

Fundamental limits on the rate of bacterial cell division

Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G. Garcia⁴, Jane Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}

¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA; ³Department of Chemical Engineering, Stanford University, Stanford, CA, USA; ⁴Department of Molecular Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA, USA; ⁵Department of Physics, Brandeis University, Waltham, MA, USA; ⁶Department of Applied Physics, Stanford University, Stanford, CA, USA; ⁷Allen Institute for Cell Science, Seattle, WA, USA; ⁸Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA; ⁹Department of Physics, California Institute of Technology, Pasadena, CA, USA; *Co-corresponding authors. Address correspondence to phillips@pboc.caltech.edu and jtheriot@uw.edu; [†]These authors contributed equally to this work

14

Abstract Recent years have seen a deluge of experiments dissecting the relationship between bacterial growth rate, cell size, and protein content, quantifying the abundance of proteins across growth conditions with unprecedented resolution. However, we still lack a rigorous understanding of what sets the scale of these quantities and when protein abundances should (or should not) depend on growth rate. Here, we seek to quantitatively understand this relationship across a collection of *Escherichia coli* proteomic data covering ≈ 4000 proteins and 36 growth rates. We estimate the basic requirements for steady-state growth by considering key processes in nutrient transport, energy generation, cell envelope biogenesis, and the central dogma. From these estimates, ribosome biogenesis emerges as a primary determinant of growth rate. We expand on this assessment by exploring a model of proteomic regulation as a function of the nutrient supply, revealing a mechanism that ties cell size and growth rate to ribosomal content.

25

Introduction

The observed range of bacterial growth rates is enormously diverse. In natural environments, some microbial organisms may double only once per year (*Mikucki et al., 2009*) while in comfortable laboratory conditions, growth can be rapid with several divisions per hour (*Schaechter et al., 1958*). This six order-of-magnitude difference in time scales encompasses different microbial species and lifestyles, yet even for a single species such as *Escherichia coli*, the growth rate can be modulated over a comparably large scale by tuning the type and amount of nutrients in the growth medium. This remarkable flexibility in growth rate illustrates the intimate relationship between environmental conditions and the rates at which cells convert nutrients into new cellular material – a relationship that has remained a major topic of inquiry in bacterial physiology for over a century (*Jun et al., 2018*).

Jacques Monod once remarked that “the study of the growth of bacterial cultures does not constitute a specialized subject or branch of research, it is the basic method of Microbiology.” Those words ring as true today as they did when they were written 70 years ago (*Monod, 1949*) with the quantitative power of this “method” recently undergoing renaissance. Many of the key questions addressed by the pioneering efforts in the middle of the last century can be revisited by examining them through the lens of the increasingly refined molecular census that is available for bacteria such as the microbial workhorse *E. coli*.

Several of the evergreen questions about bacterial growth and physiology that were originally raised by

42 microbiologists in the middle of the 20th century can now be reframed in light of this newly available data. For
43 example, what biological processes are the primary determinants for how quickly bacterial cells can grow and
44 reproduce? How do cells modulate the absolute numbers and relative ratios of their molecular constituents as
45 a function of changes in growth rate or nutrient availability? In this paper, we begin by considering these two
46 questions from two distinct angles. First, as a result of an array of high-quality proteome-wide measurements of *E.*
47 *coli* under diverse growth conditions, we have a census that allows us to explore how the number of key molecular
48 players change as a function of growth rate. Here, we have assembled a singular data set using measurements
49 collected over the past decade via mass spectrometry (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al.,
50 2013*) or ribosomal profiling (*Li et al., 2014*) of the composition of the *E. coli* proteome across 36 unique growth
51 rates (see Appendix Experimental Details Behind Proteomic Data for a further discussion of the data). Second,
52 by compiling molecular turnover rate measurements for many of the fundamental processes associated with
53 bacterial growth, we make quantitative estimates of key cellular processes (schematized in *Figure 1*) to determine
54 whether our current understanding of the dynamics of these processes are sufficient to explain the magnitude of
55 the observed protein copy numbers across conditions (see *Box 1* describing the philosophy behind this approach).
56 The census, combined with these estimates, provide a window into the question of whether the rates of central
57 processes such as energy generation or DNA synthesis are regulated systematically as a function of cell growth
58 rate by altering protein copy number.

59 Throughout our estimates, we consider an archetypal growth rate of $\approx 0.5 \text{ hr}^{-1}$ corresponding to a doubling
60 time of ≈ 5000 seconds, as the data sets examined here heavily sample this growth regime. While we formulate
61 point estimates for the protein abundances at this division time, we also consider how these values will vary
62 at other growth rates due to changes in cell size, surface area, and chromosome copy number (*Taheri-Araghi
63 et al., 2015; Harris and Theriot, 2018*). For the majority of the processes considered, we find that the protein copy
64 numbers appear tuned for the task of cell doubling across a continuum of growth rates. Thus, our understanding
65 of the kinetics of myriad biological processes is sufficient to quantitatively explain the observed abundances of
66 these proteins.

67 From these estimates, it emerges that translation, particularly the synthesis of ribosomal proteins, is a plausible
68 candidate that limits the rate of cell division in *E. coli*. We reach this conclusion by considering that ribosome
69 synthesis is 1) a rate limiting step for the *fastest* bacterial division, and 2) a major determinant of bacterial growth
70 across the nutrient conditions we have considered under steady state, exponential growth. This enables us to
71 suggest that the long-observed correlation between growth rate and cell size (*Schaechter et al., 1958; Si et al.,
72 2017*) can be simply attributed to the increased absolute number of ribosomes per cell under conditions supporting
73 extremely rapid growth. To better understand how the observed alterations in absolute protein abundances, and
74 in particular, changes in ribosome copy number, influence growth rate across different nutrient conditions we
75 consider a minimal model of cellular growth. Our conclusions from these analyses provide important insight
76 into how *E. coli* regulates growth across conditions of differing nutrient availability and identifies fundamental
77 constraints in bacterial growth more broadly.

118 Nutrient Transport

119 We begin by considering the critical transport processes diagrammed in *Figure 1(A)*. In order to build new cellular
120 mass, the molecular and elemental building blocks must be scavenged from the environment in different forms.
121 Carbon, for example, is acquired via the transport of carbohydrates and sugar alcohols with some carbon sources
122 receiving preferential treatment in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the
123 other hand, are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate, and ammonia (*Jun
124 et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al., 1977; Willsky et al.,
125 1973*). All of these compounds have different membrane permeabilities (*Phillips, 2018*) and most require some
126 energetic investment either via ATP hydrolysis or through the proton electrochemical gradient to bring the material
127 across the hydrophobic cell membrane.

128 The elemental composition of *E. coli* has received much quantitative attention over the past half century
129 (*Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv, 1976*), providing us with
130 a starting point for estimating how many atoms of each element must be scavenged from the environment. A
131 synthesis of these studies presents an approximate dry mass composition of $\approx 50\%$ carbon (BNID: 100649; see

79 **Box 1. The Rules of Engagement for Order-Of-Magnitude Estimates**

80 This work relies heavily on "back-of-the-envelope" estimates to understand the growth-rate dependent
81 abundances of molecular complexes. This moniker arises from the limitation that any estimate should be
82 able to fit on the back of a postage envelope. As such, we must draw a set of rules governing our precision
83 and sources of key values.

84 **The rule of "one, few, and ten".** The philosophy behind order-of-magnitude estimates is to provide a estimate
85 of the appropriate scale, not a prediction with many significant digits. We therefore define three different
86 scales of precision in making estimates. The scale of "one" is reserved for values that range between 1 and 2.
87 For example, If a particular process has been experimentally measured to transport 1.87 protons for a process
88 to occur, we approximate this process to require 2 protons per event. The scale of "few" is reserved for values
89 ranging between 3 and 7. For example, we will often use Avogadro's number to compute the number of
90 molecules in a cell given a concentration and a volume. Rather than using Avogadro's number as 6.02214×10^{23} ,
91 we will approximate it as 5×10^{23} . Finally, the scale of "ten" is reserved for values which we know within an
92 order of magnitude. If a particular protein complex is present at 883 copies per cell, we say that it is present
93 in approximately 10^3 copies per cell. These different scales will be used to arrive at simple estimates that
94 report the expected scale of the observed data. Therefore, the estimates presented here should not be viewed
95 as hard-and-fast predictions of precise copy numbers, but as approximate lower (or upper) bounds for the
96 number of complexes that may be needed to satisfy some cellular requirement.

97 Furthermore, we use equality symbols (=) sparingly and frequently defer to approximation (\approx) or scaling (\sim)
98 symbols when reporting an estimate. When \approx is used, we are implicitly stating that we are confident in this
99 estimate within a factor of a few. When a scaling symbol \sim is used, we are stating that we are confident in our
100 estimate to within an order of magnitude.

101 **The BioNumbers Database as a source for values.** In making our estimates, we often require approximate
102 values for key cellular properties, such as the elemental composition of the cell, the average dry mass, or
103 approximate rates of synthesis. We rely heavily on the BioNumbers Database (bionumbers.hms.harvard.edu,
104 *Milo et al. (2010)*) as a repository for such information. Every value we draw from this database has an
105 associated BioNumbers ID number, abbreviated as BNID, and we provide this reference in grey-boxes in each
106 figure.

107 **Uncertainty in the data sets and the accuracy of an estimate.** The data sets presented in this work are the
108 products of careful experimentation with the aim to report, to the best of their ability, the absolute copy
109 numbers of proteins in the cell. These data, collected over the span of a few years, come from different
110 labs and use different internal standards, controls, and even techniques (discussed further in Appendix
111 Experimental Details Behind Proteomic Data). As a result, there is notable disagreement in the measured copy
112 numbers for some complexes across data sets. In assessing whether our estimates could explain the observed
113 scales and growth-rate dependencies, we also considered the degree of variation between the different data
114 sets. For example, say a particular estimate undercuts the observed data by an order of magnitude. If all
115 data sets agree within a factor of a few of each other, we revisit our estimate and consider what me may
116 have missed. However, if the data sets themselves disagree by an order of magnitude, we determine that our
117 estimate is appropriate given the variation in the data.

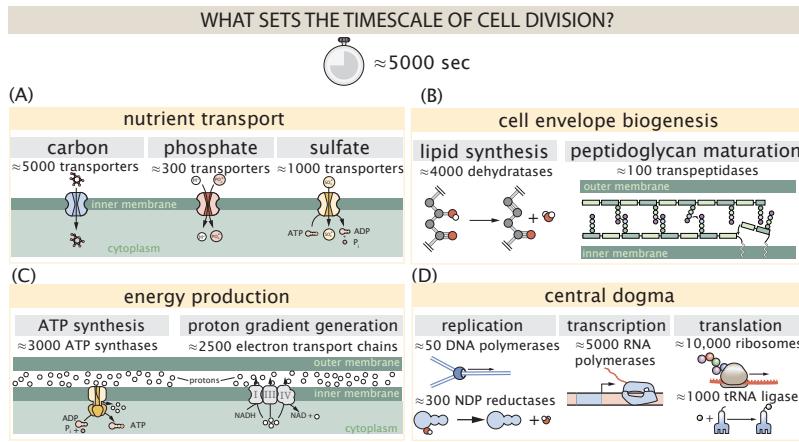


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) cell envelope biogenesis, (C) energy production (namely, ATP synthesis), and (D) processes associated with the central dogma. Numbers shown are the approximate number of complexes of each type observed at a growth rate of 0.5 hr^{-1} , or a cell doubling time of $\approx 5000 \text{ s}$.

132 **Box 1** for explanation of BNID references), $\approx 15\%$ nitrogen (BNID: 106666), $\approx 3\%$ phosphorus (BNID: 100653), and
 133 1% sulfur (BNID: 100655) with remainder being attributable to oxygen, hydrogen, and various transition metals.
 134 We use this stoichiometric breakdown to estimate the abundance and growth rate dependence of a variety of
 135 transporters responsible for carbon uptake, and provide more extensive investigation of the other critical elements
 136 – phosphorus, sulfur, and nitrogen – in the Appendix Additional Estimates of Fundamental Biological Processes.

137 Using $\approx 0.3 \text{ pg}$ as the typical *E. coli* dry mass at a growth rate of $\approx 0.5 \text{ hr}^{-1}$ (BNID: 103904), coupled with an
 138 approximation that $\approx 50\%$ of this mass is carbon, we estimate that $\sim 10^{10}$ carbon atoms must be brought into
 139 the cell in order to double all of the carbon-containing molecules (**Figure 2(A)**, top). Typical laboratory growth
 140 conditions provide carbon as a single class of sugar (such as glucose, galactose, or xylose) often transported cross
 141 the cell membrane by a transporter complex specific to that particular sugar. One such mechanism of transport
 142 is via the PTS system, which is a highly modular system capable of transporting a diverse range of sugars with
 143 high specificity (*Escalante et al., 2012*). The glucose-specific component of this system transports ≈ 200 glucose
 144 molecules (≈ 1200 carbon atoms) per second per transporter (BNID: 114686). Making the assumption that this is a
 145 typical sugar transport rate for the PTS system, coupled with the need to transport $\sim 10^{10}$ carbon atoms, we then
 146 expect on the order of ≈ 1000 transporters must be expressed per cell in order to bring in enough carbon atoms
 147 (**Figure 2(A)**, top).

148 We find, however, that this estimate is exceeded by several fold by experimental measurements (**Figure 2(A)**,
 149 bottom), implying that the cell is capable of transporting more carbon atoms than strictly needed for biosynthesis.
 150 We can also abstract this calculation to consider any particular growth rate given knowledge of the cell density and
 151 volume as a function of growth rate and direct the reader to the Appendix Extending Estimates to a Continuum
 152 of Growth Rates for more information. This abstraction, shown as a grey line in **Figure 2(A)**, reveals an excess of
 153 transporters even at faster growth rates. This contrasts with our observations for uptake of phosphorus and sulfur,
 154 which turn out to align well with our expectations across different growth conditions (**Figure 2–Figure Supplement 1**
 155 and discussed further in Appendix Additional Estimates of Fundamental Biological Processes).

156 It is important to note that so far we have neglected any specifics of the regulation of the carbon transport
 157 system. Using the diverse array of growth conditions available in the data, we can explore how individual carbon
 158 transport systems depend on specific carbon availability. In **Figure 2(B)**, we show the total number of carbohydrate
 159 transporters specific to different carbon sources. A striking observation, shown in the top-left plot of **Figure 2(B)**, is
 160 the constancy in the expression of the glucose-specific transport systems, an observation that stands in contrast
 161 with other species of transporters. Additionally, we note that the total number of glucose-specific transporters is
 162 tightly distributed at $\approx 10^4$ per cell, the approximate number of transporters needed to sustain rapid growth of
 163 several divisions per hour. This illustrates that *E. coli* maintains a substantial number of complexes present for

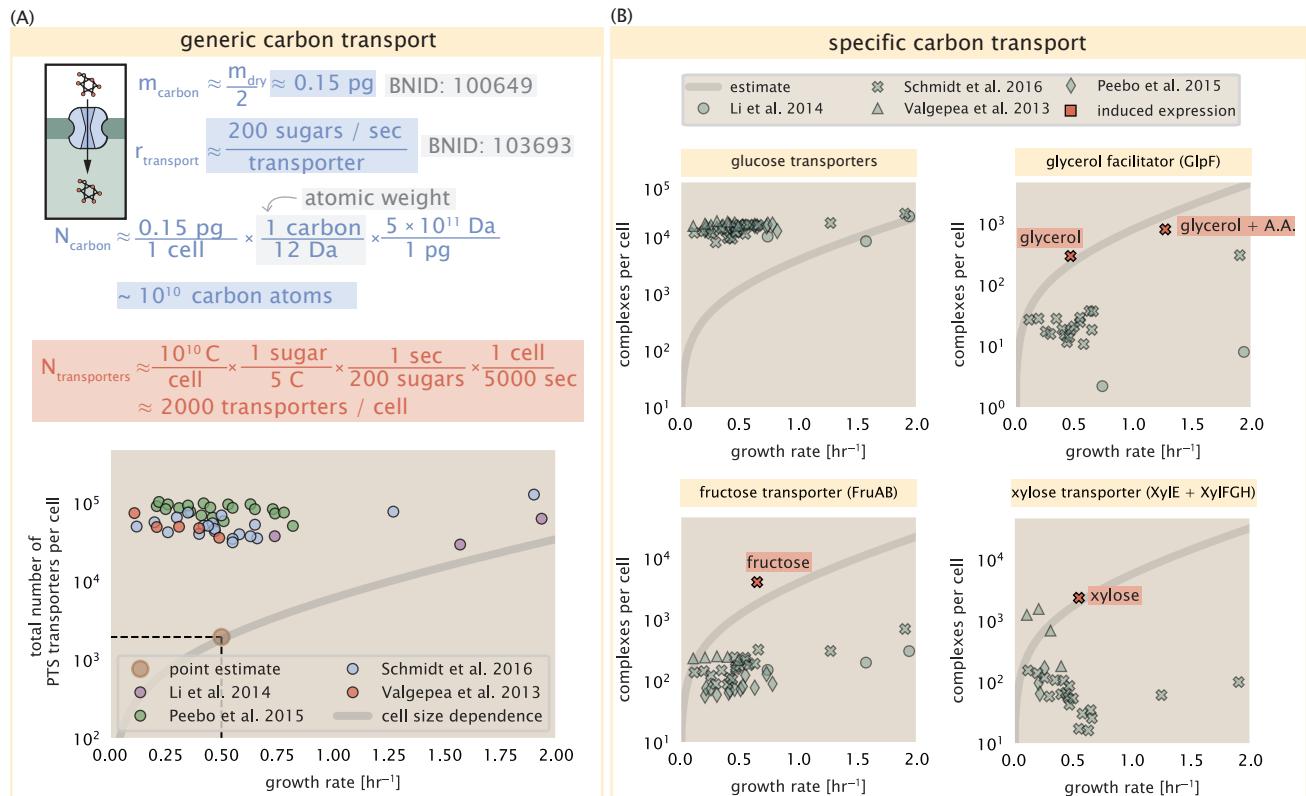


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules at a rate of ≈ 200 per second. Bottom plot shows the estimated number of transporters needed at a growth rate of ≈ 0.5 per hr (light-brown point and dashed lines). Colored points correspond to the mean number of complexes involved in carbohydrate import (complexes annotated with the Gene Ontology terms GO:0009401 and GO:0098704) for different growth conditions across different published datasets. (B) The abundance of various specific carbon transport systems plotted as a function of the population growth rate. The rates of substrate transport differ between these transporter species. To compute the continuum growth rate estimate (grey line), we used the following transport rates for each transporter species: 200 glucose- s^{-1} (BNID: 103693), 2000 glycerol- s^{-1} (*Lu et al., 2003*), 200 fructose- s^{-1} (assumed to be similar to PtsI, BNID: 103693), and 50 xylose- s^{-1} (assumed to be comparable to LacY, BNID:103159). Red points and highlighted text indicate conditions in which the only source of carbon in the growth medium induces expression of the transport system. Grey lines in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

Figure 2-Figure supplement 1. Estimates and observed abundances of phosphate and sulfate transporters.

164 transporting glucose regardless of growth condition, which is known to be the preferential carbon source (*Monod*,
165 *1947; Liu et al., 2005; Adelberg et al., 2014*).

166 Many metabolic operons are regulated with dual-input logic gates that are only expressed when glucose
167 concentrations are low and the concentration of other carbon sources are elevated (*Gama-Castro et al., 2016;*
168 *Zhang et al., 2014b; Gama-Castro et al., 2016; Belliveau et al., 2018; Ireland et al., 2020*). Points colored in red in
169 *Figure 2(B)* (labeled by red text-boxes) correspond to growth conditions in which the specific carbon source (glycerol,
170 xylose, or fructose) is present as the sole source of carbon. The grey lines in *Figure 2(B)* show the estimated number
171 of transporters needed at each growth rate to satisfy the cellular carbon requirement, adjusted for the specific
172 carbon source in terms of number of carbon atoms per molecule and the rate of transport for the particular
173 transporter species. These plots show that, even in the absence of the particular carbon source, expression of the
174 transporters is maintained on the order of $\sim 10^2$ per cell. The low but non-zero abundances may reflect the specific
175 regulatory logic involved, requiring that cells are able to transport some minimal amount of an alternative carbon
176 source in order to induce expression of these alternative carbon-source systems.

177 **Limits on Transporter Expression**

178 If acquisition of nutrients was a limiting process in cell division under the typical growth conditions explored here,
179 the growth rate could be theoretically increased simply by expressing more transporters, but is this feasible at a
180 physiological level? A way to approach this question is to compute the amount of space in the bacterial membrane
181 that could be occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of about
182 $5 \mu\text{m}^2$ (BNID: 101792), we expect an areal density for 1000 transporters to be approximately 200 transporters/ μm^2 .
183 For a typical transporter occupying about 50 nm^2 , this amounts to about only $\approx 1\%$ of the total inner membrane
184 surface area (*Szenk et al., 2017*). Bacterial cell membranes typically have densities of 10^5 proteins/ μm^2 (*Phillips*,
185 *2018*), with roughly 60 % of the surface area occupied by protein (BNID: 100078), implying that the cell could easily
186 accommodate more transporters. There are, however, additional constraints on the space that can be devoted to
187 nutrient uptake due to occupancy by proteins involved in processes like cell wall synthesis and energy production,
188 and we will consider this further in the coming sections.

189 **Cell Envelope Biogenesis**

190 In contrast to nutrient transporters, which support the synthesis of biomolecules throughout the cell and therefore
191 need to scale with the cell size, here we must consider the synthesis of components that will need to scale with
192 the surface area of the cell. *E. coli* is a rod-shaped bacterium with a remarkably robust length-to-width aspect
193 ratio of $\approx 4:1$ (*Harris and Theriot, 2018; Ojkic et al., 2019*). Assuming this surface area is approximately the same
194 between the inner and outer membranes of *E. coli*, and the fact that each membrane is itself a lipid bilayer, our
195 rule-of-thumb of $5 \mu\text{m}^2$ per surface suggests a total membrane surface area of $\approx 20 \mu\text{m}^2$ (see Appendix Estimation
196 of Cell Size and Surface Area for a description of the calculation of cell surface area as a function of cell size). In
197 this section, we will estimate the number of key protein complexes needed to synthesize the lipids as well as the
198 complexes involved in assembling the peptidoglycan scaffold that make up the cell envelope.

199 **Lipid Synthesis**

200 The dense packing of the membrane with proteins means that the cell membranes are not composed entirely of
201 lipid molecules, with only $\approx 40\%$ of the membrane area is occupied by lipids (BNID: 100078). Using a rule-of-thumb
202 of 0.5 nm^2 as the surface area of the typical lipid (BNID: 106993), we can estimate $\sim 2 \times 10^7$ lipids per cell, which is
203 in close agreement with experimental measurements (BNID: 100071, 102996).

204 The membranes of *E. coli* are composed of a variety of different lipids, each of which are unique in their
205 structures and biosynthetic pathways (*Sohlenkamp and Geiger, 2016*). Recently, a combination of stochastic kinetic
206 modeling (*Ruppe and Fox, 2018*) and *in vitro* kinetic measurements (*Ranganathan et al., 2012; Yu et al., 2011*)
207 have revealed remarkably slow steps in the fatty acid synthesis pathways which may serve as the rate limiting
208 reactions for making new membrane phospholipids in *E. coli*. One such step is the removal of hydroxyl groups
209 from the fatty-acid chain by ACP dehydratase that leads to the formation of carbon-carbon double bonds. This
210 reaction, catalyzed by proteins FabZ and FabA (*Yu et al., 2011*), have been estimated to have kinetic turnover
211 rates of ≈ 1 dehydration per second per enzyme (*Ruppe and Fox, 2018*). Thus, given this rate and the need to

212 synthesize $\approx 2 \times 10^7$ lipids over 5000 seconds, one can estimate that a typical cell requires ≈ 4000 ACP dehydratases.
213 This is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA (**Figure 3(A)**).
214 Furthermore, we can extend this estimate to account for the change in membrane surface area as a function of the
215 growth rate (grey line in **Figure 3(A)**), which captures the observed growth rate dependent expression of these two
216 enzymes.

217 Peptidoglycan Synthesis

218 Bacterial cells demonstrate exquisite control over their cell shape. This is primarily due to a stiff, several nanometer
219 thick meshwork of polymerized disaccharides that makes up the cell wall. The formation of the peptidoglycan
220 is an intricate process involving many macromolecular players (*Shi et al., 2018; Morgenstein et al., 2015*), whose
221 coordinated action maintains cell shape and integrity even in the face of large-scale perturbations (*Harris and*
222 *Theriot, 2018; Shi et al., 2018*). The peptidoglycan alone comprises $\approx 3\%$ of the cellular dry mass (BNID: 1019360),
223 making it the most massive molecule in *E. coli*. The polymerized unit of the peptidoglycan is a N-acetylglucosamine
224 and N-acetylmuramic acid disaccharide, of which the former is functionalized with a short pentapeptide. With a
225 mass of ≈ 1000 Da, this unit, which we refer to as a murein monomer, is polymerized to form long strands in the
226 periplasm which are then attached to each other via their peptide linkers. Together, these quantities provide an
227 estimate of $\approx 5 \times 10^6$ murein monomers per cell.

228 The crosslinking of the pentapeptides between adjacent glycan strands is responsible for maintaining the
229 structural integrity of the cell wall and, in principle, each murein monomer can be involved in such a crosslink.
230 In some microbes, such as in gram-positive bacterium *Staphylococcus aureus*, the extent of crosslinking can be
231 large with $> 90\%$ of pentapeptides forming a connection between glycan strands. In *E. coli*, however, a much
232 smaller proportion ($\approx 20\%$) of the peptides are crosslinked, resulting in a weaker and more porous cell wall (*Vollmer*
233 *et al., 2008; Rogers et al., 1980*). The formation of these crosslinks occurs primarily during the polymerization
234 of the murein monomers and is facilitated by a family of enzymes called transpeptidases. The four primary
235 transpeptidases of *E. coli* have only recently been quantitatively characterized *in vivo*, via liquid chromatography
236 mass spectrometry, which revealed a notably slow kinetic turnover rate of ≈ 2 crosslinking reactions formed per
237 second per enzyme (*Catherwood et al., 2020*).

238 Assembling these quantities permits us to make an estimate that on the order of ≈ 100 transpeptidases per cell
239 are needed for complete maturation of the peptidoglycan, given a division time of ≈ 5000 seconds; a value that is
240 comparable to experimental observations (**Figure 3(B)**). Expanding this estimate to account for the changing mass
241 of the peptidoglycan as a function of growth rate (grey line in **Figure 3(B)**) also qualitatively captures the observed
242 dependence in the data, though systematic disagreements between the different data sets makes the comparison
243 more difficult.

244 Limits on Cell Wall Biogenesis

245 While the processes we have considered represent only a small portion of proteins devoted to cell envelope
246 biogenesis, we find it unlikely that they limit cellular growth in general. The relative amount of mass required
247 for lipid and peptidoglycan components decrease at faster growth rates due to a decrease in their surface area
248 to volume (S/V) ratio (*Ojkic et al., 2019*). Furthermore, despite the slow catalytic rate of FabZ and FabA in lipid
249 synthesis, experimental data and recent computational modeling has shown that the rate of fatty-acid synthesis
250 can be drastically increased by increasing the concentration of FabZ (*Yu et al., 2011; Ruppe and Fox, 2018*). With
251 a proteome size of $\approx 3 \times 10^6$ proteins, a hypothetical 10-fold increase in expression from 4000 to 40,000 ACP
252 dehydratases would result in a paltry $\approx 1\%$ increase in the size of the proteome. In the context of peptidoglycan
253 synthesis, we note that our estimate considers only the transpeptidase enzymes that are involved lateral and
254 longitudinal elongation of the peptidoglycan. This neglects the presence of other transpeptidases that are present
255 in the periplasm and also involved in remodeling and maturation of the peptidoglycan. It is therefore possible that
256 if this was setting the speed limit for cell division, the simple expression of more transpeptidases may be sufficient
257 to maintain the structural integrity of the cell wall.

CELL ENVELOPE BIOSYNTHESIS

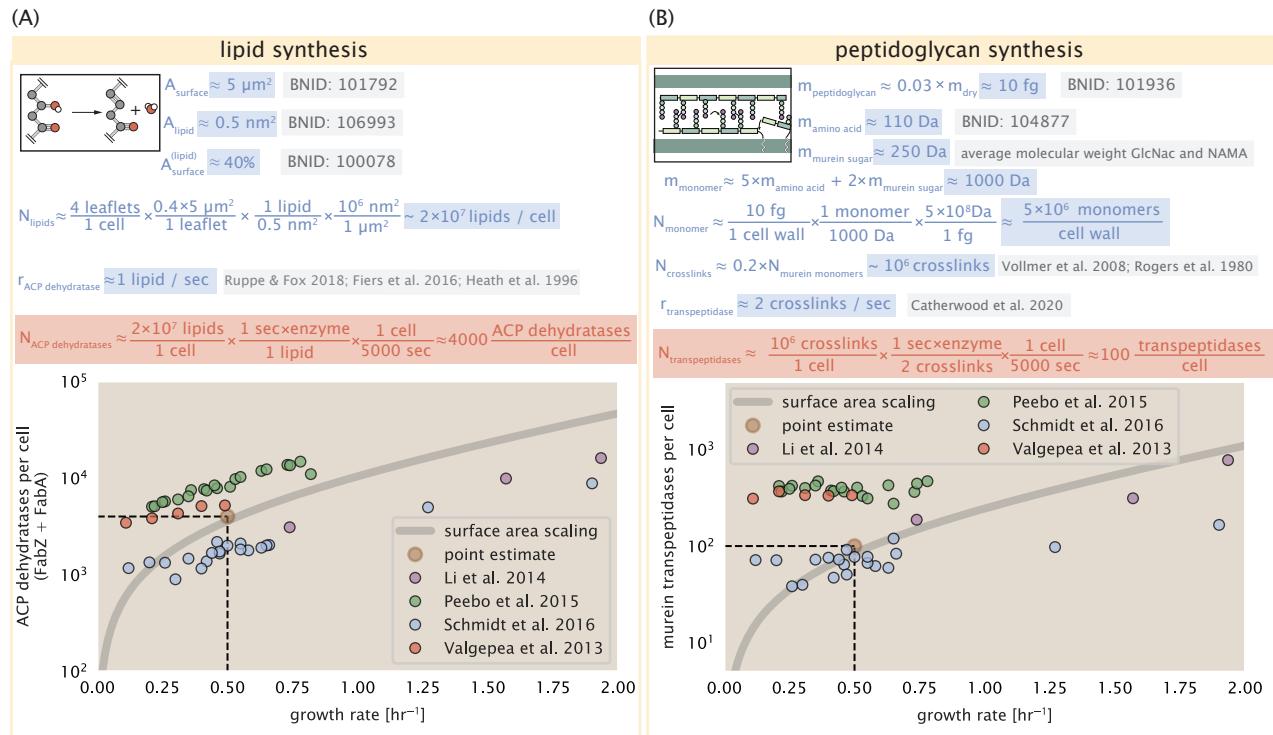


Figure 3. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabA}]_2$ and $[\text{FabZ}]_2$. (B) An estimate for the number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan. The mass of the murein monomer was estimated by approximating each amino acid in the pentapeptide chain as having a mass of $\approx 110 \text{ Da}$ and each sugar in the disaccharide having a mass of $\approx 250 \text{ Da}$. The *in vivo* rate of transpeptidation $r_{\text{transpeptidase}}$ in *E. coli* was taken from recent analysis by [Catherwood et al. \(2020\)](#). The bottom panel shows experimental measurements of the transpeptidase complexes in *E. coli* following the stoichiometries $[\text{MrcA}]_2$, $[\text{MrcB}]_2$, $[\text{MrdA}]_1$, and $[\text{MrdB}]_1$. Grey curves in each plot show the estimated number of complexes needed to satisfy the synthesis requirements scaled by the surface area as a function of growth rate.

258 **Energy Production**

259 Cells consume and generate energy predominantly in the form of nucleoside triphosphates (NTPs) in order to
260 grow. The high-energy phosphodiester bonds of (primarily) ATP power a variety of cellular processes that drive
261 biological systems away from thermodynamic equilibrium. Here we now turn to the synthesis of ATP as a potential
262 process that may limit growth, which also requires us to consider the maintenance of the electrochemical proton
263 gradient which powers it.

264 **ATP Synthesis**

265 Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic
266 phosphate provides the thermodynamic driving force in a wide array of biochemical reactions. One such reaction
267 is the formation of peptide bonds during translation, which requires ≈ 2 ATPs for the charging of an amino acid
268 to the tRNA and ≈ 2 GTP for the formation of each peptide bond. Assuming the ATP costs associated with error
269 correction and post-translational modifications of proteins are negligible, we can make the approximation that each
270 peptide bond has a net cost of ≈ 4 ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from ATP is achieved
271 via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment
272 (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to be functionally equivalent to
273 being exclusively ATP. In total, the energetic costs of peptide bond formation consume $\approx 80\%$ of the cells ATP
274 budget (BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015); Stouthamer (1973)*). The pool of ATP is
275 produced by the F_1 - F_0 ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield \approx
276 300 ATP per second (BNID: 114701; *Weber and Senior (2003)*).

277 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make the approximation
278 that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300 peptide bonds (BNID: 115702;
279 108986; 104877). Taking these values together, coupled with an estimate of ≈ 4 ATP equivalents per peptide bond,
280 we find that the typical *E. coli* cell consumes $\sim 5 \times 10^9$ ATP per cell cycle on protein synthesis alone. Assuming
281 that each ATP synthases operates at its maximal speed (300 ATP per second per synthase), ≈ 3000 ATP synthases
282 are needed to keep up with the energy demands of the cell. This estimate is comparable with the experimental
283 observations, shown in *Figure 4* (A). We note that this estimate assumes all ATP is synthesized via ATP synthase
284 and neglects synthesis via fermentative metabolism. This simplification may explain why at the fastest growth
285 rates ($\approx 2 \text{ hr}^{-1}$), our continuum estimate predicts more synthase than is experimentally observed (gray line in
286 *Figure 4*). At rapid growth rates, *E. coli* enters a type of overflow metabolism where fermentative metabolism
287 becomes pronounced (*Szenk et al., 2017*).

288 **Generating the Proton Electrochemical Gradient**

289 In order to produce ATP, the F_1 - F_0 ATP synthase itself must consume energy. Rather than burning through its
290 own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the
291 electrochemical potential established across the inner membrane through cellular respiration. This electrochemical
292 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as
293 they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner
294 membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in the
295 intermembrane space. However, each rotation of an ATP synthase shuttles ≈ 4 protons into the cytosol (BNID:
296 103390). With a few thousand ATP synthases producing ATP at their maximal rate, the potential difference would
297 be rapidly abolished in a few milliseconds if it were not being actively maintained.

298 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical
299 and biophysical study (*Ingledew and Poole, 1984; Khademian and Imlay, 2017; Cox et al., 1970; Henkel et al., 2014*).
300 A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E. coli* electron transport complexes
301 using structural and biochemical data, revealing that each electron transport chain rapidly pumps protons into the
302 intermembrane space at a rate of ≈ 1500 protons per second (BNID: 114704; 114687). Using our estimate of the
303 number of ATP synthases required per cell [*Figure 4(A)*], coupled with these recent measurements, we estimate
304 that ≈ 3000 electron transport complexes would be necessary to facilitate the $\sim 5 \times 10^6$ protons per second diet of
305 the cellular ATP synthases. This estimate is in agreement with the number of complexes identified in the proteomic

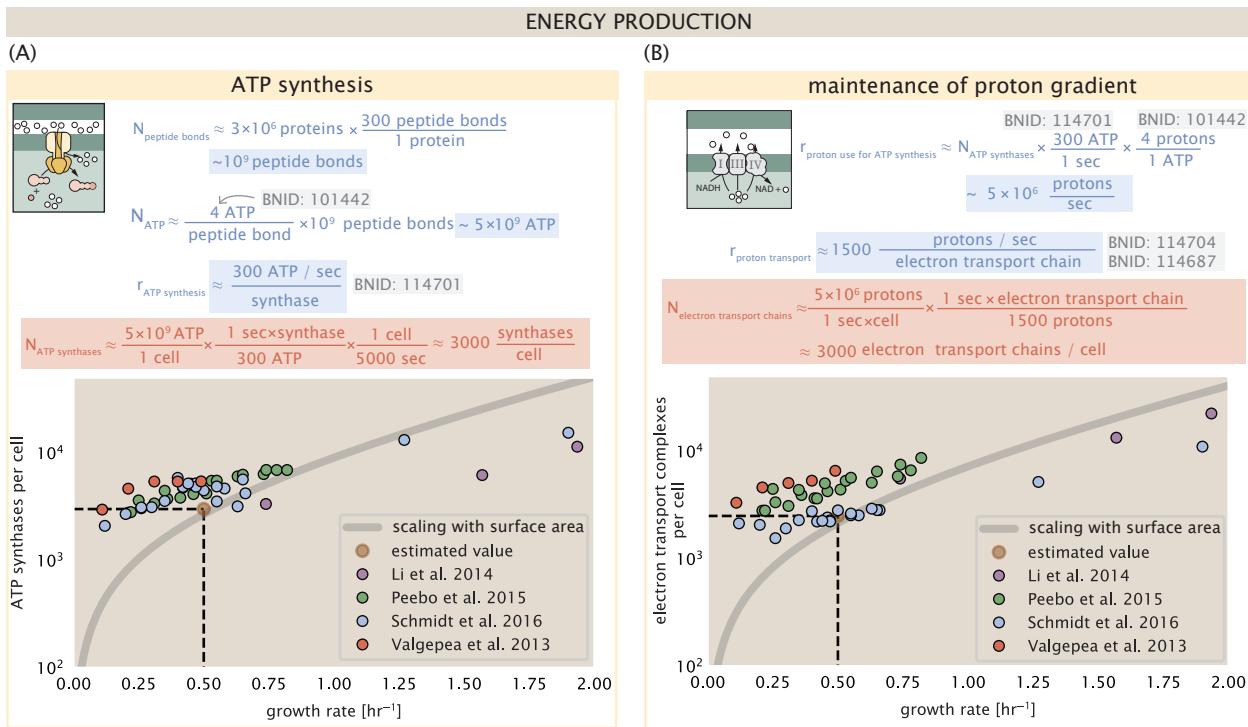


Figure 4. The abundance of F₁-F₀ ATP synthases and electron transport chain complexes as a function of growth rate.

(A) Estimate of the number of F₁-F₀ ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F₁-F₀ ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry [AtpE]₁₀[AtpF]₂[AtpB][AtpC][AtpH][AtpA]₃[AtpG][AtpD]₃. (B) Estimate of the number of electron transport chain complexes needed to maintain a membrane potential of -200 mV given estimate of number of F₁-F₀ ATP synthases from (A). Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646 that could be formed given proteomic observations. These complexes include cytochromes *bd1* ([CydA][CydB][CydX][CydH]), *bdII* ([AppC][AppB]), *bo₃*,([CyoD][CyoA][CyoB][CyoC]) and NADH:quinone oxioreductase I ([NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]) and II ([Ndh]). Grey lines in both (A) and (B) correspond to the estimate procedure described, but applied to a continuum of growth rates. We direct the reader to the Supporting Information for a more thorough description of this approach.

306 datasets (plot in **Figure 4(B)**). This suggests that every ATP synthase must be accompanied by ≈ 1 functional electron
307 transport chain.

308 Limits on Biosynthesis in a Crowded Membrane

309 Our estimates thus far have focused on biochemistry at the periphery of the cell. Since surface area and volume
310 do not scale identically as cell size changes, in order to better understand the physical constraints on transport and
311 energy production it is necessary to consider the consequence of a changing S/V ratio, which will decrease at faster
312 growth rates. Here we use our analysis of ATP production to consider this constraint.

313 In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell cycle or 10^6
314 ATP/s. With a cell volume of roughly 1 fL (BNID: 100004), this corresponds to about 2×10^{10} ATP per fL of cell volume,
315 in line with previous estimates (**Stouthamer and Bettenhausen, 1977; Szenk et al., 2017**). In **Figure 5 (A)** we plot
316 this ATP demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from
317 spherical to rod-shaped with an aspect ratio (length/width) equal to 4. In order to consider the maximum ATP
318 that could be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP
319 synthase and electron transport complexes and a maximal production rate of about 3 ATP / (nm²·s) (**Szenk et al.,**
320 **2017**). This is shown in blue in **Figure 5(A)**, which shows that at least for the growth rates observed (right column
321 in plot), the energy demand is roughly an order of magnitude less. Interestingly, **Szenk et al. (2017)** found that
322 ATP production by respiration is less efficient than by fermentation on a per membrane area basis, due to the

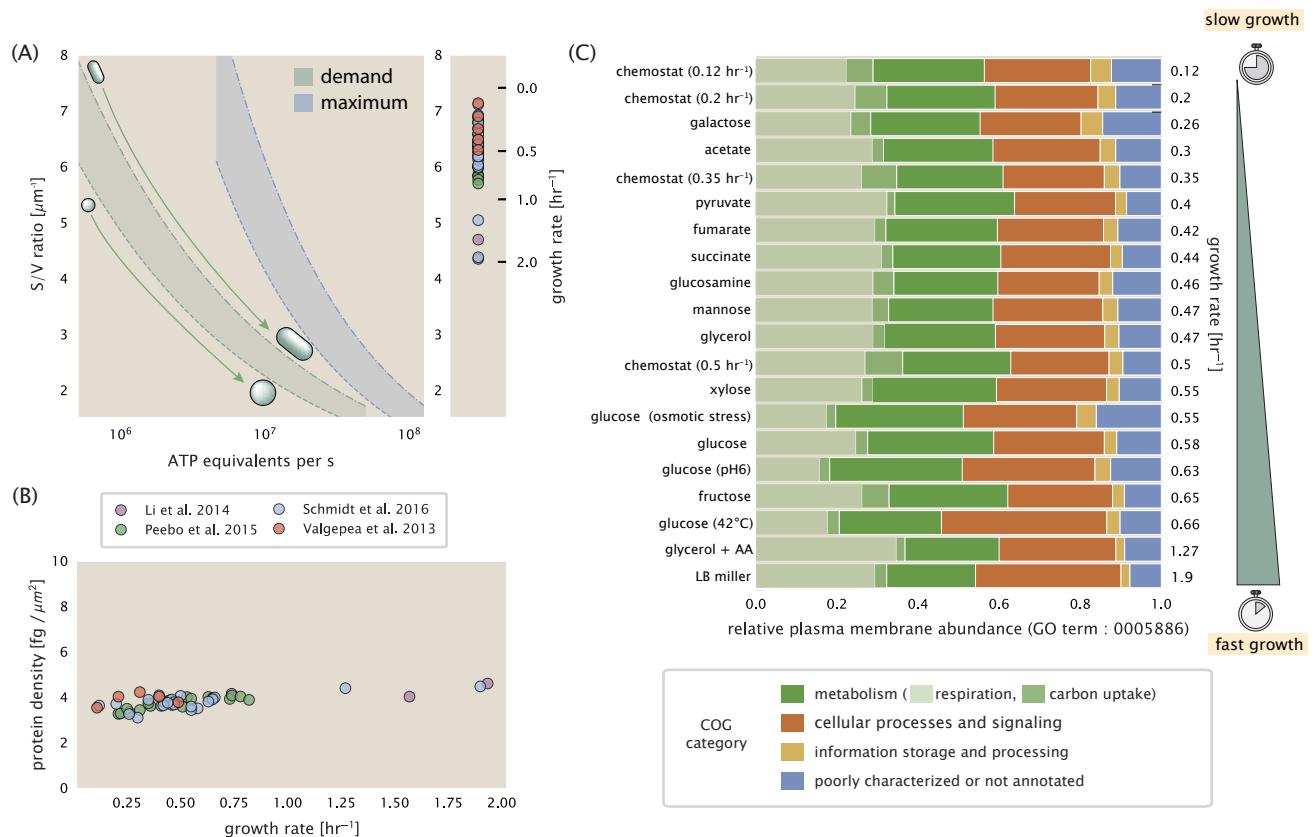


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---) reflecting a rod-shaped bacterium like *E. coli* with a typical aspect ratio (length / width) of 4 (Shi et al., 2018). Approximately 50% of the bacterial inner membrane is assumed to be protein, with the remainder lipid. The right plot shows the measured growth rates, and their estimated S/V ratio using growth rate dependent size measurements from Si et al. (2017) (See Appendix Estimation of Cell Size and Surface Area on calculation). (B) Total protein mass per μm^2 calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances by mass based on COG annotation. Metabolic proteins are further separated into respiration (F_1-F_0 ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome bo₃ ubiquinol oxidase, cytochrome bd-I ubiquinol oxidase) and carbohydrate transport (GO term: GO:0008643). Note that the elongation factor EF-Tu can also associate with the inner membrane, but was excluded in this analysis due to its high relative abundance (roughly identical to the summed protein shown in part (B)).

323 additional proteins of the electron transport chain. This suggests that, even under anaerobic growth, cells will have
324 sufficient membrane space for ATP production.

325 The analysis highlights that there will indeed be a maximum attainable cell size due to a diminishing capacity
326 to provide resources as the cell increases in size. The maximum energy production in Figure 5(A), however, does
327 represent a somewhat unachievable limit since the inner membrane must also include other proteins such as
328 those we've considered for nutrient transport and cell wall biogenesis. To better understand the overall proteomic
329 makeup of the inner membrane, we therefore used Gene Ontology (GO) annotations (Ashburner et al., 2000;
330 The Gene Ontology Consortium, 2018) to identify all proteins embedded or peripheral to the inner membrane
331 (GO term: 0005886). Those associated but not membrane-bound include proteins like MreB and FtsZ and must
332 nonetheless be considered as a vital component occupying space on the membrane. In Figure 5(B), we find that the
333 total protein mass per μm^2 is nearly constant across growth rates. Interestingly, when we consider the distribution
334 of proteins grouped by their Clusters of Orthologous Groups (COG) (Tatusov et al., 2000), the relative abundance
335 for those in metabolism (including ATP synthesis via respiration) is also relatively constant across growth rates,
336 suggesting that no one process (energy production, nutrient uptake, etc.) is particularly dominating even at fast
337 growth rates Figure 5(C).

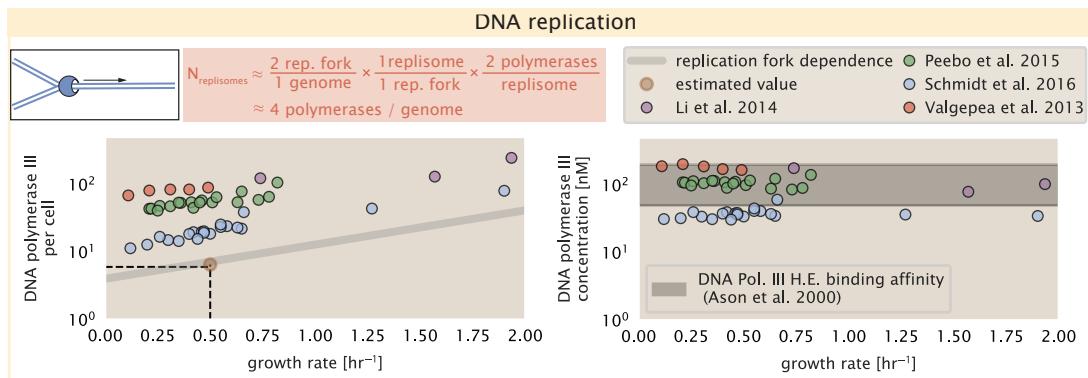


Figure 6. Complex abundance estimates for dNTP synthesis and DNA replication. An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the left-hand plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[\text{DnaE}]_3[\text{DnaQ}]_3[\text{HolE}]_3[\text{DnaX}]_5[\text{HolB}][\text{HolA}][\text{DnaN}]_4[\text{HolC}]_4[\text{HolD}]_4$) per cell. Right-hand plot shows the effective concentration of DNA polymerase III holoenzyme (See Appendix Estimation of Cell Size and Surface Area for calculation of cell size). Grey lines in left-hand panel show the estimated number of complexes needed as a function of growth, the details of which are described in the Supplemental Information.

Figure 6-Figure supplement 1. Estimate and observations of the abundance of ribonucleotide reductase, a key component in dNTP synthesis.

338 Processes of the Central Dogma

339 Up to this point, we have considered a variety of transport and biosynthetic processes that are critical to acquiring
340 and generating new cell mass. While there are of course many other metabolic processes we could consider, we
341 now turn our focus to some of the most central processes which *must* be undertaken irrespective of the growth
342 conditions – those of the central dogma.

343 DNA Replication

344 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids
345 up to ~ 100 kbp in length. While we consider the starting material dNTPs in **Figure 6-Figure Supplement 1** and
346 discussed further in Appendix Additional Process of the Central Dogma, here we focus our quantitative thinking on
347 the chromosome of *E. coli*, which harbors ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs.

348 To successfully divide and produce viable progeny, this chromosome must be faithfully replicated and seg-
349 regated into each nascent cell. Replication is initiated at a single region of the chromosome termed the *oriC*
350 locus where a pair of replisomes, each consisting of two DNA polymerase III, begin their high-fidelity replication
351 of the genome in opposite directions (Fijalkowska et al., 2012). *In vitro* measurements have shown that DNA
352 Polymerase III copies DNA at a rate of ≈ 600 nucleotides per second (BNID: 104120). Therefore, to replicate a single
353 chromosome, two replisomes moving at their maximal rate would copy the entire genome in ≈ 4000 s. Thus, with
354 a division time of 5000 s, there is sufficient time for a pair of replisomes complexes to replicate the entire genome.

355 In rapidly growing cultures, bacteria like *E. coli* can initiate as many as 10 - 12 replication forks at a given time
356 (Bremer and Dennis, 2008; Si et al., 2017), we expect only a few DNA polymerases (≈ 10) are needed. However, as
357 shown in **Figure 6** DNA polymerase III is nearly an order of magnitude more abundant. This discrepancy can be
358 understood by considering its binding constant to DNA. *In vitro* characterization has quantified the K_D of DNA
359 polymerase III holoenzyme to single-stranded and double-stranded DNA to be 50 and 200 nM, respectively (Ason
360 et al., 2000). The right-hand plot in **Figure 6** shows that the concentration of DNA polymerase III across all data
361 sets is within this range. Thus, its copy number appears to vary such that its concentration is approximately
362 equal to the dissociation constant to the DNA. While the processes regulating the initiation of DNA replication are
363 complex and involve more than just the holoenzyme, these data indicate that the kinetics of replication rather than
364 the explicit copy number of the DNA polymerase III holoenzyme is the more relevant feature of DNA replication
365 to consider. In light of this, the data in **Figure 6** suggests that for bacteria like *E. coli*, DNA replication does not
366 represent a rate-limiting step in cell division. However, it is worth noting that for bacterium like *C. crescentus*

367 whose chromosomal replication is initiated only once per cell cycle (*Jensen et al., 2001*), the time to double their
368 chromosome indeed represents an upper limit to their growth rate.

369 RNA Synthesis

370 We now turn our attention to the next stage of the central dogma – the transcription of DNA to form RNA. We
371 consider three major groupings of RNA, namely the RNA associated with ribosomes (rRNA), the RNA encoding the
372 amino-acid sequence of proteins (mRNA), and the RNA which links codon sequence to amino-acid identity during
373 translation (tRNA).

374 rRNA serves as the catalytic and structural component of the ribosome, comprising approximately 2/3 of the
375 total ribosomal mass, and is decorated with \approx 50 ribosomal proteins. Each ribosome contains three rRNA molecules
376 of lengths 120, 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides
377 overall. *In vivo* measurements of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded
378 onto the promoter of an rRNA gene at a rate of \approx 1 per second (BNID: 111997, 102362). If RNA polymerases are
379 constantly loaded at this rate, then we can assume that \approx 1 functional rRNA unit is synthesized per second per
380 rRNA operon. While *E. coli* possesses 7 of these operons per chromosome, the fact that chromosome replication
381 can be parallelized means that the average dosage of rRNA genes can be substantially higher (up to \approx 70 copies) at
382 fast growth rates (*Dennis et al., 2004*). At a growth rate of \approx 0.5 hr⁻¹, however, the average cell has \approx 1 copy of its
383 chromosome and therefore approximately \approx 7 copies of the rRNA operons, therefore producing \approx 7 rRNA units per
384 second. With a 5000 second division time, this means the cell is able to generate around 3×10^4 functional rRNA
385 units, comparable within an order of magnitude to the number of ribosomes per cell.

386 How many RNA polymerases are then needed to constantly transcribe the required rRNA? If one polymerase is
387 loaded per second, and the transcription rate is \approx 40 nucleotides per second (BNID: 101094), then the typical
388 spacing between polymerases will be \approx 40 nucleotides. However, we must note that the polymerase itself has a
389 footprint of \approx 40 nucleotides (BNID: 107873), meaning that one could expect to find one RNA polymerase per 80
390 nucleotide stretch of an rRNA gene. With a total length of \approx 4500 nucleotides per operon and 7 operons per cell,
391 the number of RNA polymerases transcribing rRNA at any given time is then \approx 500 per cell.

392 As outlined in *Figure 7*, and discussed further the Appendix Additional Process of the Central Dogma, synthesis
393 of mRNA and tRNA together require on the order of \approx 400 RNAP. Thus, in total, one would expect the typical
394 cell to require \approx 1000 RNAP to satisfy its transcriptional demands. As is revealed in *Figure 7(B)*, this estimate
395 is about an order of magnitude below the observed number of RNA polymerase complexes per cell (\approx 5000 -
396 7000). The difference between the estimated number of RNA polymerase needed for transcription and these
397 observations, however, are consistent with recent literature revealing that \approx 80 % of RNA polymerases in *E. coli* are
398 not transcriptionally active (*Patrick et al., 2015*).

399 Our estimates also neglect other mechanistic features of transcription and transcriptional initiation more
400 broadly. For example, we acknowledge that some fraction of the RNAP pool is nonspecifically bound to DNA
401 during its search for promoters from which to begin transcription. Furthermore, we ignore the obstacles that RNA
402 polymerase and DNA polymerase present to each other as they move along the DNA (*Finkelstein and Greene,*
403 *2013*). Finally, we neglect the fact that RNA polymerase also require σ -factors for promoter recognition and
404 transcription initiation (*Browning and Busby, 2016*).

405 While they represent the core machinery for transcription, RNA polymerase is not sufficient to initiate transcrip-
406 tion. Initiation of transcription is dependent on the presence of σ -factors, protein cofactors that bind directly to the
407 polymerase (*Browning and Busby, 2016*) and aid in promoter recognition. In *Figure 7–Figure Supplement 1*, we
408 show that the predicted RNA polymerase copy number indeed is more comparable with the abundance of σ -70
409 (RpoD), the primary sigma factor in *E. coli*. There therefore remains more to be investigated as to what sets the
410 observed abundance of RNA polymerase in these proteomic data sets. However, we conclude that the observed
411 RNA polymerase abundances are generally in excess of what appears to be needed for growth, suggesting that the
412 synthesis of RNA polymerase themselves are not particularly limiting.

413 Protein Synthesis

414 We conclude our dialogue between back-of-the-envelope estimates and comparison with the proteomic data by
415 examining the final process in the central dogma – translation. In doing so, we will begin with an estimate of the

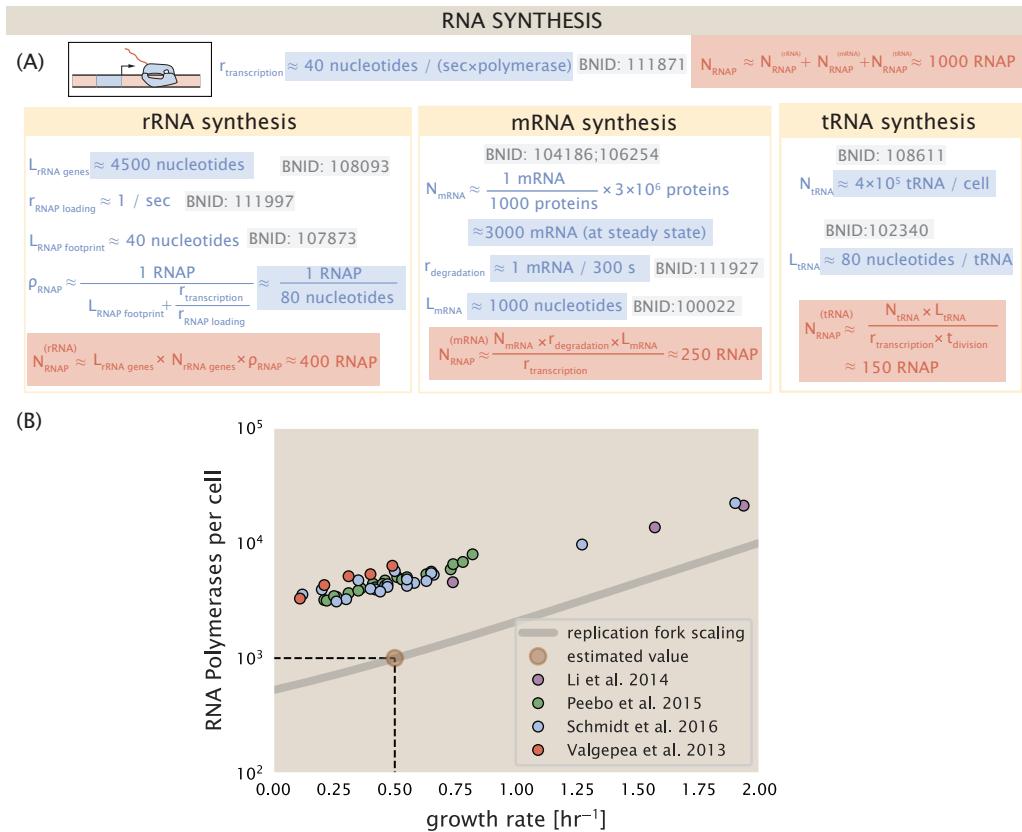


Figure 7. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively.(B) The RNA polymerase core enzyme copy number as a function of growth rate. Colored points correspond to the average number RNA polymerase core enzymes that could be formed given a subunit stoichiometry of $[RpoA]_2[RpoC][RpoB]$.

Figure 7-Figure supplement 1. Abundance and growth rate dependence of σ -70.

416 number of ribosomes needed to replicate the cellular proteome. While the rate at which ribosomes translate is
417 well known to be dependent on the growth rate (*Dai et al. (2018)*, a phenomenon we consider later in this work) we
418 will make the approximation that translation occurs at a modest rate of \approx 15 amino acids per second per ribosome
419 (BNID: 100233) Under this approximation and our previous estimate of 10^9 peptide bonds per cell at a growth rate
420 of 0.5 hr^{-1} , we can easily arrive at an estimate of $\approx 10^4$ ribosomes needed per cell to replicate the entire protein
421 mass (*Figure 8(A)*, top). This point estimate, as well as the corresponding estimate across a continuum of growth
422 rates, proves to be notably comparable to the experimental observations, shown in the bottom panel of *Figure 8(A)*.
423 While the ribosome is responsible for the formation of peptide bonds, we do not diminish the importance of
424 charging tRNAs with their appropriate amino acid, a process which occurs with remarkable fidelity. In the Appendix
425 and in *Figure 8–Figure Supplement 1*, we consider the process of ligating tRNAs to their corresponding amino acid
426 and again find notable accord between the data and our quantitative expectations.

427 Having completed our circuit through key processes of cellular growth outlined in *Figure 1*, we can now take stock
428 of our understanding of the observed growth rate dependence and abundances of various protein complexes. We
429 note that, broadly speaking, these simple estimates have been reasonably successful in quantitatively describing
430 the observations in the proteomic data, suggesting that the proteome is tuned in composition and absolute
431 abundance to match the growth rate requirements without any one process representing a singular bottleneck or
432 rate limiting step in division. However, in our effort to identify key limitations on growth, there are two notable
433 observations that we wish to emphasize.

434 The first is a recurring theme throughout the estimates investigated here, which is that any inherent biochemical
435 rate limitation can be overcome by expressing more proteins. We can view this as a parallelization of each
436 biosynthesis task, which helps explain why bacteria tend to increase their protein content (and cell size) as growth
437 rate increases (*Ojkic et al., 2019*). The second, and ultimately the most significant in defining the cellular growth
438 rate, is that the synthesis of ribosomal proteins presents a special case where parallelization is *not* possible and
439 thereby imposes a limit on the fastest possible growth rate. Each ribosome has ≈ 7500 amino acids across all of its
440 protein components which must be strung together as peptide bonds through the action of another ribosome.
441 Once again using a modest elongation rate of ≈ 15 amino acids per second, we arrive an estimate of ≈ 500 seconds
442 or ≈ 7 minutes to replicate a single ribosome. This limit, as remarked upon by others (*Dill et al., 2011*), serves as a
443 hard theoretical boundary for how quickly a bacterium like *E. coli* can replicate. As each ribosome would therefore
444 need to copy itself, this 7 minute speed limit is independent of the number of ribosomes per cell (*Figure 8(B)*), yet
445 assumes that the only proteins that need to be replicated for division to occur are ribosomal proteins, an unrealistic
446 regime not met in biological reality. This poses an optimization problem for the cell – how are the translational
447 demands of the entire proteome met without investing resources in the production of an excess of ribosomes?

448 This question, more frequently presented as a question of optimal resource allocation, has been the target of
449 an extensive dialogue between experiment and theory over the past decade. In a now seminal work, *Scott et al.*
450 (2010) present an elegant treatment of resource allocation through partitioning of the proteome into sectors – one
451 of which being ribosome-associated proteins whose relative size ultimately defines the total cellular growth rate. In
452 more recent years, this view has been more thoroughly dissected experimentally (*Klumpp and Hwa, 2014; Basan*
453 *et al., 2015; Dai et al., 2018, 2016; Erickson et al., 2017*) and together have led to a paradigm-shift in how we think
454 of cellular physiology at the proteomic-level. However, the quantitative description of these observations is often
455 couched in terms of phenomenological constants and effective parameters with the key observable features of
456 expression often computed in relative, rather than absolute, abundances. Furthermore, these approaches often
457 exclude or integrate away effects of cell size and chromosome content, which we have found through our estimates
458 to have important connections to the observed cellular growth rate.

459 In the closing sections of this work, we explore how ribosomal content, total protein abundance, and chromo-
460 somal replication are intertwined in their control over the cellular growth rate. To do so, we take a more careful
461 view of ribosome abundance, increasing the sophistication of our analysis by exchanging our order-of-magnitude
462 estimates for a minimal mathematical model of growth rate control. This is defined by parameters with tangible
463 connections to the biological processes underlying cellular growth and protein synthesis. Using this model, we
464 interrogate how the size of the ribosome pool and its corresponding translational capacity enable cells to maintain
465 a balance between the of amino acids via metabolism and catabolism and their consumption through the peptide
466 bond formation required for growth.

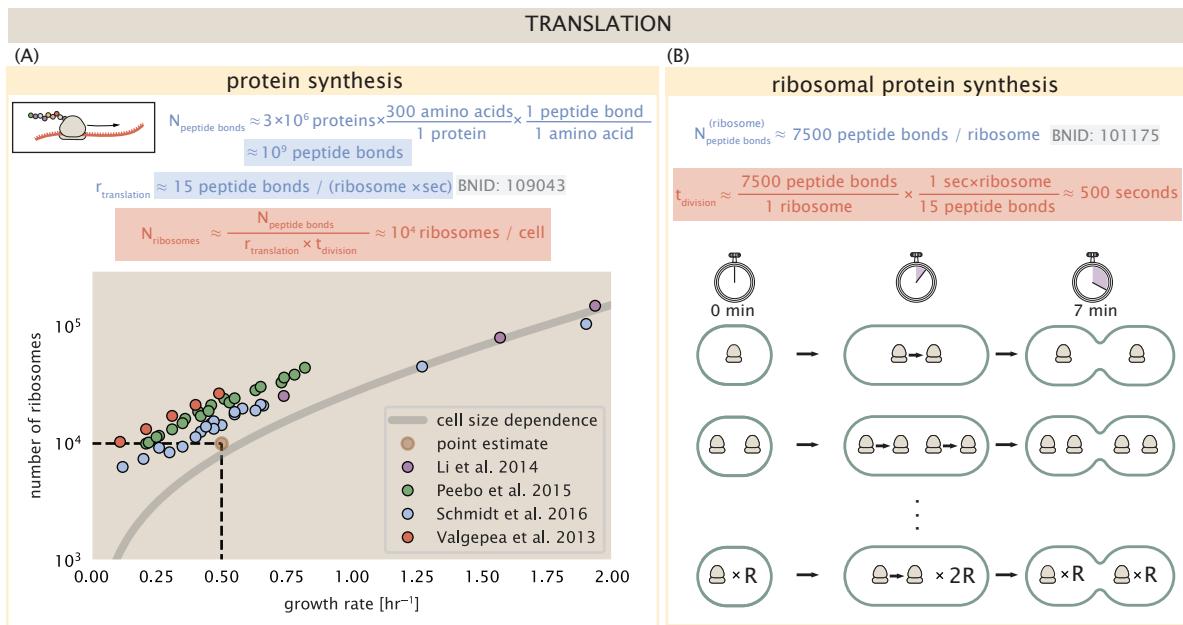


Figure 8. Estimation of the required number of ribosomes and the speed limit for bacterial replication. (A) Estimation of the number of ribosomes required to synthesize 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr^{-1} . Grey lines correspond to the estimated complex abundance calculated at different growth rates. (B) Estimation for the time to replicate a ribosome. This rate is independent of the number of ribosomes R and instead is limited by the time required to double an individual ribosome.

Figure 8-Figure supplement 1. Estimate and observed abundance and growth rate dependence of tRNA ligases.

467 Maximum Growth Rate is Determined by the Ribosomal Mass Fraction

468 The 7 minute speed limit shown in **Figure 8(B)** assumes all proteins in the cell are ribosomes. In order to connect
 469 this to the experimental data (and physiological reality more broadly), we first need to relax this assumption and
 470 determine a translation-limited growth rate. Here, we will assume that the cell is composed of N_{pep} peptide bonds
 471 and R ribosomes, whose precise values will depend on the growth rate λ . The protein subunits of each ribosomal
 472 peptide sum to a total of ≈ 7500 amino acids as noted earlier, which we denote by L_R . With an average mass of an
 473 amino acid of $m_{\text{AA}} \approx 110 \text{ Da}$ (BNID: 104877), the total ribosomal mass fraction Φ_R is given by

$$\Phi_R = \frac{m_{\text{ribosomes}}}{m_{\text{proteome}}} \approx \frac{m_{\text{AA}} \times R \times L_R}{m_{\text{AA}} \times N_{\text{pep}}} = \frac{R \times L_R}{N_{\text{pep}}}. \quad (1)$$

474 For exponentially growing cells (**Godin et al., 2010**), the rate of cellular growth will be related to the rate of protein
 475 synthesis via

$$\lambda N_{\text{pep}} = r_t \times R \times f_a, \quad (2)$$

476 where r_t is the translation rate. Here, we've introduced a multiplicative factor f_a which represents the fraction of the
 477 ribosomes that are actively translating. This term allows us to account for immature or non-functional ribosomes
 478 or active sequestration of ribosomes through the action of the secondary messenger alarmone (p)ppGpp in poorer
 479 nutrient conditions **Hauryliuk et al. (2015)**.

480 Combining **Equation 1** and **Equation 2** results in an expression for a translation-limited growth rate, which is
 481 given by

$$\lambda_{\text{translation-limited}} = \frac{r_t \times \Phi_R \times f_a}{L_R}. \quad (3)$$

482 This result, derived in a similar manner in **Klumpp et al. (2013)**, reflects mass-balance under steady state growth
 483 and has long provided a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function
 484 of growth rate (**Maaløe, 1979; Dennis et al., 2004; Scott et al., 2010**). The left-hand panel of **Figure 9(A)** shows this
 485 growth rate plotted as a function of the ribosomal mass fraction. In the regime where all ribosomes are active

486 ($f_a = 1$) and the entire proteome is composed of ribosomal proteins ($\Phi_R = 1$), indeed, we arrive at the maximum
487 theoretical growth rate of r_t/L_R , and ≈ 7 min for *E. coli*.

488 Connecting **Equation 3** to the proteomic data serving as the centerpiece of our work, however, requires
489 knowledge of f_a at each growth rate as proteomic measurements only provide a measure of Φ_R . Recently, *Dai et al.*
490 (2016) determined f_a as a function of the growth rate (**Figure 9(A)**, right-hand panel, inset), revealing that $f_a \approx 1$ at
491 growth rates above 0.75 hr^{-1} and $f_a < 1$ as the growth rate slows. Using these data, we inferred the approximate
492 active fraction (see Appendix Calculation of active ribosomal fraction) at each growth rate and used this to compute
493 $\Phi_R \times f_a$ (**Figure 9(A)**, colored points in right-hand panel). In general, these data skirt the translation-limited growth
494 rate determined using **Equation 3**, where we have taken r_t to be the maximal elongation rate of 17 amino acids per
495 second measured by *Dai et al.* (2016). There is a notable discrepancy between the data collected in *Schmidt et al.*
496 (2016); *Li et al.* (2014) and that collected from *Valgepea et al.* (2013); *Peebo et al.* (2015). When compared to other
497 measurements (non-proteomic based) of the active ribosome mass fraction (**Figure 9(B)**, grey points in right-hand
498 panel), the data from *Valgepea et al.* (2013) and *Peebo et al.* (2015) are notably aberrant, suggesting a systematic
499 error in these data. These additional measurements come from a number of recent studies and are determined
500 from measurements of total RNA to total protein mass ratios (**Figure 9–Figure Supplement 1**).

501 Together, these results illustrate that the growth rates observed across the amalgamated data sets are close
502 to the translation-limited growth rate determined through ribosomal activity, at least for the data reported in
503 *Schmidt et al.* (2016) and *Li et al.* (2014). While this is a useful framework to consider how the relative abundance of
504 ribosomes (compared to all other proteins) defines the growth rate, it is worth noting that as growth rate increases,
505 so does the cell size and therefore so will the total proteomic mass (*Basan et al.*, 2015). With a handle on how
506 elongation rate and the total number of peptide bonds per proteome is related to the growth rate, we now expand
507 this description to account for the increasing cell size and ribosome copy number at faster growth rates, enabling
508 us to identify a potential bottleneck in the synthesis of rRNA.

509 **rRNA Synthesis Presents a Potential Bottleneck During Rapid Growth**

510 Even under idealized experimental conditions, *E. coli* rarely exhibits growth rates above 2 hr^{-1} (*Bremer and Dennis*,
511 2008), which is still well-below the synthesis rate of a single ribosome, and below the maximum growth rates
512 reported for several other bacteria (*Roller et al.*, 2016). While we have considered potential limits imposed by
513 translation of ribosomal proteins, here we consider potential limiting regimes specific to the synthesis of rRNA.

514 Due to multiple initiations of chromosomal replication per cell doubling, the effective number of rRNA operons
515 increases with growth rate and will do so in proportion to the average number of origins per cell, $\langle \# \text{ ori} \rangle$. This latter
516 parameter is set by how often replication must be initiated in order to keep up with cell doubling times τ , whose
517 time may be shorter than the cell cycle time τ_{cyc} (referring to the time from replication initiation to cell division)
518 (*Dennis et al.*, 2004). This is quantified by

$$\langle \# \text{ ori} \rangle = 2^{\tau_{\text{cyc}}/\tau} = 2^{\tau_{\text{cyc}}\lambda/\log(2)}. \quad (4)$$

519 We used the experimental measurements of τ_{cyc} (the timescale of chromosome replication and cell division) and τ
520 (the timescale of a cell doubling) from *Si et al.* (2017) (**Figure 9–Figure Supplement 1(B)**) to calculate $\langle \# \text{ ori} \rangle$ with
521 **Equation 4** as a function of growth rates. For growth rates above about 0.5 hr^{-1} , t_{cyc} is approximately constant
522 at about 70 minutes, implying that $\langle \# \text{ ori} \rangle$ will grow exponentially with growth rates beyond 0.5 hr^{-1} . As the
523 rRNA operons are predominantly located close to origin of replication (BNID: 100352), we make the simplifying
524 assumption that that the number of rRNA operons will be directly proportional to $\langle \# \text{ ori} \rangle$.

525 Returning to our rule-of-thumb of 1 functional rRNA unit per second per transcribing operon, we estimate the
526 maximum number of ribosomes that could be made as a function of growth rate (**Figure 9(B)**, blue curve). Although
527 we expect this estimate to significantly overestimate rRNA abundance at slower growth rates ($\lambda < 0.5 \text{ hr}^{-1}$), this
528 provides a useful reference alongside the proteomic measurements, particularly in the regime of fast growth. For
529 growth rates above about 1 hr^{-1} , for example, we find that cells will need to transcribe rRNA near their maximal
530 rate. As a counter example, if *E. coli* did not initiate multiple rounds of replication, but managed to replicate their
531 chromosome within the requisite time limit, they would be unable to make enough rRNA for the observed number
532 of ribosomes (dashed blue curve in **Figure 9(C)**). The convergence between the maximum rRNA production and

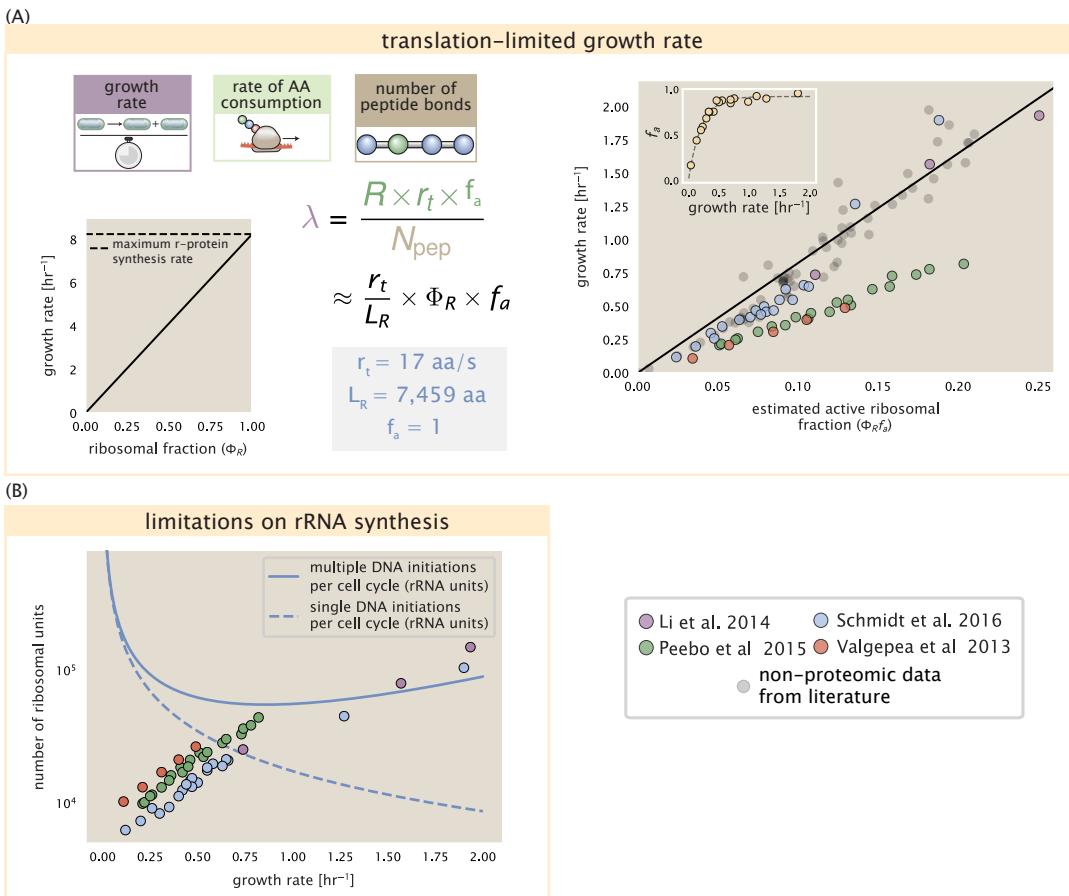


Figure 9. Translation-limited growth rate. (A) *left*: Translation-limited growth as a function of the ribosomal fraction. The solid line is calculated for an elongation rate of 17 aa per second. The dashed line corresponds to the maximum rate of ribosomal protein synthesis (≈ 7 min). *right*: Translation-limited growth rate as a function of the actively translating ribosomal fraction. The actively translating ribosomal fraction is calculated using the estimated values of f_a from [Dai et al. \(2016\)](#) (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). Gray data points show additional measurements from literature and consider further in the supplemental figure part (A). (B) Maximum number of rRNA units that can be synthesized as a function of growth rate. Solid curve corresponding to the rRNA copy number is calculated by multiplying the number of rRNA operons by the estimated number of (# ori) at each growth rate. The quantity (# ori) was calculated using Equation 4 and the measurements from [Si et al. \(2017\)](#). The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

Figure 9–Figure supplement 1. Comparison of $\Phi_R f_a$ with literature and estimation of (# ori).

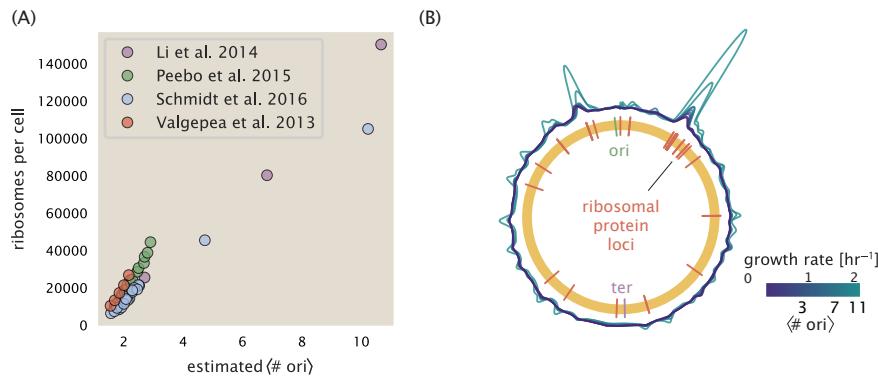


Figure 10. Cells increase both absolute ribosome abundance and Φ_R with (# ori). (A) Plot of the ribosome copy number estimated from the proteomic data against the estimated (# ori) (see Appendix Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ for additional details). (B) A running Gaussian average (20 kbp st. dev.) of protein copy number is calculated for each growth condition considered by (Schmidt et al., 2016) based on each gene's transcriptional start site. Since total protein abundance increases with growth rate, protein copy numbers are median-subtracted to allow comparison between growth conditions. (# ori) are estimated using the data in (A) and Equation 4.

533 measured ribosome copy number suggests rRNA synthesis may begin to present a bottleneck at the fastest growth
534 rates due to the still-limited copies of rRNA genes.

535 Rapid Growth Requires *E. coli* to Increase Both Cell Size and Ribosomal Mass Fraction

536 In Figure 9(B, right-hand side) we also find that above about 0.75 hr^{-1} , the growth rate is determined solely by the
537 ribosomal mass fraction Φ_R , since f_a is close to 1, and r_i is near its maximal rate (Dai et al., 2016). While Φ_R will
538 need to increase in order for cells to grow faster, the fractional dependence in Equation 3 gives little insight into
539 how this scaling is actually achieved by the cell.

540 It is now well-documented that *E. coli* cells add a constant volume per origin of replication, which is robust
541 to a remarkable array of cellular perturbations (Si et al., 2017). Given the proteomic measurements featured in
542 this work, we find that the ribosome copy number also scales in proportion to (# ori) (Figure 10(A)). However, an
543 increase in ribosome abundance alone is not necessarily sufficient to increase growth rate and we also need to
544 consider how Φ_R varies with (# ori). Importantly, as shown in Figure 10(B), we find that the deviations in protein
545 expression with (# ori) are largely restricted to regions of ribosomal protein genes Figure 10(B). Here we have
546 calculated the position-dependent protein expression across the chromosome by a running Gaussian average of
547 protein copy number (20 kbp st. dev. averaging window) based on each gene's transcriptional start site. These were
548 median-subtracted to account for the change in total protein abundance with (# ori). This result suggests that
549 Φ_R is also being tuned in proportion to (# ori) under nutrient-limited growth, and in particular, it is through this
550 additional dependence on Φ_R , combined with the exponential increase in (# ori), that *E. coli* exhibits an exponential
551 increase in cell size with growth rate.

552 A Minimal Model of Nutrient-Mediated Growth Rate Control

553 While the preceding subsections highlight a dominant role for ribosomes in setting the growth rate, our analysis
554 on the whole emphasizes how the total proteomic content also changes in response to variable growth conditions
555 and growth rate. In this final section we use a minimal model of growth rate control to better understand how this
556 interconnection between ribosomal abundance and total protein abundance influences the observed growth rate.

557 Here we propose that cells modulate their protein abundance in direct response to the availability of nutrients
558 in their environment. As noted earlier, bacteria can modulate ribosomal activity through the secondary messenger
559 molecules like (p)ppGpp in poorer nutrient conditions (Figure 9(C) - inset; Dai et al. (2016)). Importantly, these
560 secondary messengers also cause global changes in transcriptional and translational activity (Hauryliuk et al., 2015;
561 Zhu and Dai, 2019; Büke et al., 2020). In *E. coli*, amino acid starvation leads to the accumulation of de-acetylated
562 tRNAs at the ribosome's A-site and a strong increase in (p)ppGpp synthesis activity by the enzyme RelA (Hauryliuk
563 et al., 2015). Along with this, there is increasing evidence that (p)ppGpp also acts to inhibit the initiation of DNA

564 replication (*Kraemer et al., 2019*), providing a potential mechanism for cells to lower (# ori) and maintain a smaller
565 cell size in poorer nutrient conditions (*Fernández-Coll et al., 2020*).

566 To consider this quantitatively, we assume that cells modulate their proteome (N_{pep} , R , Φ_R) to better maximize
567 their rate of peptide elongation r_t . The elongation rate r_t will depend on how quickly the ribosomes can match
568 codons with their correct amino-acyl tRNA, along with the subsequent steps of peptide bond formation and
569 translocation. This ultimately depends on the cellular concentration of amino acids, which we treat as a single
570 effective species, $[AA]_{\text{eff}}$. In our model, we determine the the rate of peptide elongation r_t and achievable growth
571 rate as simply depending on the supply of amino acids (and, therefore, also amino-acyl tRNAs), through a parameter
572 r_{AA} in units of AA per second, and the rate of amino acid consumption by protein synthesis ($r_t \times R \times f_a$). This is
573 shown schematically in *Figure 11(A)* and derived in Appendix Derivation of Minimal Model for Nutrient-Mediated
574 Growth Rate Control. Given our observation that protein synthesis and energy production are not limiting, we
575 assume that other molecular players required by ribosomes such as elongation factors and GTP are available in
576 sufficient abundance.

577 In *Figure 11(B)*, we illustrate how the elongation rate will depend on the ribosomal copy number. Here, we
578 have considered an arbitrarily chosen $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$ and $f_a = 1$ for a unit cell volume $V = 1\text{fL}$. At
579 low ribosome copy numbers, the observed elongation rate is dependent primarily on $[AA]_{\text{eff}}$ through r_{AA} [as
580 $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in *Figure 11(B)*]. As the ribosome copy number is increased such that the amino acid
581 supply rate and consumption rate are nearly equal [point (2) in *Figure 11(B)*], the observed elongation rate begins
582 to decrease sharply. When the ribosome copy number is increased even further, consumption at the maximum
583 elongation rate exceeds the supply rate, yielding a significantly reduced elongation rate [point (3) in *Figure 11B*].
584 While the elongation rate will always be dominated by the amino acid supply rate at sufficiently low ribosome
585 copy numbers, the elongation rate at larger ribosome abundances can be increased by tuning f_a such that not all
586 ribosomes are elongating, reducing their total consumption rate.

587 Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity
588 To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic
589 measurements; namely, we restrict the values of R , N_{pep} , and cell size V to those associated with the amalgamated
590 proteomic data (described in Appendix Estimation of Total Protein Content per Cell). We then consider how changes
591 in the nutrient conditions, through the parameter r_{AA} , influence the maximum growth rate as determined by
592 *Equation 3*. *Figure 11(C)* shows how the growth rate depends on the rate of amino acid supply r_{AA} as a function of
593 the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal growth rate
594 whose dependence on R (and consequently, the cell size) increases with increasing r_{AA} . Importantly, however, there
595 is an optimum set of R , N_{pep} , and V that are strictly dependent on the value of r_{AA} . This shows that increasing the
596 ribosomal concentration beyond the cell's metabolic capacity will have the adverse consequence of depleting the
597 supply of amino acids and a concomitant decrease in the elongation rate r_t (*Figure 11(B)*) and growth rate.

598 Also of note is the growth rate trends obaserved at low values of r_{AA} [purple and blue lines in *Figure 11(C)*],
599 representative of growth in nutrient-poor media. In these conditions, there no longer exists a peak in growth, at
600 least within the range of physiologically-relevant ribosome copy numbers. This is a regime, associated with slower
601 growth rates, where cells limit their pool of actively translating ribosomes by decreasing f_a (*Dai et al., 2016*), which
602 we find would help maintain the pool of available amino acids $[AA]_{\text{eff}}$ and increase the achievable elongation rate.
603 This observation is in agreement with the central premise of the cellular resource allocation principle proposed by
604 *Scott et al. (2010); Klumpp et al. (2009); Klumpp and Hwa (2014)* and *Hui et al. (2015)*.

605 Discussion

606 Continued experimental and technological improvements have led to a treasure trove of quantitative biological
607 data (*Hui et al., 2015; Schmidt et al., 2016; Si et al., 2017; Gallagher et al., 2020; Peebo et al., 2015; Valgepea et al.,
608 2013*), and an ever advancing molecular view and mechanistic understanding of the constituents that support
609 bacterial growth (*Taheri-Araghi et al., 2015; Morgenstein et al., 2015; Si et al., 2019; Karr et al., 2012; Kostinski
610 and Reuveni, 2020*). In this work we have compiled what we believe to be the state-of-the-art knowledge on
611 proteomic copy number across a broad range of growth conditions in *E. coli*. We have made this data accessible
612 through a [GitHub repository](#), and an [interactive figure](#) that allows exploration of specific protein and protein

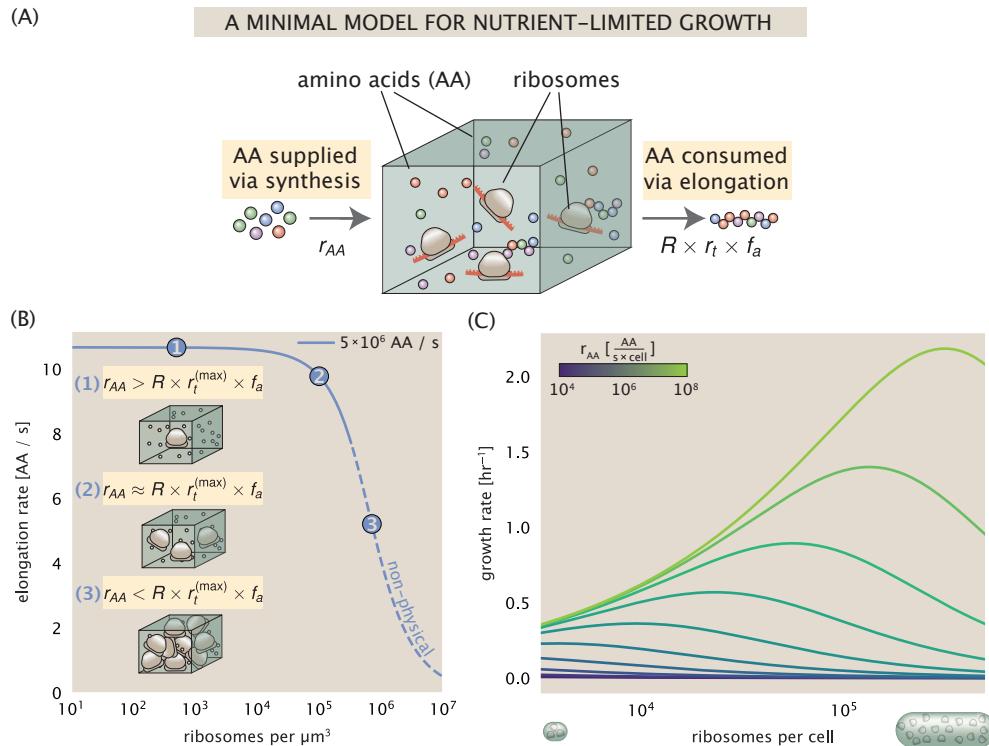


Figure 11. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, where r_t is the elongation rate, R is the ribosome copy number in the unit volume, and f_a is the fraction of those ribosomes actively translating. (B) The observed elongation rate is plotted as a function of ribosomes. The three points correspond to three regimes of ribosome copy numbers and are shown schematically on the left-hand side. The region of the curve shown as dashed lines represents a non-physical copy number, but is shown for illustrative purposes. This curve was generated using an amino acid supply rate of $5 \times 10^6 \text{ AA} / \text{s}$, a maximal elongation rate of $17.1 \text{ AA} / \text{s}$, $f_a = 1$, and a unit cell volume of $V = 1 \text{ fL}$. See Appendix Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control for additional model details. (C) The cellular growth rate is plotted as a function of total cellular ribosome copy number for different cellular amino acid supply rates, with blue and green curves corresponding to low and high supply rates, respectively. As the ribosome copy number is increased, so too is the cell size and total protein abundance N_{pep} . We direct the reader to the Supplemental Information for discussion on the inference of the relationship between cell size, number of peptide bonds, and ribosome copy number.

Figure 11–Figure supplement 1. An interactive figure for exploration of the model parameter space.

613 complex copy numbers. Through a series of order-of-magnitude estimates that traverse key steps in the bacterial
614 cell cycle, this proteomic data has been a resource to guide our understanding of two key questions: what biological
615 processes limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents
616 as a function of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates
617 provide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the
618 transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous
619 steps of the central dogma.

620 In general, the copy numbers of the complexes involved in these processes were in reasonable agreement
621 with our order-of-magnitude estimates. Since many of these estimates represent soft lower-bound quantities, this
622 suggests that cells do not express proteins grossly in excess of what is needed for a particular growth rate. Several
623 exceptions, however, also highlight the dichotomy between a proteome that appears to "optimize" expression
624 according to growth rate and one that must be able to quickly adapt to environments of different nutritional quality.
625 Take, for example, the expression of carbon transporters. Shown in *Figure 2(B)*, we find that cells always express a
626 similar number of glucose transporters irrespective of growth condition. At the same time, it is interesting to note
627 that many of the alternative carbon transporters are still expressed in low but non-zero numbers (\approx 10-100 copies
628 per cell) across growth conditions. This may relate to the regulatory configuration for many of these operons,
629 which require the presence of a metabolite signal in order for alternative carbon utilization operons to be induced
630 (*Monod, 1949; Laxhuber et al., 2020*). Furthermore, upon induction, these transporters are expressed and present
631 in abundances in close agreement with a simple estimate.

632 Of the processes illustrated in *Figure 1*, we arrive at a ribosome-centric view of cellular growth rate control.
633 This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary
634 their ribosomal abundance as a function of growth conditions and growth rate (*Scott et al., 2010; Metzl-Raz et al.,
635 2017*). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates
636 to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize
637 many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA
638 replication, this is not so for synthesis of ribosomes (*Figure 9(A)*). The translation time for each ribosome [\approx 7 min,
639 *Dill et al. (2011)*] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase
640 their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The
641 second point relates to the long-observed correlations between growth rate and cell size (*Schaechter et al., 1958; Si
642 et al., 2017*), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous
643 curiosity and driven substantial amounts of research in their own regards, these relationships are themselves
644 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions
645 of rapid growth that require cells to also grow in size. Further experiments are needed to test the validity of this
646 hypothesis. In particular, we believe that the change in growth rate in response to translation-inhibitory drugs
647 (such as chloramphenicol) could be quantitatively predicted, given one had precision measurement of the relevant
648 parameters, including the fraction of actively translating ribosomes f_a and changes in the metabolic capacity of the
649 cell (i.e. the rate that amino acids can be made available) for a particular growth condition.

650 While the generation of new ribosomes plays a dominant role in growth rate control, there exist other physical
651 limits to the function of cellular processes. One of the key motivations for considering energy production was the
652 physical constraints on total volume and surface area as cells vary their size (*Harris and Theriot, 2018; Ojikic et al.,
653 2019*). As *E. coli* get larger at faster growth rates, an additional constraint begins to arise in energy production and
654 nutrient uptake due to the relative decrease in total surface area, where ATP is predominantly produced (*Szenk
655 et al., 2017*). Specifically, the cell interior requires an amount of energy that scales cubically with cell size, but the
656 available surface area only grows quadratically (*Figure 5(A)*). While this threshold does not appear to be met for
657 *E. coli* cells growing at 2 hr^{-1} or less, it highlights an additional constraint on growth given the apparent need to
658 increase cell size in order to grow faster. This limit is relevant even to eukaryotic organisms, whose mitochondria
659 exhibit convoluted membrane structures that nevertheless remain bacteria-sized organelles (*Guo et al., 2018*). In
660 the context of bacterial growth and energy production more generally, we have mainly limited our analysis to
661 the aerobic growth conditions associated with the proteomic data and further consideration will be needed for
662 anaerobic growth.

663 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control, and there are

many aspects of the bacterial proteome and growth that we neglected to consider. For example, other recent work (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*) has explored how the proteome is structured and how that structure depends on growth rate. In the work of *Hui et al. (2015)*, the authors coarse-grained the proteome into six discrete categories being related to either translation, catabolism, anabolism, and others related to signaling and core metabolism. The relative mass fraction of the proteome occupied by each sector could be modulated by external application of drugs or simply by changing the nutritional content of the medium. While we have explored how the quantities of individual complexes are related to cell growth, we acknowledge that higher-order interactions between groups of complexes or metabolic networks at a systems-level may reveal additional insights into how these growth-rate dependences are mechanistically achieved. Furthermore, while we anticipate the conclusions summarized here are applicable to a wide collection of bacteria with similar lifestyles as *E. coli*, other bacteria and archaea may have evolved other strategies that were not considered. Further experiments with the level of rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to learn more about how regulation of proteomic composition and growth rate control has evolved over the past 3.5 billion years.

Methods

Data Analysis and Availability

All proteomic measurements come from the experimental work of *Schmidt et al. (2016); Peebo et al. (2015); Valgepea et al. (2013)* (mass spectrometry) and *Li et al. (2014)* (ribosomal profiling). Data curation and analysis was done programmatically in Python, and compiled data and analysis files are accessible through a [GitHub repository] (DOI:XXX) associated with this paper as well as on the associated [paper website](#). An interactive figure that allows exploration of specific protein and protein complex copy numbers is available at [link].

Acknowledgements

We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding their data. We also thank members of the Phillips, Theriot, Kondev, and Garcia labs for useful discussions. We thank Suzannah M. Beeler, Avi Flamholz, Soichi Hirokawa, and Manuel Razo-Mejia for reading and providing comments on drafts of this manuscript. R.P. is supported by La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and the NIH 1R35 GM118043 (MRA). J.A.T. is supported by the Howard Hughes Medical Institute, and NIH Grant R37-AI036929. N.M.B is a HHMI Fellow of The Jane Coffin Childs Memorial Fund.

Competing Interests

The authors declare no competing interests.

694 Appendix for: Fundamental limits on the 695 rate of bacterial cell division

696 **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G. Garcia⁴, Jane
697 Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}**

698 ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of Applied Physics,
699 California Institute of Technology, Pasadena, CA, USA; ³Department of Chemical Engineering, Stanford
700 University, Stanford, CA, USA; ⁴Department of Molecular Cell Biology and Department of Physics,
701 University of California Berkeley, Berkeley, CA, USA; ⁵Department of Physics, Brandeis University,
702 Waltham, MA, USA; ⁶Department of Applied Physics, Stanford University, Stanford, CA, USA; ⁷Allen
703 Institute for Cell Science, Seattle, WA, USA; ⁸Division of Biology and Biological Engineering, California
704 Institute of Technology, Pasadena, CA, USA; ⁹Department of Physics, California Institute of Technology,
705 Pasadena, CA, USA; *Co-corresponding authors. Address correspondence to phillips@pboc.caltech.edu
706 and jtheriot@uw.edu; [†]These authors contributed equally to this work

707	Contents	
708	Introduction	1
709	Nutrient Transport	2
710	Limits on Transporter Expression	6
711	Cell Envelope Biogenesis	6
712	Lipid Synthesis	6
713	Peptidoglycan Synthesis	7
714	Limits on Cell Wall Biogenesis	7
715	Energy Production	9
716	ATP Synthesis	9
717	Generating the Proton Electrochemical Gradient	9
718	Limits on Biosynthesis in a Crowded Membrane	10
719	Processes of the Central Dogma	12
720	DNA Replication	12
721	RNA Synthesis	13
722	Protein Synthesis	13
723	Maximum Growth Rate is Determined by the Ribosomal Mass Fraction	16
724	rRNA Synthesis Presents a Potential Bottleneck During Rapid Growth	17
725	Rapid Growth Requires <i>E. coli</i> to Increase Both Cell Size and Ribosomal Mass Fraction	19
726	A Minimal Model of Nutrient-Mediated Growth Rate Control	19
727	Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity	20
728	Discussion	20
729	Methods	23
730	Data Analysis and Availability	23
731	Acknowledgements	23
732	Competing Interests	23
733	Additional Estimates of Fundamental Biological Processes	27
734	Nutrient Transport	27
735	Nitrogen	27
736	Phosphorus	27
737	Sulfur	27
738	Additional Process of the Central Dogma	28
739	dNTP synthesis	28
740	mRNA and tRNA Synthesis	28
741	tRNA Charging	29
742	Experimental Details Behind Proteomic Data	30
743	Fluorescence based measurements	30
744	Ribosomal profiling measurements	31
745	Mass spectrometry measurements	31
746	Summary of Proteomic Data	31
747	Estimation of Cell Size and Surface Area	32

748	Estimation of Total Protein Content per Cell	33
749	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	34
750	Additional Considerations of Schmidt <i>et al.</i> Data Set	34
751	Effect of cell volume on reported absolute protein abundances	38
752	Relaxing assumption of constant protein concentration across growth conditions	38
753	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	38
754	Calculation of Complex Abundance	39
755	Extending Estimates to a Continuum of Growth Rates	40
756	Estimation of the total cell mass	40
757	Complex Abundance Scaling With Cell Volume	42
758	A Relation for Complex Abundance Scaling With Surface Area	43
759	Number of Lipids	43
760	Number of Murein Monomers	43
761	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	43
762	Calculation of active ribosomal fraction.	44
763	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	44
764	Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control	45

765 **Additional Estimates of Fundamental Biological Processes**

766 In the main text of this work, we present estimates for a significant number of fundamental biological processes
767 that are necessary for cell division. While we believe the estimates provided in the main text provide a succinct
768 summary of the corresponding process, we left out additional estimates of related processes for brevity. In this
769 section of the appendix, we present these additional estimates in full.

770 **Nutrient Transport**

771 In the main text, we make passing mention that while transport carbon often comes in the form of carbohydrates
772 and sugar alcohols while other critical elements – such as nitrogen, sulfur, and phosphorus – are transported as
773 inorganic ions. Below, we present estimates for the transport requirements of these materials.

774 **Nitrogen**

775 We must first address which elemental sources must require active transport, meaning that the cell cannot acquire
776 appreciable amounts simply via diffusion across the membrane. The permeability of the lipid membrane to a large
777 number of solutes has been extensively characterized over the past century. Large, polar molecular species (such
778 as various sugar molecules, sulfate, and phosphate) have low permeabilities while small, non-polar compounds
779 (such as oxygen, carbon dioxide, and ammonia) can readily diffuse across the membrane. Ammonia, a primary
780 source of nitrogen in typical laboratory conditions, has a permeability on par with water ($\sim 10^5$ nm/s, BNID:110824).
781 In nitrogen-poor conditions, *E. coli* expresses a transporter (AmtB) which appears to aid in nitrogen assimilation,
782 though the mechanism and kinetic details of transport are still a matter of debate (*van Heeswijk et al., 2013*;
783 *Khademi et al., 2004*). Beyond ammonia, another plentiful source of nitrogen come in the form of glutamate,
784 which has its own complex metabolism and scavenging pathways. However, nitrogen is plentiful in the growth
785 conditions examined in this work, permitting us to neglect nitrogen transport as a potential rate limiting process in
786 cell division in typical experimental conditions.

787 **Phosphorus**

788 Phosphorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds making
789 up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational modification of
790 proteins and defining the polar-heads of lipids. In total, phosphorus makes up $\approx 3\%$ of the cellular dry mass which
791 in typical experimental conditions is in the form of inorganic phosphate. The cell membrane has remarkably low
792 permeability to this highly-charged and critical molecule, therefore requiring the expression of active transport
793 systems. In *E. coli*, the proton electrochemical gradient across the inner membrane is leveraged to transport
794 inorganic phosphate into the cell (*Rosenberg et al., 1977*). Proton-solute symporters are widespread in *E. coli*
795 (*Ramos and Kaback, 1977; Booth et al., 1979*) and can have rapid transport rates of 50 to 100 molecules per second
796 for sugars and other solutes (BNID: 103159; 111777). As a more extreme example, the proton transporters in
797 the F₁-F₀ ATP synthase, which use the proton electrochemical gradient for rotational motion, can shuttle protons
798 across the membrane at a rate of ≈ 1000 per second (BNID: 104890; 103390). In *E. coli* the PitA phosphate
799 transport system has been shown to be very tightly coupled with the proton electrochemical gradient with a 1:1
800 proton:phosphate stoichiometric ratio (*Harris et al., 2001; Feist et al., 2007*). Taking the geometric mean of the
801 aforementioned estimates gives a plausible rate of phosphate transport on the order of 300 per second. Illustrated
802 in *Figure 2–Figure Supplement 1(A)*, we can estimate that ≈ 200 phosphate transporters are necessary to maintain
803 an $\approx 3\%$ dry mass with a 5000 s division time. This estimate is consistent with observation when we examine
804 the observed copy numbers of PitA in proteomic data sets (plot in *Figure 2–Figure Supplement 1(A)*). While our
805 estimate is very much in line with the observed numbers, we emphasize that this is likely a slight overestimate of the
806 number of transporters needed as there are other phosphorous scavenging systems, such as the ATP-dependent
807 phosphate transporter Pst system which we have neglected.

808 **Sulfur**

809 Similar to phosphate, sulfate is highly-charged and not particularly membrane permeable, requiring active transport.
810 While there exists a H⁺/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well characterized
811 (*Zhang et al., 2014a*). Sulfate is predominantly acquired via the ATP-dependent ABC transporter CysUWA system

which also plays an important role in selenium transport (*Sekowska et al., 2000; Sirko et al., 1995*). While specific kinetic details of this transport system are not readily available, generic ATP transport systems in prokaryotes transport on the order of 1 to 10 molecules per second (BNID: 109035). Combining this generic transport rate, measurement of sulfur comprising 1% of dry mass, and a 5000 second division time yields an estimate of ≈ 1000 CysUWA complexes per cell (*Figure 2–Figure Supplement 1(B)*). Once again, this estimate is in notable agreement with proteomic data sets, suggesting that there are sufficient transporters present to acquire the necessary sulfur. In a similar spirit of our estimate of phosphorus transport, we emphasize that this is likely an overestimate of the number of necessary transporters as we have neglected other sulfur scavenging systems that are in lower abundance.

Additional Process of the Central Dogma

In the main text, we consider the processes underlying the backbone of the central dogma, namely DNA replication, RNA transcription, and protein translation. In this section we turn our attention to additional processes related to the central dogma, primarily dNTP synthesis for DNA replication and amino-acyl tRNA synthesis for translation. Additionally, we explore in more detail the estimates shown in *Figure 7(A)* for the RNA polymerase requirements of mRNA and tRNA synthesis.

dNTP synthesis

The four major dNTPs (dATP, dTTP, dCTP, and dGTP) serve as the fundamental units of the genetic code. Thus, to faithfully replicate the chromosome, the cell must be able to synthesize enough of these bases in the first place. All dNTPs are synthesized *de novo* in separate pathways, requiring different building blocks. However, a critical step present in all dNTP synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via the removal of the 3' hydroxyl group of the ribose ring (*Rudd et al., 2016*). This reaction is mediated by a class of enzymes termed ribonucleotide reductases, of which *E. coli* possesses two aerobically active complexes (termed I and II) and a single anaerobically active enzyme. Due to their peculiar formation of a radical intermediate, these enzymes have received much biochemical, kinetic, and structural characterization. One such work (*Ge et al., 2003*) performed a detailed *in vitro* measurement of the steady-state kinetic rates of these complexes, revealing a turnover rate of ≈ 10 dNTP per second.

Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the abundance of these complexes is a measure of the total dNTP production in *E. coli*. Illustrated schematically in *Figure 6 (A)*, we consider the fact that to replicate the cell's genome, on the order of $\approx 10^7$ dNTPs must be synthesized. Assuming a production rate of 10 per second per ribonucleotide reductase complex and a cell division time of 5000 seconds, we arrive at an estimate of ≈ 200 complexes needed per cell. As shown in the bottom panel of *Figure Supplement 1 (A)*, this estimate agrees with the experimental measurements of these complexes abundances within $\approx 1/2$ an order of magnitude. Extension of this estimate across a continuum of growth rate, including the fact that multiple chromosomes can be replicated at a given time, is shown as a grey transparent line in *Figure Supplement 1*. Similarly to our point estimate, this refinement agrees well with the data, accurately describing both the magnitude of the complex abundance and the dependence on growth rate.

Recent work has revealed that during replication, the ribonucleotide reductase complexes coalesce to form discrete foci colocalized with the DNA replisome complex (*Sánchez-Romero et al., 2011*). This is particularly pronounced in conditions where growth is slow, indicating that spatial organization and regulation of the activity of the complexes plays an important role.

mRNA and tRNA Synthesis

In *Figure 7* of the main text, we present with limited explanation estimates for the number of RNA polymerases needed to synthesize enough mRNA and tRNA molecules. Here, we present a rationalization for these estimates.

To form a functional protein, all protein coding genes must first be transcribed from DNA to form an mRNA molecule. While each protein requires an mRNA blueprint, many copies of the protein can be synthesized from a single mRNA. Factors such as strength of the ribosomal binding site, mRNA stability, and rare codon usage frequency dictate the number of proteins that can be made from a single mRNA, with yields ranging from 10^1 to 10^4 (BNID: 104186; 100196; 106254). Computing the geometric mean of this range yields ≈ 1000 proteins synthesized

860 per mRNA, a value that agrees with experimental measurements of the number of proteins per cell ($\approx 3 \times 10^6$, BNID:
861 100088) and total number of mRNA per cell ($\approx 3 \times 10^3$, BNID:100064).

862 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time, there will exist
863 approximately 3000 unique mRNA molecules. To determine the *total* number of mRNA that need to be synthesized
864 over the cell's lifetime, we must consider degradation of the mRNA. In most bacteria, mRNAs are rather unstable
865 with life times on the order of several minutes (BNID: 104324; 106253; 111927; 111998). For convenience, we
866 assume that the typical mRNA in our cell of interest has a typical lifetime of ≈ 300 seconds. Using this value, we
867 can determine the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.
868 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional dynamics, we
869 state here that the total mRNA production rate must be on the order of ≈ 15 mRNA per second. In *E. coli*, the
870 average protein is ≈ 300 amino acids in length (BNID: 108986), meaning that the corresponding mRNA is ≈ 900
871 nucleotides which we will further approximate as ≈ 1000 nucleotides to account for the non-protein coding regions
872 on the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about to sustain a
873 transcription rate of $\approx 1.5 \times 10^4$ nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA
874 at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of ≈ 250 RNA polymerase complexes
875 needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that
876 required to synthesize enough rRNA, as we saw in the previous section. We find this to be a striking result as these
877 250 RNA polymerase molecules are responsible for the transcription of the ≈ 4000 protein coding genes that are
878 not ribosome associated.

879 We now turn our attention to the synthesis of tRNA. Unlike mRNA or rRNA, each individual tRNA is remarkably
880 short, ranging from 70 to 95 nucleotides each (BNID: 109645; 102340). What they lack in length, they make up
881 for in abundance, with reported values ranging from $\approx 5 \times 10^4$ (BNID: 105280) to $\approx 5 \times 10^5$ (BNID: 108611). To test
882 tRNA synthesis as a possible growth-rate limiting stage, we will err towards a higher abundance of $\approx 5 \times 10^5$ per cell.
883 Combining the abundance and tRNA length measurements, we make the estimate that $\approx 5 \times 10^7$ nucleotides are
884 sequestered in tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes *in vivo* on the order of \approx
885 48 hours (*Abelson et al., 1974; Svenningsen et al., 2017*) – well beyond the timescale of division. Once again using
886 our rule-of-thumb for the rate of transcription to be 40 nucleotides per second and assuming a division time of \approx
887 5000 seconds, we arrive at an estimate of ≈ 200 RNA polymerases to synthesize enough tRNA. This requirement
888 pales in comparison to the number of polymerases needed to generate the rRNA and mRNA pools and can be
889 neglected as a significant transcriptional burden.

890 tRNA Charging

891 In the previous subsection, we focused solely on estimating the number of RNA polymerases needed for the
892 generation of the tRNA molecule itself. We now explore the protein complex requirements for ligation of the
893 appropriate amino acid to each tRNA. We begin by again using an estimate of $\approx 3 \times 10^6$ proteins per cell at a 5000 s
894 division time (BNID: 115702) and a typical protein length of ≈ 300 amino acids (BNID: 100017), we can estimate
895 that a total of $\approx 10^9$ amino acids are stitched together by peptide bonds.

896 How many tRNAs are needed to facilitate this remarkable number of amino acid delivery events to the translating
897 ribosomes? It is important to note that tRNAs are recycled after they've passed through the ribosome and can be
898 recharged with a new amino acid, ready for another round of peptide bond formation. While some *in vitro* data
899 exists on the turnover of tRNA in *E. coli* for different amino acids, we can make a reasonable estimate by comparing
900 the number of amino acids to be polymerized to cell division time. Using our stopwatch of 5000 s and 10^9 amino
901 acids, we arrive at a requirement of $\approx 2 \times 10^5$ tRNA molecules to be consumed by the ribosome per second.

902 There are many processes which go into synthesizing a tRNA and ligating it with the appropriate amino acids.
903 As we discussed previously, there appear to be more than enough RNA polymerases per cell to synthesize the
904 needed pool of tRNAs. Without considering the many ways in which amino acids can be scavenged or synthesized
905 *de novo*, we can explore ligation the as a potential rate limiting step. The enzymes which link the correct amino acid
906 to the tRNA, known as tRNA synthetases or tRNA ligases, are incredible in their proofreading of substrates with the
907 incorrect amino acid being ligated once out of every 10^4 to 10^5 events (BNID: 103469). This is due in part to the
908 consumption of energy as well as a multi-step pathway to ligation. While the rate at which tRNA is ligated is highly
909 dependent on the identity of the amino acid, it is reasonable to state that the typical tRNA synthetase has charging

Table 1. Overview of proteomic data sets.

Author	Method	Reported Quantity
Taniguchi <i>et al.</i> (2010)	YFP-fusion, cell fluorescence	fg/copies per cell
Valgepea <i>et al.</i> (2012)	mass spectrometry	fg/copies per cell
Peebo <i>et al.</i> (2014)	mass spectrometry	fg/copies per fl
Li <i>et al.</i> (2014)	ribosomal profiling	fg/copies per cell ^a
Soufi <i>et al.</i> (2015)	mass spectrometry	fg/copies per cell
Schmidt <i>et al.</i> (2016)	mass spectrometry	fg/copies per cell ^b

a. The reported values assume that the proteins are long-lived compared to the generation time but are unable to account for post-translational modifications that may alter absolute protein abundances.

b. This mass spectrometry approach differs substantially from the others since in addition to the relative proteome-wide abundance measurements, the authors performed absolute quantification of 41 proteins across all growth conditions (see section on Additional Considerations of Schmidt *et al.* Data Set for more details on this).

910 rate of \approx 20 AA per tRNA synthetase per second (BNID: 105279).

911 We can make an assumption that amino-acyl tRNAs are in steady-state where they are produced at the same
912 rate they are consumed, meaning that 2×10^5 tRNAs must be charged per second. Combining these estimates
913 together, as shown schematically in **Figure 8–Figure Supplement 1**, yields an estimate of $\sim 10^4$ tRNA synthetases
914 per cell with a division time of 5000 s. This point estimate is in very close agreement with the observed number
915 of synthetases (the sum of all 20 tRNA synthetases in *E. coli*). This estimation strategy seems to adequately
916 describe the observed growth rate dependence of the tRNA synthetase copy number (shown as the grey line in
917 **Figure 8–Figure Supplement 1**, suggesting that the copy number scales with the cell volume).

918 In total, the estimated and observed $\sim 10^4$ tRNA synthetases occupy only a meager fraction of the total cell
919 proteome, around 0.5% by abundance. It is reasonable to assume that if tRNA charging was a rate limiting
920 process, cells would be able to increase their growth rate by devoting more cellular resources to making more tRNA
921 synthetases. As the synthesis of tRNAs and the corresponding charging can be highly parallelized, we can argue that
922 tRNA charging is not a rate limiting step in cell division, at least for the growth conditions explored in this work.

923 Experimental Details Behind Proteomic Data

924 Here we provide a brief summary of the experiments behind each proteomic data set considered. The purpose of
925 this section is to identify how the authors arrived at absolute protein abundances. In the following section (see
926 section on Summary of Proteomic Data) we will then provide a summary of the protein abundance measurements.
927 Table 1 provides an overview of the publications we considered. These are predominately mass spectrometry-based,
928 with the exception of the work from Li *et al.* (2014) which used ribosomal profiling, and the fluorescence-based
929 counting done in Taniguchi *et al.* (2010). After having compiled and comparing these measurements, we noted
930 substantial deviations in the measurements from Taniguchi *et al.* (2010) and Soufi *et al.* (2015) (shown in the
931 following section), and decided to only use the data from Taniguchi *et al.* (2010); Li *et al.* (2014); Valgepea *et al.*
932 (2013); Peebo *et al.* (2015) in the main text. For completeness, we include these additional datasets in our discussion
933 of the experimental data.

934 Fluorescence based measurements

935 In the work of Taniguchi *et al.* (2010), the authors used a chromosomal YFP fusion library where individual strains
936 have a specific gene tagged with a YFP-coding sequence. 1018 of their 1400 attempted strains were used in the work.
937 A fluorescence microscope was used to collect cellular YFP intensities across all these strains. Through automated
938 image analysis, the authors normalized intensity measurements by cell size to account for the change in size and
939 expression variability across the cell cycle. Following correction of YFP intensities for cellular autofluorescence,
940 final absolute protein levels were determined by a calibration curve with single-molecule fluorescence intensities.
941 This calibration experiment was performed separately using a purified YFP solution.

942 **Ribosomal profiling measurements**

943 The work of *Li et al. (2014)* takes a sequencing based approach to estimate protein abundance. Ribosomal
944 profiling, which refers to the deep sequencing of ribosome-protected mRNA fragments, can provide a quantitative
945 measurement of the protein synthesis rate. As long as the protein life-time is long relative to the cell doubling time,
946 it is possible to estimate absolute protein copy numbers. The absolute protein synthesis rate has units of proteins
947 per generation, and for stable proteins will also correspond to the protein copy number per cell.

948 In the experiments, ribosome-protected mRNA is extracted from cell lysate and selected on a denaturing
949 polyacrylamide gel for deep sequencing (15–45 nt long fragments collected and sequenced by using an Illumina
950 HiSeq 2000 in *Li et al. (2014)*). Counts of ribosome footprints from the sequencing data were then corrected
951 empirically for position-dependent biases in ribosomal density across each gene, as well as dependencies on
952 specific sequences including the Shine-Dalgarno sequence. These data-corrected ribosome densities represent
953 relative protein synthesis rates. Absolute protein synthesis rates are obtained by multiplying the relative rates by
954 the total cellular protein per cell. The total protein per unit volume was determined with the Lowry method to
955 quantify total protein, calibrated against bovine serum albumin (BSA). By counting colony-forming units following
956 serial dilution of their cell cultures, they then calculated the total protein per cell.

957 **Mass spectrometry measurements**

958 Perhaps not surprisingly, the data is predominantly mass spectrometry based. This is largely due to tremendous
959 improvements in the sensitivity of mass spectrometers, as well as improvements in sample preparation and
960 data analysis pipelines. It is now a relatively routine task to extract protein from a cell and quantify the majority
961 of proteins present by shotgun proteomics. In general, this involves lysing cells, enzymatically digesting the
962 proteins into short peptide fragments, and then introducing them into the mass spectrometer (e.g. with liquid
963 chromatography and electrospray ionization), which itself can have multiple rounds of detection and further
964 fragmentation of the peptides.

965 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multiple samples
966 to be measured together by the mass spectrometer. By measuring samples of known total protein abundance
967 simultaneously (i.e. one sample of interest, and one reference), it is possible to determine relative protein
968 abundances. Absolute protein abundances can be estimated following the same approach used above for
969 ribosomal profiling, which is to multiply each relative abundance measurement by the total cellular protein per
970 cell. This is the approach taken by *Valgepea et al. (2013)*; *Peebo et al. (2015)* and *Soufi et al. (2015)*, with relative
971 protein abundances determined based on the relative peptide intensities (label free quantification 'LFQ' intensities).
972 For the data of *Valgepea et al. (2013)*, total protein per cell was determined by measuring total protein by the
973 Lowry method, and counting colony-forming units following serial dilution. For the data from *Peebo et al. (2015)*,
974 the authors did not determine cell quantities and instead report the cellular protein abundances in protein per
975 unit volume by assuming a mass density of 1.1 g/ml, with a 30% dry mass fraction.

976 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide fragments of known
977 abundance. These can serve as a direct way to calibrate mass spectrometry signal intensities to absolute mass.
978 This is the approach taken by *Schmidt et al. (2016)*. In addition to a set of shotgun proteomic measurements to
979 determine proteome-wide relative abundances, the authors also performed absolute quantification of 41 proteins
980 covering over four orders of magnitude in cellular abundance. Here, a synthetic peptide was generated for each of
981 the proteins, doped into each protein sample, and used these to determine absolute protein abundances of the 41
982 proteins. These absolute measurements, determined for every growth condition, were then used as a calibration
983 curve to convert proteomic-wide relative abundances into absolute protein abundance per cell. A more extensive
984 discussion of the *Schmidt et al. (2016)* data set can be found in Section Additional Considerations of Schmidt et al.
985 Data Set.

986 **Summary of Proteomic Data**

987 In the work of the main text we only used the data from *Valgepea et al. (2013)*; *Li et al. (2014)*; *Peebo et al. (2015)*;
988 *Schmidt et al. (2016)*. As shown in *Figure 12(A)*, the reported total protein abundances in the work of *Taniguchi*
989 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi et al.*

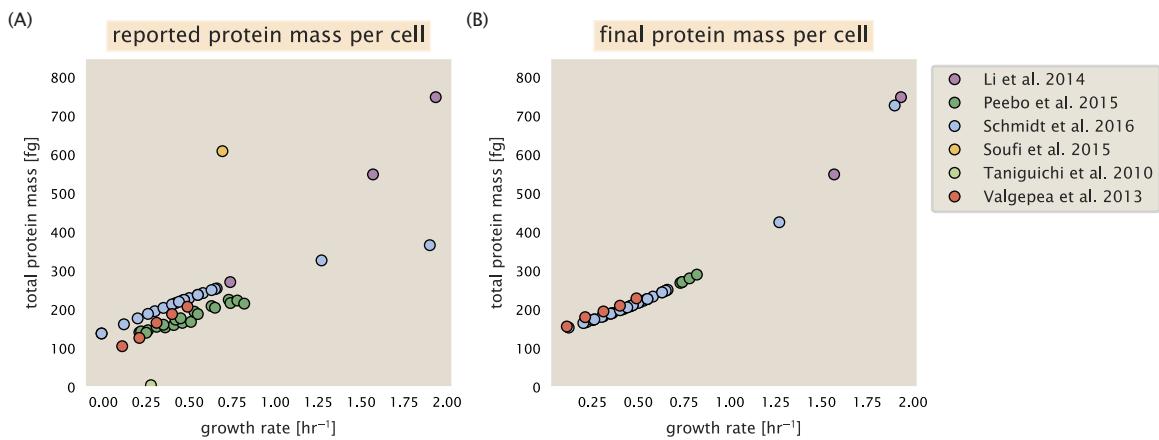


Figure 12. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as originally reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (2016)*. Note that the data from *Peebo et al. (2015)* only reported protein abundances per unit volume and total protein per cell was found by multiplying these by the growth-rate dependent cell size as determined by *Si et al. (2017)*. (B) Adjusted total protein abundances across the proteomic data sets are summarized. Protein abundances were adjusted so that all data shared a common set of growth-rate dependent total protein per cell and cellular protein concentration following the cell size expectations of *Si et al. (2017)* (see section on Estimation of Cell Size and Surface Area for further details).

990 (2010) this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that most
991 proteins appear undercounted when compared to the other data.

992 *Figure 12(B)* summarizes the total protein mass for each data set used in our final compiled data set. Our
993 inclination initially was to leave reported copy numbers untouched, but a notable discrepancy between the scaling
994 of the total protein per cell between *Schmidt et al. (2016)* and the other data sets forced us to dig deeper into
995 those measurements (compare *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 12(A)*). The particular trend
996 in *Schmidt et al. (2016)* appears to be due to assumptions made about cell size and we provide a more extensive
997 discussion and analysis of their data in Additional Considerations of Schmidt et al. Data Set. As a compromise, and
998 in an effort to treat all data equally, we instead applied a correction factor to all protein abundance values based
999 on a data-driven estimate of total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*,
1000 and an estimate of total protein content through expected dry mass. Total protein per cell was then determined
1001 using available data on total DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for
1002 the majority of dry mass in the cell. We describe these details further in sections on Estimation of Cell Size and
1003 Surface Area and Estimation of Total Protein Content per Cell that follows.

1004 Lastly, in *Figure 13* we show the total proteomic coverage and overlap of proteins quantified across each data
1005 set. Here we have used an UpSet diagram (*Lex et al., 2014*) to compare the data. Overall, the overlap in quantified
1006 proteins is quite high, with 1157 proteins quantified across all data sets. The sequencing based approach of *Li*
1007 *et al. (2014)* has substantially higher coverage compared to the mass spectrometry data sets (3394 genes versus
1008 the 2041 genes quantified in the work of *Schmidt et al. (2016)*). However, in terms of total protein mass, the data
1009 from *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (2015)* each quantify roughly equivalent total protein mass.
1010 An exception to this is in the data from *Valgepea et al. (2013)*, where we find that the total protein quantified in
1011 *Valgepea et al. (2013)* is 90-95 % of the total protein mass (when using the data from *Schmidt et al. (2016)* as a
1012 reference).

1013 Estimation of Cell Size and Surface Area

1014 Since most of the proteomic data sets lack cell size measurements, we chose instead to use a common estimate of
1015 size for any analysis requiring cell size or surface area. Since each of the data sets used either K-12 MG1655 or
1016 its derivative, BW25113 (from the lab of Barry L. Wanner; the parent strain of the Keio collection (*Datsenko and*
1017 *Wanner, 2000; Baba et al., 2006*)), below we fit the MG1655 cell size data from the supplemental material of *Si*

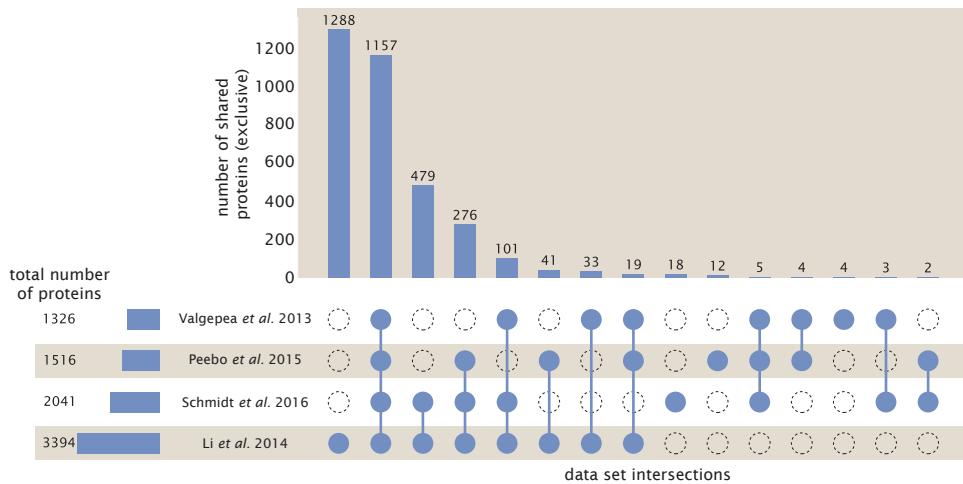


Figure 13. Comparison of proteomic coverage across different data sets. An UpSet diagram (Lex et al., 2014) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015). Bar plot on bottom left indicates the total number of genes reported in each individual data. The main bar plot summarizes the number of unique proteins identified across overlapping subsets of the data. For example, in the first column only the data from Li et al. (2014) is considered (indicated by solid blue circle) and 1288 proteins are identified as exclusive to the data set. In the second column, the intersection of all four data sets is considered, with 1157 proteins quantified across them. This follows for each additional column in the plot, with the subset under consideration denoted by the solid blue circles.

et al. (2017, 2019) using the optimize.curve_fit function from the Scipy python package (Virtanen et al., 2020). A quick comment on nomenclature: throughout the text, we usually refer to cell size, in units of μm^3 ; however, on occasion we will mention size as a volume in units of fL.

The average size measurements from each of their experiments are shown in Figure 14, with cell length and width shown in (A) and (B), respectively. The length data was well described by the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. In order to estimate cell size we take the cell as a cylinder with two hemispherical ends (Si et al., 2017; Basan et al., 2015). Specifically, cell size is estimated from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (5)$$

where r is half the cell width. A best fit to the data is described by $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$. Calculation of the cell surface area is given by,

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (6)$$

where η is the aspect ratio ($\eta = l/w$) (Ojkic et al., 2019).

1028 Estimation of Total Protein Content per Cell

In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the fit shown in Figure 14(C) (cell size = $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$, as noted in the previous section). We then estimate the total protein content from the total dry mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here, protein, DNA, and RNA were reported to account for at least 90 % of the dry mass (Basan et al. (2015)). The authors also found that the total dry mass concentration was roughly constant across growth conditions. Under such a scenario, we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30 \% \times 90 \% \text{ or about } [M_p] = 300 \text{ fg per fL}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how protein concentration might vary due to the substantial increase in rRNA at faster growth rates (Dai et al. (2016)). The increase in rRNA increases from the linear increase in ribosomal content with faster growth rate (Scott et al.

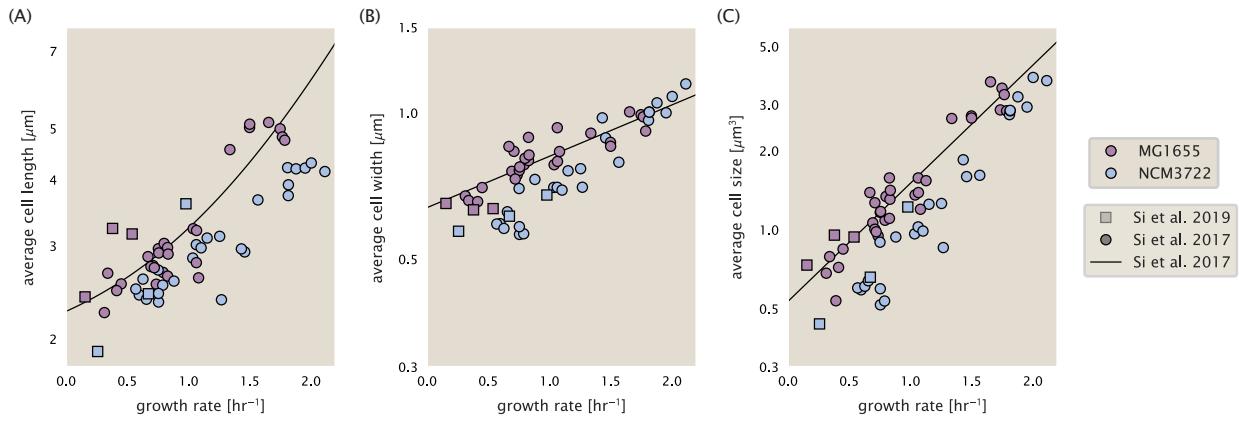


Figure 14. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 images and 5,000-30,000 for each set of samples; which likely means about 1,000-5,000 measurements per mean value reported here since they considered about 6 conditions at a time). Fits were made to the MG1655 strain data; length: $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. (C) Cell size was calculated as cylinders with two hemispherical ends (Equation 5). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

1040 (2010)), since it makes up about about 2/3 or the ribosomal mass. To proceed we therefore relied on experimental
 1041 measurements of total DNA content per cell from Basan et al. (2015), and RNA to protein ratios that were measured
 1042 in Dai et al. (and cover the entire range of growth conditions considered here). These are reproduced in Figure 15(A)
 1043 and (B), respectively.

1044 Assuming that the protein, DNA, and RNA account for 90 % of the total dry mass, the protein mass can then
 1045 determined by first subtracting the experimentally measured DNA mass, and then using the experimental estimate
 1046 of the RNA to protein ratio. The total protein per cell is will be related to the summed RNA and protein mass by,

$$M_P = \frac{[M_P + M_{RNA}]}{1 + (RP_{ratio})}. \quad (7)$$

1047 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai et al.. In Figure 15(C) we plot the estimated cellular
 1048 concentrations for protein, DNA, and RNA from these calculations, and in Figure 15(D) we plot their total expected
 1049 mass per cell. This later quantity is the growth rate-dependent total protein mass that was used to estimate total
 1050 protein abundance across all data sets (and summarized in Figure 12(B)).

1051 Estimating Volume and Number of Amino Acids from Ribosome Copy Number

1052 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A key point
 1053 in our analysis was to consider how elongation rate r_i and growth rate λ vary with respect to the experimentally
 1054 observed changes in cell size, total number of peptide bonds per cell N_{pep} , and ribosomal content. In order to
 1055 restrict parameters to those observed experimentally, but otherwise allow us to explore the model, we performed
 1056 a phenomenological fit of N_{pep} and V as a function of the measured ribosomal copy number R . As has been
 1057 described in the preceding sections of this supplement, we estimate cell volume for each growth condition using
 1058 the size measurements from Si et al. (2017, 2019), and N_{pep} is approximated by taking the total protein mass and
 1059 dividing this number by the average mass of an amino acid, 110 Da (BNID: 104877).

1060 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log
 1061 transform of these parameters as a function of the log transform of the ribosome copy number. Using optimization
 1062 by minimization, we estimated the best-fit values of the intercept and slope for each regression. Figure 16 shows
 1063 the result of each regression as a dashed line.

1064 Additional Considerations of Schmidt et al. Data Set

1065 While the data set from Schmidt et al. (2016) remains a heroic effort that our labs continue to return to as a resource,
 1066 there were steps taken in their calculation of protein copy number that we felt needed further consideration. In

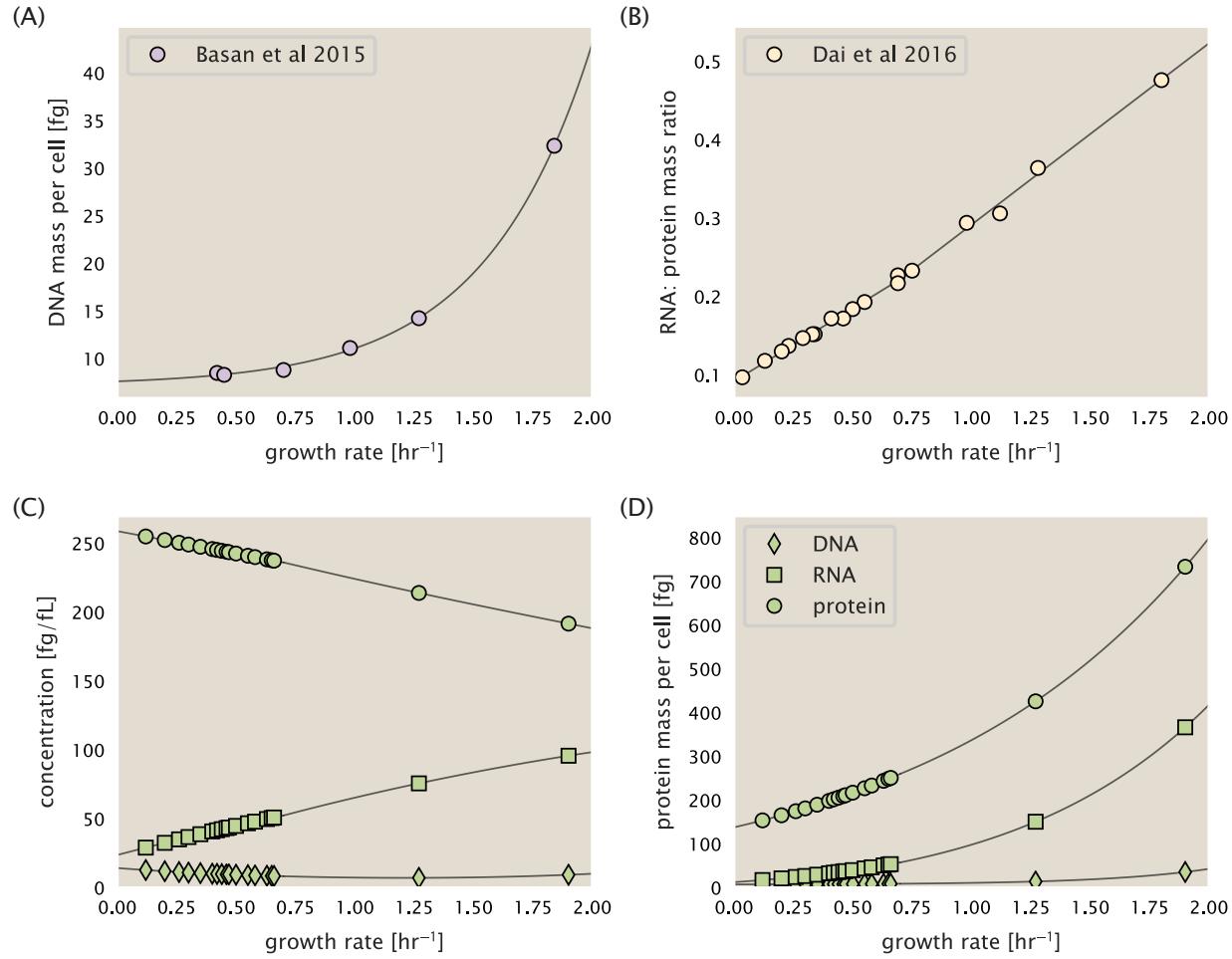


Figure 15. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ fg per cell, where λ is the growth rate in hr^{-1}). (B) RNA to protein measurements as a function of growth rate. The data was fit to two lines (shown in black) due to the change in slope at slower growth rates *Neidhardt et al.* (1991); *Dai et al.* (2016). For growth rates below 0.7 hr^{-1} , the RNA/protein ratio is $0.18 \cdot \lambda + 0.093$, while for growth rates faster than 0.7 hr^{-1} the RNA/protein ratio is given by $0.25 \cdot \lambda + 0.035$. For (A) and (B) cells are grown under varying levels of nutrient limitation, with cells grown in minimal media with different carbon sources for the slowest growth conditions, and rich-defined media for fast growth rates. (C) Estimation of cellular protein, DNA, and RNA concentration. (D) Total cellular mass estimated for protein, DNA, and RNA using the cell size calculated in *Estimation of Cell Size and Surface Area*. Symbols (diamond: DNA, square: RNA, circle: protein) show estimated values of mass concentration and mass per cell for the specific growth rates in *Schmidt et al.* (2016).

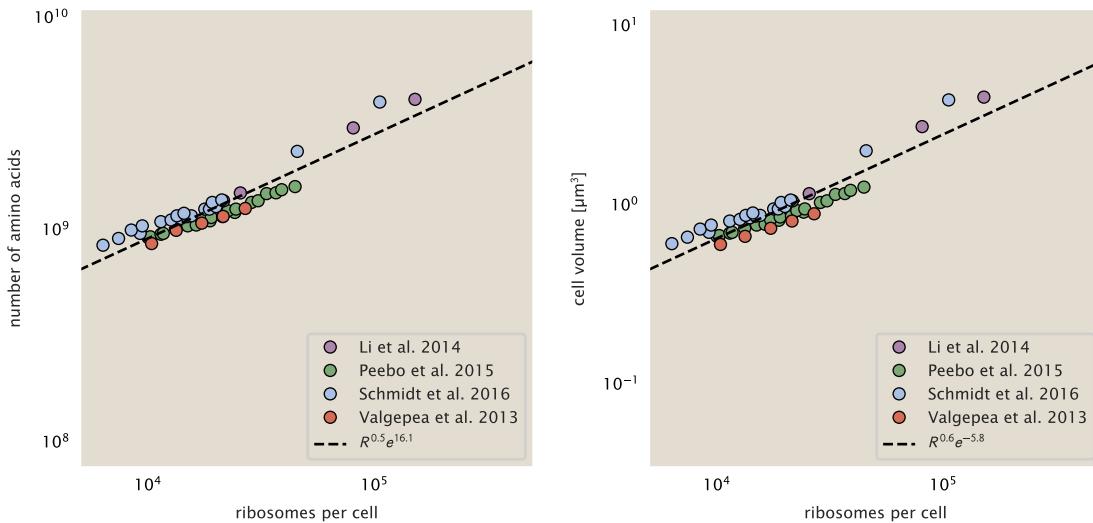


Figure 16. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows the result of the fit, with the functional form of the equation given in the legend with R representing the ribosome copy number.

particular, the authors made an assumption of constant cellular protein concentration across all growth conditions and used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al. (2017)*).

We begin by looking at their cell volume measurements, which are shown in blue in Figure *Figure 17*. As a comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and then calculated cell size by treating the cell as a cylinder with two hemispherical ends, as we considered in the previous section. While there is notable discrepancy between the two research groups, which are both using strain NCM3722, Basan *et al.* found that this came specifically from uncertainty in determining the cell width. This is prone to inaccuracy given the small cell size and optical resolution limits (further described in their supplemental text). Perhaps the more concerning point is that while each of these alternative measurements show an exponential increase in cell size at faster growth rates, the measurements used by Schmidt *et al.* appear to plateau. This resulted in an analogous trend in their final reported total cellular protein per cell as shown in *Figure 18* (purple data points), and is in disagreement with other measurements of total protein at these growth rates (*Basan et al., 2015*).

Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the following subsections we begin by considering how the authors determined total protein mass per cell. We then consider three different approaches to estimate the growth-rate dependent total protein mass and compare these estimates with those reported by Schmidt *et al. (2016)*. Those results are summarized in *Figure 17(B)*, with the original values from both Schmidt *et al. (2016)* and Li *et al. (2014)* shown in *Figure 17(A)* for reference. For most growth conditions, we find reasonable agreement between our estimates and the reported total protein per cell. However, for the fastest growth conditions, with glycerol + supplemented amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main reason why we chose to readjust protein abundance as shown in *Figure 12(B)* (with the calculation described in section Estimation of Total Protein Content per Cell).

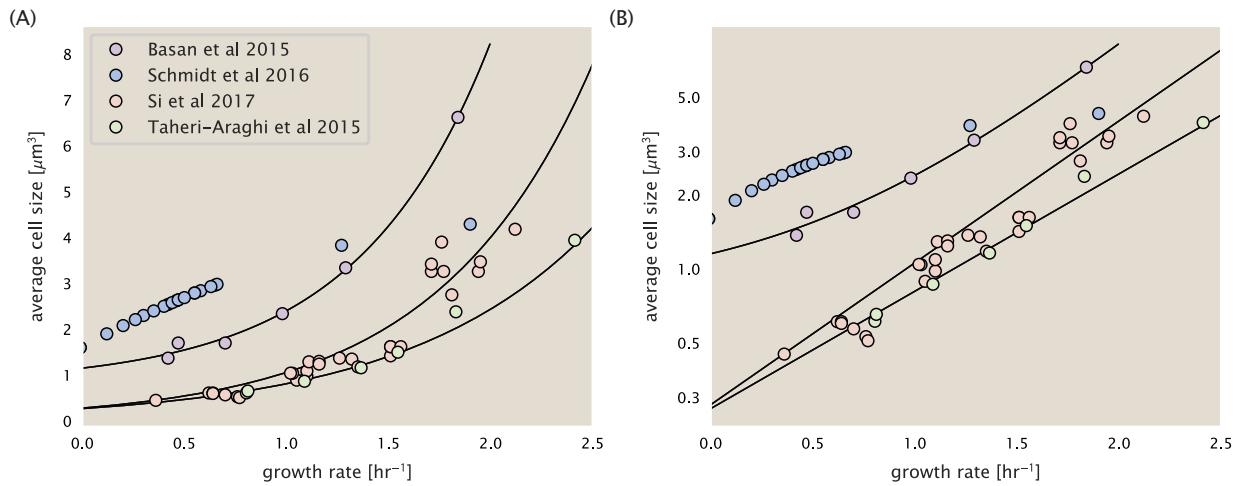


Figure 17. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are shown in purple (*Basan et al. (2015)*), while the two data sets shown in green and light red come from the lab of Suckjoon Jun (*Taheri-Araghi et al. (2015); Si et al. (2017)*). (B) Same as in (A) but with the data plotted on a logarithmic y-axis to highlight the exponential scaling that is expected for *E. coli*.

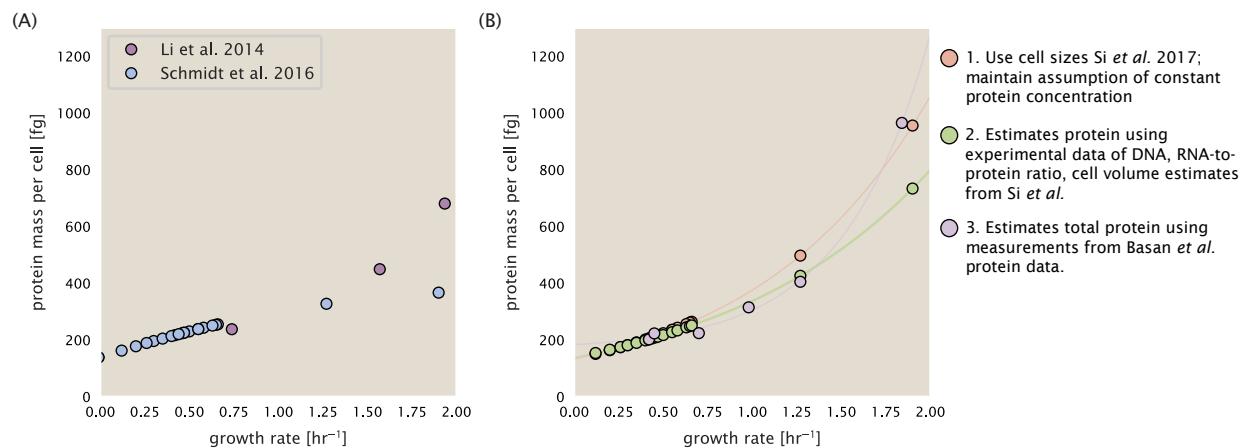


Figure 18. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt *et al.* (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: Rescaling of total protein mass assuming a growth rate independent protein concentration and cell volumes estimated from Si *et al.* 2017. 2. light green: Rescaling of total protein mass using estimates of growth rate-dependent protein concentrations and cell volumes estimated from Si *et al.* 2017. Total protein per cell is calculated by assuming a 1.1 g/ml cellular mass density, 30% dry mass, with 90% of the dry mass corresponding to DNA, RNA, and protein (*Basan et al., 2015*). See Estimation of Total Protein Content per Cell for details on calculation. 3. light purple: Rescaling of total protein mass using the experimental measurements from Basan *et al.* 2015.

1094 **Effect of cell volume on reported absolute protein abundances**

1095 As noted in Experimental Details Behind Proteomic Data, the authors from the work in *Schmidt et al. (2016)*
1096 calculated proteome-wide protein abundances by first determining absolute abundances of 41 pre-selected
1097 proteins, which relied on adding synthetic heavy reference peptides into their protein samples at known abundance.
1098 This absolute quantitation was performed in replicate for each growth condition. Separately, the authors also
1099 performed a more conventional mass spectrometry measurement for samples from each growth condition,
1100 which attempted to maximize the number of quantified proteins but only provided relative abundances based
1101 on peptide intensities. Finally, using their 41 proteins with absolute abundances already determined, they then
1102 created calibration curves with which to relate their relative intensity to absolute protein abundance for each
1103 growth condition. This allowed them to estimate absolute protein abundance for all proteins detected in their
1104 proteome-wide data set. Combined with their flow cytometry cell counts, they were then able to determine
1105 absolute abundance of each protein detected on a per cell basis.

1106 While this approach provided absolute abundances, another necessary step to arrive at total cellular protein
1107 was to account for any protein loss during their various protein extraction steps. Here the authors attempted
1108 to determine total protein separately using a BCA protein assay. In personal communications, it was noted that
1109 determining reasonable total protein abundances by BCA across their array of growth conditions was particularly
1110 troublesome. Instead, they noted confidence in their total protein measurements for cells grown in M9 minimal
1111 media + glucose and used this as a reference point with which to estimate the total protein for all other growth
1112 conditions.

1113 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per cell was measured.
1114 Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentration of $[M_p]_{orig} = M_p/V_{orig} =$
1115 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,
1116 they could then estimate the total protein mass for all other growth conditions from,

$$M_{p_i} = [M_p]_{orig} \cdot V_i \quad (8)$$

1117 where M_{p_i} represents the total protein mass per cell and V_i is the cell volume for each growth condition i as
1118 measured in Volkmer and Heinemann, 2011. Here the thinking is that the values of M_{p_i} reflects the total cellular
1119 protein for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due
1120 to protein loss during sample preparation. The protein abundances from their absolute abundance measurements
1121 noted above were therefore scaled to their estimates and are shown in Figure *Figure 18* (purple data points).

1122 If we instead consider the cell volumes predicted in the work of Si *et al.*, we again need to take growth in M9
1123 minimal media + glucose as a reference with known total mass, but we can follow a similar approach to estimate
1124 total protein mass for all other growth conditions. Letting $V_{Si_glu} = 0.6$ fl be the predicted cell volume, the cellular
1125 protein concentration becomes $[M_p]_{Si} = M_p/V_{Si_glu} = 400$ fg/fl. The new total protein mass per cell can then be
1126 calculated from,

$$M'_{p_i} = [M_p]_{Si} \cdot V_{Si_i} \quad (9)$$

1127 where M'_{p_i} is the new protein mass prediction, and V_{Si_i} refers to the new volume prediction for each condition i ,
1128 These are shown as red data points in Figure *Figure 18(B)*.

1129 **Relaxing assumption of constant protein concentration across growth conditions**

1130 We next relax the assumption that cellular protein concentration is constant and instead, attempt to estimate
1131 it using experimental data. Here we use the estimation of total protein mass per cell detailed in Estimation of
1132 Total Protein Content per Cell for all data points in the *Schmidt et al. (2016)* data set. The green data points in
1133 *Figure 18(B)* show this prediction, and this represents the approach used to estimate total protein per cell for all
1134 data sets.

1135 **Comparison with total protein measurements from Basan *et al.* 2015.**

1136 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration and
1137 cell volumes. These are inherently difficult to measure accurately due to the small size of *E. coli*. Indeed, for all

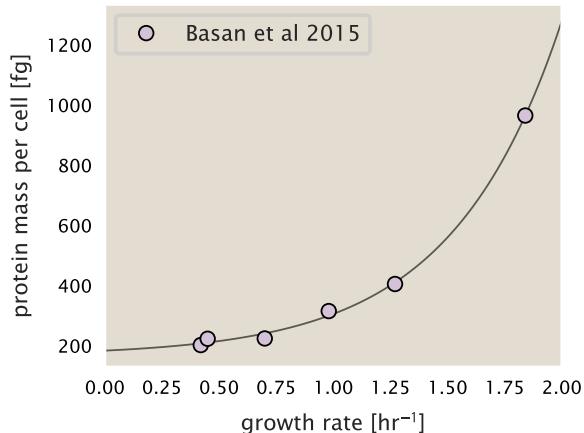


Figure 19. Total cellular protein reported in Basan *et al.* 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given by $14.65 e^{2.180 \cdot \lambda} + 172$ fg per cell, where λ is the growth rate in hr $^{-1}$.

the additional measurements of cell volume included in Figure *Figure 17*, no measurements were performed for cells growing at rates below 0.5 hr^{-1} . It therefore remains to be determined whether our extrapolated cell volume estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower growth.

In our last approach we therefore attempt to estimate total protein using experimental data that required no estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*, the authors measured total protein per cell for a broad range of growth rates (reproduced in Figure *Figure 19*). These were determined by first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method (*You et al. (2013)*), and then abundance per cell was calculated from cell counts from either plating cells or a Coulter counter. While it is unclear why Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more consistent with our expectation that cell mass will increase exponentially with faster growth rates. In addition, although they do not consider growth rates below about 0.5 hr^{-1} , it is interesting to note that the protein mass per cell appears to plateau to a minimum value at slow growth. In contrast, our estimates using cell volume so far have predicted that total protein mass should continue to decrease slightly for slower growing cells. By fitting this data to an exponential function dependent on growth rate, we could then estimate the total protein per cell for each growth condition considered by Schmidt *et al.* (2016). These are plotted as red data points in *Figure 18(B)*.

Calculation of Complex Abundance

All protein data quantified the abundance of individual proteins per cell. However, this work requires estimates on the abundance of individual protein *complexes*, rather than the copy number of individual proteins. In our analysis of the protein copy number data, it became clear that the reported copy numbers do not always align with those based on reported stoichiometry. As one example of this, the F-O subunit of ATP synthase consists of three protein subunits with a stoichiometry of $[AtpB][AtpF]_2[AtpE]_{10}$ (also referred to as subunits a, b, and c, respectively). In the experimental data of Schmidt *et al.* (2016), the values deviate from this quite substantially, with approximately 1000 AtpB, 9000 AtpF, and 300 AtpE reported per cell (minimal media + glucose growth condition). This highlights the technical challenges that still remain in our ability to quantify cellular composition, particularly for membrane-bound proteins like the ATP synthase complex considered here. In this section, we outline the approach we used to annotate proteins as part of each macromolecular complex and how we used averaging across the individual protein measurements to estimate an absolute complex abundances per cell.

Protein complexes, and proteins individually, often have a variety of names, both longform and shorthand. As individual proteins can have a variety of different synonyms, we sought to ensure that each protein annotated in the data sets used the same synonym. To do use, we relied heavily on the EcoCyc Genomic Database (*Keseler*

1169 **et al., 2017**). Each protein in available data sets included an annotation of one of the gene name synonyms as
1170 well as an accession ID – either a UniProt or Blattner "b-number". We programmatically matched up individual
1171 accession IDs between the proteins in different data sets. In cases where accession IDs matched but the gene
1172 names were different, we manually verified that the gene product was the same between the datasets and chose a
1173 single synonym. All code used in the data cleaning and unification procedures can be found on the associated
1174 [GitHub repository] (DOI:XXX) associated with this paper as well as on the associated [paper website](#).

1175 With each protein conforming to a single identification scheme, we then needed to identify the molecular
1176 complexes each protein was a member of. Additionally, we needed to identify how many copies of each protein
1177 were present in each complex (i.e. the subunit copy number) and compute the estimated abundance complex that
1178 accounted for fluctuations in subunit stoichiometry. To map proteins to complexes, we accessed the EcoCyc *E.*
1179 *coli* database **Keseler et al. (2017)** using PathwayTools version 23.0 **Karp et al. (2019)**. With a license for PathWay
1180 Tools, we mapped each unique protein to its annotated complexes via the BioCyc Python package. As we mapped
1181 each protein with *all* of its complex annotations, there was redundancy in the dataset. For example, ribosomal
1182 protein L20 (RplT) is annotated to be a component of the 50S ribosome (EcoCyc complex CPLX-03962) as well as a
1183 component of the mature 70S ribosome (EcoCyc complex CPLX-03964).

1184 In addition to the annotated complex, we collected information on the stoichiometry of each macromolecular
1185 complex. For a complex with N_{subunits} protein species, for each protein subunit i we first calculate the number of
1186 complexes that *could* be formed given the measured protein copy numbers per cell,

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (10)$$

1187 Here, $P_{\text{subunit } i}^{(\text{measured})}$ refers to the measured protein copy number of species i , and m refers to the number of monomers
1188 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,
1189 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each ribosomal
1190 protein, we then calculate the maximum number of complexes that could be formed using **Equation 10**. This
1191 example, along with example from 5 other macromolecular complexes, can be seen in **Figure 20**.

1192 It is important to note that measurement noise, efficiency of protein extraction, and differences in protein
1193 stability will mean that the precise value of each calculation will be different for each component of a given complex.
1194 Thus, to report the total complex abundance, we use the arithmetic mean of across all subunits in the complex,

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (11)$$

1195 in **Figure 20**, we show this mean value as a grey line for a variety of different complexes. Additionally, we have built
1196 an interactive figure accessible on the [paper website](#) where the validity of this approach can be examined for any
1197 complex with more than two subunits (thus, excluding monomers and dimers).

1198 Extending Estimates to a Continuum of Growth Rates

1199 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of the various protein
1200 complexes considered. In addition to point estimates, we also showed the estimate as a function of growth rate as
1201 transparent grey curves. In this section, we elaborate on this continuum estimate, giving examples of estimates
1202 that scale with either cell volume, cell surface area, or number of origins of replication.

1203 Estimation of the total cell mass

1204 For many of the processes estimated in the main text we relied on a cellular dry mass of ≈ 300 fg from which we
1205 computed elemental and protein fractions using knowledge of fractional composition of the dry mass. At modest
1206 growth rates, such as the 5000 s doubling time used in the main text, this is a reasonable number to use as the
1207 typical cell mass is ≈ 1 pg and *E. coli* cells can approximated as 70% water by volume. However, as we have shown
1208 in the preceding sections, the cell size is highly dependent on the growth rate. This means that a dry mass of 300 fg
1209 cannot be used reliably across all growth rates.

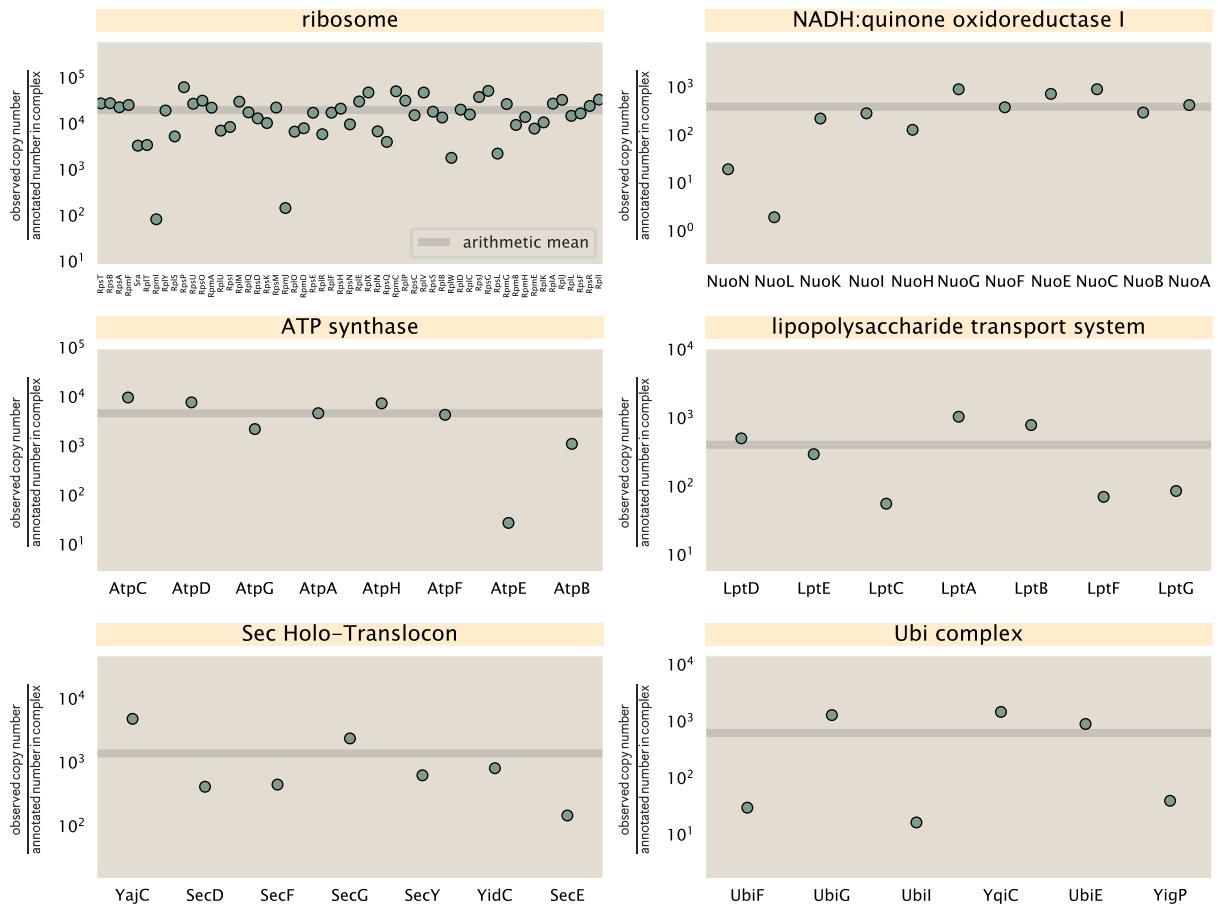


Figure 20. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds to the arithmetic mean across all subunits. These data correspond to measurements from Schmidt *et al.* (2016) in a glucose-supplemented minimal growth medium.

1210 Rather, using the phenomenological description of cell volume scaling exponentially with growth rate, and
 1211 using a rule-of-thumb of a cell buoyant density of $\approx 1.1 \text{ pg} / \text{fL}$ (BNID: 103875), we can calculate the cell dry mass
 1212 across a range of physiological growth rates as

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{\lambda * b} \quad (12)$$

1213 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can be estimated
 1214 from the careful volume measurements performed by *Si et al. (2017,?)*, as considered in Appendix Estimation of
 1215 Cell Size and Surface Area earlier.

1216 Complex Abundance Scaling With Cell Volume

1217 Several of the estimates performed in the main text are implicitly dependent on the cell volume. This includes
 1218 processes such as ATP utilization and, most prominently, the transport of nutrients, whose demand will be
 1219 proportional to the volume of the cell. Of the latter, we estimated the number of transporters that would be
 1220 needed to shuttle enough carbon, phosphorus, and sulfur across the membrane to build new cell mass. To do so,
 1221 we used elemental composition measurements combined with a 300 fg cell dry mass to make the point estimate.
 1222 As we now have a means to estimate the total cell mass as a function of volume, we can generalize these estimates
 1223 across growth rates.

1224 Rather than discussing the particular details of each transport system, we will derive this scaling expression in
 1225 very general terms. Consider that we wish to estimate the number of transporters for some substance X , which
 1226 has been measured to be made up some fraction of the dry mass, θ_X . If we assume that, irrespective of growth
 1227 rate, the cell dry mass is relatively constant (*Basan et al., 2015*) and $\approx 30\%$ of the total cell mass, we can state that
 1228 the total mass of substance X as a function of growth rate is

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (13)$$

1229 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in *Equation 12*. To convert this to the
 1230 number of units N_X of substance X in the cell, we can use the formula weight w_X of a single unit of X in conjunction
 1231 with *Equation 13*,

$$N_X \approx \frac{m_X}{w_X}. \quad (14)$$

1232 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
 1233 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
 1234 can transport substance X at a rate r_X . As this rate r_X is in units of X per time per transporter, we must provide
 1235 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
 1236 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
 1237 transport scaling relation of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (15)$$

1238 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 1239 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 1240 point estimate outlined in the main text. This is due to the choice of $\approx 300 \text{ fg}$ total dry mass per cell for the point
 1241 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 1242 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 1243 and not the predict the exact values of the abundances.

1244 *Equation 15* is a very general relation for processes where the cell volume is the "natural variable" of the
 1245 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 1246 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 1247 dry mass attributable to the substance, and the formula mass of the substance is known, *Equation 15* can be used
 1248 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 1249 *Equation 15* can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (16)$$

1250 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per amino acid for
 1251 translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The grey curves in Figure 4 of the
 1252 main text were made using this type of expression.

1253 A Relation for Complex Abundance Scaling With Surface Area

1254 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used
 1255 a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and the fraction of dry mass attributable to
 1256 peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values come from glucose-fed *E. coli* in balanced growth. As we
 1257 are interested in describing the scaling as a function of the growth rate, we must also consider how these values
 1258 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,
 1259 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1260 and murein monomers that need to be synthesized and crosslinked, respectively.

1261 Number of Lipids

1262 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 1263 the surface area of a single lipid ($A_{lipid} \approx 0.5 \text{ nm}^2$, BNID: 106993) and the total fraction of the membrane composed
 1264 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 1265 **Equation 6**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 1266 lipids as a function of growth rate as

$$N_{lipids}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta\pi \left(\frac{\eta\pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{lipid}} \quad (17)$$

1267 where η is the length-to-width aspect ratio and V is the cell volume.

1268 Number of Murein Monomers

1269 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1270 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murien monomer is
 1271 $m_{murein} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1272 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1273 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1274 number of murein monomers as a function of growth rate can be calculated as

$$N_{murein \text{ monomers}}(\lambda) \approx \frac{\rho_{pg} w \eta \pi \left(\frac{\eta\pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{murein}}, \quad (18)$$

1275 where ρ_{pg} is the density of peptidoglycan.

1276 Complex Abundance Scaling With Number of Origins, and rRNA Synthesis

1277 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
 1278 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
 1279 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
 1280 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in
 1281 *Si et al. (2017)* and reproduced in the main text and Appendix Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ below, the
 1282 average number of origins of replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{cyc}\lambda / \ln 2} \quad (19)$$

1283 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{cyc} \approx 70 \text{ min}$, which is
 1284 the time from the initiation of chromosomal replication until division. This time corresponds to the sum of the
 1285 so-called C and D periods of the cell cycle, which correspond to the time it takes to replicate the entire chromosome
 1286 (C period) and the time from completion to eventual division (D period) *Helmstetter and Cooper (1968)*.

1287 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1288 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{\text{rRNA}}(\lambda) \approx N_{\text{rRNA operons}} \times 2^{t_{\text{cyc}}\lambda/\ln 2}. \quad (20)$$

1289 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons are
1290 some distance away from the origin, making **Equation 20** an approximation (*Dennis et al., 2004*).

1291 In the main text, we stated that at a growth rate of 0.5 hr^{-1} , there is ≈ 1 chromosome per cell. While a fair
1292 approximation, **Equation 19** illustrates that is not precisely true, even at slow growth rates. In estimating the
1293 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1294 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1295 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{\text{rRNA}} \rho_{\text{RNA polymerase}}, \quad (21)$$

1296 where L_{operon} is the total length of an rRNA operon ($\approx 4500 \text{ bp}$) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA
1297 polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1298 Calculation of active ribosomal fraction.

1299 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai et al. (2016)* to
1300 estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to
1301 consider how f_a should vary with growth rate, and instead find that the data is well-approximated by fitting to an
1302 exponential curve ($f_a = -0.889 e^{4.6\lambda} + 0.922$; dashed line in inset of **Figure 9(C)**). We use this function to estimate f_a
1303 for each of the data points shown in **Figure 9(C)**.

1304 Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$.

1305 *E. coli* shows robust scaling of cell size with the average number of origins per cell, $\langle \# \text{ori} \rangle$ (*Si et al., 2017*). Since
1306 protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes in
1307 proteomic composition and total abundance across growth conditions. Given the potential constraints on rRNA
1308 synthesis and changes in ribosomal copy number with $\langle \# \text{ori} \rangle$, it becomes important to also consider how protein
1309 copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make sense of
1310 the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a mechanistic
1311 level. As considered in the main text, it is becoming increasingly apparent that regulation through the secondary
1312 messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer nutrient
1313 conditions. In this context, both $\langle \# \text{ori} \rangle$, as well as the $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ ratio become important parameters to consider
1314 and keep track of. An increase in $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ ratio in particular, causes a relatively higher gene dosage in rRNA
1315 and r-protein genes due to skew in genes near the origin, where the majority of these are located

1316 In the main text we estimated the change in $\langle \# \text{ori} \rangle$ with growth rate using the nutrient-limited wild-type cell
1317 data from *Si et al. (2017)*. We consider their measurements of DNA replication time (t_C , 'C' period of cell division),
1318 total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across
1319 a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ ratio
1320 from their data. We begin by considering $\langle \# \text{ori} \rangle$. If the cell cycle time takes longer than the time of cell division, the
1321 cell will need to initiate DNA replication more often than its rate of division, $2^{\lambda t} = 2^{\ln(2)\cdot t/\tau}$ to maintain steady state
1322 growth. Cells will need to do this in proportion to the ratio $\lambda_{\text{cyc}}/\lambda = t_{\text{cyc}}/\tau$, and the number of origins per cell (on
1323 average) is then given by $2^{t_{\text{cyc}}/\tau}$. The average number of termini will in contrast depend on the lag time between
1324 DNA replication and cell division, t_D , with $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ ratio = $2^{t_{\text{cyc}}/\tau - t_D/\tau} = 2^{t_C/\tau}$.

1325 In Figure 21(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from *Si et al. (2017)*.
1326 The authors estimated t_C by marker frequency analysis using qPCR, while $t_{\text{cyc}} = t_C + t_D$ were inferred from t_C and
1327 τ . In the plots we see that both t_C and t_{cyc} reach a minimum at around 40 and 75 minutes, respectively. For a C
1328 period of 40 minutes, this would correspond to a maximum rate of elongation of about 1,000 bp/sec. Since we
1329 lacked a specific model to describe how each of these parameters vary with growth condition, we assumed that
1330 they were linearly dependent on the doubling time. For each parameter, t_C and t_{cyc} , we split them up into two

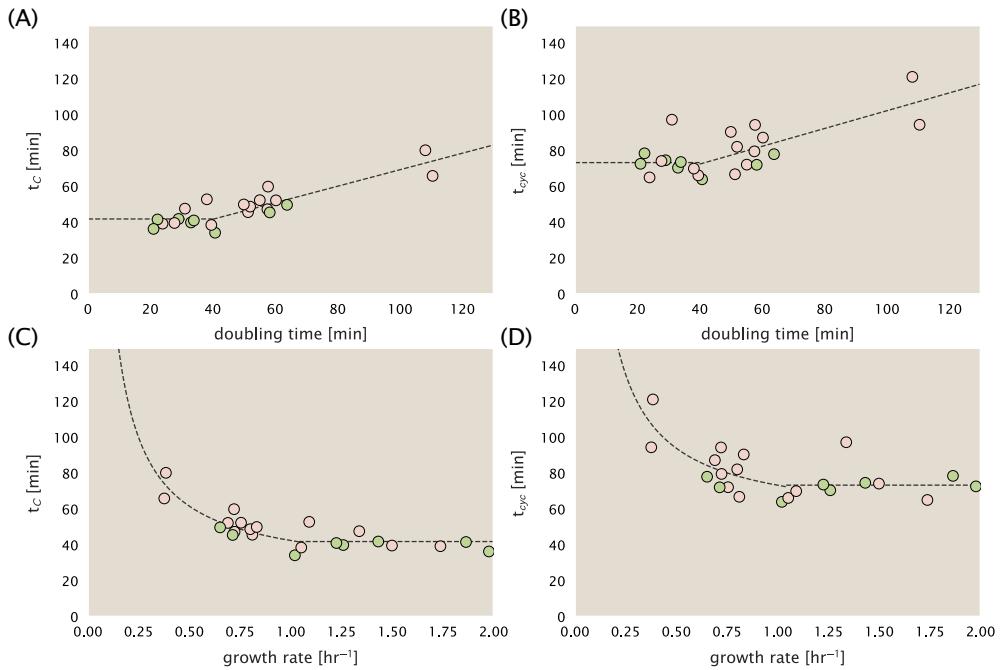


Figure 21. Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant ($t_C = 42$ minutes, $t_{cyc} = 73$ minutes). At the transition, taken to occur at 40 minutes, the dashed line corresponds to an assumed proportional increase in each parameter as a function of the doubling time ($t_C = 0.46\tau + 23.3$ minutes, $t_{cyc} = 0.50\tau + 52.7$ minutes). (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = \ln(2)/\tau$.

domains corresponding to poorer nutrient conditions and rich nutrient conditions (cut off at $\tau \approx 40$ minutes where chromosomal replication becomes nearly constant). The fit lines are shown as solid black lines. In Figure 21(C) and (D) we also show t_C and t_{cyc} as a function of growth rate λ along with our piecewise linear fits, which match the plots in the main text.

Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control

Here we provide a derivation of the minimal model for growth rate control under nutrient-limited growth. By growth rate control, we are specifically referring to the ability of bacteria to modulate their proteome (N_{pep} , R , Φ_R) and cell size as nutrient conditions change, with slower growing cells generally being smaller in size (Ojkic et al., 2019). This capability provides bacteria with a particular benefit when nutrients are more scarce since it will mean there is a smaller net demand on carbon, phosphorus, sulfur, and nitrogen. The specific goal of developing this model is to help us better explore the overall constraints on growth that follow from 1) our observation that many of the cellular processes we've considered require increased protein abundance at faster growth rates, and 2) a strict limit on growth rate that is governed by the ribosomal synthesis rate and ribosomal mass fraction Φ_R .

In Figure 11(A) of the main text we provide a schematic of the model, where we consider growth as simply governed by the rate of protein synthesis ($r_t \times R \times f_a$). In order to grow rapidly, at least to the extent possible, these three parameters need to be maximized (with $r_t \leq 17$ amino acids per second, and $f_a \leq 1$ reported in the work of Dai et al. (2016)). The elongation rate r_t will depend on how quickly ribosomes can match codons with their correct amino-acyl tRNA, along with the subsequent steps of peptide bond formation and translocation. This ultimately depends on the cellular concentration amino acids, which we treat as a single effective species, $[AA]_{eff}$.

In our model, we need to determine the rate of peptide elongation r_t , which we consider as simply depending on the supply of amino acids (and, therefore, also amino-acyl tRNAs) through a parameter r_{AA} in units of AA per second, and the rate of amino acid consumption by protein synthesis ($r_t \times R \times f_a$). The balance between these two rates will determine the effective amino acid concentration in the cell $[AA]_{eff}$. An important premise for this formulation

is growing evidence that cells are able to modulate their biosynthesis activity according to nutrient availability (i.e. extent of chromosomal replication, transcriptional, and translation activity) through secondary-messenger molecules like (p)ppGpp (*Hauryliuk et al., 2015; Zhu and Dai, 2019; Kraemer et al., 2019; Fernández-Coll et al., 2020; Büke et al., 2020*). Given our observation that protein synthesis and energy production are not limiting, we assume that other molecular players required by ribosomes like elongation factors and GTP are available in sufficient abundance. In addition, experimentally, the relative number of tRNA and elongation factor EF-Tu per ribosome have been found to increase in poorer nutrient conditions *Pedersen (1978); Dong et al. (1996); Klumpp et al. (2013)*.

We begin by considering a coarse-grained description of peptide elongation, which includes 1) the time required to find and bind each correct amino-acyl tRNA, and 2) the remaining steps in peptide elongation that will not depend on the amino acid availability. These time scales will be related to the inverse of the elongation rate r_t ,

$$\frac{1}{r_t} = \frac{1}{k_{on}\alpha[AA]_{eff}} + \frac{1}{r_t^{\max}}. \quad (22)$$

where we have assumed that the rate of binding by amino-acyl tRNA k_{on} is proportional to $[AA]_{eff}$ by a constant α . r_t^{\max} refers to the maximum elongation rate. This leads to a Michaelis-Menten dependence of the elongation rate r_t on the effective amino acid concentration $[AA]_{eff}$ (*Klumpp et al., 2013; Dai et al., 2016*). We can re-write this more succinctly in terms of an effective dissociation constant,

$$K_D = \frac{r_t^{\max}}{\alpha k_{on}}, \quad (23)$$

where the elongation rate r_t is now given by

$$r_t = \frac{r_t^{\max}}{1 + K_D/[AA]_{eff}}. \quad (24)$$

The rate of amino acid supply r_{AA} will vary with changing nutrient conditions and the cell can maintain $[AA]_{eff}$ by tuning the rate of amino acid consumption, $r_t \times R \times f_a$. Thus, $[AA]_{eff}$ is determined by the difference in the rate of amino acid synthesis (or import, for rich media) and/or tRNA charging, r_{AA} , and the rate of consumption, $r_t \times R \times f_a$. Over an arbitrary length of time t of cellular growth, the cell will grow in volume, requiring us to consider these rates in terms of concentration rather than absolute numbers, with $[AA]_{eff}$ given by,

$$\int_0^t \frac{d[AA]_{eff}}{dt} dt = \int_0^t ([r_{AA}] - [r_t \times R \times f_a]) dt. \quad (25)$$

This considers the net change in amino acid concentration over a time from 0 to t , with the square brackets indicating concentrations per unit time. Integrating **Equation 25** yields.

$$[AA]_{eff} = t([r_{AA}] - [r_t \times R \times f_a]). \quad (26)$$

Alternatively, to connect to the experimental data in terms of absolute ribosome copy number R we can consider a unit volume V ,

$$[AA]_{eff} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V \times N_A}, \quad (27)$$

where r_{AA} is in units of AA per unit time and r_t is in units of AA per unit time per ribosome. N_A refers to Avogadro's number and is needed to convert between concentration and absolute numbers per cell. With an expression for $[AA]_{eff}$ in hand, we can now solve **Equation 24** for r_t , which is a quadratic function with a physically-meaningful root of

$$r_t = \frac{t(r_{AA} + r_t^{(\max)} R f_a) + K_D V N_A - \sqrt{(r_{AA} + r_t^{(\max)} R f_a + K_D V N_A)^2 - 4(R f_a t)(r_t^{(\max)} r_{AA} t)}}{2R f_a t}. \quad (28)$$

This is the key equation that allows us to calculate growth rate for any combination of N_{pep} , R , f_a , and cell size V as a function of amino acid supply r_{AA} (**Equation 3** of the main text). We refer the reader to A Minimal Model of Nutrient-Mediated Growth Rate Control of the main text for our exploration of this model in the context of the proteomic data.

1387 We end this section by noting several distinctions of this formulation with previous work. The first, as noted
1388 in the main text, relates to the now seminal work of *Scott et al. (2010)*, which provides a treatment of resource
1389 allocation that partitions of the proteome into sectors – including one for ribosome-associated proteins and one
1390 for metabolic proteins. As cells grow faster, there is a notable change in the mass fraction of these sectors, with
1391 an increase in ribosomal content that is predominantly achieved at the expense of a decrease in the metabolic
1392 sector. By including an additional constraint through the phenomenological parameter ν , which characterizes the
1393 quality of the growth medium *Scott et al. (2010); Klumpp et al. (2013); Klumpp and Hwa (2014)*, the authors derive
1394 a model of growth rate, dependent on optimal resource allocation. Here we have developed a model that considers
1395 the effect of changes in absolute protein abundance and ribosomal content, and consider how these influence the
1396 achievable growth rate. In addition, by accounting for the metabolic supply of amino acids directly through their
1397 availability in the cell (i.e. $[AA]_{\text{eff}}$), we are able to consider how the balance between translation-specific metabolic
1398 capacity and translational capacity influences both the elongation rate r_e and growth rate λ .

1399 The second and last point we note is that the recent works from *Dai et al. (2016)* and *Klumpp et al. (2013)* also
1400 employ a similar coarse-graining of translation elongation as we've considered above. Here, however, a notable
1401 distinction is that the authors consider the entire ternary complex (i.e. the complex of amino-acyl tRNA, EF-Tu,
1402 and GTP) as rate limiting. Further, through an assumed proportionality between ternary complex and ribosome
1403 abundance, they arrive at a formulation of elongation rate r_e that exhibits a Michaelis-Menten dependence on the
1404 ribosomal fraction Φ_R . They demonstrate that all their measurements of elongation rate, even upon addition of
1405 sublethal doses of chloramphenicol (which cause an increase in both r_e and Φ_R), can be collapsed onto a single
1406 curve described by this Michaelis-Menten dependence. There is always a benefit to increase their ribosomal
1407 fraction Φ_R on growth rate when nutrient conditions allow (see Maximum Growth Rate is Determined by the
1408 Ribosomal Mass Fraction), and this trend in the data in part follows from the tendency for cells to increase Φ_R
1409 and better maximize r_e as nutrient conditions improve. In addition, it does not account for the decrease in the
1410 fraction of actively translating ribosome f_a that was strikingly apparent at slow growth rates or in sublethal doses of
1411 chloramphenicol in the work of *Dai et al. (2016)*. Through **Equation 28** we also account for changes in the fraction
1412 of actively translating ribosomes. Ultimately, we find that cells are able to maximize both Φ_R , r_e , and their growth
1413 rate only to the extent allowed by the nutrient conditions (i.e. via r_{AA}) and through the maintenance of the cellular
1414 pool of amino acids $[AA]_{\text{eff}}$, amino-acyl tRNA, GTP, as well as the synthesis of other key molecular constituents like
1415 EF-Tu.

1416 **References**

- 1417 Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast: II. The lifetime
1418 of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 1419 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1420 *coli*. *BMC Systems Biology*, 8(1):133.
- 1421 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by
1422 ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1423 Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris,
1424 M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Mateescu, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G.
1425 (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1426 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1427 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1428 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1429 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH₃ Channel.
1430 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1431 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1432 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1433 2(1):2460.
- 1434 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1435 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1436 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1437 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1438 Belliveau, N. M., Barnes, S. L., Ireland, W. T., Jones, D. L., Sweredoski, M. J., Moradian, A., Hess, S., Kinney, J. B., and Phillips, R.
1439 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings of
1440 the National Academy of Sciences*, 115(21):E4796–E4805.
- 1441 Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport systems. The lactose
1442 permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 1443 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential
1444 Growth Rates. *EcoSal Plus*, 3(1).
- 1445 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews
1446 Microbiology*, 14(10):638–650.
- 1447 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1448 266:2020.06.16.154187.
- 1449 Catherwood, A. C., Lloyd, A. J., Tod, J. A., Chauhan, S., Slade, S. E., Walkowiak, G. P., Galley, N. F., Punekar, A. S., Smart, K., Rea, D.,
1450 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1451 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1452 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1453 *Biochemical Journal*, 117(3):551–562.
- 1454 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational
1455 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1456 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1457 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1458 Microbiology*, 2(2):16231.
- 1459 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1460 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1461 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1462 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.

- 1463 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*,
1464 108(44):17876–17882.
- 1465 Dong, H., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA Abundance and Codon Usage in *Escherichia coli* at Different
1466 Growth Rates. *Journal of Molecular Biology*, 260(5):649–663.
- 1467 Erickson, D. W., Schink, S. J., Patsalo, V., Williamson, J. R., Gerland, U., and Hwa, T. (2017). A global resource allocation strategy
1468 governs growth transition kinetics of *Escherichia coli*. *Nature*, 551(7678):119–123.
- 1469 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyru-
1470 vate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation.
1471 *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1472 Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson,
1473 B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and
1474 thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 1475 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M.
1476 (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth
1477 Rates. *mBio*, 11(2):45.
- 1478 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair.
1479 *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1480 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1481 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of
1482 the National Academy of Sciences*.
- 1483 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández,
1484 K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-
1485 Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D.,
1486 Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif
1487 clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1488 Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide
1489 Reductase. *Biochemistry*, 42(34):10071–10083.
- 1490 Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D., Kirschner, M. W., and
1491 Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1492 Guo, Y., Li, D., Zhang, S., Yang, Y., Liu, J.-J., Wang, X., Liu, C., Milkie, D. E., Moore, R. P., Tulu, U. S., Kiehart, D. P., Hu, J., Lippincott-
1493 Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution
1494 on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1495 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1496 microbiology*, 26(10):815–832.
- 1497 Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from *Escherichia coli*. *Journal of
1498 Bacteriology*, 183(17):5008–5014.
- 1499 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of
1500 (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1501 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content
1502 of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1503 Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal of
1504 Molecular Biology*, 31(3):507–518.
- 1505 Henkel, S. G., Beek, A. T., Steinsiek, S., Stagge, S., Bettenbrock, K., de Mattos, M. J. T., Sauter, T., Sawodny, O., and Ederer, M. (2014).
1506 Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1507 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic
1508 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- 1509 Ingledeew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.

- 1510 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1511 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1512 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1513 20(17):4952–4963.
- 1514 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1515 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1516 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong,
1517 Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1518 *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1519 Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W.
1520 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- 1521 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari,
1522 A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez,
1523 D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about
1524 *Escherichia coli*K-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1525 Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004). Mechanism of Ammonia
1526 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–1594.
- 1527 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1528 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1529 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1530 *Current Opinion in Biotechnology*, 28:96–102.
- 1531 Klumpp, S., Scott, M., Pedersen, S., and Hwa, T. (2013). Molecular crowding limits translation and cell growth. *Proceedings of the
1532 National Academy of Sciences*, 110(42):16754–16759.
- 1533 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–
1534 1375.
- 1535 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1536 125(2):028103.
- 1537 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli*by
1538 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1539 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and
1540 Biomembranes*, 32(3):237–246.
- 1541 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1542 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1543 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions
1544 on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1545 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1546 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1547 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1548 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1549 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1550 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1551 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator
1552 GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1553 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1554 112(51):15690–15695.
- 1555 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On. Gene Expression*. Springer.

- 1556 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by
1557 condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1558 Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and Lee, P. A. (2009). A
1559 Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1560 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1561 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1562 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1563 *Growth Symposium*, 9:223–289.
- 1564 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1565 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall
1566 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–
1567 12515.
- 1568 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1569 Ojikic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.
1570 *eLife*, 8:642.
- 1571 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1572 Pedersen, S. (1978). Patterns of protein synthesis in *E. coli*: a catalog of the amount of 140 individual proteins at different growth
1573 rates. *Cell*, 14(1):179–190.
- 1574 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with
1575 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1576 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1577 Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active transport in *Escherichia coli*
1578 membrane vesicles. *Biochemistry*, 16(5):854–859.
- 1579 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An
1580 integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*,
1581 14(6):687–704.
- 1582 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1583 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive
1584 strategies. *Nature microbiology*, 1(11):1–7.
- 1585 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of
1586 Bacteriology*, 131(2):505–511.
- 1587 Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome stability. *DNA Repair*,
1588 44:193–204.
- 1589 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–
1590 11734.
- 1591 Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside diphosphate reductase during
1592 multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–2225.
- 1593 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical
1594 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1595 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,
1596 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1597 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:
1598 origins and consequences. *Science*, 330(6007):1099–1102.
- 1599 Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction.
1600 *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.

- 1601 Shi, H., Bratton, B. P., Gitai, Z., and Huang, K. C. (2018). How to Build a Bacterial Cell: MreB as the Foreman of *E. coli* Construction.
1602 *Cell*, 172(6):1294–1305.
- 1603 Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control and Homeostasis in
1604 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1605 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
1606 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1607 Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for a
1608 functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*, 177(14):4134–4136.
- 1609 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews*,
1610 40(1):133–159.
- 1611 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth
1612 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1613 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- 1614 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van
1615 Leeuwenhoek*, 39(1):545–565.
- 1616 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.
1617 *Archives of Microbiology*, 113(3):185–189.
- 1618 Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly unstable
1619 during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.
- 1620 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1621 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1622 Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control and
1623 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1624 Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X. S. (2010). Quantifying *E. coli* proteome and
1625 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1626 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1627 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1628 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1629 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1630 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1631 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1632 47(D1):D330–D338.
- 1633 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1634 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1635 van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*: Putting Molecular Data
1636 into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.
- 1637 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1638 Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson,
1639 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1640 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1641 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1642 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data
1643 Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1644 Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*,
1645 32(2):149–167.
- 1646 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.

- 1647 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1648 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1649 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
1650 bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1651 Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from
1652 *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- 1653 Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent sulfate transporter that
1654 can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(7):1809–1816.
- 1655 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite repressor/activator, cra, in
1656 *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1657 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1658 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.

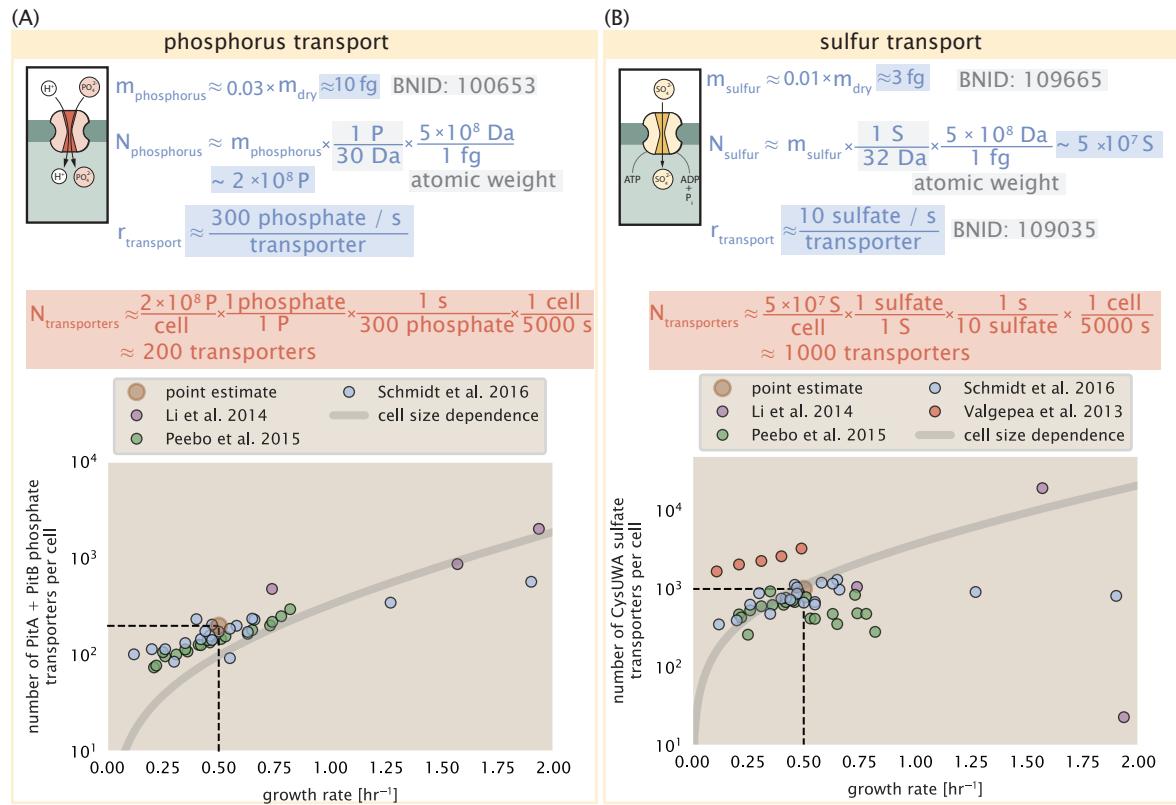


Figure 2-Figure supplement 1. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (B) Estimate of the number of CysUWA complexes necessary to maintain a 1% sulfur *E. coli* dry mass. Points in plot correspond to average number of CysUWA transporter complexes that can be formed given the transporter stoichiometry $[CysA]_2[CysU][CysW][Sbp/CysP]$. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

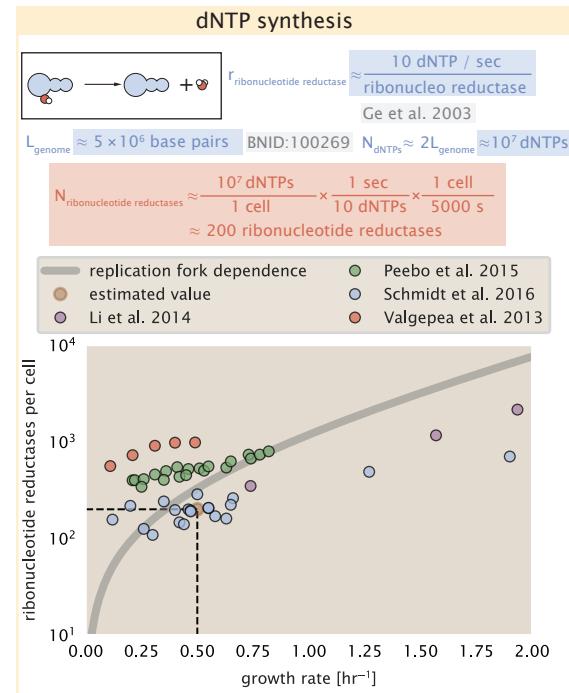


Figure 6–Figure supplement 1. Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[NrdA]_2[NrdB]_2$) and ribonucleotide reductase II ($[NrdE]_2[NrdF]_2$) complexes. Grey lines in top panel show the estimated number of complexes needed as a function of growth, the details of which are described in the Appendix.

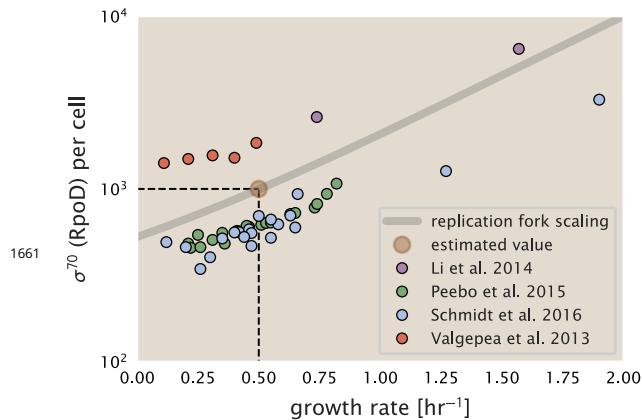


Figure 7–Figure supplement 1. The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown as a translucent brown point and grey line.

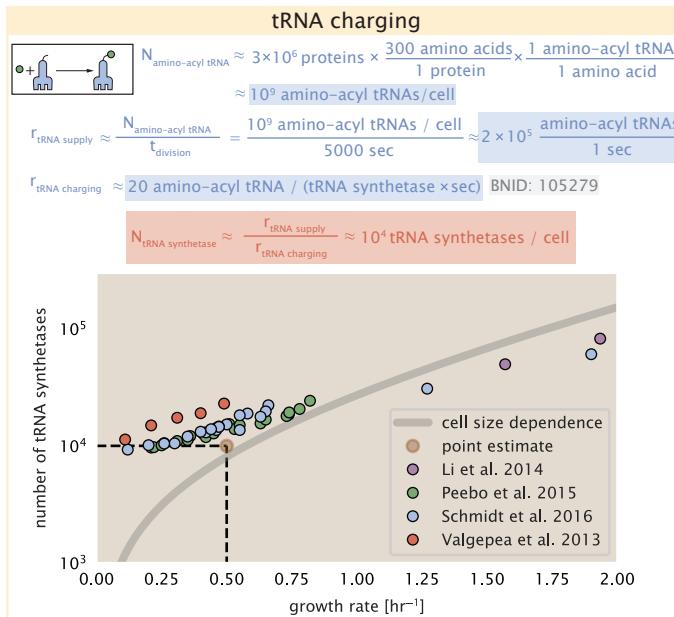


Figure 8-Figure supplement 1. Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [lysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂).

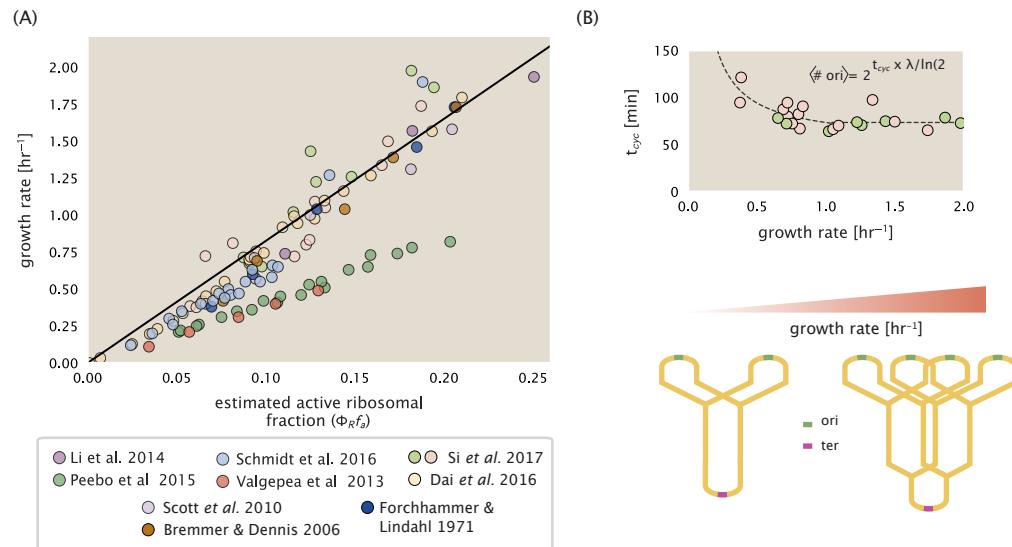


Figure 9-Figure supplement 1. (A) Actively translating ribosomal fraction versus growth rate. The actively translating ribosomal fraction is calculated using the estimated values of f_a from **Dai et al. (2016)** (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). Additional measurements in addition to the proteomic measurements are based on measurements of cellular RNA to protein ratio, with $\Phi_R \approx$ the cellular RNA to protein ratio divided by 2.1 (**Dai et al., 2016**). (B) Experimental measurements of the cell doubling time τ and cell cycle time t_{cyc} from **Si et al. (2017)**. Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times greater than 40 minutes, and reach a minimum value of 73 minutes. See Appendix Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ for additional details exact estimation of rRNA copy number. Red data points correspond to measurements in strain MG1655, while light green points are for strain NCM3722. Schematic shows the expected increase in replication forks (or number of ori regions) as *E. coli* cells grow faster.

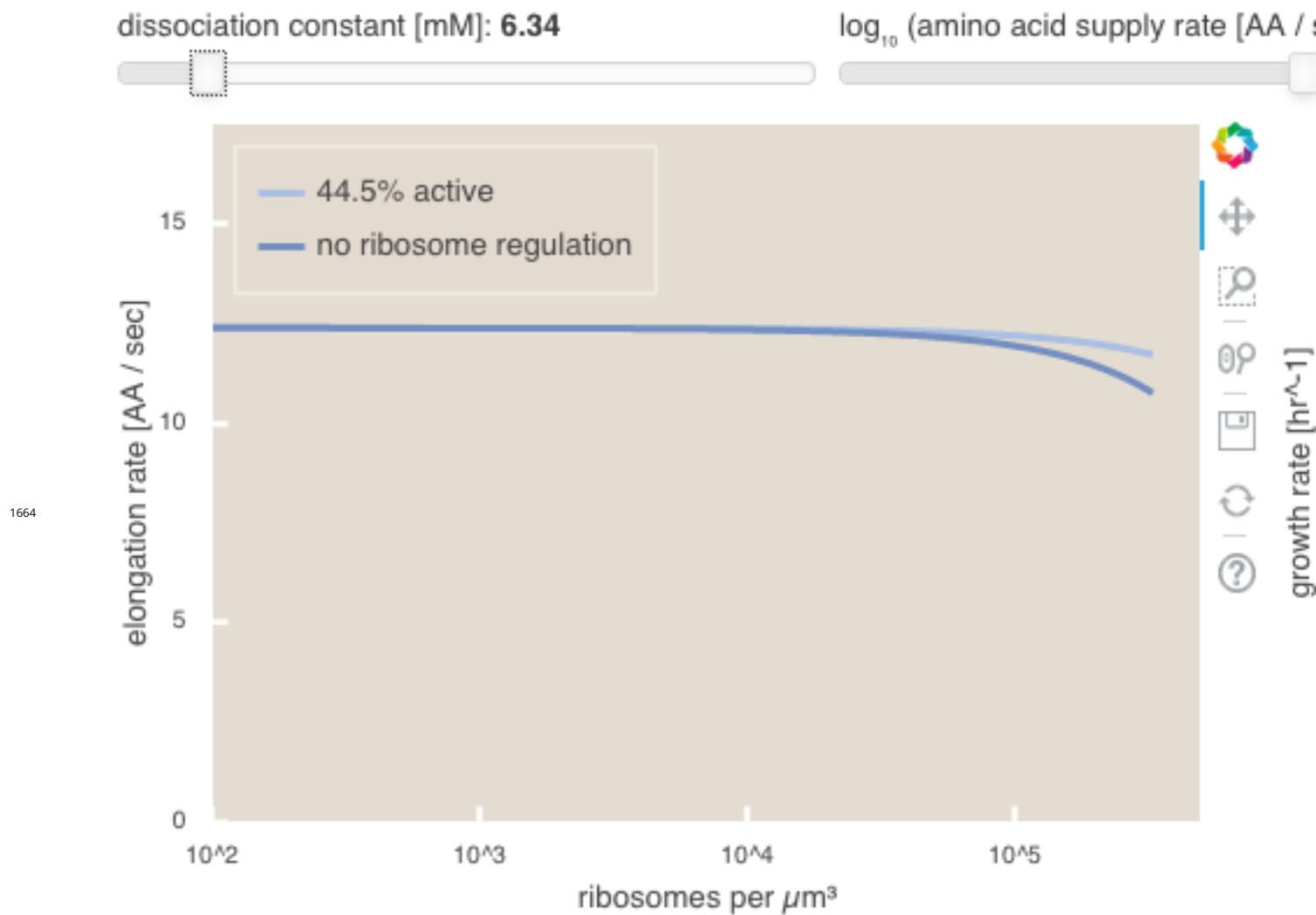


Figure 11–Figure supplement 1. An interactive version of parts (B) and (C) of **Figure 11** which permit the user to modulate the rate of amino acid supply, the dissociation constant of amino acids to the ribosome, and the fraction of the ribosome pool that is actively translating. This interactive figure, and the code used to generate it, is available on the [paper website](#).