bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

How total mRNA influences cell growth

Ludovico Calabrese[1*], Luca Ciandrini [2,3*◕], Marco Cosentino Lagomarsino[4,5*◕],

1 IFOM-ETS, The AIRC Institute of Molecular Oncology, via Adamello 16, Milan, Italy
2 Centre de Biologie Structurale (CBS), Univ Montpellier, CNRS, INSERM, Montpellier 34090, France
3 Institut Universitaire de France (IUF)
4 Dipartimento di Fisica, Università degli Studi di Milano, via Celoria 16, Milan, Italy
5 INFN sezione di Milano, via Celoria 16, Milan, Italy

◕These authors contributed equally to this work.
* For correspondence: ludovico.calabrese@ifom.eu (LCa); luca.ciandrini@umontpellier.fr (LCi); marco.cosentino-lagomarsino@ifom.eu (MCL)

Abstract

While the conventional wisdom is that growth rate is prominently set by ribosome amounts, in many biologically relevant situations the levels of mRNA and RNA polymerase can become a bottleneck for growth. Here, we construct a quantitative model of biosynthesis providing testable predictions for these situations. Assuming that RNA polymerases compete for genes and ribosomes for transcripts, the model gives general expressions relating growth rate, mRNA concentrations, ribosome and RNAP levels. On general grounds, the model predicts how the fraction of ribosomes in the proteome depends on total mRNA concentration, and inspects an underexplored regime in which the trade-off between transcript levels and ribosome abundances sets the cellular growth rate. In particular, we show that the model predicts and clarifies three important experimental observations, in budding yeast and *E. coli* bacteria: (i) that the growth-rate cost of unneeded protein expression can be affected by mRNA levels, (ii) that resource optimization leads to decreasing trends in mRNA levels at slow growth, and (iii) that ribosome allocation may increase, stay constant, or decrease, in response to transcription-inhibiting antibiotics.

Introduction

Understanding cell growth is a classic and central question in biology (Neidhardt and Magasanik, 1960; Koch, 1988; Zhu and Thompson, 2019). A striking progress of the recent years is the development of mathematical theories that explain quantitative relationships between global parameters of biosynthesis and growth rate, initially found classically for bacteria (Bremer and Dennis, 2008; Dennis and Bremer, 1974; Schaechter et al., 1958). Such growth laws can be understood as "resource allocation" constraints, as they reflect the presence of a finite pool of resources allocated to the production of different protein classes. For instance, allocating ribosomes to production of one protein class implies withdrawing ribosomes from the production of another class. Crucially, this approach has proved important to generate predictive models, for example, of the global regulation of gene expression across growth conditions (Scott and Hwa, 2011; You et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2013; Qian et al., 2017), the growth response to different classes of translation-targeting antibiotics (Scott et al., 2010; Greulich et al., 2012) and the cost of producing unnecessary proteins (Scott et al., 2010). While most of the recent work in this area has been performed on *E. coli* (You et al., 2013; Dai et al., 2016; Erickson et al., 2017; Scott and Hwa, 2011; Hu et al., 2020; Serbanescu et al., 2020; Belliveau et al., 2021), there are strong indications that different quantitative relationships hold across bacteria, in budding yeast and across lower eukaryotes (Scott et al., 2010; Scott and Hwa, 2011; Maitra and Dill, 2014; Borkowski et al., 2016; Metzl-Raz et al., 2017,?; Kostinski and Reuveni, 2021; Hu and Lercher, 2021). This suggests the presence of strong unifying principles (likely reflecting universal aspects due to resource-allocation trade-offs) in biosynthesis across kingdoms, despite the profound architectural and regulatory differences between organisms (Bruggeman et al., 2020; Kostinski and Reuveni, 2021; Kafri et al., 2016b; Shahrezaei and Marguerat, 2015; Weiße et al., 2015).

A key aspect for sustaining growth is autocatalysis from ribosome self replication, which is also a primary ingredient of growth-laws theories (Roy et al., 2021; Koch, 1988; Kafri et al., 2016b; Kostinski and Reuveni, 2020). More broadly, growth laws originate from the presence of a finite pool of cellular resources needed for biosynthesis. As such, their specific form depends on which factors are limiting for biosynthesis, and indeed many experiments show that in most conditions ribosomes are a major limiting factor for growth. However, there are both physiological and perturbed situations where this ceases to be the case. In particular, a growing body of evidence suggests that in several circumstances transcription, RNA polymerase and mRNA levels become relevant for setting growth rate (Espinosa et al., 2022; Balakrishnan et al., 2022; Neurohr et al., 2019; Kafri et al., 2016b,a; Lin and Amir, 2018; Liang et al., 2000).

Therefore, a critical next question for theory is how to produce theoretical frameworks that evolve the ribosome-centered view, preserving its transparency, while also accounting for the experimental observations that call for a more comprehensive description of biosynthesis, and in particular for the general observation that gene transcription plays a major role in setting the growth rate (Kafri et al., 2016b; Balakrishnan et al., 2022).

Here, our goal is to provide precisely this extended framework. In order to introduce our approach, let us revise the body of the experimental evidence that inspires it, and the previous theories the form the basis of our work. Experimentally, we started from several lines of evidence. The first is the production of unnecessary proteins, which leads to a reduction in growth rate (a "growth cost"). The established view (based on *E. coli* data and ribosome-centric models), is that the cost of unneeded protein expression comes from the mass fraction of the proteome occupied by this unneeded protein (Scott et al., 2010). In *S. cerevisiae*, a study by Kafri and collegues (Kafri et al., 2016a) has shown that in some conditions the growth cost of protein overexpression also comes from a transcriptional burden (but we lack a quantitative framework that captures this cost). Turning to drugs perturbing the global transcription rate in *E. coli*, several studies (Si et al., 2017; Scott et al., 2010; Izard et al., 2015) have shown that the growth rate decreases smoothly under exposure to rifampicin, a well-known antibiotic that targets transcription elongation by binding to RNA polymerase (RNAP). Intriguingly, Scott and coworkers (Scott et al., 2010) also found a non obvious (and yet unexplained) nutrient-dependent ribosome re-allocation under rifampicin treatment, which may suggest that cells possess specific mechanisms in place to deal with limited transcriptional capacity. From a theoretical perspective, two recent studies have included the central dogma into a theory of biosynthesis. Lin and Amir (Lin and Amir, 2018) have used the framework to explore the limits of full saturation, for example when ribosomes fully saturate mRNAs. Roy and coworkers (Roy et al., 2021) defined a more complex

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

model of growth stemming from a universal 'autocatalytic network', which includes several layers of biosynthesis. The model identifies several limiting regimes beyond translation-limited growth, including transcription, but it does not analyze specifically the situation of competition for transcripts. Additionally, both of these frameworks do not consider the problem of optimal allocation of resources, which was shown to underline multiple growth laws (Scott et al., 2014; Chure and Cremer, 2023).

Our framework integrates biosynthesis and the central dogma at both the transcriptional and translational level along the lines of the studies by Lin and Amir and Roy and coworkers (Roy et al., 2021; Lin and Amir, 2018), but also adds a description of growth laws with growth-rate optimization (Scott et al., 2014; Chure and Cremer, 2023). We analyze this model in both growth-optimized and non-optimized conditions. As we will argue in the following, this combination results in a theory able to predict the key experimental findings described above. More widely, it describes the dependency of growth rate on both transcription and translation, through allocation of resources of the biosynthetic machinery, mainly RNAPs and ribosomes. Crucially, we show that mRNA levels can affect growth rate well outside of the saturation regime analyzed by Lin and Amir, in a regime where ribosomes within a finite pool compete for transcripts. In this regime the ribosome-mRNA complex formation limits biosynthesis: ribosome autocatalysis is still the engine of cell growth, but its throttle are mRNA levels.

Results

A minimal biosynthesis model including transcription

Our first step is to propose a general framework (Figure 1A) that links cellular growth physiology with two layers of the central dogma: transcription (for which we will use the suffix TX in our notation) and translation (suffix TL). The equations describing the expression of different classes of genes representing distinct coarse-grained biological functions (indexed by i) are similar to ref. (Lin and Amir, 2018). Each class is populated by g_i gene copies. Transcription synthesizes m_i mRNA copies of each gene class, which are then translated into proteins with abundances P_i. The abundance of transcripts m_i and proteins P_i belonging to gene type i evolves according to the following equations,

$$\frac{dm_i}{dt} = g_i \; J_i^{TX}([N]) - d_i^m \; m_i \tag{1}$$

$$\frac{dP_i}{dt} = m_i \; J_i^{TL}([R]) - d_i^P \; P_i \, , \tag{2}$$

where the fluxes J_i^{TX} and J_i^{TL} are the RNAP flux per gene (corresponding to the amount of mRNA produced per unit time per gene) and the ribosome flux per mRNA (protein produced per unit time per mRNA). The previous equations consider the general case in which transcripts are degraded with rate d_i^m and proteins with rate d_i^p. Regardless of their particular form, the transcriptional and translational fluxes J_i^{TX} and J_i^{TL} are functions of the total RNAPs and ribosome concentrations $[N]$ and $[R]$ respectively. Figure 1B graphically summarises the main elements of the model.

The reservoirs of RNAPs and ribosomes couple the expression of all gene types. Such 'global coupling' is a crucial element of a class of biosynthesis models able to link gene expression and growth physiology (Klumpp et al., 2009; Lin and Amir, 2018; Roy et al., 2021). Note that the expression of proteins constituting RNAPs and ribosomes will also need to obey Eqs.[1] and [2] above. This gives rise to autocatalytic cycles able to sustain exponential growth (Roy et al., 2021).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

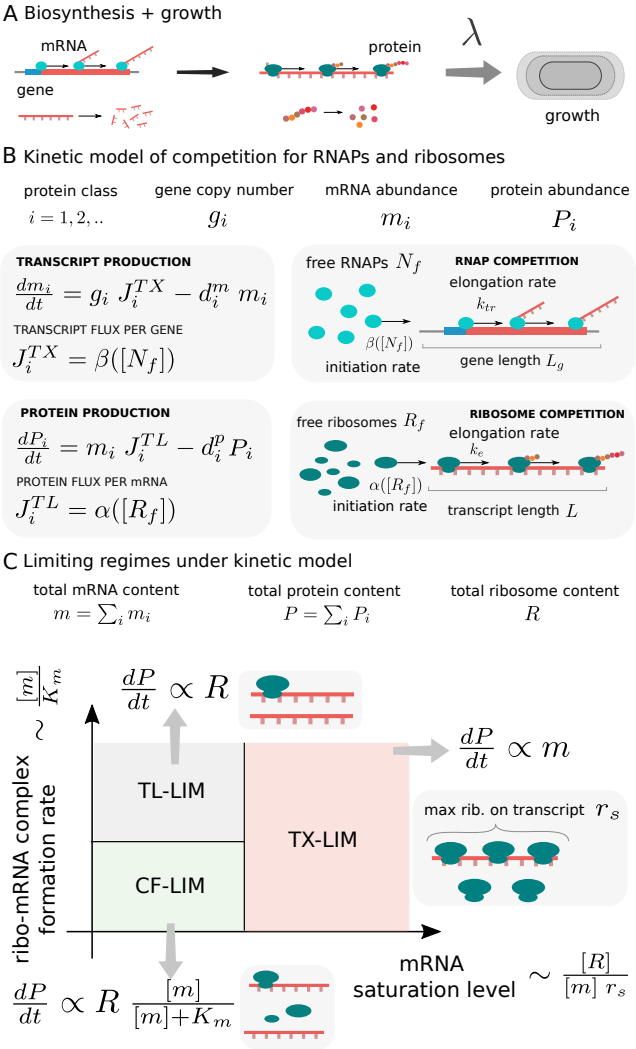

Fig 1. Illustration of the cell growth model including resource allocation and transcription. (A) The model predicts cellular growth rates taking into account translation and transcription. (B) In the model, transcription and translation are treated on the same footing (transcript and protein production boxes). The core of the model is encoded in the transcript flux per gene and the protein flux per gene. These fluxes are determined by the respective initiation and elongation rates, which describe the recruitment of RNAPs and ribosomes. In the model, different gene types compete for recruiting free RNAPs (RNAP competition box) and transcripts compete for free ribosomes (ribosome competition box). Consequently, the synthesis of all proteins is coupled through the availability of RNAPs and ribosomes. (C) Qualitative sketch of the different expected regimes of growth. Our model focuses on the regime in which growth is driven by both ribosome levels and mRNA concentration (green region). In this regime the formation of the ribosome-mRNA complex is the limiting step (complex-formation limiting regime, CF-LIM). Instead, if translation or transcription alone are the limiting step, protein production depends only on either the number of ribosomes (TL-LIM) or the number of mRNAs (TX-LIM), according to the level of mRNA saturation by ribosomes (see grey and red regions).

In typical applications, e.g. for bacteria, mRNA degradation rates are considered

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to be high (Lin and Amir, 2018; Balakrishnan et al., 2022), and protein degradation rates are neglected (Scott and Hwa, 2011). Although these approximations should be taken with caution (Kafri et al., 2016a; Calabrese et al., 2022), they are reasonable in regimes where the time scales of growth and protein degradation are separated (Calabrese et al., 2022). Thus, for the sake of simplicity, in the following we will neglect the role of protein degradation. Instead, mRNA degradation rates are fast (typical half lives of 5-20 minutes, depending on the organism), and will play a crucial role in the following. We will consider homogeneous transcript degradation rates, $d_i^m = d$.

The total amounts of mRNAs and proteins, which will be the main quantities of interest for us, are given by the sum of transcripts and proteins over all the gene types, $m := \sum_i m_i$, and $P := \sum_i P_i$. The model produces steady-state exponential ("balanced") growth (see SI Appendix), where extensive quantities (defined as the quantities that increase linearly with total biomass) all grow exponentially with the same rate.

A suitable parameter to characterise cellular physiology is thus the protein-specific growth rate, defined as

$$\lambda := \frac{1}{P}\frac{dP}{dt} \,. \tag{3}$$

We have singled out conceptually protein production over mRNA production because measurements of total protein content are much more prevalent in the literature. At the same time, the definition above becomes the growth rate of every molecular species at the steady state, including mRNAs (Scott et al., 2010). Therefore, in balanced growth regimes (to which we will restrict ourselves here) such rate still fully characterizes the growth process.

Finally, while we are very aware that organism-specific features are often very important, we will keep our framework organism-agnostic, aiming to describe a general economy of cellular trade-offs (at the level of protein synthesis) with a reduced set of parameters.

mRNA-ribosome complex formation impacts growth rate

To evaluate the impact of the transcription machinery on cellular physiology we have derived a general expression that relates the exponential growth rate λ, the ribosome allocation, and the total mRNA concentration $[m]$, and is valid in a particular regime of "limiting complex formation" (CF-LIM, see below and SI Appendix, sec. 3). Both *a priori* arguments and comparison of model predictions with data lead us to hypothesize that this regime is relevant in many experimental situations. Such an expression takes the particularly transparent form under a few further simplifying assumptions (see Materials and Methods and SI Appendix)

$$\lambda = \gamma \phi_R \, \frac{[m]}{K_m + [m]} \,, \tag{4}$$

where γ is an effective translation rate, which can be interpreted as the inverse of the time needed to assemble the amino acids constituting a ribosome, $\phi_R = P_R/P$ is the ribosomal protein fraction[1], and K_m is an effective mRNA concentration scale set by the ratio of the translation initiation and elongation rates (see below).

Among the simplifying assumptions used to derive Eq. [4] (see Methods and SI Appendix), it is worth noting that we assumed that all genes have identical properties (transcript length and degradation rate, promoter and ribosome binding site strengths,

[1]Note that for simplicity of notation we use ϕ_X to denote proteome *number* fractions, and not mass fractions as some previous literature, but the two quantities can be converted into each other, using the masses of the proteins belonging to each sector.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

etc.). Moreover, we have neglected gene expression regulation at the translational level (so that genes are differentially expressed only through differences at the transcriptional level (Balakrishnan et al., 2022)). It is simple to relax all these assumptions and obtain more precise expressions, at the cost of some transparency and ease of interpretation.

Similarly to previous frameworks (Scott et al., 2010; Scott and Hwa, 2011; Scott et al., 2014; Roy et al., 2021), Eq. [4] linearly relates the growth rate to the ribosomal protein fraction ϕ_R with a proportionality factor associated with the translation capacity γ. However, differently from other frameworks, Eq. [4] also includes a Michaelis-Menten factor $[m]/(K_m + [m])$ reflecting the competition of transcripts for ribosomes. This term comes from the assumption of a regime where the formation of the ribosome-transcript complex is the limiting step of the process (Fig. 1C). The competition between transcripts and ribosomes, and thus the dependence of growth rate on both mRNA and ribosome levels, becomes relevant when $[m] \approx K_m$ (see SI Appendix). In the remainder of this work, we will focus on the 'strong' complex-formation limiting regime (CF-LIM) when $[m] \ll K_m$, but we will also address the 'weak' case, when $[m] \approx K_m$.

The Michaelis-Menten-like factor emerges from a a kinetic model of competition for ribosomes between mRNAs. In brief, we first assume a translation initiation rate α proportional to the concentration $[R_f]$ of free, unbound ribosomes. Mathematically, this can be written as $\alpha = \alpha_0[R_f]$, where α_0 is a rate constant representing the affinity between ribosomes and transcripts. We combined such model of initiation with a model of protein production (Greulich et al., 2012), which allows us to re-write the free ribosome concentration in terms of the total ribosome concentration $[R]$, total mRNA concentration $[m]$ and translation elongation rate k_{tl}. In the regime where the translation flux J^{TL} is proportional to the initiation rate, we obtain equation [4], with $K_m := \frac{k_{tl}}{\alpha_0 L_p}$where L_p is the average protein length.

Our hypothesis (motivated by the experimental evidence on both budding yeast and *E. coli*) is that the CF-LIM regime illustrated in Fig. 1C may be relevant in different situations. The Michaelis-Menten-like factor of Eq. [4] describes the non-neglibile competition of mRNAs for ribosomes. If mRNAs are abundant, they strongly compete for ribosomes, which become the main limiting factor. In this situation the ratio $K_m/[m]$ becomes sufficiently small, and the growth rate λ depends on the ribosomal fraction ϕ_R only. This is the conventional translation-limited regime (TL-LIM in Fig. 1C), in which transcripts are abundant, $[m] \gg K_m$. At the other extreme, if transcript abundance is low and ribosomes are in excess, our framework predicts that protein production is proportional to both the amount of mRNA $[m]$ and ϕ_R (see equation [4]). Thus, when $[m] \ll K_m$, $\lambda \propto \phi_R[m]$ and the formation of the ribosome/transcript complex still limits the protein synthesis process in this regime (CF-LIM). Contrary to the saturated transcript regime explored in previous studies (Lin and Amir, 2018; Roy et al., 2021) where protein production is proportional to only transcript abundances (TX-LIM in Fig. 1C), our framework predicts a new regime of growth driven by mRNA/ribosome complex formation where growth rate depends on both ribosome and transcript concentrations.

A back-of-the-envelope estimate (see Materials and Methods) supports our hypothesis. We get $K_m \approx 0.05 - 0.15\mu M$ for *E. coli* and $K_m \approx 0.25 - 0.5\mu M$ for *S. cerevisiae*. This range of values is roughly an order of magnitude smaller than the typical mRNA concentration at fast growth (Balakrishnan et al., 2022), placing ourself in between the translation limited (TL-LIM) and the complex-limited (CF-LIM) competition regimes. This suggests that small perturbations of mRNA levels do not impact growth, but a ten-fold reduction of mRNA levels compared to fast growth could significantly impact the growth rate. While such perturbations could be considered quite strong, mRNA levels can indeed span an order of magnitude in *E. coli*, and they can reach $[m] \approx K_m$

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

physiologically at slow growth (Balakrishnan et al., 2022). We will see in the following that the CF-LIM regime can explain key experimental trends, further supporting our hypothesis.

All the above arguments lead us to focus (differently from previous studies) in particular on the CF-LIM regime, where the association mRNA/ribosome becomes limiting. For simplicity (and testability of our model), we have considered a theory where ribosome-mRNA binding (competition) is the only determinant of ribosome activity. While a full description of growth limitations beyond the initiation-limited assumption of the translation flux is beyond the scope of this work, SI Appendix sec. 3 provides further arguments regarding the wider phase space of growth-limiting regimes and the relations of our assumptions with other studies. We are also aware that other contributions, including ribosome sequestration or hibernation, degradation, variations in translation rates at slow growth, drugs perturbing transcription and translation etc. can contribute to Eq. [4]. This is particularly expected at slow growth (Dai and Zhu, 2020; Dai et al., 2016; Calabrese et al., 2022), where consequently Eq. [4] is less precise. Some variants of the model are discussed below. The following subsection shows how the relation between fraction of ribosomes in the proteome and growth rate depends on RNAP allocation through a factor that is set by the ratio of the translation-initiation rate to the translation-elongation rate.

mRNA availability couples the translational and transcriptional machineries

Equation [4] shows that our framework relates growth and ribosome allocation taking into account transcript levels. It is possible to probe the interdependence between the translational and transcriptional machineries by decoupling the contributions of ribosomes and RNAP polymerases. Indeed, the model connects mRNA concentration with RNAP-proteins and with key parameters related to mRNA production and degradation (see Materials and Methods for a derivation):

$$[m] = \frac{\gamma_{tx}}{d} f_{bn} \, [P_N] \tag{5}$$

The above expression encodes an intuitive picture of mRNA production: transcript concentration is proportional to the protein fraction of RNAP and to a "transcription capacity" γ_{tx} (which is defined by the ratio between transcription elongation rate k_{tx} and the total length of the genes forming RNAPs $\gamma_{tx} := k_{tx}/L_{\mathrm{N}}$), while it is inversely proportional to the transcript degradation rate d. Note that this expression is equivalent to the equation $[m] = \frac{\gamma_{tx}}{d} f_{bn} \, \frac{L_N}{L_g} \, [N]$ with $[N]$ the RNA polymerase concentration, and the term $\frac{L_N}{L_g}$ is a conversion factor allowing to pass with ease from abundances of proteins of the polymerase class N to the number of polymerases, assuming perfect stochiometry (see Materials and Methods). The parameter f_{bn} represents the fraction of bound RNAPs, i.e., those that are actively transcribing. This term is the analogous of the Michaelis-Menten factor $[m]/(K_m + [m])$ of Eq. [4], which describes the fraction of ribosomes bound to transcripts. However, the expression of f_{bn} is lengthy and not particularly transparent, and we provide it in SI Appendix.

Equivalently, Eq. (5) can also be written as

$$[m] = \Gamma \, [P] \, \phi_N \ ,$$

with $\Gamma := \frac{\gamma_{tx}}{d} f_{bn}$ depending on the transcription efficiency $\gamma_{tx} f_{bn}$ and transcript degradation d, and $[P]$ being the overall protein concentration.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Combining these considerations with Eq. (4) leads to the following expression of the growth rate in terms of protein fractions only,

$$\lambda = \gamma \phi_R \frac{\phi_N}{\phi_N + \epsilon^{-1}}\,, \tag{6}$$

where

$$\epsilon := \frac{\Gamma[P]}{K_m}\,. \tag{7}$$

While Eq. [4] couples growth and mRNA concentration, Eq. [6] unfolds the link between growth and the fraction of RNAP-associated proteins ϕ_N. The parameter ϵ incorporates a purely transcriptional "supply" term Γ, which represents the mRNA concentration per unit of RNA polymerase fraction), and a purely translational "demand" term K_m, discussed above, which corresponds to the typical amount of transcripts needed by translation. Equivalently, ϵ can also be seen as the increase in mRNA concentration (in units of K_m) per units of added RNAP fraction as $\frac{[m]}{K_m} = \epsilon\, \phi_N$. Therefore, this parameter can also be interpreted as the mRNA transcription capacity rescaled by the the amount of needed mRNA. The term ϵ^{-1} represents the proteome fraction of RNAP-associated proteins around which the CF-LIM regime becomes relevant: if ϕ_N is significantly smaller than this quantity, then growth is constrained by the available fraction of RNAPs (TX-LIM). At the opposite situation, if the cell allocates an RNAP fraction $\epsilon\phi_N \gg 1$, then ribosomes become scarce (TL-LIM).

Thus, the ratio ϵ can be interpreted as a "supply-demand trade-off" between transcription and translation. This parameter can be experimentally perturbed in different ways, acting on both transcription (and RNA pools) and translation (and ribosome pools). Here, we will focus on changes in transcriptional supply that vary Γ (as it happens for example for transcription-targeting drugs). Instead, in the following section we will exploit its dependence on the transcripts degradation rates to quantify transcription limitation under protein overexpression of unnecessary proteins.

Competition for transcripts increases the cost of unneeded protein production

This section explores the consequences of transcriptional limitation on the cost of unnecessary protein expression in the competition regime *versus* translation-limited growth (Fig. 2A). Overexpression of unnecessary "burden" proteins that do not contribute to growth helps determining which factors are limiting for growth (Shachrai et al., 2010; Dekel and Alon, 2005; Scott and Hwa, 2011; Scott et al., 2010; Kafri et al., 2016a). Experiments can induce this burden by different methods. For example, Scott *et al.* (Scott et al., 2010) over-expressed β-Galactosidase from an IPTG-inducible gene on a medium-copy plasmid in *E. coli*. Kafri *et al.* (Kafri et al., 2016a) generated a library of *S. cerevisiae* expressing mCherry by integrating multiple gene copies into the genome.

In presence of an unnecessary "type-U" gene, inspired by the experimental approach of ref. (Kafri et al., 2016a), we considered two ways of modulating protein overexpression (Fig. 2B): (i) by varying the copy number of the gene g_U within the cell and (ii) by changing the degradation rate d_U of the corresponding transcript. The former perturbation globally tunes the growth cost of protein overexpression. Increasing the copies of the genes g_U always increases the growth cost, regardless of which factors are limiting. Instead, tuning the degradation rate modifies the relative contribution of transcription and translation burden to the overall growth cost. To understand how this works, we can consider the limit of "infinitely fast" unnecessary-mRNA degradation. In this case, there are no unnecessary transcripts, and the presence of unnecessary genes does not affect at all

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ribosome allocation, i.e., there is no burden on translation. In general, increasing the

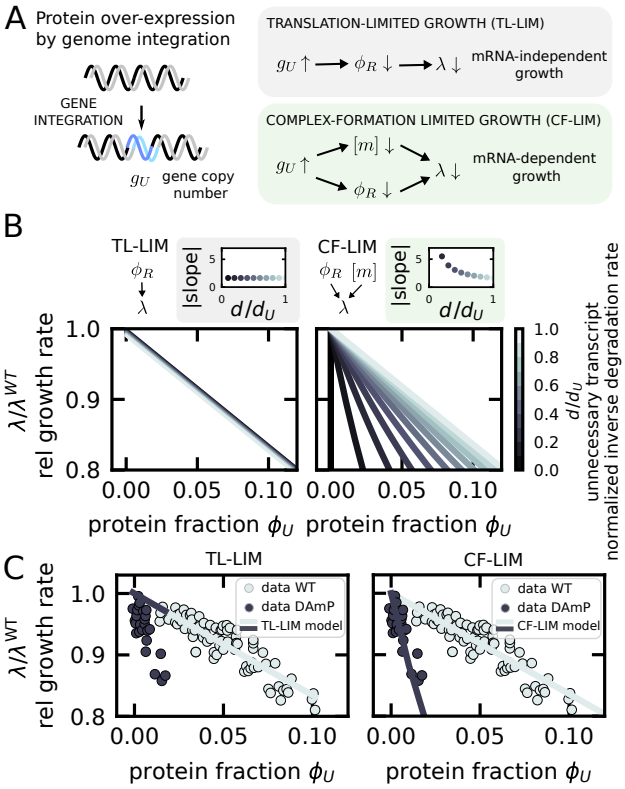

Fig 2. If formation of the ribosome-mRNA complex is limiting, mRNA levels contribute to the protein expression cost. (A) The expression cost of an unneeded protein was probed experimentally in refs. (Scott et al., 2010; Kafri et al., 2016a) by integrating into the genome highly-expressed unnecessary genes (possibly in multiple copies), labelled by U. In the CF-LIM regime, the presence of such genes reduces both mRNA concentration $[m]$ and ribosomal protein fraction ϕ_R, and the decrease of total mRNA level contributes to a decrease in growth rate. Conversely, in TL-LIM only a reduction of ribosomal levels causes a reduction of the growth rate λ. (B) The model predicts a drop in growth rate (quantified by the relative growth rate λ/λ^{WT}) as a function of protein fraction ϕ_U of the unnecessary proteins, and how such trend changes as the degradation rate of the unnecessary transcript d_U varies (right panel, where d/d_U is the inverse degradation rate of the unnecessary transcript normalized to the average degradation rate of the other transcripts). The left panel shows the prediction in TL-LIM. The right panel shows the prediction in the CF-LIM regime. Above the two plots, the sheded boxes show how the slopes of the curve $\lambda/\lambda^{WT}(\phi_U)$ change with d/d_U under translation limitation and complex-formation limitation respectively. (C) *S. cerevisiae* data from ref. (Kafri et al., 2016a) falsify a scenario of translation-limited growth and show the trends predicted by the CF-LIM regime. The left panel shows the comparison between the data (circles) and a TL-LIM model (solid lines). Light-grey circles represent data corresponding to stable transcripts ($d/d_U \approxeq 1$). Dark-grey circles represent unstable trascripts ($d/d_U \approxeq 0.08$). Dark- and light-grey lines are model predictions for the two conditions. The right panel shows the comparison between the data (circles) and our model of the complex-formation limiting regime (CF-LIM, solid lines), using the same color-code.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

degradation rate d_U always decreases the translation burden because fewer unnecessary transcripts remove active ribosomes from the pool of available ones. However, the presence of unnecessary genes will still affect overall mRNA content, because they remove active RNAPs from the pool of all the other transcribed genes (Klumpp et al., 2009). If the overall mRNA content limits growth, the contribution to the growth cost coming from translation remains unchanged.

In our framework, ribosomes are not necessarily the only growth-limiting component. Indeed, Eq. [4] implies that unnecessary proteins may affect growth both through their impact on ribosomes and through their effect on total mRNA content (as well as through RNAP allocation, as shown by Eq. [6]). To predict the burden of unnecessary proteins, it is therefore necessary to determine their effect on mRNA as well as on ribosomes. For the reasons explained above, the modulation of the unnecessary transcript degradation rate (Kafri et al., 2016a) is an ideal testing ground for our model as it allows to tune the relative contribution of translation and transcription to the growth cost. Increasing the degradation rate of unnecessary transcripts decreases the translation cost (as fewer ribosome are allocated to unnecessary transcripts) while also increasing the transcription cost (as some transcripts degrade faster, depressing the overall average lifetime of the transcripts).

In order to pursue this question mathematically, let us first consider the two limit cases $[m]/K_m \to \infty$ (complete translation limitation) and $[m]/K_m \to 0$ (complete complex-formation limitation). By labelling as wild-type (WT) all quantities without overexpression burden, we find,

$$\frac{\lambda}{\lambda^{WT}} \approx \begin{cases} \frac{\phi_R}{\phi_R^{\mathrm{WT}}} & \text{TL-LIM} \quad [m]/K_m \to \infty \\ \frac{\phi_R}{\phi_R^{\mathrm{WT}}} \frac{[m]}{[m]^{\mathrm{WT}}} & \text{CF-LIM} \quad [m]/K_m \to 0\,. \end{cases} \tag{8}$$

Equation [8] shows that, in these two limiting cases, the relative growth is determined by the ratios $\frac{\phi_R}{\phi_R^{\mathrm{WT}}}$ and $\frac{[m]}{[m]^{\mathrm{WT}}}$. Our framework provides an expression for these ratios under the assumption that unnecessary proteins compete for resources equally with the other protein sectors.

For the ribosomal protein ratio, we recover the well-known result (Scott et al., 2010)

$$\frac{\phi_R}{\phi_R^{\mathrm{WT}}} = \left(1 - \frac{\phi_U}{1-\phi_Q}\right) , \tag{9}$$

where ϕ_U is the fraction of the proteome taken up by the unnecessary protein, while ϕ_Q is the fractional size of a "Q sector" whose size remains constant across growth perturbations (by negative auto-regulation for instance) (Scott et al., 2010; Scott and Hwa, 2011).

Regarding the Q sector we need to distinguish the following two possible scenarios, which make a difference for the mRNA ratio: (i) RNAP proteins may not be part of the Q sector, therefore their fraction changes across growth conditions, or (ii) RNAP proteins may be part of the sector Q, and consequently they remain constant. Although the first hypothesis may appear more conservative, as it does not imply the existence of a circuit that keeps constant the fraction of RNAPs, Balakrishnan et al. (Balakrishnan et al., 2022) show experimentally that RNAP proteins are part of the Q sector in *E. coli*. As different organisms may not regulate RNAP proteins in the same way, we will nonetheless consider both scenarios.

We begin by considering scenario (i) where RNAP is not a part of the Q sector. In

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

this case, our model provides the following expression for the mRNA ratio

$$\frac{[m]}{[m]^{\mathrm{WT}}} = \left(1 - \frac{\phi_U}{1-\phi_Q}\right)\left[1 - \frac{\phi_U\left(\frac{d_U}{d}-1\right)}{1+\phi_U\left(\frac{d_U}{d}-1\right)}\right]\frac{f_{bn}}{f_{bn}^{\mathrm{WT}}} \ . \tag{10}$$

In Eq. [10], the first term $\left(1 - \frac{\phi_U}{1-\phi_Q}\right)$ represents the decrease in RNAP fraction due to the burden from the unnecessary protein. In scenario (ii), RNAP fraction cannot change, and consequenlty this term is simply 1. The second term $\left[1 - \frac{\phi_U\left(\frac{d_U}{d}-1\right)}{1+\phi_U\left(\frac{d_U}{d}-1\right)}\right]$ corresponds to a decrease in mRNA content due to the decreased mRNA stability of the unnecessary gene (hence the term is 1 if $d_U = d$). The third term of Eq. [10], $\frac{f_{bn}}{f_{bn}^{\mathrm{WT}}}$ represents the relative fraction of bound RNAPs. As the exact expression of the ratio $\frac{f_{bn}}{f_{bn}^{\mathrm{WT}}}$ in terms of the model parameters is not particularly transparent, we provide it in SI Appendix. We proved (see SI Appendix) that such ratio increases as more unnecessary genes are added to the genome if RNAP is not saturated, and that mRNA concentration always decreases with unnecessary protein expression ($\frac{[m]}{[m]^{\mathrm{WT}}} < 1$). Biologically, this is a consequence of two basic mechanisms in our model. First, a fraction of RNAPs needs to be allocated to transcribe unnecessary genes, decreasing the activity of RNAP genes. Second, the presence of fast-degrading unnecessary transcripts decreases the lifetime of the average transcript. Note that, if the fraction of RNAPs remains constant as in the scenario (ii), mRNA decreases only due to the second mechanism.

To illustrate the effect of transcriptional limitation on the cost of protein overexpression, we now consider the two limits of Eq. [8] in conjunction with Eq. [9] and [10]. Fig. 2B shows the main results of the model in the scenario where RNAP is constant across growth conditions. SI Fig. S1 shows the results for the scenario of variable RNAP sector. Both scenarios can reproduce the data, but the latter can match the data only in the regime where the housekeeping Q-sector is absent, which we deem to me less realistic. We also stress that the qualitative results are independent of whether the RNAP sector is constant or variable across growth conditions. The plots in Fig. 2B show the relative growth rate (the ratio $\frac{\lambda}{\lambda^{\mathrm{WT}}}$) against the unnecessary gene protein fraction ϕ_U in the two limiting regimes. Under translation-limitation (TL-LIM), Eq. [9] predicts that the growth rate decreases linearly with the protein fraction ϕ_U. In addition, increasing the degradation rate d_U of unnecessary transcripts has no effect on the slope of the curve, and consequently all the curves collapse onto the same curve (as observed for *E. coli* in ref. (Scott et al., 2010)). Under competition for transcripts (CF-LIM), the growth rate still decreases linearly with the protein fraction ϕ_U. Yet, the response of the system to decreasing the unnecessary transcript lifetime is very different: the slope of the plot of relative growth rate *vs* unneeded protein fraction does not stay flat but increases with decreasing stability, indicating that higher growth cost per protein fraction stems from the decrease in total mRNA content.

The results above imply that the different behavior of overexpression experiments under increasing transcript instability provides a way to probe whether the regime of competition for transcripts described by our model is in place. In brief, the lack of curve-collapse under a change of transcript lifetime would would be a signature of a transcriptional burden. We found that the study of Kafri and coworkers (Kafri et al., 2016a) has performed precisely such an experiment in *S. cerevisiae*.

Specifically, they have generated *S. cerevisiae* strains inegrating an mCherry reporter into the genome in multiple copies, and modulated the lifetime of the transcript by an antibiotic resistance cassette that inhibits termination in the mCherry gene, a method

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

termed DAmP (decreased abundance by mRNA perturbation). They do not quantify directly the reduction in the lifetime of the DAmP transcript, but they show that protein levels of the construct are reduced roughly by a factor of 10. We estimated the reduction of transcript lifetime from these measurement (SI Appendix) and found a corresponding 10-fold decrease. Importantly, such estimate does not assume any specific limitation regime. Next, we used this estimated parameter in our model and observe its predictions under different limitation regimes. We note that we also analyzed directly data from ref. (Kafri et al., 2016a) in order to be able to compare with our model, and in particular we quantified the protein fraction of the unneeded protein (see SI Appendix and SI Fig. S2). Fig. 2C shows the comparison of their data to the predictions of our model (which do not entail any adjustable parameter). A model of purely translation-limited growth cannot describe these data, as the two curves corresponding to different transcript stability do not collapse onto one another (Kafri et al., 2016a). Instead, a model describing the complex-limited regime (CF-LIM) reproduces quantitatively the experimental trends. Finally, we note that the study of Kafri and co-workers (Kafri et al., 2016a) also quantified the trend of the relative growth rate against the gene copy number. Our framework also makes correct predictions for this quantity (see SI Fig. S3). We note that a 10-fold reduction of unnecessary mRNA lifetime can also be obtained directly by fitting the data of growth cost as a function of gene copy number from ref. (Kafri et al., 2016a) (SI Fig. S4), leading to the same independent prediction of Fig. 2C. In their work, Kafri and colleagues also state that perturbing the transcript stability of the unneeded genes induces a 30-fold decrease in the unneeded mRNA in correspondence to the 10-fold decrease of protein levels used in our estimate. Instead, our model would predict the same fold change for unnecessary mRNA and protein levels. In the SI Appendix, we discuss more carefully the origin of this discrepancy. The possible explanations include a larger-than-predicted decrease in total mRNA content, reduced translation burden due to reduced misfolding and protein damage (a mechanism not taken into account by our model, but observed in similar experiment in yeast (Farkas et al., 2018)) or the importance of post-transcriptional regulation of gene expression.

The model formulated thus far outputs the cellular growth rate given ribosome fraction and total mRNA concentration. However, we did not put forward any hypothesis on how the cell chooses such inputs. If such quantities contribute setting growth rates, given that growth rate is a crucial component of evolutionary fitness, they are likely to be strongly regulated. Indeed, it is well-known that ribosomal proteins are tightly regulated in fast-growing organisms such as *E. coli* and *S. cerevisiae* (Scott et al., 2010; Kafri et al., 2016a). Our framework can describe regulation of gene expression in a simple way. The next section describes how our model can encode different regulatory strategies, and examines the regulation strategy that maximizes the growth rate (Scott et al., 2014), formulating predictions for changes in mRNA and RNAP abundance across nutrient conditions under such strategy.

Optimization of growth rate leads to decrease of total mRNA in poor nutrient conditions

Our model outputs the cellular growth rate given the ribosome allocation and total mRNA concentration via Eq. [4], which can be rewritten in terms of RNAP allocation, as shown in Eq. [6]). However, we did not discuss the biological mechanisms responsible for fixing these allocation parameters. This section describes how our model encodes different regulatory strategies, and examines the optimal allocation strategy maximising the growth rate (Scott et al., 2014), and which predicts changes in mRNA and RNAP abundance across nutrient conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In order to do that, we focus on the proteome composition $\Phi = \{\phi_1, \ldots, \phi_i, \ldots, \phi_S\}$, where S is the number of sectors. As nutrient conditions change, cells adjust their growth

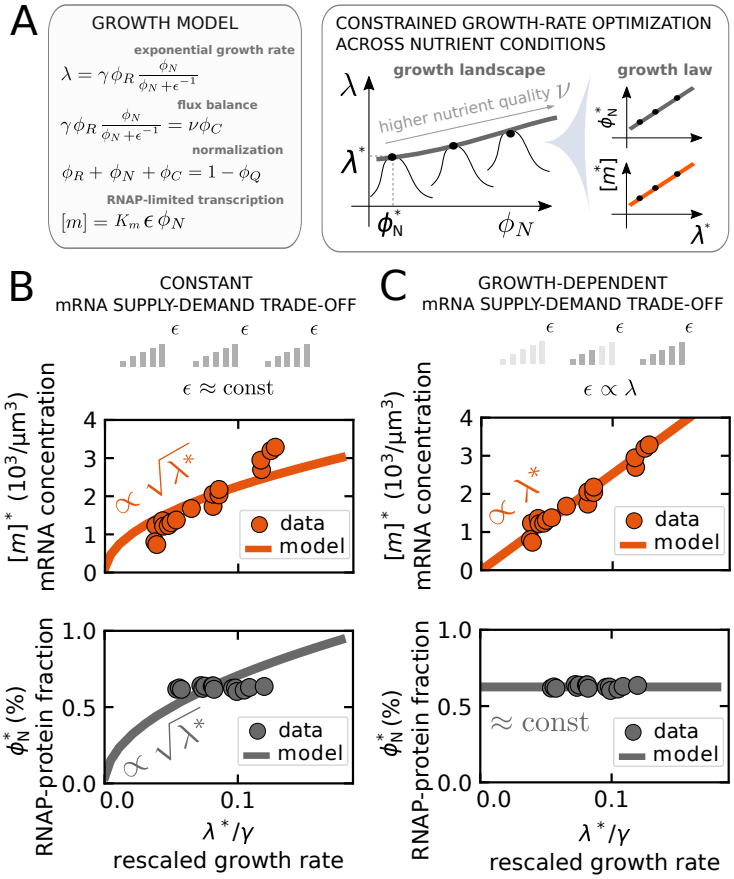

Fig 3. The mRNA levels that maximize growth rate vary across nutrient conditions. (A) The model predicts proteome composition and mRNA levels across nutrient conditions under the assumption that RNAP allocation optimizes growth rate. The left panel illustrates the basic mathematical ingredients of the model, including the dependency of the growth rate on ribosomes and RNAPs, as well as the flux-balance constraint between protein-precursors metabolism and the production of housekeeping proteins. The right panel shows how the model ingredients lead to a nutrient-dependent growth "landscape" where the growth rate is a bell-shaped function of RNAP levels. By taking the maximum of the curve, the model predicts the optimal RNAP and mRNA levels. Such predictions can take the form of a growth law (Scott et al., 2010, 2014) by plotting such levels against the corresponding optimal growth rate. (B) Under growth-rate optimization and constant mRNA supply-demand trade-off ϵ across nutrient conditions, the model predicts that both the total mRNA concentration (upper panel, orange solid line) and the fraction of RNAP proteins in the proteome (lower panel, grey solid line) increase as the square root of the growth rate. (C) Under growth-rate optimization and linearly increasing trade-off parameter ϵ with growth rate, the model predicts that the total mRNA concentration increases linearly with the growth rate (upper panel, orange solid line) while the fraction of RNAP proteins remains constant (lower panel, grey solid line). Both panel (B) and (C) show *E. coli* mRNA and RNAP data from ref. (Balakrishnan et al., 2022) across nutrient conditions (orange and grey circles), validating the scenario of increasing transcriptional activity (panel C) for this model organism.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

rate by changing their proteome composition (Scott and Hwa, 2011), which we can summarise with the remodelling $\Phi \to \Phi'$. Such changes require the presence of specific molecular pathways able to sense the nutrient conditions and relay the signal to the gene expression machinery (Zhu et al., 2019; Potrykus and Cashel, 2008; Nomura, 1999; Ramirez et al., 1991). While a detailed mechanistic understanding of such architecture is only at its dawn (Wu et al., 2022), previous work has shown that often simplifying assumptions such as flux balance and growth rate optimization allow to formulate effective predictions. Such coarse-grained models bypass an explicit description of the sensing-regulation mechanisms, whose function must be at least in part to enforce growth optimality (Scott et al., 2010, 2014; Erickson et al., 2017).

In the following, we will assume that cells regulate resource allocation to achieve an optimal regime for growth, and look for the proteome allocation assignment Φ^* that maximizes the growth rate under different perturbations. Since at steady state proteome composition reflects proteome allocation, this is equivalent to looking for the RNAP allocation assignment that maximizes the growth rate under different perturbations. We discuss in more detail the correspondence, valid in balanced exponential growth, between transcript and protein fraction $\frac{m_i}{m}$ and $\phi_i = \frac{P_i}{P}$ with the fraction ω_i of RNAP polymerases transcribing a gene of type i in SI Appendix.

We proceed by analyzing the predictions of our model under the assumption of growth optimality. We have already discussed how the model links growth rate to ribosome and RNAP allocation via Eq. [6]. In order to optimize the growth rate, we need to find an assignment of the RNAP allocation ϕ_N and ribosome allocation ϕ_R such that λ is maximal. Importantly, ϕ_N and ϕ_R also need to satisfy a number of constraints (Fig. 3A). A trivial constraint is normalization (see above), ϕ_i are protein fractions, their sum $\sum_i \phi_i$ must add to unity, and such sum includes ϕ_N and ϕ_R. In addition, we assume a "flux balance" condition (Scott et al., 2010), according to which regulatory constraints must match the influx of protein precursors such as amino acids to the flux of protein synthesis. Specifically, if ϕ_C is a protein sector that synthesizes and imports precursors, the precursors influx (per unit of time) is $dA_{in} = \nu\phi_C\, dP$. The parameter ν represents the efficiency of converting nutrients to precursors, often called nutrient quality (Scott et al., 2010; Erickson et al., 2017). Biologically, this parameter depends both on the nutrient type and on the metabolic efficiency of the cell. The outflux of precursors due to conversion into proteins is $dA_{out} = -\gamma\phi_R \frac{\phi_N}{\phi_N+\epsilon^{-1}}\, dP$. By the assumption of flux matching, we impose $dA_{in} = dA_{out}$ and obtain

$$\nu\phi_C = \gamma\phi_R \frac{\phi_N}{\phi_N + \epsilon^{-1}} \,. \tag{11}$$

Combining our framework with these constraints, we considered a minimal model with only four protein sectors: the RNAP sector (proteome fraction ϕ_N), the ribosomal sector (ϕ_R), a catabolic sector (ϕ_C) and a housekeeping sector (ϕ_Q) that does not change with growth conditions (Scott et al., 2010; Hui et al., 2015), and which effectively sets the maximum level $\phi^{max} = 1 - \phi_Q$ that any other sector can reach. Therefore, the normalization condition reads

$$\phi_N + \phi_R + \phi_C + \phi_Q = 1 \,. \tag{12}$$

By combining Eq. [6], [11], and [12], we solved analytically the constrained optimization problem as illustrated in Fig. 3A.

A key quantity to describe the role of transcription in growth is the proteome fraction of RNA polymerases. Such quantity takes the following expression in the solution of the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

following optimization problem,

$$\frac{\phi_N^*}{\phi_{max}} = \frac{\epsilon^{-1}}{\phi^{max}} \frac{\bar{\nu}}{1+\bar{\nu}} \left[\sqrt{1 + \left(\frac{\epsilon^{-1}}{\phi^{max}} \frac{\bar{\nu}}{1+\bar{\nu}} \right)^{-1}} - 1 \right], \tag{13}$$

where we indicate with the superscript * the optimized variables, and $\bar{\nu} = \frac{\nu}{\gamma}$ is a dimensionless nutrient quality. Equation [13] shows that $\frac{\phi_N^*}{\phi^{max}}$ solely depends on a single composite parameter $p = \frac{\epsilon^{-1}}{\phi^{max}} \frac{\bar{\nu}}{1+\bar{\nu}}$, which can be tuned by changing either the dimensionless nutrient quality ν or the mRNA trade-off parameter ϵ. The Materials and Methods and SI Appendix provide details of these calculations.

We note that if the supply-demand trade-off parameter ϵ is a function of the nutrient quality ν, the expression above still holds. Motivated by the findings of ref. (Balakrishnan et al., 2022), we considered two particular cases, which we can be compared with data. In the first scenario, ϵ is a constant that does not depend on the nutrient quality. Conversely, in the second scenario ϵ takes the form $\epsilon = \epsilon^{max} \frac{\bar{\nu}}{1+\bar{\nu}}$, which is equivalent to stating that ϵ is linearly proportional to the growth rate (see SI Appendix and SI Fig. S5).

Fig. 3B shows the model predictions for overall mRNA concentration in the first scenario. As nutrient conditions improve (increasing $\bar{\nu}$), the RNAP protein fraction ϕ_N increases under the optimal solution. As total mRNA levels are proportional to ϕ_N, they also increase with increasing nutrient quality. Interestingly, the model predicts that such quantities have a square-root dependency on the growth rate as opposed to the typical linear laws (see also SI appendix for an extended discussion). Fig. 3B also shows *E. coli* data from ref. (Balakrishnan et al., 2022). These data agree qualitatively, but not quantitatively, with the mRNA trend in the data and fail to capture the measured constant RNAP protein fraction. Finally, Fig. 3C shows the result of the optimization under the assumption that the transcription-translation trade-off is proportional to growth rate $\epsilon \propto \lambda^*$. In this case, the model predicts that mRNA increases linearly with the growth rate, while the the fraction of RNAP proteins stays constant. Crucially, the model also predicts that there is no need to tune the fraction of RNAP in this scenario, as it happens in *E. coli*. Specifically, Balakrishanan and coworkers show that ϕ_N stays constant across growth conditions (gray circles), and that the expression of σ-70 sequestration factor Rsd decreases with growth rate, modulating mRNA synthesis fluxes at constant fraction of RNAP (Balakrishnan et al., 2022). In our framework, the RNAP sequestration at slow growth is achieved by the linear reduction of the transcription-translation trade-off parameter ϵ, which can be interpreted as a lowering of the fraction of RNAPs bound to genes f_{bn} in order to maintain the linear density of ribosomes on transcripts constant across conditions (Balakrishnan et al., 2022). Indeed, our analysis also shows that such density remains constant under growth-dependent ϵ, but increases if ϵ stays constant (SI Fig. S6 and S7).

An important (more general) aspect is that whatever the architecture causing this trend, and whatever the scenario, in both data and model mRNA concentration decreases with decreasing growth rate, suggesting that at slow growth competition for transcripts is likely relevant as the mRNA concentration becomes closer to K_m, the concentration scale that determines to what extent mRNA affect the growth rate.

We note again that this slow-growth regime is also the one where the simplified model considered here for illustration becomes approximate. Specifically, it neglects documented processes in *E. coli* leading to a reduction of the ribosomes that are effectively available for translation (Dai et al., 2016; Dai and Zhu, 2020; Calabrese et al., 2022).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Under transcription-targeting drugs, RNAP levels must increase to achieve optimal growth, but ribosome levels may not

The expressions obtained in the previous section can be used to investigate the effect of transcription-targeting drugs on growth (Scott et al., 2010; Si et al., 2017; Roy et al., 2021). More precisely, we analyzed the optimal re-arrangement of coarse-grained proteome composition. We modeled a transcription-targeting drug that affects the global mRNA concentration $[m]$. Our model shows trivially that in the regime where growth is dependent on such concentration, transcription-targeting drugs also directly affect growth. We can ask the further question of how the proteome responds to such perturbations.

Such drugs might target different steps of the transcription process to reduce overall mRNA concentration. From the perspective of our model, we focus on transcription-targeting drugs that modify the trade-off parameter ϵ (Eq. [7]).

We recall that $[m]/K_m = \epsilon\phi_N$, which implies that in principle transcription-inhibitory drugs could also decrease mRNA levels by reducing RNAP expression. Note that ϵ is a composite parameter shaped by several distinct biological processes. According to our model, drugs attacking distinct aspects of transcription can cause quantitatively similar results. For instance, drugs targeting transcriptional elongation decrease ϵ by lowering the elongation rate k_{tx}; drugs targeting transcriptional initiation decrease ϵ by reducing the fraction of gene-bound RNAP f_{bn}; drugs targeting transcript stability decrease ϵ by curbing the transcript degradation rate d. Hence, by focusing generically on how ϵ changes, our model makes generic predictions about multiple situations.

To derive the effect of transcription-inhibitory drugs in the model, we follow the same formal steps of the previous section. The growth rate is a function of the coarse-grained proteome composition made of ribosomal, RNAP and catabolic proteins. In addition, it also depends on the mRNA supply-demand trade-off parameter ϵ and the nutrient quality ν. Mathematically, the growth rate λ is a function $f(\phi_R, \phi_N, \phi_C, \epsilon, \nu)$ given by equations [6] and [11]. We find the composition $\{\phi_R, \phi_N, \phi_C\}$ that maximizes the growth rate at given supply-demand trade-off ϵ. Consequently, the optimal composition is a function of the trade-off parameter ϵ. If transcription-targeting drugs effectively vary ϵ, the optimal proteome composition as a function of ϵ can be interpreted as the optimal cellular gene-regulatory control in response to transcription-targeting drugs - see also the sketch in Fig. 4A.

Biologically, there can be situations where cells can or cannot implement an optimal response strategy, hence it makes sense to compare a growth-optimized scenario with a non-optimized one. Fig. 4B shows how the growth rate slows down under transcription inhibitors when mRNA trade-off ϵ decreases with and and without the enforcement of growth-rate optimization. The dashed grey line shows the relative growth rate against the trade-off parameter ϵ without any response in proteome rearrangement (Eq. [6] plotted as a function of ϵ). The solid blue line shows the relative growth rate under optimization, following the procedure described above. The plot shows that optimal feedback always determines a faster growth rate.

Additionally, our model makes testable predictions about how the proteome re-arranges under transcription-targeting drugs. The RNAP fraction of the proteome always increases, to compensate the lower transcriptional efficiency with more RNA polymerases (see equation [13]). Despite such increase, overall mRNA concentration still decreases (Fig. 4D), although it decreases less than it would without growth-optimized response. Clearly, RNAP upregulation must come at the expense of other protein sectors. How do the other coarse-grained protein sectors behave? Intriguingly, the ribosomal proteome fraction changes qualitatively differently under treatment of transcription-targeting drugs in nutrient-poor *vs* nutrient-rich media. Fig. 4C shows that the model predicts up-

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

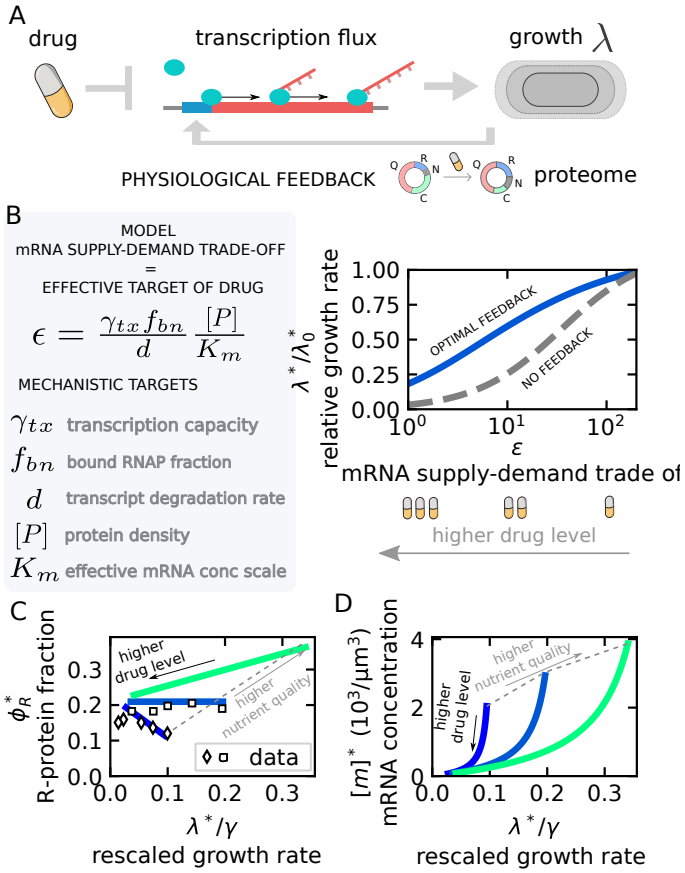

Fig 4. The predicted response to transcription-targeting drugs entails up-regulation of the RNA polymerase sector, but not necessarily of ribosomes. (A) Transcription-targeting drugs affect growth directly, but the overall effect also depends on the physiological response of the cell to reduced transcriptional capacity. Our model predicts ribosome- and RNAP-allocation rearrangements under transcription inhibitors. (B) The grey box illustrates the model expression for the transcription-translation trade-off ϵ, i.e., the ability of RNAP to produce mRNA (see Fig. 3A). This effective parameter combines several quantities, including the transcription elongation rate of RNAP on genes, the fraction of gene-bound RNAPs and the mean lifetime of transcripts. The model predicts that drugs attacking any of these parameters modify the way mRNA level affect growth rate only through changes in ϵ. The plot shows the growth rate reduction *vs* ϵ under growth-optimized (continuous blue line) and non-optimized (grey dashed line) conditions. Note that λ_0^* is the growth rate in the unpertubed condition. Optimal proteome re-arrangement "rescues" the growth rate decrease under transcription inhibition (reduction of ϵ). (C) The predicted ribosomal proteome fraction ϕ_R under growth optimization changes as transcription is inhibited (ϵ reduced). Solid lines of different colors indicate different nutrient conditions with darker colors representing poorer media. Importantly, the qualitative trend depends on the nutrient quality. Symbols are experimental data points from ref. (Scott et al., 2010) (D) The model prediction for mRNA concentration $[m]$ changes as ϵ decreases under optimal allocation. Different solid lines indicate different nutrient conditions, with darker colors representing poorer media.

regulation of ribosomal content in poorer media and down-regulation in richer media. For a "critical" value of the nutrient quality, the ribosomal change does not change at

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

all upon inhibition of transcription. These predictions are in agreement with *E. coli* experimental data under treatment with rifampicin, a well-known inhibitor of global transcription (Fig. 4C, data from ref. (Scott et al., 2010)).

The reasons for such a trend change are subtle, although easy to prove computationally and analytically. In brief, as transcription capacity decreases, the model predicts that the cell will reduce the size of the catabolic sector ϕ_C (see SI Fig. S8) because fewer precursors are being used due lower overall biosynthetic capacity. Hence, the freed fraction of the proteome $\Delta\phi_C$ becomes available to the other sectors. Quantitatively, the size change $\Delta\phi_C$ depends on the nutrient quality of the medium. In rich media $\phi_C^* \simeq 0$, while in poor media ϕ_C^* can become considerable. This is because in rich media nutrients are easily imported and catabolized, so ϕ_C^* can be small. Part of the freed $\Delta\phi_C$ will increase the fraction of the proteome occupied by ribosomes, offsetting its decrease due to the upregulation of RNAP (see Fig. 1). If ϕ_C^* is sufficiently large (which may happen in poor media), the ribosomal fraction can also *increase* as transcriptional capacity decreases.

Finally, we note that we made the crucial assumption that the RNAP fraction can vary under treatment of transcription-targeting drugs. In SI Appendix, we consider the optimal response under constant level of RNAP fraction as in the scenario described in Fig. 3.

Discussion

We presented an organism-agnostic framework describing biosynthesis, accounting explicitly for the two key steps of the central dogma, mRNA and protein production, which can describe the growth laws determined by RNAP and ribosome allocation. The fact that our model is able to formulate correct predictions for both *E. coli* and *S. cereviesiae* supports the hypothesis that unifying principles, due to simple trade-offs (allocation, flux balance, etc.) may apply across organisms and kingdoms, and therefore adds up to the thread of evidence supporting the existence of universal aspects of growth physiology (Bruggeman et al., 2020; Kostinski and Reuveni, 2021; Kafri et al., 2016b; Shahrezaei and Marguerat, 2015; Calabrese et al., 2022). Clearly, further work could investigate the role of specific aspects of different transcription-translation architectures. For example, in *E. coli* transcription and translation co-occur in the nucleoid, while in budding yeast the formation of ribosome-mRNA complexes needs nucleocytoplasmic transport of mRNA, which relies on multiple proteins and organelles. Such differences may affect the parameters leading to regime of competition for complexes, as well as the biological perturbations that affect this regime (Liang et al., 2000; Nomura, 2001). While the effect of proteome allocation on translation is clear from previous work, general feedbacks between translation and transcription capacity remain relatively unexplored. Our results show clearly that there are interesting physiological and perturbed situations where competition for transcripts sets growth rate. Interestingly, the RNAP proteome sector, while being very small, plays the crucial role of controlling mRNA levels, therefore it can be determinant to decide to which extent mRNAs set growth rate. In eukaryotes, a further extension of our framework could investigate the role of different dedicated RNAP pools (Kostinski and Reuveni, 2021).

As we have mentioned in the Introduction, two recent studies (Lin and Amir, 2018; Roy et al., 2021) have considered the impact of the transcriptional layer of protein production on growth, but our work differs in several important ways. We built our transcription-translation framework following Lin and Amir (Lin and Amir, 2018), who focus on the transcription-limited growth regime (TX-LIM) where *only* mRNA determines growth (and RNAP autocatalysis is essential for exponential growth). Conversely,

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

our model focuses on the regime where both mRNA and ribosomes are relevant for setting growth rate, as a consequence of the competition for ribosomes by mRNA. We note that cells grow exponentially in this competition regime, through ribosome autocatalysis. The study by Roy and coworkers (Roy et al., 2021) also did not consider this regime. In addition, our study also investigates the optimal RNAP/ribosomal allocation across nutrient conditions, while both previous studies focused only on non-optimized relationships. Finally our framework is very similar to the data-driven model proposed by Balakrishan and coworkers (Balakrishnan et al., 2022), with the advantage of being able to provide a mechanistic interpretation for the factor relating growth rate and total mRNA concentration, and to compare growth-optimized situations with non-optimized ones as done in ref. (Scott et al., 2014).

Growth limitations can be important across different perturbed and physiological contexts, therefore we expect that our approach can form the basis for further investigations. A first perturbation where transcriptional limitation is obviously important is treatment of cells with transcription inhibitors. The net effect of protein synthesis inhibitors on growth is mediated by the physiological feedback of the cell in response to drug treatment(Scott et al., 2010). A well-known example of this is the growth reponse of E. coli to ribosome-targeting antibiotics. Such drugs also induce up-regulation of ribosomal content which partially rescues growth rate decrease. More in detail, the response depends on the growth condition as well as the affinity of the drug to its target (Greulich et al., 2012). The structure of this feedback has been used to predict the shape of dose-response curves for different translation-targeting antibiotics (Angermayr et al., 2022). This kind of analysis remains largely open for the case of transcription inhibitors. The relationship of transcription inhibitors with growth rate (in non-optimized conditions) was considered in ref. (Roy et al., 2021). Our model adds the further step of being able to compare growth-optimized with non-optimized resonse scenarios, and to make definite predictions for ribosomal and RNAP sector response to transcription inhibitors. However, additional elements such as drug affinity and feedback mechanisms (Angermayr et al., 2022; Greulich et al., 2012) may be important to fully understand the physiological response to transcription inhibitors. Future studies could extend our framework in these directions.

A second important perturbation where transcription becomes important for growth is the expression of unnecessary proteins (Scott et al., 2010; Dekel and Alon, 2005; Shachrai et al., 2010; Kafri et al., 2016b). Unnecessary protein expression imposes a cost on growth by affecting the abundance of growth-limiting components. In the standard ribo-centric growth model, the growth rate is proportional to ribosome content (Kafri et al., 2016b). Because of finite resources, expression of unnecessary protein decreases the expression of ribosomal proteins, which reduces the number of ribosomes and slows down growth. In our framework, ribosomes are not necessarily the only growth-limiting components. We found that the growth cost depends on the decrease of total mRNA content as well as ribosomal content and provided a quantitative predictive model of this decrease and the corresponding growth cost. The transcription rate of the unneccessary gene always increases the growth cost of overexpression, while the decrease of unnecessary transcript stability decreases the growth cost, because fewer unnecessary transcripts take up ribosomes from the unnecessary mRNAs. Our model provides a quantitative framework to describe both perturbations jointly. We speculate that this quantitative understanding could be useful to predict the fitness landscape of situations with perturbed gene dosage, such as large-scale gene duplications in absence of dosage-compensation mechanisms (Pompei and Lagomarsino, 2023). In addition, our model provides a method to rigorously test the presence of transcriptional limitations by considering how the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cost of overexpression varies as the mRNA stability of the unnecessary protein changes. Hence, the problem of the growth cost of unneeded proteins may be important as an experimental testing tool for the presence of transcriptional limitations in a specific reference regime. Other situations where transcription may become limiting indirectly, as a consequence of orthogonal perturbations are inhibition of translation (as already shown in *E. coli* (Zhang et al., 2020)) and DNA dilution (as shown in yeast under G1/S arrest (Neurohr et al., 2019)).

In physiological conditions, our model predicts that under the assumption of transcript competition, optimization of growth rate leads to the observed decrease of total mRNA levels in poor nutrient conditions (in absence of any external perturbations). This prediction is in agreement with the recent experimental observation that total mRNA concentration decreases with decreasing growth rate in *E. coli* (Balakrishnan et al., 2022). Mechanistically, as the authors show, this trend is realized in *E. coli* by increasing RNAP sequestration from anti-sigma factor Rsd, indicating that growth rate is modulated by a change of the fraction of active RNAPs, instead of the total RNAP levels, a situation that, as we have shown, can be implemented in our model. We speculate that this sequestration mechanism could help the system to react faster in nutrient upshifts, as desequestration of RNAP components could be much faster than transcriptional reprogramming (Balakrishnan et al., 2022). Transcriptional limitation in physiological conditions was also reported to be relevant for phosphate limitations, in both yeast and *E.coli* (Espinosa et al., 2022; Metzl-Raz et al., 2020; Kafri et al., 2016b), and may be relevant for regulating cell-cycle dependent expression in budding yeast (Swaffer et al., 2021). On more general grounds, an understanding of the link between mRNA concentration and ribosome/RNAP allocation needs a theoretical framework able to link growth rate to mRNA levels (Lin and Amir, 2018; Roy et al., 2021; Balakrishnan et al., 2022).

Finally, we note that in transcription-limited situation, as well as under competition for transcripts, mRNA scarcity may contribute to inactive ribosomes (Dai et al., 2016; Dai and Zhu, 2020; Calabrese et al., 2022). Since mRNA concentration can be quite low for slow-growing *E. coli* (Balakrishnan et al., 2022), we speculate that this factor may play a role in setting the fraction of idle ribosomes.

Methods and Materials

This Materials and Methods section discusses the model ingredients, assumptions, and derives the main mathematical results presented in the text, in particular the expression linking the growth rate to the ribosome fraction and mRNA concentrations, Eq. [4], and the formulas for the cost of protein overexpression Eqs. [8], [9], [10]. See also the SI Appendix for more detailed information on the model.

Initiation limited fluxes

We consider a situation in which the transcription and translation fluxes $J_i^{TX}([N])$ and $J_i^{TL}([R])$ in Eqs. [1] and [2] are initiation-limited:

$$J_i^{TX}([N]) \approx \beta_0^i [N_f]; \quad J_i^{TL}([R]) \approx \alpha_0 [R_f] \tag{14}$$

where β_0^i is the rate constant of transcription initiation of the genes belonging to class i, and α_0 is the rate constant for translation initiation, assumed to be identical for all transcripts. $[N_f]$ and $[R_f]$ are the concentrations of free RNAPs and of ribosomes, respectively (see SI Appendix). Eqs. [14] relate the biosynthetic fluxes with the process of recruitment and

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

complex formation (RNAP-gene and ribosome-transcript). We approximate for simplicity the current with the initiation, neglecting RNAP and ribosome traffic (MacDonald et al., 1968; Klumpp and Hwa, 2008; Erdmann-Pham et al., 2020; Ciandrini et al., 2013). While in mRNA translation approximating the ribosomal current with the initiation rate is a crude approximation (Szavits-Nossan et al., 2018), this assumption makes the model analytically treatable (see the SI Appendix for further comments on the different limiting regimes).

Minimal derivation of mRNA-dependent growth rate

In a regime of negligible ribosome traffic (see (Calabrese et al., 2022) and SI Appendix for a quick derivation), the relation $\alpha_0[R_f] = \frac{k_{tl}}{L_p}\frac{[R_b]}{[m]}$ links the free ribosomal pool $[R_f]$ with the concentration of bound ribosomes $[R_b]$. Since $[R] = [R_b] + [R_f]$, we can write the translation flux/initiation rate in terms of the total ribosome concentration, obtaining $J_i^{TL}([R]) = \frac{k_{tl}}{L_p}\frac{[R]}{K_m+[m]}$, with $K_m := \frac{k_{tl}}{\alpha_0 L_p}$. Combining this result with Eq. [2], gives

$$\frac{dP_i}{dt} = \frac{k_{tl}}{L_p}\frac{m_i}{m}\frac{[m]}{K_m + [m]}R\,. \tag{15}$$

Each term in this equation can be interpreted in a simple way: (i) $\frac{[m]}{K_m+[m]}$ is the fraction of bound ribosomes, (ii) $\chi_i := \frac{m_i}{m}$ is the fraction of bound ribosome translating a transcript of type i and (iii) $\frac{L_p}{k_{tl}}$ is the typical time to translate a protein. Associating a symbol to each of these quantities, we write schematically

$$\frac{dP_i}{dt} = \gamma_p\,\chi_i\,R\,f_{br}\,, \tag{16}$$

where γ_p is the inverse time to translate the typical protein, , χ_i is the fraction of bound ribosomes translating transcripts of type i and f_{br} is the overall fraction of ribosomes bound to transcripts. This way to represent Eq. [2] is particularly interpretable, and our model has a parallel equation for transcript production (see below).

Assuming perfect stochiometry (i.e. all ribosomal subunits are involved in a functional ribosome), we write the number of total ribosomes R in terms of the number of ribosomal proteins P_R as $R = P_R\frac{L_p}{L_R}$, where L_R stands for the number of amino-acids present in a ribosome. Summing Eq. [16] for all classes and dividing by the total protein abundance P gives Eq. [4] after defining $\gamma := k_{tl}/L_R$.

Connecting mRNA concentration with RNAP abundance

Mirroring Eq. [16] for the transcript dynamics (see SI Appendix for details), we write

$$\frac{dm_i}{dt} = \frac{k_{tx}}{L_g}\omega_i f_{bn} N - d\,m_i \tag{17}$$

with $\frac{k_{tx}}{L_g}$, ω_i and f_{bn} being respectively the inverse typical time needed to transcribe a gene (the ratio of a typical transcription elongation rate and a typical gene length), the fraction of bound RNAPs translating genes of type i and f_{bn} the overall fraction of RNAPs bound to transcripts. SI Appendix derives the exact expressions for ω_i and f_{bn}. Crucially, they depend only on the binding constants β_0^i (promoter strength) and the gene copy number g_i. Compared to the case of translation, these expressions are more complex due to gene- or sector-specific promoter strengths, although qualitatively they behave very similarly to the corresponding translation quantities χ_i and f_{br} provided we substitute the transcripts m_i for the genes g_i.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Since transcript degradation is fast compared to dilution (growth rate), we assume that transcripts are steady state, and we sum all the amount m_i for all classes i, from Eq. [17] we obtain Eq. [5]. We convert the amount of RNAPs in number of proteins composing RNAPs: $N = P_N \frac{L_g}{L_N}$, where L_N is the total number of amino acids needed to form a RNAP complex and L_g the typical gene length (perfect stochiometry assumption). This allows us to write the concentration $[N]$ in terms of fraction of polymerases ϕ_N and total protein concentration $[P]$ as $[N] = \phi_N \frac{L_g}{L_N}[P]$. Finally, to derive Eq. [6] we plug the obtained relation $[m] = \frac{\gamma_{tx}}{d} f_{bn}\ [P]\ \phi_N$ into Eq. [4].

Regime of limiting complex formation (CF-LIM)

Eq. [4] gives the mRNA dependence of the growth rate via a Michaelis-Menten factor, which becomes relevant when $\frac{K_m}{[m]} \approx 1$. This quantity has a simple interpretation as the ratio between the ribosome-mRNA association and dissociation times

$$\frac{K_m}{[m]} = \frac{k_{tl}/L_p}{\alpha_0[m]} = \frac{\tau_{\text{on}}}{\tau_{\text{off}}}\ , \tag{18}$$

since (k_{tl}/L_p) can be seen as the inverse of a ribosome characteristic "dissociation" time, i.e., the time necessary to elongate the typical protein, while a free ribosome takes a time roughly inversely proportional to $\alpha_0[m]$ to find a transcript and form a complex, i.e., "associate" with it. The growth rate depends on total mRNA concentration when the time scale of ribosome-mRNA complex formation (τ_{on}) is comparable to the timescale of full protein elongation (τ_{off}). The slower complex formation is with respect to full protein elongation, the more dependent the growth rate becomes on mRNA. Eq. [4] provides a general relation between growth rate, ribosome fraction and transcript concentration, without assuming that ribosomes or mRNAs are limiting. Note that this trade-off depends on a "supply" of mRNA and a "demand" from trnaslation: if $\tau_{\text{on}} < \tau_{\text{off}}$ ribosomes accumulate on mRNAs, because supply is scarce, and viceversa if $\tau_{\text{on}} > \tau_{\text{off}}$ the demand for transcripts is scarce with respect to the supply.

In the competition regime described above, it is simple to show that the growth rate $\lambda := \frac{1}{P}\frac{dP}{dt}$ is proportional to a Michelis-Menten-like factor $\lambda \propto \frac{[m]}{K_m+[m]}$ (see SI Appendix). The effective concentration K_m is

$$K_m = \frac{k_{tl}}{L_p}\frac{1}{\alpha_0}. \tag{19}$$

Let us recall that we defined the translation initiation rate as $\alpha = \alpha_0[R_f]$, where $[R_f]$ is the concentration of free ribosomes. Therefore, α_0 is a binding constant that can be estimated with the ratio $\frac{\alpha}{[R_f]}$. To estimate K_m from literature data we expressed it as

$$\begin{aligned} K_m &= \frac{(k_{tl}/L_p)}{\alpha}\ [R_f] \\ &= \frac{\text{tot. translation elongation rate}}{\text{translation initiation rate}} \times \text{free ribosome conc.} \end{aligned} \tag{20}$$

Estimate of K_m for *E. coli*

To estimate K_m for *E. coli* we started from the quantities linked to translation elongation. As $k_{tl} \approx 10 - 20\ aa/s$ (Dai et al., 2016) and $L_p \approx 300\ aa$ (Milo and Philips, 2015), it follows that the ratio $(k_{tl}/L_p) \approx 0.03 - 0.06\ s^{-1} \approx 2 - 4\ \text{min}^{-1} \approx 100 - 200\ h^{-1}$.

For translation initiation, we took $\alpha \approx 0.2 - 0.3\ s^{-1} =\approx 12 - 18\ \text{min}^{-1} = 700 - 1000\ h^{-1}$ (Balakrishnan et al., 2022). By taking the ratio of the protein elongation rate

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to the initiation rate, we obtain so far $K_m = 0.1 - 0.3\ [R_f]$. A rough estimate of the total concentration of ribosomes gives $\approx 50\mu$M independent on nutrient conditions (Milo and Philips, 2015). To conclude our estimate, we needed the fraction of free ribosomes. A recent study (Metelev et al., 2022) provides direct measurements of the percentage of ribosomal subunits that are engaged in translation, which is 90% at rich nutrient conditions. This gives an upper bound $[R_f] < 5\,\mu$M. Thus, a back-of-the-envelope estimate provides a value of K_m in the range $0.05 - 0.15\,\mu$M.

S. cerevisiae estimate of K_m

In yeast, we used $k_{tl} \approx 10\ aa/s$ (Metzl-Raz et al., 2020) and $L_p \approx 370\ aa$ (Milo and Philips, 2015), it follows that their ratio is $(k_{tl}/L_p) \approx 0.03\ s^{-1} \approx 2\ \text{min}^{-1} \approx 100\ h^{-1}$. The translation initiation rate is rather broadly distributed with a mean $\alpha = 0.12\ s^{-1} = 7.2\ \text{min}^{-1} \approx 400\ h^{-1}$ (Ciandrini et al., 2013), providing $K_m \approx 0.25\ [R_f]$. We estimate the total ribosome concentration using the fact that *S. cerevisiae* contains roughly $2 - 4 \times 10^5$ ribosomes (An and Harper, 2020) and has a volume of about 40 μm^3 (Milo and Philips, 2015). Consequently, the total ribosomes concentration is roughly $\approx 25\ \mu$M. In ref. (Metzl-Raz et al., 2017) the authors use polysome profiling to estimate the fraction of inactive ribosomes, of the order of $10 - 20\%$. Using these figures, our estimate of K_m in *S. cerevisiae* gives $K_m \approx 0.25 - 0.5\ \mu$M.

Minimal derivation of the growth cost of protein overexpression

To derive the growth of cost of protein overexpression, we used an expression for the ribosomal-protein fraction fold-change $\frac{\phi_R}{\phi_R^{WT}}$ and the total mRNA fold change $\frac{[m]}{[m]^{WT}}$ in the presence of the unneeded protein type U. Subsequently, we used Eq. [4] to obtain the growth cost (see SI Appendix).

Fold change of the ribosomal protein fraction

The fold change of the ribosomal-protein fraction $\frac{\phi_R}{\phi_R^{WT}}$ was obtained directly from the normalization condition $\sum_i \phi_i = 1$ by assuming that each protein fraction changes in the same way. With this assumption $\phi_i = c\phi_i^{WT}$ for all $i \neq U, Q$ (including R-proteins). Since $\sum_{i\neq U,Q} \phi_i^{WT} = 1 - \phi_Q$, we can find from the normalization condition that $c(1 - \phi_Q) + \phi_U + \phi_Q = 1$, which means $c = 1 - \frac{\phi_U}{1-\phi_Q}$. Therefore $\frac{\phi_R}{\phi_R^{WT}} = 1 - \frac{\phi_U}{1-\phi_Q}$, as stated in the main text in Eq. [9]. To derive the change of the total mRNA, we used Eq.[1], obtaining the following expression for the mRNA transcript for sector i and U (see SI Appendix)

$$[m_i] = \begin{cases} \frac{1}{d}\frac{k_{tr}}{L_{RNAP}}\omega_i f_{bn}[P_N] & i \neq U \\ \frac{1}{d_U}\frac{k_{tr}}{L_{RNAP}}\omega_U f_{bn}[P_N] & i = U\ . \end{cases} \tag{21}$$

Note that ω_i above is the fraction of RNAPs transcribing transcript type i. the total mRNA was obtained by summing $[m_i]$ on i,

$$[m] = \frac{k_{tr}}{L_{RNAP}} f_{bn}[P_N] \left(\frac{\sum_{i\neq U}\omega_i}{d} + \frac{\omega_U}{d_U} \right)\ .$$

To further simplify this expression, we expressed the concentration in terms of the protein fraction, i.e. $[P_N] = [P]\phi_N$ with $[P]$ the protein concentration, and used the the fact that $\sum_{i\neq U}\omega_i = 1 - \omega_U$. Finally, we also used the link between ω_U and ϕ_U, i.e.,

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$\phi_U = \frac{\frac{d}{d_U}\omega_U}{\left[1-\left(1-\frac{d}{d_U}\right)\omega_U\right]}$ (see SI Appendix and Fig. S3 for the full derivation). Following this procedure, we obtained Eq. [10] in the main text.

Acknowledgments

This work was supported by Associazione Italiana per la Ricerca sul Cancro AIRC IG grant no. 23258 (MCL and LC), and by the French National Research Agency (REF: ANR-21-CE45-0009) (LCi). LC was supported by an AIRC Fellowship.

References

An, H. and Harper, J. W. (2020). Ribosome abundance control via the ubiquitin–proteasome system and autophagy. *Journal of Molecular Biology*, 432(1):170–184. Molecular Mechanisms of Selective Autophagy.

Angermayr, S. A., Pang, T. Y., Chevereau, G., Mitosch, K., Lercher, M. J., and Bollenbach, T. (2022). Growth-mediated negative feedback shapes quantitative antibiotic response. *Molecular systems biology*, 18:e10490.

Balakrishnan, R., Mori, M., Segota, I., Zhang, Z., Aebersold, R., Ludwig, C., and Hwa, T. (2022). Principles of gene regulation quantitatively connect dna to rna and proteins in bacteria. *Science (New York, N.Y.)*, 378:eabk2066.

Belliveau, N. M., Chure, G., Hueschen, C. L., Garcia, H. G., Kondev, J., Fisher, D. S., Theriot, J. A., and Phillips, R. (2021). Fundamental limits on the rate of bacterial growth and their influence on proteomic composition. *Cell Systems*, 12(9):924–944.e2.

Borkowski, O., Goelzer, A., Schaffer, M., Calabre, M., Mäder, U., Aymerich, S., Jules, M., and Fromion, V. (2016). Translation elicits a growth rate-dependent, genome-wide, differential protein production in bacillus subtilis. *Molecular systems biology*, 12(5):870.

Bremer, H. and Dennis, P. P. (2008). Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus*, 3.

Bruggeman, F. J., Planqué, R., Molenaar, D., and Teusink, B. (2020). Searching for principles of microbial physiology. *FEMS microbiology reviews*, 44:821–844.

Calabrese, L., Grilli, J., Osella, M., Kempes, C. P., Lagomarsino, M. C., and Ciandrini, L. (2022). Protein degradation sets the fraction of active ribosomes at vanishing growth. *PLoS computational biology*, 18:e1010059.

Chure, G. and Cremer, J. (2023). An optimal regulation of fluxes dictates microbial growth in and out of steady-state. *eLife*, 12:e84878.

Ciandrini, L., Stansfield, I., and Romano, M. C. (2013). Ribosome traffic on mrnas maps to gene ontology: genome-wide quantification of translation initiation rates and polysome size regulation. *PLoS computational biology*, 9:e1002866.

Dai, X. and Zhu, M. (2020). Coupling of ribosome synthesis and translational capacity with cell growth. *Trends in biochemical sciences*, 45:681–692.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T. (2016). Reduction of translating ribosomes enables escherichia coli to maintain elongation rates during slow growth. *Nature microbiology*, 2:16231.

Dekel, E. and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436:588–592.

Dennis, P. P. and Bremer, H. (1974). Macromolecular composition during steady-state growth of escherichia coli b-r. *Journal of bacteriology*, 119:270–281.

Erdmann-Pham, D. D., Dao Duc, K., and Song, Y. S. (2020). The key parameters that govern translation efficiency. *Cell Systems*, 10(2):183–192. tex.arxivid: 1803.05609.

Erickson, D. W., Schink, S. J., Patsalo, V., Williamson, J. R., Gerland, U., and Hwa, T. (2017). A global resource allocation strategy governs growth transition kinetics of escherichia coli. *Nature*, 551:119–123.

Espinosa, R., Sørensen, M. A., and Svenningsen, S. L. (2022). Escherichia coli protein synthesis is limited by mrna availability rather than ribosomal capacity during phosphate starvation. *Frontiers in Microbiology*, 13.

Farkas, Z., Kalapis, D., Bódi, Z., Szamecz, B., Daraba, A., Almási, K., Kovács, K., Boross, G., Pál, F., Horváth, P., Balassa, T., Molnár, C., Pettkó-Szandtner, A., Klement, [U+FFFD], Rutkai, E., Szvetnik, A., Papp, B., and Pál, C. (2018). Hsp70-associated chaperones have a critical role in buffering protein production costs. *eLife*, 7.

Greulich, P., Ciandrini, L., Allen, R. J., and Romano, M. C. (2012). Mixed population of competing totally asymmetric simple exclusion processes with a shared reservoir of particles. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 85:011142.

Hu, X.-P., Dourado, H., Schubert, P., and Lercher, M. J. (2020). The protein translation machinery is expressed for maximal efficiency in escherichia coli. *Nature communications*, 11:5260.

Hu, X.-P. and Lercher, M. J. (2021). An optimal growth law for rna composition and its partial implementation through ribosomal and trna gene locations in bacterial genomes. *PLoS genetics*, 17:e1009939.

Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular systems biology*, 11:784.

Izard, J., Gomez Balderas, C. D. C., Ropers, D., Lacour, S., Song, X., Yang, Y., Lindner, A. B., Geiselmann, J., and de Jong, H. (2015). A synthetic growth switch based on controlled expression of rna polymerase. *Molecular systems biology*, 11:840.

Kafri, M., Metzl-Raz, E., Jona, G., and Barkai, N. (2016a). The cost of protein production. *Cell reports*, 14(1):22—31.

Kafri, M., Metzl-Raz, E., Jonas, F., and Barkai, N. (2016b). Rethinking cell growth models. *FEMS yeast research*, 16.

Klumpp, S. and Hwa, T. (2008). Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. *Proc. Natl. Acad. Sci.*, 105(47):18159–18164.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7):1366–1375.

Koch, A. L. (1988). Why can't a cell grow infinitely fast? *Canadian journal of microbiology*, 34:421–426.

Kostinski, S. and Reuveni, S. (2020). Ribosome composition maximizes cellular growth rates in e. coli. *Physical review letters*, 125:028103.

Kostinski, S. and Reuveni, S. (2021). Growth laws and invariants from ribosome biogenesis in lower eukarya. *Phys. Rev. Res.*, 3:013020.

Liang, S. T., Xu, Y. C., Dennis, P., and Bremer, H. (2000). mrna composition and control of bacterial gene expression. *Journal of bacteriology*, 182:3037–3044.

Lin, J. and Amir, A. (2018). Homeostasis of protein and mrna concentrations in growing cells. *Nature communications*, 9:4496.

MacDonald, C. T., Gibbs, J. H., and Pipkin, A. C. (1968). Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, 6(1):1–5.

Maitra, A. and Dill, K. A. (2014). Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proceedings of the National Academy of Sciences*, 112(2):406–411.

Metelev, M., Lundin, E., Volkov, I. L., Gynnå, A. H., Elf, J., and Johansson, M. (2022). Direct measurements of mRNA translation kinetics in living cells. *Nature Communications*, 13(1):1852. Number: 1 Publisher: Nature Publishing Group.

Metzl-Raz, E., Kafri, M., Yaakov, G., and Barkai, N. (2020). Gene transcription as a limiting factor in protein production and cell growth. *G3 (Bethesda, Md.)*, 10:3229–3242.

Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife*, 6.

Milo, R. and Philips, R. (2015). *Cell Biology by the Numbers*. CRC Press.

Neidhardt, F. C. and Magasanik, B. (1960). Studies on the role of ribonucleic acid in the growth of bacteria. *Biochimica et biophysica acta*, 42:99–116.

Neurohr, G. E., Terry, R. L., Lengefeld, J., Bonney, M., Brittingham, G. P., Moretto, F., Miettinen, T. P., Vaites, L. P., Soares, L. M., Paulo, J. A., Harper, J. W., Buratowski, S., Manalis, S., van Werven, F. J., Holt, L. J., and Amon, A. (2019). Excessive cell growth causes cytoplasm dilution and contributes to senescence. *Cell*, 176:1083–1097.e18.

Nomura, M. (1999). Regulation of ribosome biosynthesis in escherichia coli and saccharomyces cerevisiae: diversity and common principles. *Journal of bacteriology*, 181:6857–6864.

Nomura, M. (2001). Ribosomal rna genes, rna polymerases, nucleolar structures, and synthesis of rrna in the yeast saccharomyces cerevisiae. *Cold Spring Harbor symposia on quantitative biology*, 66:555–565.

Pompei, S. and Lagomarsino, M. C. (2023). A fitness trade-off explains the early fate of yeast aneuploids with chromosome gains. *bioRxiv*.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Potrykus, K. and Cashel, M. (2008). (p)ppgpp: still magical? *Annual review of microbiology*, 62:35–51.

Qian, Y., Huang, H.-H., Jiménez, J. I., and Del Vecchio, D. (2017). Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology*, 6(7):1263–1272. PMID: 28350160.

Ramirez, M., Wek, R. C., and Hinnebusch, A. G. (1991). Ribosome association of gcn2 protein kinase, a translational activator of the gcn4 gene of saccharomyces cerevisiae. *Molecular and cellular biology*, 11:3027–3036.

Roy, A., Goberman, D., and Pugatch, R. (2021). A unifying autocatalytic network-based framework for bacterial growth laws. *Proceedings of the National Academy of Sciences of the United States of America*, 118.

Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced grown of salmonella typhimurium. *J. Gen. Microbiol.*, 19(3):592–606.

Scott, M. and Hwa, T. (2011). Bacterial growth laws and their applications. *Current opinion in biotechnology*, 22:559–565.

Scott, M., Klumpp, S., Mateescu, E. M., and Hwa, T. (2014). Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Molecular systems biology*, 10:747.

Scott, M., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of Cell Growth Origins and Consequences. *Science*, 330(November):1099–1102.

Serbanescu, D., Ojkic, N., and Banerjee, S. (2020). Nutrient-dependent trade-offs between ribosomes and division protein synthesis control bacterial cell size and growth. *Cell Reports*, 32(12):108183.

Shachrai, I., Zaslaver, A., Alon, U., and Dekel, E. (2010). Cost of unneeded proteins in e. coli is reduced after several generations in exponential growth. *Molecular cell*, 38:758–767.

Shahrezaei, V. and Marguerat, S. (2015). Connecting growth with gene expression: of noise and numbers. *Current opinion in microbiology*, 25:127–135.

Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017). Invariance of initiation mass and predictability of cell size in escherichia coli. *Current Biology*, 27(9):1278–1287.

Swaffer, M. P., Kim, J., Chandler-Brown, D., Langhinrichs, M., Marinov, G. K., Greenleaf, W. J., Kundaje, A., Schmoller, K. M., and Skotheim, J. M. (2021). Transcriptional and chromatin-based partitioning mechanisms uncouple protein scaling from cell size. *Molecular cell*, 81:4861–4875.e7.

Szavits-Nossan, J., Ciandrini, L., and Romano, M. C. (2018). Deciphering mRNA sequence determinants of protein production rate. *Physical Review Letters*, 120:128101. arXiv: 1708.07678 tex.arxivid: 1708.07678.

Weiße, A. Y., Oyarzún, D. A., Danos, V., and Swain, P. S. (2015). Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, 112(9).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533181; this version posted March 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Wu, C., Balakrishnan, R., Braniff, N., Mori, M., Manzanarez, G., Zhang, Z., and Hwa, T. (2022). Cellular perception of growth rate and the mechanistic origin of bacterial growth law. *Proceedings of the National Academy of Sciences of the United States of America*, 119:e2201585119.

You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C. W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of bacterial proteome with metabolism by cyclic amp signalling. *Nature*, 500:301–306.

Zhang, Q., Brambilla, E., Li, R., Shi, H., Cosentino Lagomarsino, M., and Sclavi, B. (2020). A decrease in transcription capacity limits growth rate upon translation inhibition. *mSystems*, 5.

Zhu, J. and Thompson, C. B. (2019). Metabolic regulation of cell growth and proliferation. *Nature reviews. Molecular cell biology*, 20:436–450.

Zhu, M., Pan, Y., and Dai, X. (2019). (p)ppgpp: the magic governor of bacterial growth economy. *Current genetics*, 65:1121–1125.