Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*

Sarah Kostinski\(^{1,*}\) and Shlomi Reuveni\(^{1,2,†}\)

\(^1\)School of Chemistry, Center for the Physics & Chemistry of Living Systems, Tel Aviv University, 6997801 Tel Aviv, Israel
\(^2\)Sackler Center for Computational Molecular & Materials Science, Ratner Institute for Single Molecule Chemistry, Tel Aviv University, 6997801 Tel Aviv, Israel

(Received 24 December 2019; accepted 14 May 2020; published 8 July 2020)

Bacterial ribosomes are composed of one-third protein and two-thirds RNA by mass. The predominance of RNA is often attributed to a primordial RNA world, but why exactly two-thirds remains a long-standing mystery. Here we present a quantitative analysis, based on the kinetics of ribosome self-replication, demonstrating that the 1:2 protein-to-RNA mass ratio uniquely maximizes cellular growth rates in *E. coli*. A previously unrecognized growth law, and an invariant of bacterial growth, also follow from our analysis. The growth law reveals that the ratio between the number of ribosomes and the number of polymerases making ribosomal RNA is proportional to the cellular doubling time. The invariant is conserved across growth conditions and specifies how key microscopic parameters in the cell, such as transcription and translation rates, are coupled to cellular physiology. Quantitative predictions from the growth law and invariant are shown to be in excellent agreement with *E. coli* data despite having no fitting parameters. Our analysis can be readily extended to other bacteria once data become available.

Why does the bacterial ribosome feature a 1:2 protein-to-RNA mass ratio [1]? Ribosomes synthesize all proteins in the cell, including their own and those of their RNA-synthesizers [2]. Cellular growth is limited by ribosome self-replication since cells can only double as fast as their ribosomes [3,4]. The limit imposed by ribosomal-protein (r-protein) production was previously used to explain the proportionality between growth rates and r-protein fractions in bacteria [5–9] and yeast [10]. The same limit was also used to explain central features of the ribosome [11]. In particular, it was shown that a protein-poor, i.e., RNA-rich, ribosome allows cells to grow faster. However, the production of ribosomal-RNA (rRNA) may also limit cellular growth, which leads to a trade-off. Here we show that accounting for this trade-off explains the exact protein-to-RNA mass ratio in the ribosome.

Two autocatalytic loops are required to describe ribosome self-replication (Fig. 1): One generates r-protein, while the other generates RNA. In the former, ribosomes directly synthesize the r-proteins required for their own making. In the other autocatalytic loop, rRNA is generated by RNA polymerases (RNAPs), whose proteins are synthesized by ribosomes. Here we obtain two bounds on the cellular growth rate based on r-protein and rRNA production. We show that maximization of cellular growth rates in *E. coli*, subject to these bounds, is uniquely attained when RNA constitutes two-thirds the ribosome mass. Furthermore, we demonstrate that *E. coli* in fact achieves the maximal growth rates permitted by these bounds, which allows them to be recast as simple quantitative laws with strong predictive power.

The autocatalytic process of r-protein production (right loop, Fig. 1) is described by [12]

\[
\frac{dr\text{-}protein}{dt} = k_{\text{ribo}} \phi_{\text{ribo}}^{r\text{-}prot} f_{\text{active}}^{\text{ribo}} N_{\text{ribo}},
\]

where \(N_{\text{ribo}}\) is the total number of ribosomes in the cell, \(f_{\text{active}}^{\text{ribo}}\) the fraction of ribosomes actively engaged in translation, \(\phi_{\text{ribo}}^{r\text{-}prot}\) the time fraction an active ribosome spends synthesizing r-proteins (or, equivalently, the fraction of active ribosomes synthesizing r-proteins, see Supplemental FIG. 1. Autocatalytic production of the ribosome. Bacterial ribosomes are composed of two-thirds RNA (gray) and one-third protein (blue) by mass. r-protein is synthesized directly by ribosomes, as illustrated by the blue autocatalytic loop on the right. rRNA is synthesized by RNA polymerases, symbolized by the top left gray arrow. RNA polymerases, in turn, are made of protein that is synthesized by ribosomes (bottom left blue arrow). These autocatalytic processes impose counteracting bounds on cellular growth. In *E. coli* they are mutually satisfied when protein constitutes one-third the ribosome mass, thereby maximizing growth rates in various conditions.

DOI: 10.1103/PhysRevLett.125.028103
An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires obtaining from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.

An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.

An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.

An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.

An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.

An additional bound on the cellular growth rate can be obtained from the autocatalytic production of rRNA (left loop, Fig. 1). Describing the synthesis of rRNA requires solving for the doubling rate of rRNA[12] and ribosome and \( N_{\text{ribo}} \) are the number of r-proteins and rRNA per ribosome, and solve Eq. (1) for the doubling rate of r-protein [12]. To sustain balanced growth, cells must on average double their ribosomes before division, leading to an upper bound on the cellular growth rate

\[
\mu \equiv \frac{\ln(2)}{T_d} \leq \frac{k_{\text{ribo}} \phi_{\text{ribo}} f_{\text{active}}}{N_{\text{a.a.}}_{\text{ribo}}},
\]

(2)

with \( T_d \) standing for the cellular doubling time.
from 12 to 22 a.a./sec. Conversely, RNAP activity \( f_{\text{RNAP}}^{\text{active}} \) ranges from 13.2% to 31.0% while transcription rate \( k_{\text{RNAP}} \) remains approximately constant at 85 nts/sec. In balanced growth, \( \phi_{\text{ribo}}^{\text{r-prot}} \) and \( \phi_{\text{ribo}}^{\text{RNA}} \) correspond to their respective proteome mass fractions (r-protein/total protein) and (RNAP-protein/total protein), which range between 7.8%–23.1% and 0.93%–1.66% (Supplemental Material [12]). The number of amino acids and nucleotides per ribosome is known from crystallographic data and are given by \( N_{\text{nucl}} = 7536 \) and \( N_{\text{RNAP}} = 3498 \) excepting the \( \sigma^70 \) factor [21].

Note, however, that including \( \sigma^70 \) self-consistently, i.e., using \( N_{\text{RNAP}}^{\text{a.a.}} = 4111 \) and its corresponding \( \phi_{\text{RNAP}}^{\text{a.a.}} \), does not change the results which follow (Table S1 [12]). Finally, note that the number of nucleotides in a mature ribosome, \( 4566 \) nts [3], is less than that in its rRNA precursor (Fig. 3). The fraction of active RNAPs synthesizing the rRNA precursor and tRNAs ranges from 24% to 86%. Omitting those RNAPs engaged in the synthesis of unstable flanking nucleotides and tRNAs, we find that \( \phi_{\text{RNA}}^{\text{RNAP}} \) ranges from 18% to 65% (Table S1 [12]).

Using the parameter values described above, we plot the bounds of Eqs. (2) and (5) for \( E. coli \). Shown in Fig. 4(a) are six pairs of color-coded bounds, corresponding to six different growth conditions. In Fig. 4(b) we plot the minimum of the two bounds for each growth condition, which gives the envelope of allowed growth rates. The bounds intersect at the largest growth rate accessible to \( E. coli \) in each medium; these points are highlighted by black-rimmed circles. Remarkably all maximal growth rates are achieved at a protein mass fraction of about one-third, thus matching \( E. coli \)’s actual ribosome composition.

More information is gained upon comparing the upper bounds to real growth rates. For example, consider the Luria-Bertani growth media, corresponding to the top gray curve in Fig. 4(b). The bounds predict a maximal possible growth rate of \( 2.1 \text{ h}^{-1} \), which is identical to the value obtained in experiments. An almost perfect match between theory and experiment is also seen for all other growth conditions, suggesting that Eqs. (2) and (5) can, in fact, be treated as approximate equalities. Slight rearrangement then allows us to write Eq. (2) as

\[
\text{fraction of ribosomes making r-protein} \approx t_{\text{r-prot}} \mu, \quad (6)
\]

where we have identified \( f_{\text{ribo}}^{\text{active}} \phi_{\text{ribo}}^{\text{r-prot}} \) as the fraction of ribosomes that are actively engaged in r-protein synthesis, and \( t_{\text{r-prot}} = N_{\text{ribo}}^{\text{active}} / k_{\text{ribo}} \) as the time a ribosome takes to synthesize a full set of r-proteins (Supplemental Material [12]). The proportionality between the r-proteome fraction \( \phi_{\text{ribo}}^{\text{r-prot}} \), and the cellular growth rate \( \mu \), is now widely recognized as a “bacterial growth law” [23] but Eq. (6) allows for a clean, intercept-free representation that matches well with data [Fig. 4(c)].

Similarly, Eq. (5) can also be treated as an equality and combined with Eq. (6) to give (Supplemental Material [12])

\[
\frac{\text{number of RNAPs making rRNA}}{\text{number of ribosomes}} \approx \frac{\tau_{\text{RNA}}}{\mu}, \quad (7)
\]

where \( \tau_{\text{RNA}} \approx N_{\text{nucl}} / k_{\text{RNAP}} \) is the time an RNAP takes to synthesize a full set of rRNAs. Note that in \( E. coli \), \( k_{\text{RNAP}} \) is constant at 85 nt/sec, which makes \( \tau_{\text{RNA}} \approx 1 \text{ min independent of growth media} \) (Table S1 [12]). Thus, at a growth rate of, e.g., \( 1 \text{ h}^{-1} \), \( E. coli \) requires just a single RNA-making RNAP per sixty ribosomes.

More generally, the number of rRNA-making RNAPs per ribosome is approximately given by the numerical value of the growth rate \( \mu \) when the latter is measured in min\(^{-1}\). Equation (7), which yields this remarkably simple rule of thumb, is in almost perfect agreement with available data [Fig. 4(d)].

Having shown that Eqs. (2) and (5) are approximate equalities, we set their right-hand sides equal to altogether eliminate growth rate. In doing so we find [12]

\[
\frac{x_{\text{opt}}^2}{1 - x_{\text{opt}}} \approx \frac{m_{\text{a.a}}}{M_{\text{nucl}}} M_{\text{ribo}} \phi_{\text{ribo}}^{\text{r-prot}} \left[ k_{\text{ribo}} / \phi_{\text{ribo}}^{\text{active}} \right]. \quad (8)
\]

where \( x_{\text{opt}} \) is the optimal protein mass fraction found in bacterial ribosomes, \( m_{\text{a.a}} \) and \( M_{\text{nucl}} \) are the average masses of a nucleotide and an amino acid, respectively, \( M_{\text{ribo}} \) and \( M_{\text{RNAP}} \) are the masses of the ribosome and RNAP, respectively, and all other parameters are defined as before. Note that \( x_{\text{opt}} \) depends on the relative—but not absolute—masses of the ribosome and RNAP. Furthermore, the ratio enclosed in square brackets on the right-hand side of Eq. (8) demonstrates a clean and symmetric separation between
translation and transcription kinetics. The numerator contains exclusively translation-associated parameters, whereas the denominator features exclusively transcription-associated parameters.

Note that while many parameters on the right-hand side of Eq. (8) have a nontrivial dependence on growth conditions, the left-hand side is fixed since $x_{\text{opt}} \approx 0.36$ in E. coli (Supplemental Material [12]). The right-hand side of Eq. (8) is thus predicted to have a value of $\approx 1/5$ irrespective of growth conditions. This prediction is verified for E. coli in Fig. 4(e). From an experimental point of view, this allows for indirect measurement of a variety of cellular parameters. For example, if RNAP activity in the cell is not known, it can be inferred via measurement of other parameters in the formula. While ribosome composition may differ slightly across bacterial species, the invariance predicted by Eq. (8) is expected to hold even for bacteria other than E. coli, where $x_{\text{opt}}$ denotes the protein mass fraction of their corresponding ribosome.

Equations (7) and (8), which are in excellent agreement with E. coli data despite having no fitting parameters, significantly advance the current body of knowledge on quantitative bacterial physiology [23–28]. Specifically, Eq. (7) adds to existing growth laws [23,29–32] by quantifying the growth-rate dependent coupling between protein- and rRNA-making molecules in the cell. Complementing it is Eq. (8) which, to our knowledge, is the first invariant of bacterial growth to be reported. The latter reveals that a nontrivial combination of molecular and cellular parameters remains constant across growth conditions.

Changing nutrient quality of growth media, as was done to obtain the data set of Fig. 4, is just one method to modulate cellular growth. In Ref. [7], for example, the authors used antibiotic treatment to explore the interplay between metabolism and autocatalytic production of r-protein. As antibiotic concentration increased, the growth rate decreased despite a larger r-protein proteome fraction.

FIG. 4. Ribosome composition maximizes cellular growth rates in E. coli. (a) The r-protein and rRNA bounds from Eqs. (2) and (5) are plotted as a function of the protein mass fraction of the ribosome, for E. coli in six different growth media. Data are publicly available [6,20] (Table S1 [12]). (b) The envelopes of allowed growth rates as obtained from the minimum of the two bounds in (a). The maximal growth rate allowed occurs at the cusp, i.e., where bounds intersect. It can be seen that this always happens at a protein mass fraction of about one-third which closely resembles that of the bacterial ribosome (dashed black line). (c) The theoretical prediction of Eq. (6) (solid line) is in excellent agreement with several E. coli datasets [5–7] (symbols, Tables S1–S3 [12]). (d) The theoretical prediction of Eq. (7) (solid line), is in excellent agreement with data for E. coli grown in six different media (black-rimmed circles) [6]. (e) Agreement between the left-hand side (solid black line) and right-hand side (black-rimmed circles) of Eq. (8) is shown for E. coli data. As expected, invariance with respect to growth conditions is observed.
which seems to come at the expense of metabolic proteins. The accompanying effect on transcription-associated parameters is not known; such measurements would provide additional insight into parameter covariation in Eqs. (7) and (8).

Thus far we have shown that autocatalytic production of ribosomes couples ribosome composition to certain cellular parameters. However, to what extent this coupling impacted the ribosome composition over the course of evolution is an open question; it could be argued that the ribosome evolved to its current form irrespective of its autocatalytic nature. One naturally occurring example that suggests the contrary is the non-autocatalytic mitochondrial ribosome in Eukarya [1,17–19,33–35]. The mitochondrial ribosome is thought to have originated from the bacterial ribosome, and then evolved to have a protein, rather than RNA, rich composition. While this suggests that autocatalysis significantly impacted ribosome composition, direct experimental proof is lacking. To that end, we propose a controlled evolution experiment on synthetic ribosomes in bacteria.

In E. coli, synthetic ribosomes [36,37] such as stapled ribosomes and Ribo-T [38–41] were recently implemented to form an independent (orthogonal) translation system that is capable of sustaining protein synthesis and cellular growth. Orthogonal ribosomes could be made nonautocatalytic by selecting the set of (nonribosomal) proteins they synthesize via custom engineering of their anti-Shine-Dalgarno sequence and corresponding mRNA Shine-Dalgarno sequence [42]. The nonautocatalytic ribosomes could then be independently evolved in vivo alongside wild-type ribosomes. While challenging, such lab-evolution experiments could determine the extent to which autocatalysis produced impacts ribosome composition.

Finally, it should be stressed that our analysis is not specific to E. coli and we anticipate other bacteria to obey the same relations. One such candidate is the bacterium A. aerogenes, for which the proportionality between r-protein proteome fractions and cellular growth rates in Eq. (6) was already verified. In fact, the same relation was also verified for C. utilis, N. crassa, E. gracilis [7], and S. cerevisiae [10], suggesting that the relations derived herein could also be extended to Eukarya. The analysis presented above would then require modifications to account for the different types of RNA polymerases found there. Following these modifications, we expect that ribosome composition in higher-level organisms could also be understood from simple kinetic considerations of cellular growth and self-replication.

S. K. and S. R. thank Moshe Kol, Michael Urbakh, Roy Beck-Barkai, and Yuval Scher for reading and commenting on early versions of the manuscript. S. K. was supported by the Zuckerman STEM postdoctoral fellowship and the U.S. Fulbright Program in Israel. S. R. acknowledges support from the Azrieli Foundation, from the Raymond and Beverly Sackler Center for Computational Molecular and Materials Science at Tel Aviv University, and from the Israel Science Foundation (Grant No. 394/19).

*skostinski@tauex.tau.ac.il
'shlomire@tauex.tau.ac.il


