchange reactions. Besides the reactions mentioned above, we also add exchange reactions that allow the influx of free ribosome, charged tRNA, mRNA, EF-Tu, EF-Ts, and GTP into the system. We also add exchange reactions that allow efflux of GDP out of the system.

Modeling antibiotic stress. To model chloramphenicol (cm) stress, we add the exchange reaction for chloramphenicol. All forms of ribosome can be inhibited by chloramphenicol:



where Ribosome includes both free and active ribosomes. The reaction rate is described by mass action kinetics with $k_{\text{on}} = 0.00057 \mu\text{M}^{-1} \text{ s}^{-1}$ ($0.034 \mu\text{M}^{-1} \text{ min}^{-1}$) and $k_{\text{off}} = 0.0014 \text{ s}^{-1}$ (0.084 min^{-1})⁵⁷.

For simplicity, we make the following assumptions:

- (1) chloramphenicol diffuses freely across cell membranes, such that the intracellular concentration of free chloramphenicol is the same as that in the medium;
- (2) chloramphenicol bound to an active ribosome (ribosome·cm) causes the ribosome·cm complex to dissociate quickly from the mRNA, thus not affecting further translation of the mRNA.

We assume that the reaction of ribosome and chloramphenicol binding is at steady state (dynamic equilibrium), and thus the active ribosome concentration will be constant. However, not all active ribosomes will be able to successfully finish translation, as the active ribosome can be inhibited by chloramphenicol during translation. Thus, to estimate the production rate of functional proteins, we need to estimate the probability that the active ribosome can finish translation without chloramphenicol inhibition.

To calculate this probability, we use the method proposed in ref. ³². The probability of chloramphenicol binding to a ribosome in a given time unit is:

$$k_{\text{hit}} = k_{\text{on}} c_{\text{cm}}, \quad (\text{16})$$

where k_{on} is the binding constant. The probability that the ribosome is bound n times in the time interval t follows the Poisson distribution

$$P(n) = e^{-k_{\text{hit}} t} \frac{(k_{\text{hit}} t)^n}{n!}, \quad (\text{17})$$

where t_{tl} is the experimental measured translation time of the translated gene. The probability that the ribosome can finish translation without inhibition by chloramphenicol is $P(0)$:

$$P_{\text{tl}} = e^{-k_{\text{hit}} t_{\text{tl}}}. \quad (\text{18})$$

Here, t_{tl} is the time for *LacZ* and $t_{\text{tl}} = 72 \text{ s}$ (from ref. ³²). For simplicity, we assume that all codon-presenting active ribosomes (61 forms of active ribosome) have the same P_{tl} , and thus the effective concentration of the ribosome presenting *codon-*i** is

$$c_{\text{ribo-eff-}i} = P_{\text{tl}} c_{\text{ribo-}i}. \quad (\text{19})$$

Under inhibition by chloramphenicol, this effective ribosome concentration of *codon-*i** ($c_{\text{ribo-eff-}i}$) replaces the ribosome concentration of ($c_{\text{ribo-}i}$) in the model.

Model optimization. We assume translation at steady state and use a constraint-based optimization model. The constraints are given by the above equations and by the requirement of a given total cellular rate of protein synthesis (estimated as the product of growth rate and experimental proteome composition at this growth rate). For simplicity, we decomposed protein synthesis into the translation of 61 codons, and so the constraint on protein synthesis rate is implemented as 61 individual equations, each representing the translation rate of one codon. At steady state, the codon translation rate equals the dilution rate of amino acids incorporated at protein positions encoded by *codon-*i**, i.e., the growth rate μ multiplied by the concentration of amino acids coded by *codon-*i** in proteome data:

$$v_{\text{tl-codon-}i} = \mu \cdot c_{\text{codon-}i}, \quad (\text{20})$$

note that if multiple codons encode the same amino acids, then the total concentration of that amino acid in cellular proteins is the sum over the $c_{\text{codon-}i}$ values for the individual codons.

Given these constraints, we minimize the total mass concentration of the modeled translation apparatus, consisting of ribosome, EF-Tu, EF-Ts, mRNA,

GTP, GDP, and charged tRNAs (aa-tRNA),

$$\sum_{m \in C} MW_m \cdot c_m, \quad (21)$$

where the molecule types together form the set C , MW_m is the molecular weight, and c_m is the concentration of molecule type m . We do not minimize the contributions of GTP and GDP, who participate in multiple other cellular processes⁵⁸ and whose concentrations are thus unlikely to be dominated by translation; their concentrations are consequently fixed to experimentally observed values in our simulations (see Supplementary Note 3).

Let R be the set of reactions that together comprise translation. We also consider the dilution of the molecules involved in these reactions due to cellular volume growth at rate μ :

$$S \cdot v(c) - \mu \cdot c = 0, \quad (22)$$

where S is the stoichiometric matrix for the reactions in R , c is a vector of the concentrations c_m , and $v(c)$ is the corresponding vector of reaction rates v_b , with the concentration-dependent kinetics described above.

Thus, we solve the non-linear constrained optimization problem:

$$\min_c \sum_{m \in C} MW_m \cdot c_m, \quad (23)$$

subject to:

$$S \cdot v(c) - \mu \cdot c = 0$$

$$v_{ti-codon-i} = \mu \cdot c_{codon-i} \text{ for } i=1, \dots, 61$$

$$v_{term} = \mu \cdot c_{protein}$$

We formulated the optimization problem in GAMS and used the BARON global solver⁵⁹ on NEOS Server⁶⁰ with 8 h as the time limit to solve this problem. Because the problem is non-linear, it is not clear a priori if it is convex, in which case only a single optimum would exist. The problem is conceptually similar to the one studied by Noor et al.²⁴, and thus convexity is conceivable. A first optimum was returned by the solver within at most a few minutes for all optimizations performed. However, to guard against the existence of overlooked alternative optima, we allowed the search to continue for a total of 8 h in each case. In addition, we repeated all simulation with the global non-linear solver in Lingo 13 (LINDO Systems, Inc., <https://www.lindo.com/index.php/products/lingo-and-optimization-modeling>). No alternative optima were ever found; we thus have no evidence of non-convexity.

To assess the alternative costs, the molecular weight of each component was replaced by the corresponding cost measure (see the section “Alternative costs of translation machinery components”). The costs for each component are shown in Supplementary Data 2, and the corresponding predictions are shown in Supplementary Data 3.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The experimental data that support the findings of this study are available from the original publications^{5,8,28–32,36}. Source data are provided with this paper.

Code availability

Model: The optimization problem, including the model and its parameterization, is provided as an SBML file (Supplementary Data 4) and as a GAMS input file (Supplementary Data 5, with protein production requirements set to those for growth on minimal glucose medium). In addition, the model has been submitted to Biomodels (MODEL2006210001). GAMS (General Algebraic Modeling System, <https://www.gams.com/>) is a modeling platform for mathematical optimization, it provides dozens of solvers for a variety of optimization problems. The model is written in GAMS (version 25.1.2). BARON (Branch-And-Reduce Optimization Navigator) is a computational system for solving nonconvex optimization problems to global optimality. NEOS Server is a free online server that provides state-of-the-art solvers for numerical optimization. We used the BARON solver (version 19.12.7) on NEOS Server (Version 5.0) for our optimizations. sybilccFBA is a package in R that provides an improved implementation of MOMENT (Metabolic Modeling with ENzyme kineTics). We used version 3.0.1. Lingo (<https://www.lindo.com/index.php/products/lingo-and-optimization-modeling>) is a commercial optimization software. Lingo version 13 was used as the alternative solver to double check the results from BARON.

Received: 26 December 2019; Accepted: 16 September 2020;

Published online: 16 October 2020

References

- Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* 1553–1569 (ASM Press, Washington, DC, 1996).
- Russell, J. B. & Cook, G. M. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol. Rev.* **59**, 48–62 (1995).
- Schaechter, M., Maaløe, O. & Kjeldgaard, N. O. Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* **19**, 592–606 (1958).
- Koch, A. L. Why can't a cell grow infinitely fast? *Can. J. Microbiol.* **34**, 421–426 (1988).
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
- Neidhardt, F. C. & Magasanik, B. Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim. Biophys. Acta* **42**, 99–116 (1960).
- Maaløe, O. Regulation of the protein-synthesizing machinery—ribosomes, tRNA, factors, and so on. In *Biological Regulation and Development* 487–542 (Springer US, Boston, MA, 1979).
- Forchhammer, J. & Lindahl, L. Growth rate of polypeptide chains as a function of the cell growth rate in a mutant of *Escherichia coli* 15. *J. Mol. Biol.* **55**, 563–568 (1971).
- Klumpp, S., Scott, M., Pedersen, S. & Hwa, T. Molecular crowding limits translation and cell growth. *Proc. Natl Acad. Sci. USA* **110**, 16754–16759 (2013).
- Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747 (2014).
- Molenaar, D., van Berlo, R., de Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).
- Dourado, H. & Lercher, M. J. An analytical theory of balanced cellular growth. *Nat. Commun.* **11**, 1226 (2020).
- Nilsson, A., Nielsen, J. & Palsson, B. O. Metabolic models of protein allocation call for the kinetome. *Cell Syst.* **5**, 538–541 (2017).
- Nanninga, N. & Woldringh, C. Cell growth, genome duplication and cell division in *Escherichia coli*. In *Molecular Cytology of Escherichia coli* (ed. Nanninga, N.) 259–318 (Academic Press, 1985).
- Kubitschek, H. E., Baldwin, W. W., Schroeter, S. J. & Graetzer, R. Independence of buoyant cell density and growth rate in *Escherichia coli*. *J. Bacteriol.* **158**, 296–299 (1984).
- Atkinson, D. E. Limitation of metabolite concentrations and the conservation of solvent capacity in the living cell. *Curr. Top. Cell. Regul.* **1**, 29–43 (1969).
- Beg, Q. K. et al. Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA* **104**, 12663–12668 (2007).
- Vazquez, A. Optimal cytoplasmic density and flux balance model under macromolecular crowding effects. *J. Theor. Biol.* **264**, 356–359 (2010).
- Dourado, H., Maurino, V. G. & Lercher, M. J. Enzymes and substrates are balanced at minimal combined mass concentration in vivo. Preprint at, <https://doi.org/10.1101/128009> (2017).
- Ehrenberg, M. & Kurland, C. G. Costs of accuracy determined by a maximal growth rate constraint. *Q. Rev. Biophys.* **17**, 45–82 (1984).
- Beck, A. E., Hunt, K. A., Bernstein, H. C. & Carlson, R. P. Interpreting and designing microbial communities for bioprocess applications, from components to interactions to emergent properties. In *Biotechnology for Biofuel Production and Optimization* 407–432 (Elsevier B.V., Amsterdam, 2016).
- Mahmoudabadi, G., Milo, R. & Phillips, R. Energetic cost of building a virus. *Proc. Natl Acad. Sci. USA* **114**, E4324–E4333 (2017).
- Weiß, A. Y., Oyarzún, D. A., Danos, V. & Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl Acad. Sci. USA* **112**, E1038–E1047 (2015).
- Noor, E. et al. The protein cost of metabolic fluxes: prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.* **12**, e1005167 (2016).
- Tadmor, A. D. & Tlusty, T. A coarse-grained biophysical model of *E. coli* and its application to perturbation of the rRNA operon copy number. *PLoS Comput. Biol.* **4**, e1000038 (2008).
- Gromadski, K. B., Wieden, H.-J. & Rodnina, M. V. Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu. *Biochemistry* **41**, 162–169 (2002).
- Louie, A. & Jurnak, F. Kinetic studies of *Escherichia coli* elongation factor Tu-guanosine 5'-triphosphate-aminoacyl-tRNA complexes. *Biochemistry* **24**, 6433–6439 (1985).
- Schmidt, A. et al. The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
- Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* **260**, 649–663 (1996).
- Valgepea, K., Adamberg, K., Seiman, A. & Vilu, R. *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. Biosyst.* **9**, 2344 (2013).

31. Skjold, A. C., Juarez, H. & Hedgcoth, C. Relationships among deoxyribonucleic acid, ribonucleic acid, and specific transfer ribonucleic acids in *Escherichia coli* 15T- at various growth rates. *J. Bacteriol.* **115**, 177–187 (1973).
32. Dai, X. et al. Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nat. Microbiol.* **2**, 16231 (2016).
33. Mori, M., Schink, S., Erickson, D. W., Gerland, U. & Hwa, T. Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat. Commun.* **8**, 1–8 (2017).
34. Erickson, D. W. et al. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature* **551**, 119–123 (2017).
35. Dennis, P. P. & Bremer, H. Differential rate of ribosomal protein synthesis in *Escherichia coli* B/r. *J. Mol. Biol.* **84**, 407–422 (1974).
36. Hui, S. et al. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.* **11**, 784 (2015).
37. Akashi, H. & Gojobori, T. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA* **99**, 3695–3700 (2002).
38. Lynch, M. & Marinov, G. K. The bioenergetic costs of a gene. *Proc. Natl Acad. Sci. USA* **112**, 15690–15695 (2015).
39. Adadi, R., Volkmer, B., Milo, R., Heinemann, M. & Shlomi, T. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. *PLoS Comput. Biol.* **8**, e1002575 <https://doi.org/10.1371/journal.pcbi.1002575> (2012).
40. Gelius-Dietrich, G., Desouki, A., Fritzemeier, C. & Lercher, M. J. sybil – efficient constraint-based modelling in R. *BMC Syst. Biol.* **7**, 125 (2013).
41. Desouki, A. Algorithms for improving the predictive power of flux balance analysis (Doctoral dissertation) (2016).
42. Furano, A. V. Content of elongation factor Tu in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **72**, 4780–4784 (1975).
43. O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R. & Palsson, B. Ø. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* **9**, 693 (2013).
44. Goelzer, A. et al. Quantitative prediction of genome-wide resource allocation in bacteria. *Metab. Eng.* **32**, 232–243 (2015).
45. Bennett, B. D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599 (2009).
46. Ikemura, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.* **146**, 1–21 (1981).
47. Keseler, I. M. et al. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res.* **45**, D543–D550 (2017).
48. Kim, D. et al. Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet.* **8**, e1002867 (2012).
49. Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S. & Cohen, S. N. 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 (2002).
50. Monk, J. M. et al. iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* **35**, 904–908 (2017).
51. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010).
52. Milon, P. & Rodnina, M. V. Kinetic control of translation initiation in bacteria. *Crit. Rev. Biochem. Mol. Biol.* **47**, 334–348 (2012).
53. Borkowski, O. et al. Translation elicits a growth rate-dependent, genome-wide, differential protein production in *Bacillus subtilis*. *Mol. Syst. Biol.* **12**, 870–870 (2016).
54. Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **42**, 2646–2659 (2014).
55. Rudolf, S., Thommen, M., Rodnina, M. V. & Lipowsky, R. Deducing the kinetics of protein synthesis in vivo from the transition rates measured in vitro. *PLoS Comput. Biol.* **10**, e1003909 <https://doi.org/10.1371/journal.pcbi.1003909> (2014).
56. Vieira, J. P., Racle, J. & Hatzimanikatis, V. Analysis of translation elongation dynamics in the context of an *Escherichia coli* cell. *Biophys. J.* **110**, 2120–2131 (2016).
57. Harvey, R. J. & Koch, A. L. How partially inhibitory concentrations of chloramphenicol affect the growth of *Escherichia coli*. *Antimicrob. Agents Chemother.* **18**, 323–337 (1980).
58. Feist, A. M. et al. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121 (2007).
59. Tawarmalani, M. & Sahinidis, N. V. A polyhedral branch-and-cut approach to global optimization. *Math. Program.* **103**, 225–249 (2005).
60. Czyzyk, J., Mesnier, M. P. & More, J. J. The NEOS Server. *IEEE Comput. Sci. Eng.* **5**, 68–75 (1998).

Acknowledgements

We thank Uri Alon, Matteo Mori, and Terence Hwa for helpful discussions and Abdelmoneim Desouki for providing the sybilccFBA R package. This work was supported by the Volkswagen Foundation under the “Life?” initiative, and by the German Research Foundation (DFG) through grants IRTG 1525, CRC 680, CRC 1310, and, under Germany’s Excellence Strategy, through grant EXC 2048/1 (Project ID: 390686111).

Author contributions

H.D. conceived of the study. X.P.H. developed, implemented, and parameterized the model and performed the analyses. P.S. prepared Fig. 1 and the S.B.M.L. model description. M.J.L. supervised the study. X.P.H. and M.J.L. interpreted the results and wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare that they have no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-18948-x>.

Correspondence and requests for materials should be addressed to M.J.L.

Peer review information *Nature Communications* thanks Wolfram Liebermeister and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020