

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 sets covering \approx 4000 proteins and 31 growth conditions. 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 which ribosome biogenesis emerges as a primary determinant of growth rate. We conclude by exploring a model of ribosomal 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 might 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 similar 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*).

As was noted by Jacques Monod, “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*). Indeed, the study of bacterial growth has undergone a 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*. In this work, we explore an amalgamation of recent proteomic data sets to explore fundamental limits of bacterial growth.

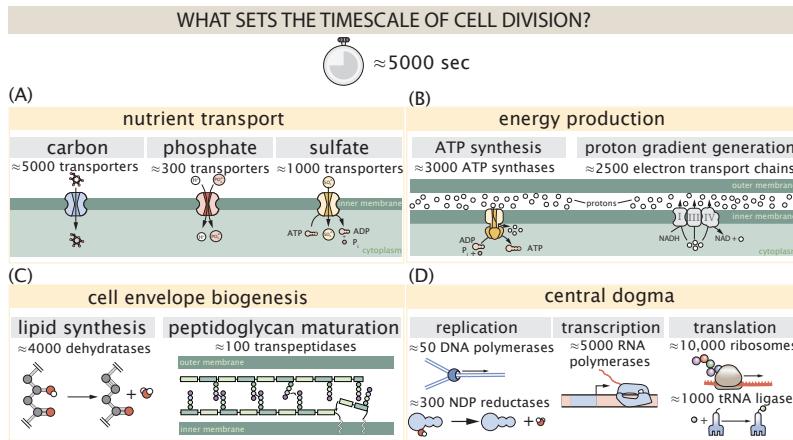


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP synthesis), (C) cell envelope biogenesis, 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 $\approx 5000 \text{ s}$.

42 Several of the evergreen questions about bacterial growth that were originally raised by microbiologists in the
 43 middle of the 20th century can now be reframed in light of this newly available data. For example, what biological
 44 processes set the absolute speed limit for how fast bacterial cells can grow and reproduce? How do cells alter the
 45 absolute numbers and relative ratios of their molecular constituents as a function of changes in growth rate or
 46 nutrient availability? In this paper, we address these two questions from two distinct angles. First, as a result of an
 47 array of high-quality proteome-wide measurements of the *E. coli* proteome under myriad growth conditions, we
 48 have a census that allows us to explore how the number of key molecular players change as a function of growth
 49 rate. Here, we have compiled a combination of *E. coli* proteomic data sets collected over the past decade using
 50 either mass spectrometry ([Schmidt et al., 2016](#); [Peebo et al., 2015](#); [Valgepea et al., 2013](#)) or ribosomal profiling
 51 ([Li et al., 2014](#)) across 31 unique growth conditions (see Appendix Experimental Details Behind Proteomic Data
 52 for further discussion of these data sets). Second, by compiling molecular turnover rate measurements for many
 53 of the fundamental processes associated with bacterial growth, we make quantitative estimates of key cellular
 54 processes (schematized in [Figure 1](#)) to determine whether the observed protein copy numbers under varying
 55 conditions appear to be in excess of what would be minimally required to support cell growth at the observed
 56 rates. The census, combined with these estimates, provide a window into the question of whether the rates of
 57 central processes such as energy generation or DNA synthesis are regulated systematically as a function of cell
 58 growth rate by altering protein copy number in individual cells.

59 Throughout our estimates, we consider a modest growth rate of $\approx 0.5 \text{ hr}^{-1}$ corresponding to a doubling time
 60 of ≈ 5000 seconds, as the the data sets heavily sample this regime. While we formulate point estimates for the
 61 complex abundances at this division time, we consider how these values will vary at other growth rates due to
 62 changes in cell size, surface area, and chromosome copy number ([Taheri-Araghi et al., 2015](#)). Broadly, we find that
 63 for the majority of these estimates, the protein copy numbers appear well-tuned for the task of cell doubling at a
 64 given growth rate. It emerges that translation, particularly of ribosomal proteins, is the most plausible candidate
 65 for a molecular bottleneck. We reach this conclusion by considering that translation is 1) a rate limiting step for the
 66 *fastest* bacterial division, and 2) a major determinant of bacterial growth across the nutrient conditions we have
 67 considered under steady state, exponential growth. This enables us to suggest that the long-observed correlation
 68 between growth rate and cell size ([Schaechter et al., 1958](#); [Si et al., 2017](#)) can be simply attributed to the increased
 69 absolute number of ribosomes per cell under conditions supporting extremely rapid growth, a hypothesis which
 70 we formally mathematize and explore.

72

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

73

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

74

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

75

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

76

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

77

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

110 **Nutrient Transport**

111 We begin our series of estimates by considering the critical transport processes diagrammed in **Figure 1(A)**. In order
112 to build new cellular mass, the molecular and elemental building blocks must be scavenged from the environment
113 in different forms. Carbon, for example, is acquired via the transport of carbohydrates and sugar alcohols with
114 some carbon sources receiving preferential treatment in their consumption (**Monod, 1947**). Phosphorus, sulfur,
115 and nitrogen, on the other hand, are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate,
116 and ammonia (**Jun et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al.,**
117 **1977; Willsky et al., 1973**). All of these compounds have different permeabilities across the cell membrane (**Phillips,**
118 **2018**) and most require some energetic investment either via ATP hydrolysis or through the proton electrochemical
119 gradient to bring the material across the hydrophobic cell membrane.

120 The elemental composition of *E. coli* has received much quantitative attention over the past half century
121 (**Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv, 1976**), providing us with a
122 starting point for estimating how many atoms of each element must be scavenged from the environment. While
123 there is some variability in the exact elemental percentages (with different uncertainties), we can approximate the
124 dry mass of an *E. coli* cell to be $\approx 45\%$ carbon (BioNumber ID: 100649, see **Box 1**), $\approx 15\%$ nitrogen (BNID: 106666),
125 $\approx 3\%$ phosphorus (BNID: 100653), and 1% sulfur (BNID: 100655). Here we use this stoichiometric breakdown to
126 estimate the abundance and growth rate dependence of a variety of transporters responsible for carbon uptake,
127 and provide more extensive investigation of the other critical elements – phosphorus, sulfur, and nitrogen – in the
128 Appendix ??.

129 Using ≈ 0.3 pg as the typical *E. coli* dry mass (BNID: 103904) at a growth rate of $\approx 0.5 \text{ hr}^{-1}$, and nearly half
130 of the dry mass consisting of carbon, we estimate that $\sim 10^{10}$ carbon atoms must be brought into the cell in
131 order to double all of the carbon-containing molecules (**Figure 2(A, top)**). Typical laboratory growth conditions
132 provide carbon as a single class of sugar such as glucose, galactose, or xylose to name a few. *E. coli* has evolved
133 myriad mechanisms by which these sugars can be transported across the cell membrane. One such mechanism
134 of transport is via the PTS system which is a highly modular system capable of transporting a diverse range of
135 sugars (**Escalante et al., 2012**). The glucose-specific component of this system transports ≈ 200 glucose molecules
136 per second per transporter (BNID: 114686). Making the assumption that this is a typical sugar transport rate,
137 coupled with the need to transport $\sim 10^{10}$ carbon atoms, we then expect on the order of 1000 transporters must
138 be expressed in order to bring in enough carbon atoms, diagrammed in the top panel of **Figure 2(A)**. While we
139 estimate 1500 transporters are needed with a 5000 s division time, we can abstract this calculation to consider any
140 particular growth rate given knowledge of the cell density and volume as a function of growth rate and direct the
141 reader to the Appendix Extending Estimates to a Continuum of Growth Rates for more information. As revealed in
142 **Figure 2(A)**, experimental measurements exceed the estimate by several fold, suggesting that transport of carbon
143 into the cell is not rate limiting for cell division. Abstracting this point estimate at 5000 s to a continuum of growth
144 rates (grey line in **Figure 2(A)**) reveals an excess of transporters even at faster growth rates. This contrasts with our
145 observations for uptake of phosphorus and sulfur, which align well with our expectations across different growth
146 conditions (**Figure 2–Figure Supplement 1** and discussed further in Appendix ??).

147 The estimate presented in **Figure 2(A)** neglects any specifics of the regulation of the carbon transport system
148 and the data shows how many carbohydrate transporters are present on average. Using the diverse array of
149 growth conditions available in the data, we also explore how individual carbon transport systems depend on
150 specific carbon availability. In **Figure 2(B)**, we show the total number of carbohydrate transporters specific to
151 different carbon sources. A striking observation, shown in the top-left plot of **Figure 2(B)**, is the constancy in the
152 expression of the glucose-specific transport systems. Additionally, we note that the total number of glucose-specific
153 transporters is tightly distributed at $\approx 10^4$ per cell, the approximate number of transporters needed to sustain
154 rapid growth of several divisions per hour. This illustrates that *E. coli* maintains a substantial number of complexes
155 present for transporting glucose regardless of growth condition, which is known to be the preferential carbon
156 source (**Monod, 1947; Liu et al., 2005; Aidelberg et al., 2014**).

157 Many metabolic operons are regulated with dual-input logic gates that are only expressed when glucose
158 concentrations are low (mediated by cyclic-AMP receptor protein CRP) and the concentration of other carbon
159 sources are elevated (**Gama-Castro et al., 2016; Zhang et al., 2014; Gama-Castro et al., 2016; Belliveau et al., 2018;**

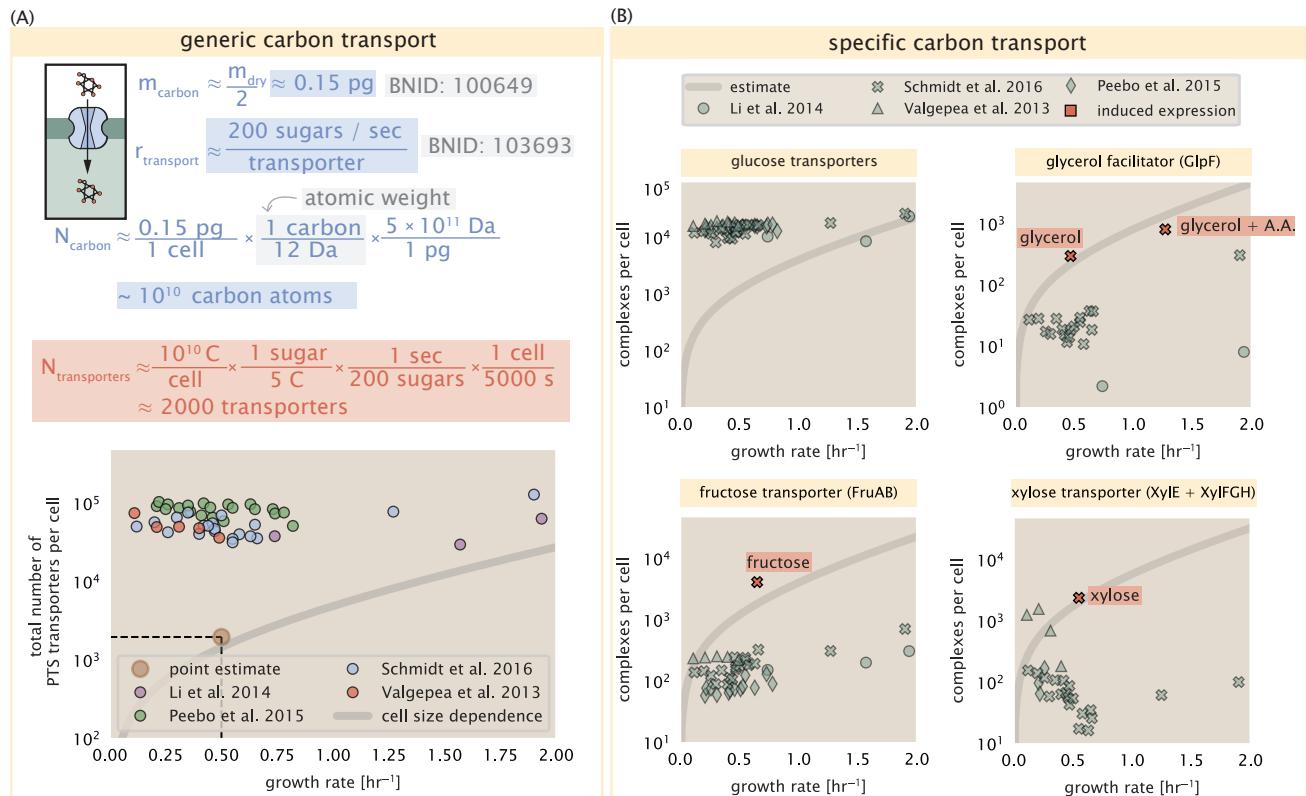


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules 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 to compute the continuum growth rate estimate (grey line) were 200 glucose- s^{-1} (BNID: 103693), 2000 glycerol- s^{-1} (**Liu 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 line 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.

160 *Ireland et al., 2020*). Several examples are shown in **Figure 2(B)**. Points colored in red (labeled by red text-boxes)
161 correspond to growth conditions in which the specific carbon source (glycerol, xylose, or fructose) is present. The
162 grey lines in **Figure 2(B)** show the estimated number of transporters needed at each growth rate to satisfy the
163 cellular carbon requirement, adjusted for the specific carbon source. These plots show that, in the absence of
164 the particular carbon source, expression of the transporters is maintained on the order of $\sim 10^2$ per cell. The
165 low but non-zero abundances may reflect the specific regulatory logic involved, requiring that cells are able to
166 transport some minimal amount of an alternative carbon source in order to induce expression of these alternative
167 carbon-source systems.

168 If acquisition of nutrients was the limiting process in cell division under typical growth conditions, could
169 expression simply be increased to accommodate faster growth? A way to approach this question is to compute
170 the amount of space in the bacterial membrane that could be occupied by nutrient transporters. Considering a
171 rule-of-thumb for the surface area of *E. coli* of about $5 \mu\text{m}^2$ (BNID: 101792), we expect an areal density for 1000
172 transporters to be approximately 200 transporters/ μm^2 . For a typical transporter occupying about $50 \text{ nm}^2/\text{dimer}$,
173 this amounts to about only 1 percent of the total inner membrane (*Szenk et al., 2017*). In addition, bacterial cell
174 membranes typically have densities of 10^5 proteins/ μm^2 (*Phillips, 2018*), implying that the cell could accommodate
175 more transporters if any one was rate limiting.

176 Energy Production

177 While the uptake of nutrients provides essential resources, metabolic pathways must then consume and generate
178 energy in the form of NTPs to build new cell mass. The high-energy phosphodiester bonds of (primarily) ATP
179 power a variety of cellular processes that drive biological systems away from thermodynamic equilibrium. The next
180 set of processes we consider as a potential bottleneck in growth is the synthesis of ATP from ADP and inorganic
181 phosphate as well as maintenance of the electrochemical proton gradient which powers it.

182 ATP Synthesis

183 Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic
184 phosphate provides the kinetic driving force in a wide array of biochemical reactions. One such reaction is the
185 formation of peptide bonds during translation, which requires ≈ 2 ATPs for the charging of an amino acid to
186 the tRNA and ≈ 2 GTP for the formation of each peptide bond. Assuming the ATP costs associated with error
187 correction and post-translational modifications of proteins are negligible, we can make the approximation that each
188 peptide bond has a net cost of ≈ 4 ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from ATP is achieved
189 via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment
190 (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to be functionally equivalent to
191 being exclusively ATP. In total, the energetic costs of peptide bond formation consume $\approx 80\%$ of the cells ATP
192 budget (BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015); Stouthamer (1973)*). The pool of ATP is
193 produced by the F₁-F₀ ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield \approx
194 300 ATP per second (BNID: 114701; *Weber and Senior (2003)*).

195 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make the approximation
196 that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300 peptide bonds (BNID: 115702;
197 108986; 104877). Taking these values together, we find that the typical *E. coli* cell consumes $\sim 5 \times 10^9$ ATP per
198 cell cycle on protein synthesis alone. Assuming that each ATP synthases provides 300 ATP per second, ≈ 3000
199 ATP synthases are needed to keep up with the energy demands of the cell. This estimate and a comparison with
200 the data are shown in **Figure 3 (A)**. Despite our assumption of maximal ATP production rate per synthase and
201 approximation of all NTP consuming reactions being the same as ATP, we find that an estimate of a few thousand
202 complete synthases per cell agree well with the experimental data. Much as we did for the estimates of transporter
203 copy numbers, we can generalize this estimate across different growth rates (indicated by the gray line in **Figure 3**),
204 which appears to be in reasonable agreement. At the fastest growth rates, we actually predict that cells need more
205 ATP synthases than are measured. However, this may reflect an increase in fermentive synthesis of ATP at faster
206 growth rates (*Szenk et al., 2017*).

207 **Generating the Proton Electrochemical Gradient**

208 In order to produce ATP, the F_1 - F_0 ATP synthase itself must consume energy. Rather than burning through its
209 own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the
210 electrochemical potential established across the inner membrane through cellular respiration. This electrochemical
211 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as
212 they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner
213 membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in the
214 intermembrane space.

215 However, the constant rotation of the ATP synthases would rapidly abolish this potential difference if it were
216 not being actively maintained. To undergo a complete rotation (and produce a single ATP), the F_1 - F_0 ATP synthase
217 must shuttle ≈ 4 protons across the membrane into the cytosol (BNID: 103390). With ≈ 3000 ATP synthases each
218 generating 300 ATP per second, the 2×10^4 protons establishing the 200 mV potential would be consumed in only
219 a few milliseconds. This brings us to our next estimate: how many electron transport complexes are needed to
220 support the consumption rate of the ATP synthases?

221 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical
222 and biophysical study over the past half century (*Ingle dew and Poole, 1984; Khademian and Imlay, 2017; Cox*
223 *et al., 1970; Henkel et al., 2014*). A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E.*
224 *coli* electron transport complexes using structural and biochemical data, revealing that each electron transport
225 chain rapidly pumps protons into the intermembrane space at a rate of ≈ 1500 protons per second (BNID: 114704;
226 114687). Using our estimate of the number of ATP synthases required per cell [*Figure 3(A)*], coupled with these
227 recent measurements, we estimate that ≈ 1000 electron transport complexes would be necessary to facilitate the
228 $\approx 5 \times 10^6$ protons per second diet of the cellular ATP synthases. This estimate is in agreement with the number of
229 complexes identified in the proteomic datasets (plot in *Figure 3(B)*). This suggests that every ATP synthase must
230 be accompanied by ≈ 1 functional electron transport chain. Again, to consider whether energy production may
231 become a rate-limiting step in growth, we need to consider whether cells could simply increase their abundances
232 in the cell membrane, and we turn our attention towards the available space in the membrane next.

233 **Biosynthesis in a Crowded Membrane**

234 For each protein considered so far, the data shows that their numbers generally increase with growth rate. This is
235 in part a consequence of the increase in cell length and width with growth rate that is common to many rod-shaped
236 bacteria (*Ojikic et al., 2019; Harris and Theriot, 2018*). For *E. coli*, the total cellular protein and cell size increase
237 exponentially with growth rate (*Schaechter et al., 1958; Si et al., 2017*). Of particular concern here is that while
238 demand for energy will then also grow exponentially, these complexes require more space on a bacterial membrane
239 whose surface area-to-volume (S/V) ratio is actually decreasing at faster growth rates.

240 In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell cycle or 10^6
241 ATP/s. With a cell volume of roughly 1 fL, this corresponds to about 2×10^{10} ATP per fL of cell volume, in line with
242 previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al., 2017*). In *Figure 4 (A)* we plot this ATP
243 demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from spherical
244 to rod-shaped with an aspect ratio (length/width) equal to 4. In order to consider the maximum ATP that could
245 be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP synthase and
246 electron transport complexes, which provides a maximal production of about 3 ATP / (nm²·s) (*Szenk et al., 2017*).
247 This is shown in blue in *Figure 4(A)*, which shows that at least for the growth rates observed (right column in plot),
248 the energy demand is roughly an order of magnitude less. Interestingly, *Szenk et al. (2017)* also found that ATP
249 production by respiration is less efficient than by fermentation per membrane area occupied due to the additional
250 proteins of the electron transport chain. This suggests that, even under anaerobic growth, there will be sufficient
251 membrane space for ATP production.

252 The analysis highlights the diminishing capacity to provide resources as the cell increases in size. However,
253 the maximum energy production in *Figure 4(A)* does represent a somewhat unachievable limit since the inner
254 membrane must also include other proteins including those required for lipid and membrane synthesis. Here we
255 consider two key steps in lipid and peptidoglycan synthesis, and similarly find protein abundances that largely

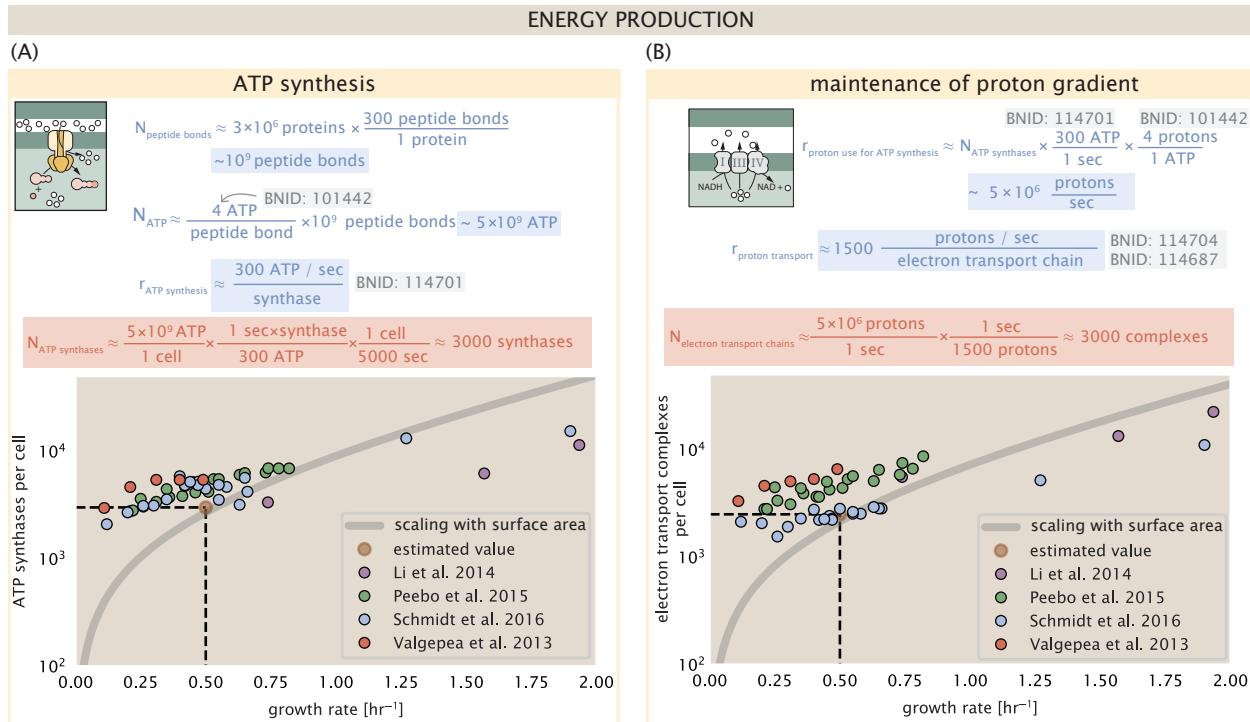


Figure 3. The abundance of F_1 - F_0 ATP synthases and electron transport chain complexes as a function of growth rate.

(A) Estimate of the number of F_1 - F_0 ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F_1 - F_0 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 F_1 - F_0 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.

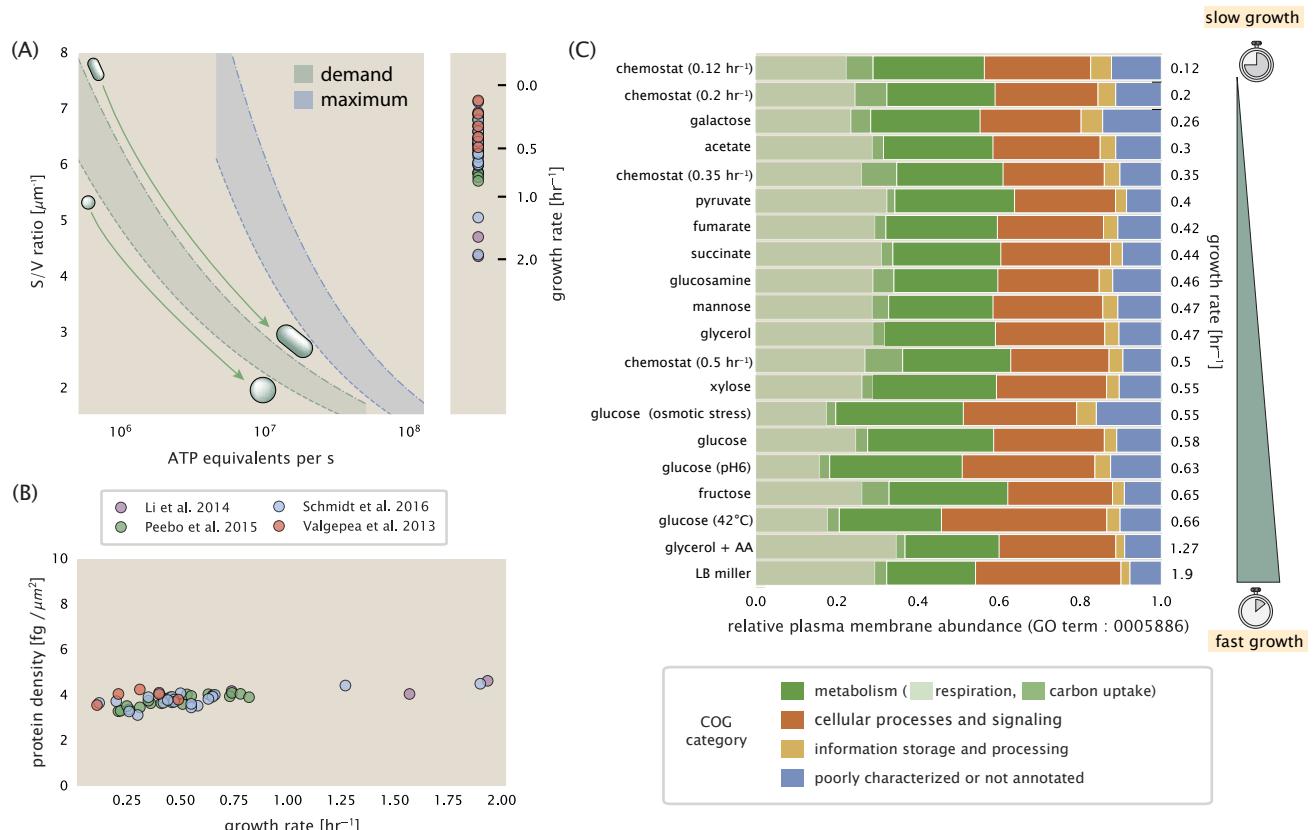


Figure 4. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (–) line corresponding to a sphere and the dash-dot line (–.) 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)).

Figure 4–Figure supplement 1. Estimation of the key components involved in cell envelope biosynthesis.

256 scale according to our estimates (Figure 4–Figure Supplement 1, and more fully discussed in Appendix ??). To
 257 better understand the overall proteomic makeup of the inner membrane, we therefore used Gene Ontology (GO)
 258 annotations (Ashburner et al., 2000; The Gene Ontology Consortium, 2018) to identify all proteins embedded or
 259 peripheral to the inner membrane (GO term: 0005886). Those associated but not membrane-bound include
 260 proteins like MreB and FtsZ, that traverse the inner membrane by treadmilling and must nonetheless be considered
 261 as a vital component occupying space on the membrane. In Figure 4(B), we find that the total protein mass per
 262 μm^2 is nearly constant across growth rates. Interestingly, when we consider the distribution of proteins grouped by
 263 their Clusters of Orthologous Groups (COG) (Tatusov et al., 2000), the relative abundance for those in metabolism
 264 (including ATP synthesis via respiration) is also relatively constant across growth rates, suggesting that no one
 265 process (energy production, nutrient uptake, etc.) is particularly dominating even at fast growth rates Figure 4(C).

266 Function of the Central Dogma

267 Up to this point, we have considered a variety of transport and biosynthetic processes that are critical to acquiring
 268 and generating new cell mass. While there are of course many other metabolic processes we could consider and
 269 perform estimates of, we now turn our focus to some of the most central processes which *must* be undertaken

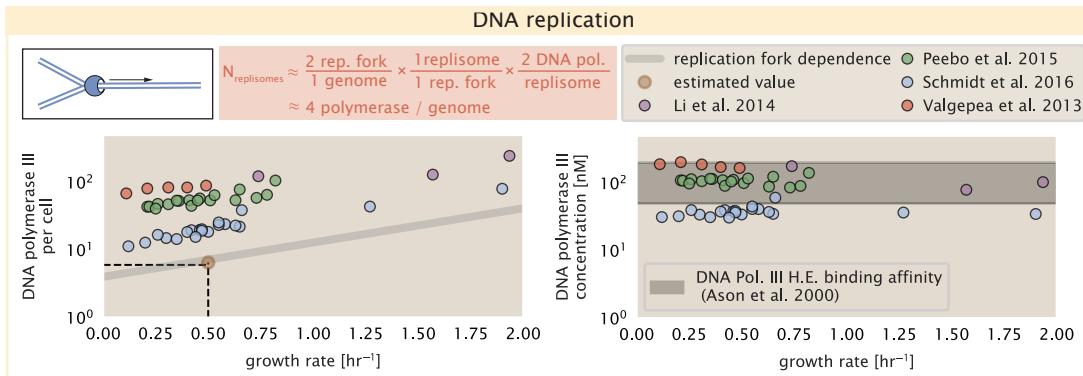


Figure 5. Complex abundance estimates for dNTP synthesis and DNA replication. An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the left-hand plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[DnaE]_3[DnaQ]_3[HolE]_3[DnaX]_5[HolB][HolA][DnaN]_4[HolC]_4[HolD]_4$) per cell. Right-hand plot shows the effective concentration of DNA polymerase III holoenzyme (See Appendix Estimation of Cell Size and Surface Area for calculate of cell volumes). 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 5-Figure supplement 1. Estimate and observations of the abundance of ribonucleotide reductase, a key component in dNTP synthesis.

270 irrespective of the growth conditions – the processes of the central dogma.

271 DNA Replication

272 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids
273 up to ~ 100 kbp in length. We will focus our quantitative thinking solely on the chromosome of *E. coli* which harbors
274 ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs. To successfully divide and produce viable progeny, this chromosome must
275 be faithfully replicated and segregated into each nascent cell. We again rely on the near century of literature
276 in molecular biology to provide some insight on the rates and mechanics of this replicative feat, as well as the
277 production of the required starting materials dNTPs (Figure 5-Figure Supplement 1, and discussed in Appendix ??).

278 Replication is initiated at a single region of the chromosome termed the *oriC* locus at which a pair of DNA
279 polymerases bind and begin their high-fidelity replication of the genome in opposite directions. Assuming
280 equivalence between the two replication forks, this means that the two DNA polymerase complexes (termed
281 replisomes) meet at the midway point of the circular chromosome termed the *ter* locus. The kinetics of the five
282 types of DNA polymerases (I – V) have been intensely studied, revealing that DNA polymerase III performs the
283 high fidelity processive replication of the genome with the other "accessory" polymerases playing auxiliary roles
284 (Fijalkowska et al., 2012). *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate of ≈ 600
285 nucleotides per second (BNID: 104120). Therefore, to replicate a single chromosome, two replisomes (containing
286 two DNA polymerase III each) moving at their maximal rate would copy the entire genome in ≈ 4000 s. Thus, with a
287 division time of 5000 s (our "typical" growth rate for the purposes of this work), there is sufficient time for a pair of
288 replisomes complexes to replicate the entire genome. However, this estimate implies that 4000 s would be the
289 upper-limit time scale for bacterial division which is at odds with the familiar 20 minute (1200 s) doubling time of *E.*
290 *coli* in rich medium.

291 Even in rapidly growing cultures, where bacteria like *E. coli* parallelize its DNA replication with as many as 10 - 12
292 replication forks at a given time (Bremer and Dennis, 2008; Si et al., 2017), we expect only a few polymerases (≈ 10)
293 are needed. However, as shown in Figure 5 DNA polymerase III is nearly an order of magnitude more abundant.
294 This discrepancy can be understood by considering its binding constant to DNA. DNA polymerase III is highly
295 processive, facilitated by a strong affinity of the complex to the DNA. *In vitro* biochemical characterization has
296 quantified the K_D of DNA polymerase III holoenzyme to single-stranded and double-stranded DNA to be 50 and
297 200 nM, respectively (Ason et al., 2000). The right-hand plot in Figure 5 shows that the concentration of the DNA
298 polymerase III across all data sets and growth conditions is within this range. Thus, while the copy number of the

299 DNA polymerase III is in excess of the strict number required to replicate the genome, its copy number appears to
300 vary such that its concentration is approximately equal to the dissociation constant to the DNA. While the processes
301 regulating the initiation of DNA replication are complex and involve more than just the holoenzyme, these data
302 indicate that the kinetics of replication rather than the explicit copy number of the DNA polymerase III holoenzyme
303 is the more relevant feature of DNA replication to consider. In light of this, the data in *Figure 5* suggests that for
304 bacteria like *E. coli*, DNA replication does not represent a rate-limiting step in cell division. However, it is worth
305 noting that for bacterium like *C. crescentus* whose chromosomal replication is initiated only once per cell cycle
306 (*Jensen et al., 2001*), the time to double their chromosome indeed represents an upper limit to their growth rate.

307 RNA Synthesis

308 With the machinery governing the replication of the genome accounted for, we now turn our attention to the next
309 stage of the central dogma – the transcription of DNA to form RNA. We consider three major groupings of RNA,
310 namely the RNA associated with ribosomes (rRNA), the RNA encoding the amino-acid sequence of proteins (mRNA),
311 and the RNA which links codon sequence to amino-acid identity during translation (tRNA).

312 rRNA serves as the catalytic and structural framework along with myriad ribosomal proteins as part of a
313 complete ribosomal complex. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904
314 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides overall. As the *E. coli* RNA
315 polymerase transcribes DNA to RNA at a rate of \approx 40 nucleotides per second (BNID: 101904), it takes a single RNA
316 polymerase \approx 100 s to synthesize the RNA needed to form a single functional ribosome. To elucidate the *maximum*
317 number of rRNA units that can be synthesized given a single copy of each rRNA gene, we will consider a hypothesis
318 in which the rRNA operon is completely tiled with RNA polymerase. *In vivo* measurements of the kinetics of rRNA
319 transcription have revealed that RNA polymerases are loaded onto the promoter of an rRNA gene at a rate of \approx 1
320 per second (BNID: 111997, 102362). If RNA polymerases are constantly loaded at this rate, then we can assume
321 that \approx 1 functional rRNA unit is synthesized per second. With a 5000 second division time, this hypothesis leads to
322 a maximal value of 5000 functional rRNA units, actually undershooting the observed number of 10^4 ribosomes per
323 cell.

324 *E. coli*, like many other bacterium, have evolved a clever mechanism to surpass this kinetic limit for the rate
325 of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli* has seven copies of the operon
326 (BIND: 100352) four of which are localized directly adjacent to the origin of replication (*Birnbaum and Kaplan,*
327 **1971**). As fast growth, with parallelized chromosomal replication, the total number of rRNA genes can be on the
328 order of \approx 10 – 70 copies (*Stevenson and Schmidt, 2004*). Given a 5000 second division time, we can make the
329 lower-bound estimate that the typical cell will have \approx 7 copies of the rRNA operon. Synthesizing one functional
330 rRNA unit per second per rRNA operon, a total of 5×10^4 rRNA units can be synthesized, comfortably above the
331 observed number of ribosomes per cell.

332 How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA genes? We approach
333 this estimate by considering the maximum number of RNA polymerases tiled along the rRNA genes with a loading
334 rate of 1 per second and a transcription rate of 40 nucleotides per second. Considering that a RNA polymerase has
335 a physical footprint of approximately 40 nucleotides (BNID: 107873), we can expect \approx 1 RNA polymerase per 80
336 nucleotides. With a total length of \approx 4500 nucleotides per operon and 7 operons per cell, the maximum number of
337 RNA polymerases that can be transcribing rRNA at any given time is \approx 500.

338 The synthesis of rRNA demands the lions share of the required RNA polymerase. As outlined in *Figure 6* and in
339 discusses further the Appendix ??, synthesis of mRNA and tRNA, together require on the order of \approx 400 RNAP. Thus,
340 on the order of \approx 1000 RNAP is needed to satisfy the cells' transcriptional demand. As is revealed in *Figure 6(B)*,
341 this estimate is about an order of magnitude below the observed number of RNA polymerase complexes per cell
342 (\approx 5000 - 7000). The difference between the estimated number of RNA polymerase needed for transcription and
343 and these observations are consistent with recent literature revealing that \approx 80 % of RNA polymerases in *E. coli* are
344 not transcriptionally active (*Patrick et al., 2015*). Our estimate ignores the possibility that some fraction is only
345 nonspecifically bound to DNA, as well as the obstacles that RNA polymerase and DNA polymerase present for each
346 other as they move along the DNA (*Finkelstein and Greene, 2013*).

347 It is also vital to consider the role of σ -factors which help RNA polymerase identify and bind to transcriptional
348 start sites (*Browning and Busby, 2016*). Here we consider σ^{70} (RpoD) which is the dominant "general-purpose"

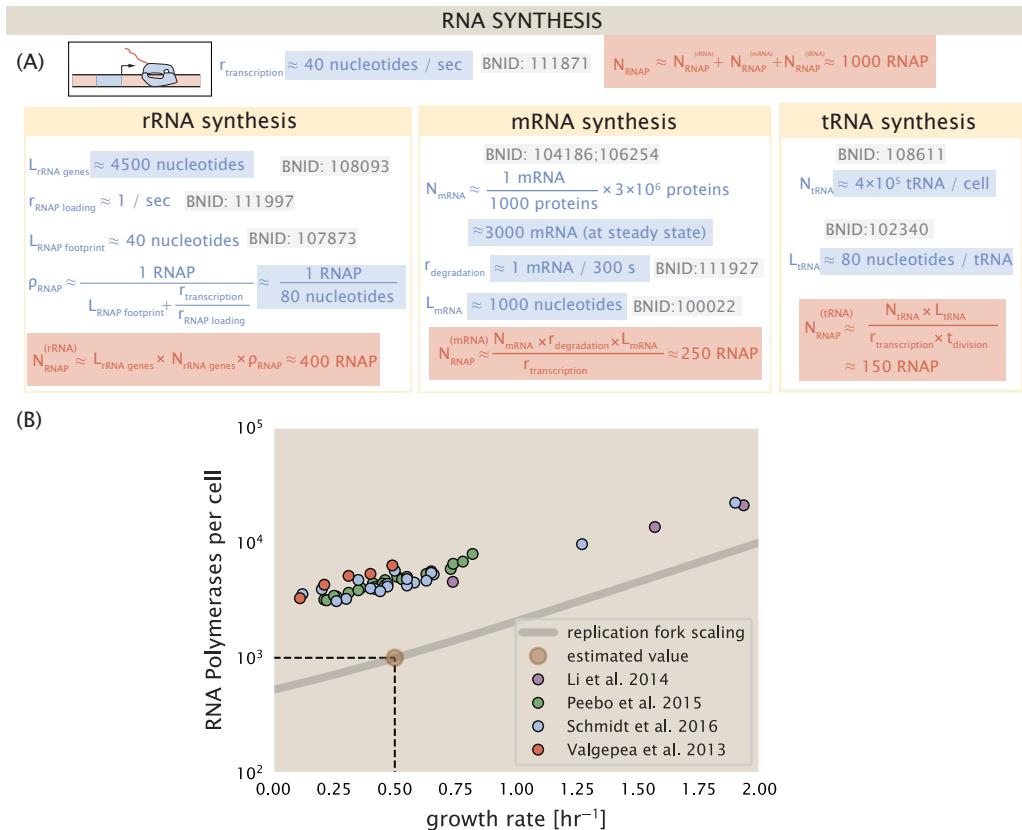


Figure 6. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of 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 6–Figure supplement 1. Abundance and growth rate dependence of σ -70.

349 σ -factor in *E. coli*. While initially thought of as being solely involved in transcriptional initiation, the past two
350 decades of single-molecule work has revealed a more multipurpose role for σ^{70} including facilitating transcriptional
351 elongation (*Kapanidis et al., 2005; Goldman et al., 2015; Perdue and Roberts, 2011; Mooney and Landick, 2003;*
352 *Mooney et al., 2005*). **Figure 6–Figure Supplement 1** is suggestive of such a role as the number of σ^{70} proteins per
353 cell is in close agreement with our estimate of the number of transcriptional complexes needed.

354 These estimates provide insight as to the observed magnitude of both RNA polymerase and the σ -70 factor.
355 While there remains some disagreement in the magnitude of copy numbers, protein abundances are generally in
356 excess of what appears to be needed for growth, suggesting that the abundance of RNA polymerase itself is not
357 particularly limiting.

358 Protein Synthesis

359 Lastly, we consider the translation and estimate the number of ribosomes needed to replicate the proteome.
360 While the rate at which ribosomes translates is well known to have a growth rate dependence *Dai et al. (2018)*
361 and we consider this more carefully in the sections that follow, here we make the approximation that translation
362 occurs at a rate of ≈ 15 amino acids per second per ribosome (BNID: 100233). Under this approximation and
363 assuming a division time of 5000 s, we can arrive at an estimate of $\approx 10^4$ ribosomes are needed to replicate the
364 cellular proteome, shown in **Figure 7**. This point estimate proves to be notably accurate when compared to the
365 experimental observations (**Figure 7(B)**). In the Appendix and in **Figure 7–Figure Supplement 1**, we consider the
366 process of ligating tRNAs to their corresponding amino acid. While this is a critical step in protein synthesis whose
367 efficiency reflects the nutritional richness of the growth medium, the ability to parallelize this process by expressing
368 more tRNA ligases makes it unlikely to be a bottleneck for cell division.

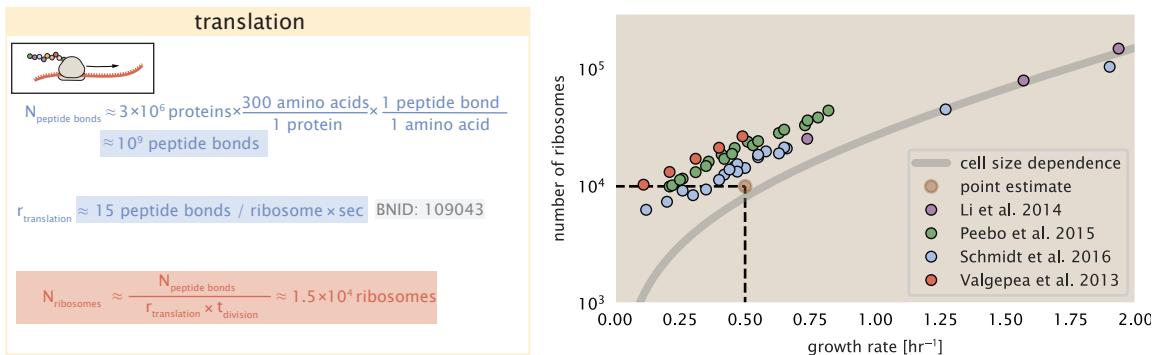


Figure 7. Estimation of the required number of ribosomes. Estimation of the number of ribosomes required to synthesize 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr^{-1} . Grey lines correspond to the estimated complex abundance calculated at different growth rates.

Figure 7–Figure supplement 1. Estimate and observed abundance and growth rate dependence of tRNA ligases.

369 Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step

370 Thus far, the general back-of-the-envelope estimates have been reasonably successful in predicting the scale of
 371 absolute protein copy number as well as their observed dependence on the cellular growth rate. A recurring
 372 theme across these varied biological processes is the ability of cells to parallelize tasks through the expression of
 373 additional proteins. Even when that is not possible, like in chromosomal replication which requires a minimum
 374 of ≈ 40 minutes, *E. coli* and many other bacteria surpass this limit by initiating additional rounds of replication
 375 per doubling. However, the synthesis of ribosomal proteins presents a special case where parallelization is not
 376 possible and must be doubled in quantity on average with every cell division (*Figure 8(A)*).

377 To gain some intuition into how translation and ribosomal synthesis may limit bacterial growth, we again
 378 consider the total number of peptide bonds that must be synthesized, which we denote as N_{pep} . With cells growing
 379 exponentially in time (*Godin et al., 2010*), the rate of cellular growth will be related to the rate of protein synthesis
 380 by

$$381 N_{\text{pep}} \lambda = r_t R f_a, \quad (1)$$

382 where λ is the cell growth rate in s^{-1} , r_t is the maximum elongation rate in $\text{AA} \cdot \text{s}^{-1}$, and R is the average ribosome
 383 copy number per cell. The addition factor f_a refers to the fraction of actively translating ribosomes, and allows us to
 384 account for the possibility of nonfunctional, immature ribosomes or active sequestration of ribosomes, mediated
 385 by the secondary-messenger molecule alarmones, guanosine pentaphosphate [(p)pGpp] at slow growth (*Dennis et al., 2004; Dai et al., 2016*). Knowing the number of peptide bonds formed per cell permits us to compute the
 386 translation-limited growth rate as

$$387 \lambda_{\text{translation-limited}} = \frac{r_t R f_a}{N_{\text{pep}}}. \quad (2)$$

388 Alternatively, since N_{pep} is related to the total protein mass through the molecular weight of each protein, we
 389 can also consider the growth rate in terms of the fraction of the total proteome mass dedicated to ribosomal
 390 proteins. By making the approximation that an average amino acid has a molecular weight of 110 Da (BNID:
 391 104877), the total protein mass m_{protein} is related to N_A by $(m_{\text{protein}}/110 \text{ Da}) \times N_A$, where N_A is Avogadro's number.
 392 Similarly, R is related to the ribosomal protein mass by $R \approx (m_R/800 \text{ Da}) \times N_A$, where 800 Da reflects the summed
 393 molecular weight of all ribosomal subunits. This allows us to approximate $R/N_{\text{pep}} \approx \Phi_R/L_R$, where Φ_R is the
 394 ribosomal mass fraction m_{protein}/m_R , and L_R the ratio of 800 kDa / 110 Da per amino acid or, alternatively, the total
 395 length in amino acids that make up a ribosome. The translation-limited growth rate can then be written in the
 form

$$396 \lambda_{\text{translation-limited}} \approx \frac{r_t}{L_R} \Phi_R f_a. \quad (3)$$

397 This is plotted as a function of ribosomal fraction Φ_R in *Figure 8(B)*, where we take $L_R = 7459 \text{ AA}$, corresponding
 398 to the length in amino acids for all ribosomal subunits of the 50S and 30S complex (BNID: 101175), and $f_a =$

398 1. In **Figure 8(C)** we use the recent measurements of f_a from **Dai et al. (2016)** to estimate the active fraction of
399 ribosomal protein across the proteomic data sets and number of other recent measurements. We see that cells
400 are consistently skirting the limit in growth rate set by **Equation 3** as nutrient conditions vary.

401 The growth rate defined by **Equation 3** reflects mass-balance under steady-state growth and has long provided
402 a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate (**Maaløe,**
403 **1979; Scott et al., 2010**). The maximum rate, when $\Phi_R = 1$, could only be achieved if a cell contained only
404 ribosomes. This corresponds to the synthesis time of all ribosomal subunits, $L_R/r_i \approx 7$ minutes (**Dill et al., 2011**)
405 and interestingly, is independent of the absolute number of ribosomes. This is because, in order to double the
406 cell's ribosomal mass, each ribosome must produce a second ribosome; a process which cannot be parallelized.
407 Unless elongation rate increased, or cells could trim their total ribosomal protein mass, this dependency limits both
408 the maximum growth rate (when $\Phi_R = 1$), and also the achievable growth rate under more realistic values of Φ_R .

409 This strain of *E. coli* rarely exhibits growth rates above 2 hr^{-1} (**Bremer and Dennis, 2008; Roller et al., 2016**)
410 and in **Figure 8(C)** we consider ribosomal generation from the perspective of rRNA synthesis. Here we use our
411 rule-of-thumb of 1 functional rRNA unit per second per operon and estimate the maximum number of ribosomes
412 that could be made as a function of growth rate (blue curve). Although we expect this estimate to drastically
413 overestimate rRNA abundance at slower growth rates ($\lambda < 0.5 \text{ hr}^{-1}$), it provides a useful reference alongside the
414 proteomic measurements. For growth rates above about 1 hr^{-1} , we find that cells will need to transcribe rRNA
415 near their maximal rate. As a counter example, if *E. coli* did not initiate multiple rounds of replication, they would
416 be unable to make enough rRNA for the observed number of ribosomes (dashed blue curve in **Figure 8(C)**). The
417 convergence between the maximum rRNA production and measured ribosome copy number suggests rRNA
418 synthesis may begin to present a bottleneck at the fastest growth rates due to the limited copies of rRNA genes.

419 **Relationship Between Cell Size and Growth Rate**

420 The relationship between cell size and growth rate has long been of interest in the study of bacterial physiology,
421 particularly following the now six decade-old observation that cell volume appears to increase exponentially with
422 growth rate; known as Schaechter's growth law (**Schaechter et al., 1958; Taheri-Araghi et al., 2015**). However, the
423 mechanism that governs this relationship, and even the question of whether the change in average cell size is truly
424 exponential, has remained under debate (**Harris and Theriot, 2018**). Here we examine the influence of ribosomal
425 content and total protein abundance on cell size.

426 Cells grow at a near-maximal rate dictated by their total ribosomal mass fraction Φ_R , at least at moderate growth
427 rates above 0.5 hr^{-1} (where f_a is close to 1, and r_i is near its maximal rate). Here, growth rate can be increased
428 only by increasing Φ_R , though the simple addition of more ribosomes is likely constrained by aspects physical
429 constrains like macromolecular crowding (**Delarue et al., 2018; Soler-Bistué et al., 2020**). As *E. coli* grows faster,
430 large swaths of its proteome increase in absolute abundance. It is now well-documented that *E. coli* cells add a
431 constant volume per origin of replication (termed a "unit cell" or "initiation mass"), which is robust to a remarkable
432 array of cellular perturbations (**Si et al., 2017**). To consider this dependency in the context of the proteomic data,
433 we used measurements from **Si et al. (2017)** (**Figure 9(A)**) to estimate the average number of origins per cell $\langle \#$
434 ori \rangle at different growth rates. $\langle \# \text{ ori} \rangle$ is set by how often replication must be initiated per cell doubling under
435 steady-state growth. This can be quantified as

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

436 where τ_{cyc} is the cell cycle time (referring to the time from replication initiation to cell division), and τ is the cell
437 doubling time. For ribosomal synthesis, we find an approximately linear correlation between ribosome copy
438 number and $\langle \# \text{ ori} \rangle$ (**Figure 9(B)**).

439 For a constant cell cycle time, observed at growth rates above about 0.5 hr^{-1} (**Helmstetter and Cooper, 1968**),
440 **Equation 4** states that $\langle \# \text{ ori} \rangle$ will need to increase exponentially with the growth rate. While this says nothing
441 of the observed scaling with cell size, the additional dependency on ribosomal content, which increases with $\langle \#$
442 ori \rangle , provides an additional link. In **Figure 9(D)**, we consider the position-dependent protein expression across
443 the chromosome by calculating a running Gaussian average of protein copy number (20 kbp st. dev. averaging
444 window) based on each gene's transcriptional start site, which were then median-subtracted to account for the
445 differences in total protein abundance. Importantly, major deviations in protein copy number are largely restricted

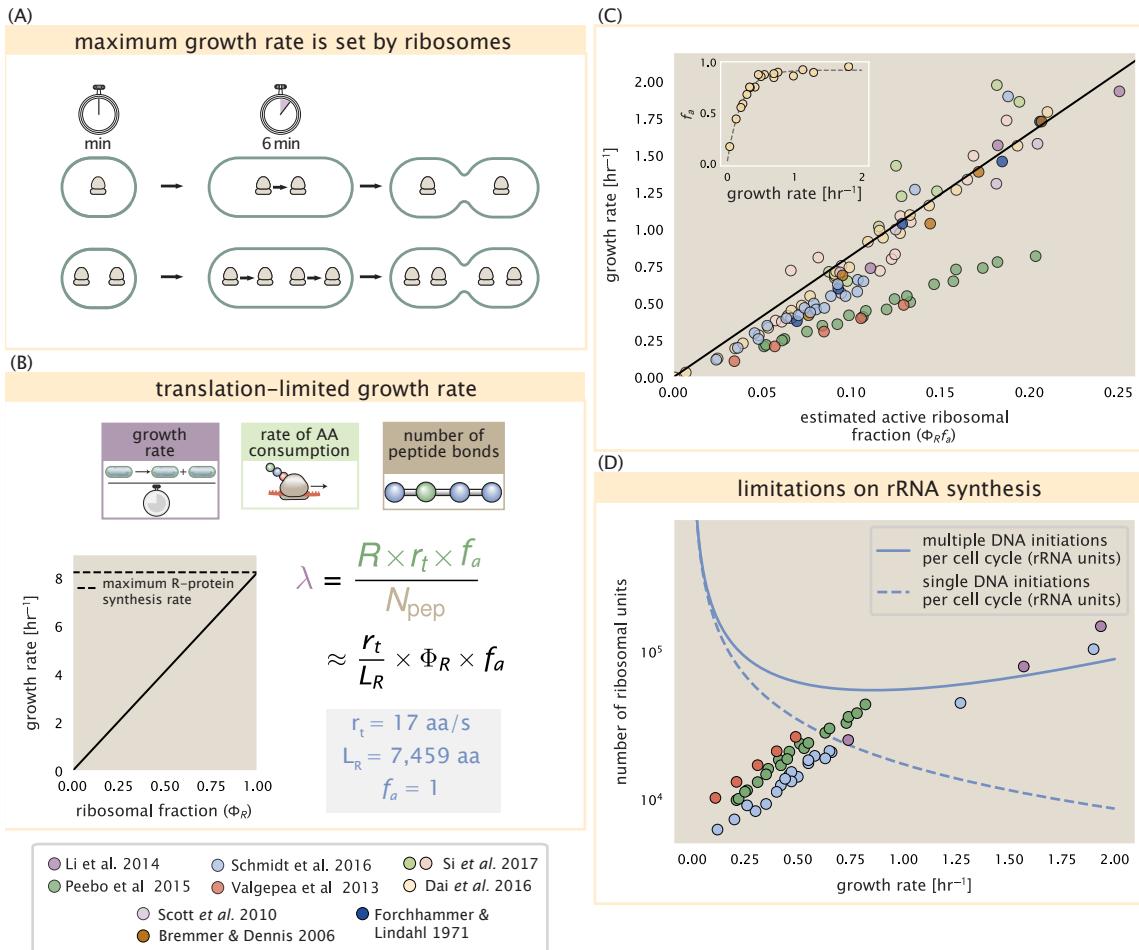


Figure 8. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate is independent of the number of ribosomes; instead limited serial requirement for each ribosome to double itself. (B) 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 (R-protein) synthesis considered in part (A). (C) 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). (D) Maximum number of rRNA units that can be synthesized as a function of growth rate. Solid curve corresponds 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)* that are plotted in *Figure 9(A)*. The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

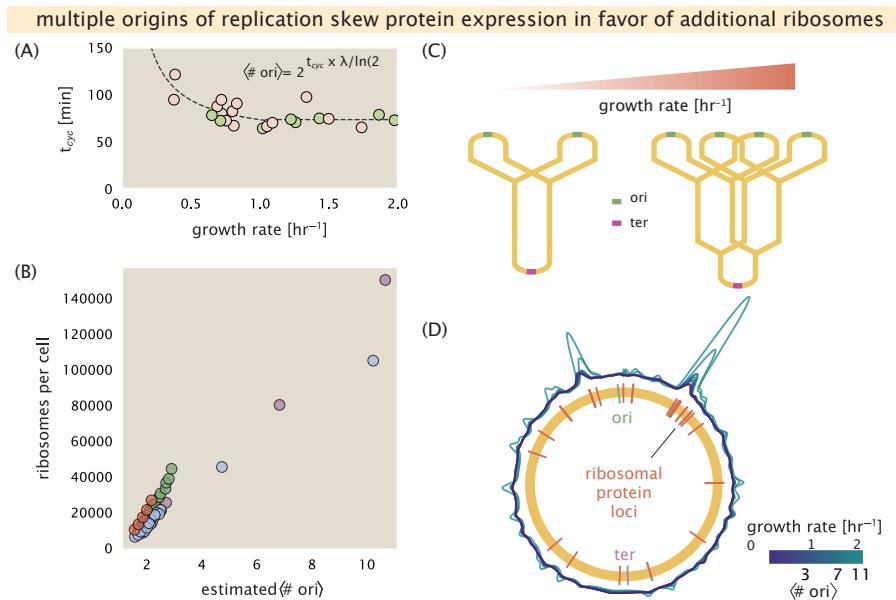


Figure 9. Cells increase absolute ribosome abundance with (# ori). (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times greater than 40 minutes, 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). Red data points correspond to measurements in strain MG1655, while light green points are for strain NCM3722. (B) Plot of the ribosome copy number estimated from the proteomic data against the estimated $\langle \# \text{ori} \rangle$. (C) Schematic shows the expected increase in replication forks (or number of ori regions) as *E. coli* cells grow faster. (D) A running Gaussian average (20 kbp st. dev.) of protein copy number is calculated for each growth condition considered by (Schmidt *et al.*, 2016). Since total protein abundance increases with growth rate, protein copy numbers are median-subtracted to allow comparison between growth conditions. $\langle \# \text{ori} \rangle$ are estimated using the data in (A) and Equation 4.

to regions of ribosomal protein genes. This suggests that the relative ribosomal abundance Φ_R is also being tuned in proportion to $\langle \# \text{ori} \rangle$, with the exponential relationship between cell size and growth rate following from how *E. coli* varies its number of ribosomes per cell.

449 Nutrient-Mediated Model of Proteomic Composition and Growth Rate Control

450 As we have seen, cell size, total proteomic content, and the number of ribosomes are all interconnected and
 451 influence the achievable growth rate. The drastic change in these parameters across different growth conditions
 452 suggests a hypothesis that each is being tuned to better match the cell's biosynthetic capacity given the specific
 453 environment. In this final, we explore the interconnection between cell size, ribosome content, and growth rate by
 454 formulating a minimal model of growth rate control. We use it to quantitatively show how tuning these parameters
 455 help cells maximize their growth rate for a particular environment.

456 To react to changes in nutrient conditions, bacteria rely on the synthesis or degradation of secondary-messenger
 457 molecules like (p)ppGpp, which cause global changes in transcriptional and translational activity. In *E. coli*, amino
 458 acid starvation causes the accumulation of de-acylated tRNAs at the ribosome's A-site and leads to a strong increase
 459 in (p)ppGpp synthesis activity by the enzyme RelA (Hauryliuk *et al.*, 2015). Cells also accumulate (p)ppGpp during
 460 steady-state growth in poorer growth conditions, which leads to a decrease in the fraction of actively translating
 461 ribosomes, f_a (with $f_a \approx 0.5$ at a growth rate of $\approx 0.3 \text{ hr}^{-1}$; Figure 8(C) - inset).

462 Furthermore, (p)ppGpp can inhibit the initiation of DNA replication by mediating a change in transcriptional
 463 activity and the supercoiling state of the origin of replication (Kraemer *et al.*, 2019). These observations all raise
 464 the possibility that it is through (p)ppGpp that cells mediate the growth-rate dependent changes in $\langle \# \text{ori} \rangle$, cell size,
 465 and ribosomal abundance and activity (Zhu and Dai, 2019; Büke *et al.*, 2020). Indeed, recent work in a (p)ppGpp
 466 deficient strain of *E. coli* found that cells exhibited a high ratio of $\langle \# \text{ori} \rangle$ to $\langle \# \text{ter} \rangle$, and cell sizes that were more
 467 consistent with a fast growth state where (p)ppGpp levels are normally low (Fernández-Coll *et al.*, 2020).

468 **Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity**

469 Here we consider a mode of regulation in which the rate of peptide elongation r_t depends only on the availability
 470 of amino acids (and, therefore, also amino-acyl tRNAs). It is through the elongation rate r_t that we assume cells
 471 adjust their ribosomal content (R , Φ_R) according to nutrient availability and for simplicity, do not explicitly model
 472 changes in (# ori) or regulation by (p)ppGpp.

473 The rate of elongation r_t will depend on how quickly the ribosomes can match codons with their correct amino-
 474 acyl tRNA, along with the subsequent steps of peptide bond formation and translocation. We therefore coarse-grain
 475 the steps of elongation to two time-scales, 1) the time required to find and bind each correct amino-acyl tRNA,
 476 and 2) the remaining steps in peptide elongation that will not depend on the amino acid availability. Under this
 477 model, other molecular players required for translation like elongation factors and GTP are considered in sufficient
 478 abundance, which appear to be valid assumptions given our analysis of the proteomic data and energy production
 479 thus far. The time to translate each codon is given by the inverse of the elongation rate r_t , which can be written as,

$$\frac{1}{r_t} = \frac{1}{k_{on}\alpha[AA]_{eff}} + \frac{1}{r_t^{\max}}. \quad (5)$$

480 where we have assumed that the rate of binding by amino-acyl tRNA k_{on} is proportional to $[AA]_{eff}$ by a constant α .
 481 The second term on the right-hand side reflects our assumption that other steps in peptide elongation are not
 482 rate-limiting, with a maximum elongation rate r_t^{\max} of about 17 amino acids per second **Dai et al. (2016)**. As the
 483 rate of amino acid supply, denote by r_{AA} , varies with changing nutrient conditions, the cell can maximize the rate of
 484 protein synthesis by tuning the rate of amino acid consumption (mathematized as $r_t \times R \times f_a$), shown schematically
 485 in **Figure 10(A)**. This can be stated more succinctly in terms of an effective dissociation constant, $K_D = r_t^{\max}/\alpha k_{on}$,
 486 where the elongation rate r_t is now given by

$$r_t = \frac{r_t^{\max}}{1 + K_D/[AA]_{eff}}. \quad (6)$$

487 Under steady-state growth, the amino acid concentration is constant ($\frac{d[AA]_{eff}}{dt} = 0$) and will relate to the rate
 488 of amino acid synthesis (or import, for rich media) and/or tRNA charging, as r_{AA} , and the rate of consumption,
 489 $r_t \times R \times f_a$. We calculate $[AA]_{eff}$ by,

$$[AA]_{eff} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V}, \quad (7)$$

490 which allows us to then solve for r_t explicitly (further described in Appendix ??). Here r_{AA} is in units of AA per unit
 491 time, and V reflects the volume of the cell over a time period t .

492 In **Figure 10(B)**, we illustrate how the elongation rate depends on the ribosomal copy number. Here, we have
 493 considered a unit volume $V = 1\mu\text{m}^3$, a unit time $t = 1\text{ s}$, a $K_D = 5\text{ mM}$ (inferred from **Bennett et al. (2009)**),
 494 $f_a = 1$, and an arbitrarily chosen $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$. At low ribosome copy numbers, the observed
 495 elongation rate is dependent primarily on the ratio of K_D/Vr_{AA} [as $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in **Figure 10(B)**].
 496 As the ribosome copy number is increased such that the amino acid supply rate and consumption rate are nearly
 497 equal [point (2) in **Figure 10(B)**], the observed elongation rate begins to decrease sharply. When the ribosome
 498 copy number is increased even further, consumption at the maximum elongation rate exceeds the supply rate,
 499 yielding a significantly reduced elongation rate [point (3) in **Figure 10B**]. While the elongation rate will always be
 500 dominated by the amino acid supply rate at sufficiently low ribosome copy numbers, the elongation rate at larger
 501 ribosome abundances can be increased by tuning f_a such that not all ribosomes are elongating, reducing the total
 502 consumption rate.

503 **Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and
 504 Metabolic Capacity.**

505 To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic
 506 measurements; namely, we restrict the values of R , N_{pep} , and V to those associated with the amalgamated
 507 proteomic data (described in Appendix Estimation of Total Protein Content per Cell). We then consider how
 508 changes in the nutrient conditions, through the parameter r_{AA} , influence the maximum growth rate as determined
 509 by **Equation 2**. **Figure 10(C)** shows how the observed growth rate depends on the rate of amino acid supply r_{AA} as

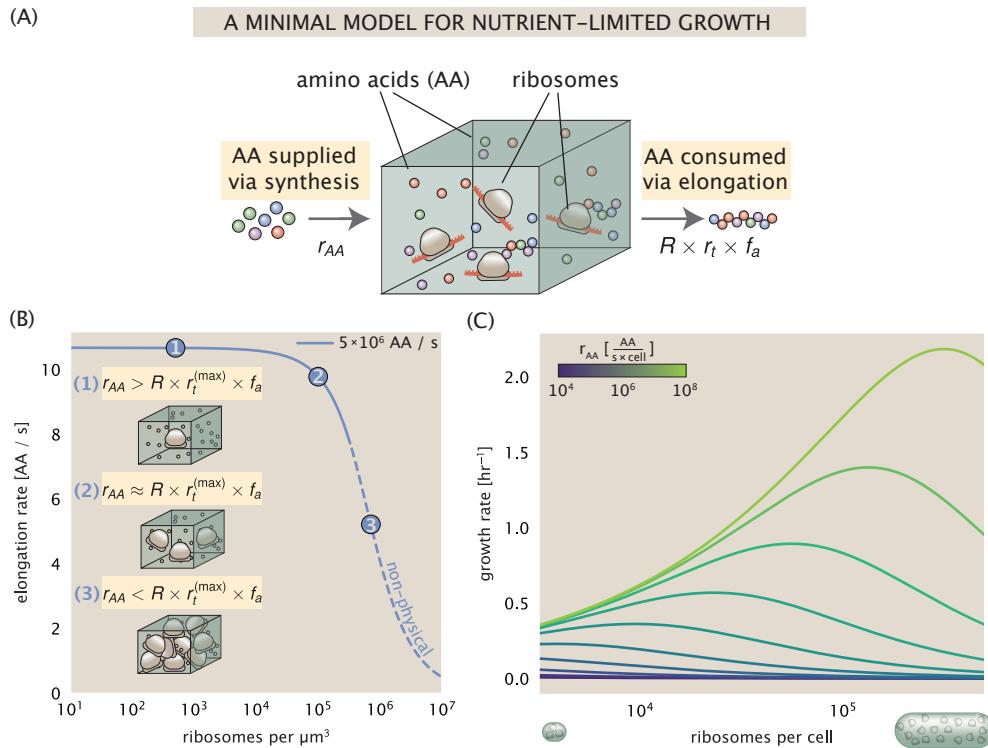


Figure 10. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, 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 in a unit volume μm^3 . 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 the parameters $r_{AA} = 5 \times 10^6 \text{ AA} / \text{s}$, $K_D = 5 \text{ mM}$, and $r_t^{(\text{max})} = 17.1 \text{ AA} / \text{s}$. (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 volume and total protein abundance. We direct the reader to the Supplemental Information for discussion on the inference of the realtionship between cell volume, number of peptide bonds, and ribosome copy number.

510 a function of the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal
511 growth rate whose dependence on R (and consequently, the cell size) increases with increasing r_{AA} . Importantly,
512 however, there is an optimum set of R , N_{pep} , and V that are strictly dependent on the value of r_{AA} . Increasing the
513 ribosomal concentration beyond the cell's metabolic capacity has the adverse consequence of depleting the supply
514 of amino acids and a concomitant decrease in the elongation rate r , [Figure 10(B)].

515 Also of note is the growth rate profiles shown for low amino acid supply rates [purple and blue lines in
516 Figure 10(C)], representing growth in nutrient-poor media. In these conditions, there no longer exists a peak in
517 growth, at least in the range of physiologically-relevant ribosome copy numbers. Instead, cells limit their pool of
518 actively translating ribosomes by decreasing f_a (Dai et al., 2016), which would help maintain the pool of available
519 amino acids [$AA]_{eff}$ and increase the achievable elongation rate. This observation is in agreement with the central
520 premise of the cellular resource allocation principle proposed by Scott et al. (2010); Klumpp et al. (2009); Klumpp
521 and Hwa (2014) and Hui et al. (2015).

522 Discussion

523 Continued experimental and technological improvements have led to a treasure trove of quantitative biological
524 data (Hui et al., 2015; Schmidt et al., 2016; Si et al., 2017; Gallagher et al., 2020; Peebo et al., 2015; Valgepea et al.,
525 2013), and an ever advancing molecular view and mechanistic understanding of the constituents that support
526 bacterial growth (Taheri-Araghi et al., 2015; Morgenstein et al., 2015; Si et al., 2019; Karr et al., 2012; Kostinski
527 and Reuveni, 2020). In this work we have compiled what we believe to be the state-of-the-art knowledge on
528 proteomic copy number across a broad range of growth conditions in *E. coli*. We have made this data accessible
529 through a [GitHub repository](#), and an interactive figure that allows exploration of specific protein and protein
530 complex copy numbers. Through a series of order-of-magnitude estimates that traverse key steps in the bacterial
531 cell cycle, this proteomic data has been a resource to guide our understanding of two key questions: what biological
532 processes limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents
533 as a function of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates
534 provide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the
535 transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous
536 steps of the central dogma.

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

549 Of the processes illustrated in Figure 1, we arrive at a ribosome-centric view of cellular growth rate control.
550 This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary
551 their ribosomal abundance as a function of growth conditions and growth rate Scott et al. (2010); Metzl-Raz et al.
552 (2017). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates
553 to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize
554 many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA
555 replication, this is not so for synthesis of ribosomes (Figure 8(A)). The translation time for each ribosome [≈ 6 min,
556 Dill et al. (2011)] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase
557 their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The
558 second point relates to the long-observed correlations between growth rate and cell size (Schaechter et al., 1958; Si
559 et al., 2017), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous

560 curiosity and driven substantial amounts of research in their own regards, these relationships are themselves
561 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions
562 of rapid growth that require cells to also grow in size. Further experiments are needed to test the validity of this
563 hypothesis. In particular, we believe that the change in growth rate in response to translation-inhibitory drugs
564 (such as chloramphenicol) could be quantitatively predicted, given one had precision measurement of the relevant
565 parameters, including the fraction of actively translating ribosomes f_a and changes in the metabolic capacity of the
566 cell (i.e. the parameter r_{AA} in our minimal model) for a particular growth condition.

567 While the generation of new ribosomes plays a dominant role in growth rate control, there exist other physical
568 limits to the function of cellular processes. One of the key motivations for considering energy production was
569 the physical constraints on total volume and surface area as cells vary their size (*Harris and Theriot, 2018; Ojicic*
570 *et al., 2019*). While *E. coli* get larger as it expresses more ribosomes, an additional constraint begins to arise in
571 energy production due to a relative decrease in total surface area where ATP is predominantly produced (*Szenk*
572 *et al., 2017*). Specifically, the cell interior requires an amount of energy that scales cubically with cell size, but the
573 available surface area only grows quadratically (*Figure 4(A)*). While this threshold does not appear to be met for
574 *E. coli* cells growing at 2 hr⁻¹ or less, it highlights an additional constraint on growth given the apparent need to
575 increase in cell size to grow faster. This limit is relevant even to eukaryotic organisms, whose mitochondria exhibit
576 convoluted membrane structures that nevertheless remain bacteria-sized organelles (*Guo et al., 2018*). In the
577 context of bacteria growth and energy production more generally, we have limited our analysis to the aerobic
578 growth conditions associated with the proteomic data and further consideration will be needed for anaerobic
579 growth.

580 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control, and there are
581 many aspects of the bacterial proteome and growth that we neglected to consider. For example, other recent work
582 (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*) has explored how the proteome is structured
583 and how that structure depends on growth rate. In the work of *Hui et al. (2015)*, the authors coarse-grained the
584 proteome into six discrete categories being related to either translation, catabolism, anabolism, and others related
585 to signaling and core metabolism. The relative mass fraction of the proteome occupied by each sector could be
586 modulated by external application of drugs or simply by changing the nutritional content of the medium. While
587 we have explored how the quantities of individual complexes are related to cell growth, we acknowledge that
588 higher-order interactions between groups of complexes or metabolic networks at a systems-level may reveal
589 additional insights into how these growth-rate dependences are mechanistically achieved. Furthermore, while we
590 anticipate the conclusions summarized here are applicable to a wide collection of bacteria with similar lifestyles as *E.*
591 *coli*, other bacteria and archaea may have evolved other strategies that were not considered. Further experiments
592 with the level of rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to learn
593 more about how regulation of proteomic composition and growth rate control has evolved over the past 3.5 billion
594 years.

595 Methods

596 Data Analysis and Availability

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

602 Acknowledgements

603 We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding their data. We
604 also thank members of the Phillips, Theriot, Kondev, and Garcia labs for useful discussions. R.P. is supported by
605 La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and the NIH 1R35 GM118043 (MIRA). J.A.T. is
606 supported by the Howard Hughes Medical Institute, and NIH Grant R37-AI036929. N.M.B is a HHMI Fellow of The
607 Jane Coffin Childs Memorial Fund.

608 **Competing Interests**

609 The authors declare no competing interests.

610 Appendix for: Fundamental limits on the 611 rate of bacterial cell division

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

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

623	Contents	
624	Introduction	1
625	Nutrient Transport	4
626	Energy Production	6
627	ATP Synthesis	6
628	Generating the Proton Electrochemical Gradient	7
629	Biosynthesis in a Crowded Membrane	7
630	Function of the Central Dogma	9
631	DNA Replication	10
632	RNA Synthesis	11
633	Protein Synthesis	12
634	Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step	13
635	Relationship Between Cell Size and Growth Rate	14
636	Nutrient-Mediated Model of Proteomic Composition and Growth Rate Control	16
637	Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity	17
638	Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic Capacity.	17
640	Discussion	19
641	Methods	20
642	Data Analysis and Availability	20
643	Acknowledgements	20
644	Competing Interests	21
645	Experimental Details Behind Proteomic Data	25
646	Fluorescence based measurements	25
647	Ribosomal profiling measurements	25
648	Mass spectrometry measurements	26
649	Summary of Proteomic Data	26
650	Estimation of Cell Size and Surface Area	27
651	Estimation of Total Protein Content per Cell	29
652	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	29
653	Additional Considerations of Schmidt <i>et al.</i> Data Set	29
654	Effect of cell volume on reported absolute protein abundances	31
655	Relaxing assumption of constant protein concentration across growth conditions	33
656	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	33
657	Calculation of Complex Abundance	34

658	Extending Estimates to a Continuum of Growth Rates	36
659	Estimation of the total cell mass	36
660	Complex Abundance Scaling With Cell Volume	36
661	A Relation for Complex Abundance Scaling With Surface Area	37
662	Number of Lipids	37
663	Number of Murein Monomers	37
664	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	38
665	Calculation of active ribosomal fraction.	38
666	Estimation of $\langle \#ori \rangle$ / $\langle \#ter \rangle$ and $\langle \#ori \rangle$.	38

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 Additional Considerations of Schmidt *et al.* Data Set for more details on this).

667 Experimental Details Behind Proteomic Data

668 Here we provide a brief summary of the experiments behind each proteomic data set. The purpose of this section
669 is to identify how the authors arrived at absolute protein abundances. In the following section (Section Summary
670 of Proteomic Data) we will then provide a summary of the final protein abundance measurements that were used
671 throughout the main text. Table 1 provides an overview of the publications we considered. These are predominately
672 mass spectrometry-based, with the exception of the work from Li *et al.* (2014) which used ribosomal profiling, and
673 the fluorescence-based counting done in Taniguchi *et al.* (2010).

674 Fluorescence based measurements

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

682 Ribosomal profiling measurements

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

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

697 **Mass spectrometry measurements**

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

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

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

726 **Summary of Proteomic Data**

727 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)*;
728 *Schmidt et al. (2016)*. As shown in *Figure 11(A)*, the reported total protein abundances in the work of *Taniguchi*
729 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi et al.*
730 (*2010*) this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that most
731 proteins appear undercounted when compared to the other data.

732 *Figure 11(B)* summarizes the total protein mass for each data point in our final compiled data set. We note that
733 protein abundances were all scaled so they followed a common growth rate-dependent change in total protein
734 mass. While our inclination initially was to leave reported copy numbers untouched, a notable discrepancy in the
735 scaling total protein per cell between *Schmidt et al. (2016)* and the other data sets forced us to dig deeper into
736 those measurements (compare *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 11(A)*). The particular trend
737 in *Schmidt et al. (2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion
738 and analysis of that data set in section Additional Considerations of Schmidt et al. Data Set. As a compromise, and
739 in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of
740 total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total
741 protein content through expected dry mass. Total protein per cell was estimated using available data on total
742 DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the
743 cell. We consider these details in sections Estimation of Cell Size and Surface Area and Estimation of Total Protein
744 Content per Cell that follows.

745 Lastly, in *Figure 12* we show the total proteomic coverage and overlap of proteins quantified across each data

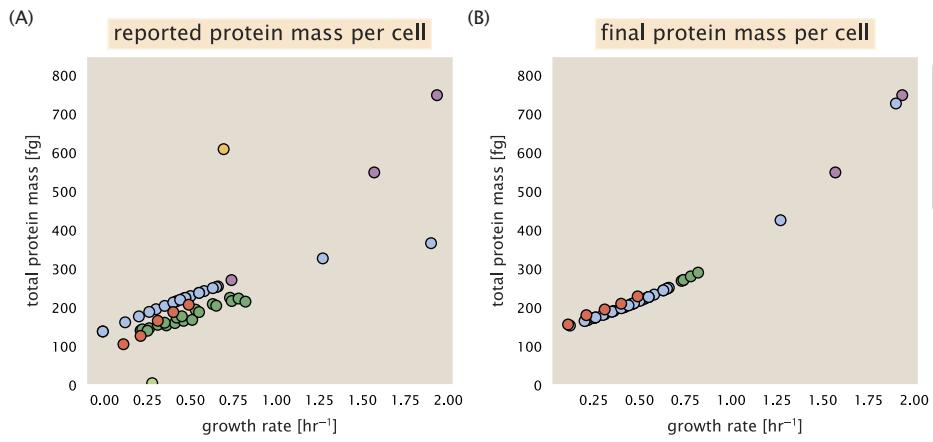


Figure 11. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (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).

746 set. Here we have used an UpSet diagram (*Lex et al., 2014*) to compare the data. Overall, the overlap in quantified
 747 proteins is quite high, with 1157 proteins quantified across all data sets. The sequencing based approach of *Li*
 748 *et al. (2014)* has substantially higher coverage compared to the mass spectrometry data sets (3394 genes versus
 749 the 2041 genes quantified in the work of *Schmidt et al. (2016)*). However, in terms of total protein mass, the data
 750 from *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (2015)* each quantify roughly equivalent total protein mass.
 751 An exception to this is in the data from *Valgepea et al. (2013)*, where we find that the total protein quantified in
 752 *Valgepea et al. (2013)* is 90-95 % of the total protein mass (when using the data from *Schmidt et al. (2016)* as a
 753 reference).

754 Estimation of Cell Size and Surface Area

755 Since most of the proteomic data sets lack cell size (i.e. volume) measurements, we chose instead to use a common
 756 estimate of size for any analysis requiring cell size or surface area. Since each of the data sets used either K-12
 757 MG1655 or its derivative, BW25113 (from the lab of Barry L. Wanner; the parent strain of the Keio collection
 758 (*Datsenko and Wanner, 2000*; *Baba et al., 2006*)), we fit the MG1655 cell size data from the supplemental material
 759 of *Si et al. (2017, 2019)* using the `optimize.curve_fit` function from the Scipy python package (*Virtanen et al., 2020*).

760 The average size measurements from each of their experiments are shown in Figure *Figure 13*, with cell length
 761 and width shown in (A) and (B), respectively. The length data was well described by the exponential function 0.5
 762 $e^{1.09 \cdot \lambda} + 1.76$ μm , while the width data was well described by $0.64 e^{0.24 \cdot \lambda}$ μm . In order to estimate cell size we take the
 763 cell as a cylinders with two hemispherical ends (*Si et al., 2017*; *Basan et al., 2015*). Specifically, cell size is estimated
 764 from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (8)$$

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

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (9)$$

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

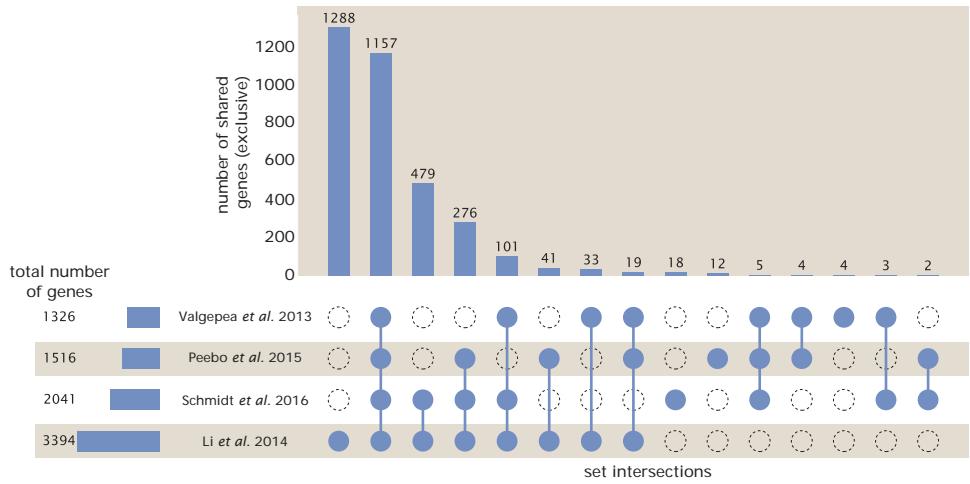


Figure 12. Comparison of proteomic coverage across different data sets. An UpSet diagram (*Lex et al., 2014*) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of *Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015)*. The total number of genes reported in each individual data set are noted on the bottom left, while their overlap is summarized in the bar chart. Each column here refers to the intersection of a set of data sets identified by blue circles.

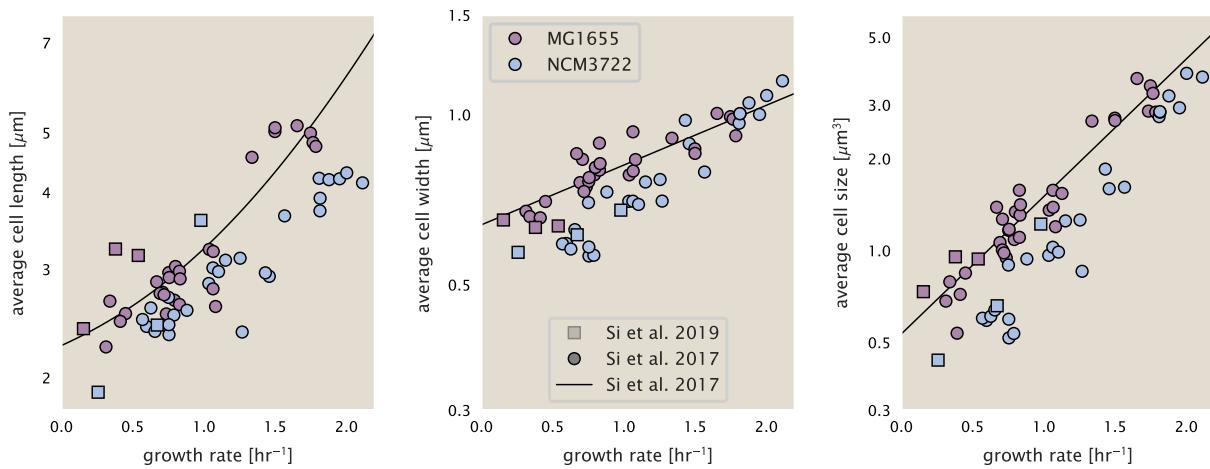


Figure 13. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 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: $0.5 e^{1.09 \cdot i} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot i} \mu\text{m}$. (C) Cell size, V , was calculated as cylinders with two hemispherical ends (Equation 8). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot i} \mu\text{m}^3$.

768 Estimation of Total Protein Content per Cell

769 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the
770 fit shown in Figure *Figure 13(C)* ($0.533 e^{1.037\lambda} \mu\text{m}^3$). We then estimate the total protein content from the total dry
771 mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here, protein,
772 DNA, and RNA were reported to account for at least 90 % of the dry mass (*Basan et al. (2015)*). The authors also
773 found that the total dry mass concentration was roughly constant across growth conditions. Under such a scenario,
774 we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30\% \times 90\%$
775 % or about $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

776 However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how
777 protein concentration might vary due to the substantial increase in rRNA at faster growth rates (*Dai et al. (2016)*).
778 This is a well-documented result that arises from an increase in ribosomal abundance at faster growth rates (*Scott*
779 *et al. (2010)*). To proceed therefore rely on experimental measurements of total DNA content per cell that also
780 come from Basan *et al.*, and RNA to protein ratios that were measured in Dai *et al.* (and cover the entire range of
781 growth conditions considered here). These are reproduced in Figure *Figure 14(A)* and (B), respectively.

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

$$785 M_p = \frac{[M_p + M_{RNA}]}{1 + (RP_{ratio})}. \quad (10)$$

786 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai *et al.*). In Figure *Figure 14(C)* we plot the estimated
787 cellular concentrations for protein, DNA, and RNA from these calculations, and in Figure *Figure 14(D)* we plot their
788 total expected mass per cell. This later quantity is the growth rate-dependent total protein mass that was used to
estimate total protein abundance across all data sets (and summarized in *Figure 11(B)*).

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

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

798 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log
799 transform of these parameters as a function of the log transform of the ribosome copy number. Using optimization
800 by minimization, we estimated the best-fit values of the intercept and slope for each regression. *Figure 15* shows
801 the result of each regression as a dashed line.

802 Additional Considerations of Schmidt *et al.* Data Set

803 While the data set from *Schmidt et al. (2016)* remains a heroic effort that our labs continue to return to as a resource,
804 there were steps taken in their calculation of protein copy number that we felt needed further consideration. In
805 particular, the authors made an assumption of constant cellular protein concentration across all growth conditions
806 and used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size
807 with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al.*
808 *(2017)*).

809 We begin by looking at their cell volume measurements, which are shown in blue in Figure *Figure 16*. As a
810 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*
811 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence
812 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and
813 then calculated cell size by treating the cell is a cylinder with two hemispherical ends, as we considered in the

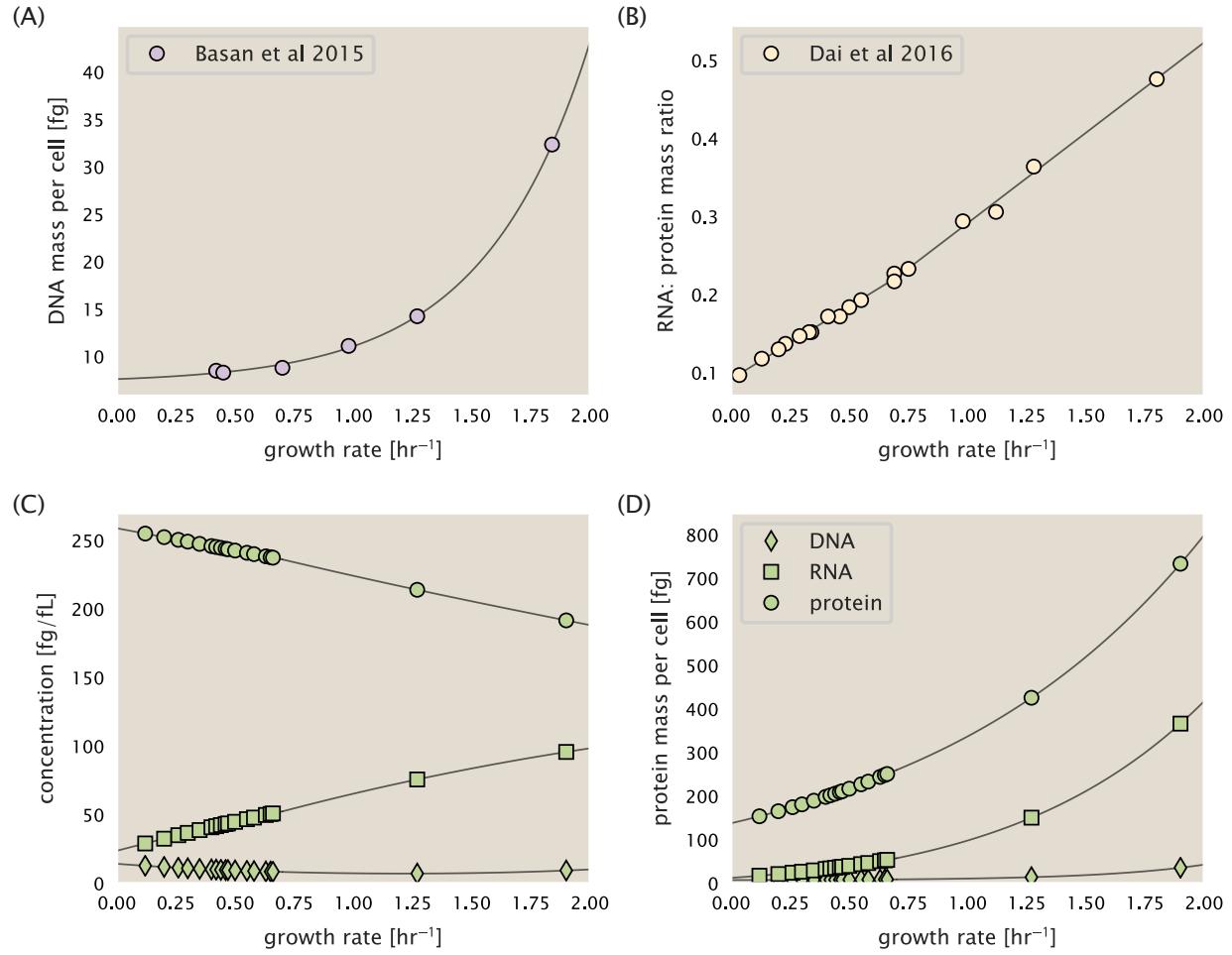


Figure 14. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ 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: 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) Predictions of cellular protein, DNA, and RNA concentration. (D) Total cellular mass predicted for protein, DNA, and RNA using the cell size predictions from Si *et al.*. 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).

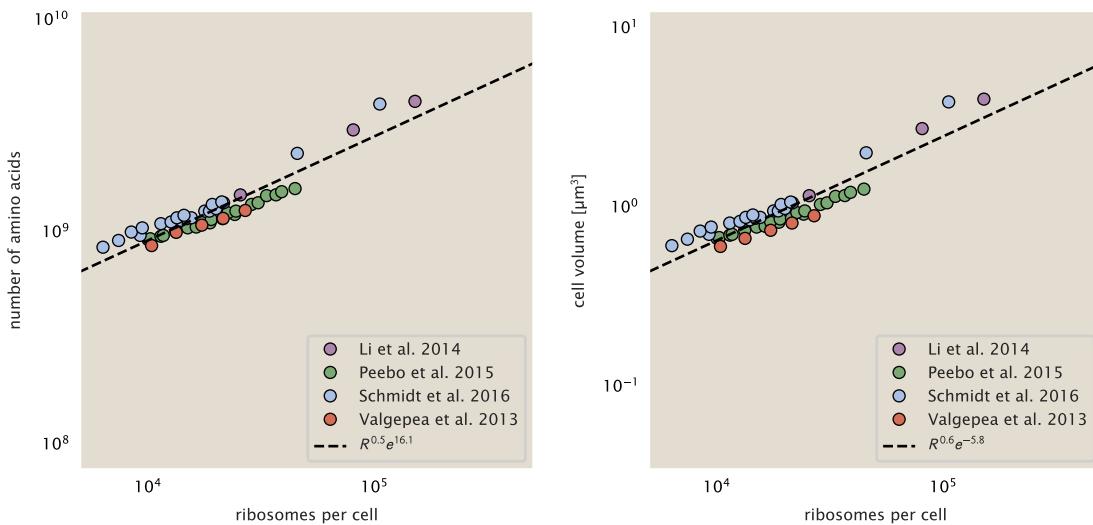


Figure 15. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows the result of the fit, with the functional form of the equation given in the legend with R representing the ribosome copy number.

814 previous section. While there is notable discrepancy between the two research groups, which are both using strain
 815 NCM3722, Basan *et al.* found that this came specifically from uncertainty in determining the cell width. This is prone
 816 to inaccuracy given the small cell size and optical resolution limits (further described in their supplemental text).
 817 Perhaps the more concerning point is that while each of these alternative measurements show an exponential
 818 increase in cell size at faster growth rates, the measurements used by Schmidt *et al.* appear to plateau. This
 819 resulted in an analogous trend in their final reported total cellular protein per cell as shown in Figure **Figure 17**
 820 (purple data points), and is in disagreement with other measurements of total protein at these growth rates (**Basan**
 821 *et al.*, 2015).

822 Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the
 823 following subsections we begin by considering this calculation. We then consider three different approaches to
 824 estimate the growth-rate dependent total protein mass to compare with those values reported from **Schmidt *et al.***
 825 (2016). The results of this are summarized in **Figure 16(B)**, with the original values from both **Schmidt *et al.* (2016)**
 826 and **Li *et al.* (2014)** shown in **Figure 16(A)** for reference. For most growth conditions, we find that total protein per
 827 cell is still in reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
 828 amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main
 829 reason why we chose to readjusted protein abundance as shown in **Figure 11(B)** (with the calculation described in
 830 section Estimation of Total Protein Content per Cell).

831 Effect of cell volume on reported absolute protein abundances

832 As noted in section Experimental Details Behind Proteomic Data, the authors calculated proteome-wide protein
 833 abundances by first determining absolute abundances of 41 pre-selected proteins, which relied on adding syn-
 834 synthetic heavy reference peptides into their protein samples at known abundance. This absolute quantitation was
 835 performed in replicate for each growth condition. Separately, the authors also performed a more conventional
 836 mass spectrometry measurement for samples from each growth condition, which attempted to maximize the
 837 number of quantified proteins but only provided relative abundances based on peptide intensities. Finally, using
 838 their 41 proteins with absolute abundances already determined, they then created calibration curves with which
 839 to relate their relative intensity to absolute protein abundance for each growth condition. This allowed them to
 840 estimate absolute protein abundance for all proteins detected in their proteome-wide data set. Combined with
 841 their flow cytometry cell counts, they were then able to determine absolute abundance of each protein detected

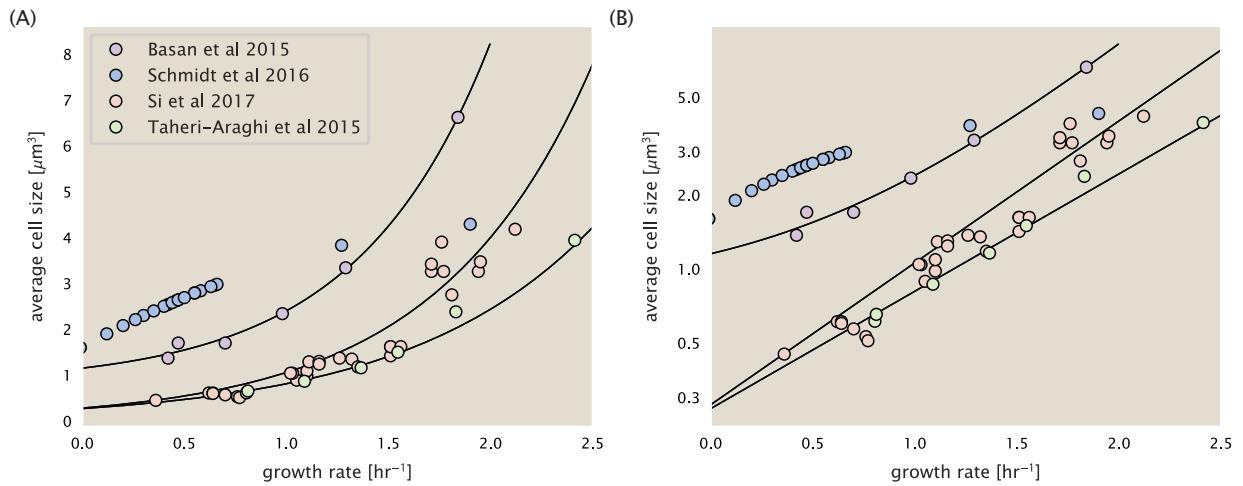


Figure 16. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are shown in purple (*Basan et al. (2015)*), while the two data sets shown in green and 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*.

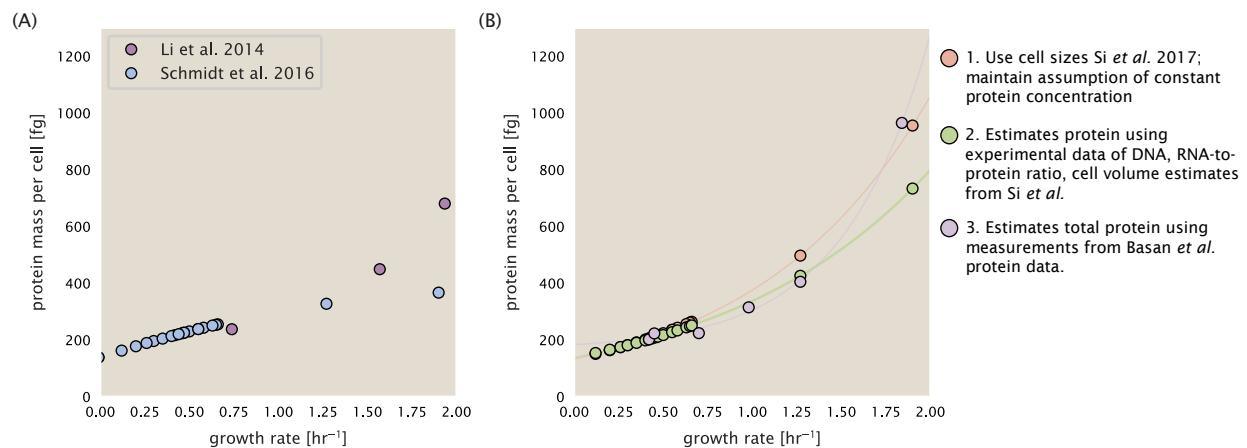


Figure 17. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt *et al.* (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: 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.

842 on a per cell basis.

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

850 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per cell was measured.
851 Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentration of $[M_p]_{orig} = M_p/V_{orig} =$
852 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,
853 they could then estimate the total protein mass for all other growth conditions from,

$$M_{P_i} = [M_p]_{orig} \cdot V_i \quad (11)$$

854 where M_{P_i} represents the total protein mass per cell and V_i is the cell volume for each growth condition i as
855 measured in Volkmer and Heinemann, 2011. Here the thinking is that the values of M_{P_i} reflects the total cellular
856 protein for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due
857 to protein loss during sample preparation. The protein abundances from their absolute abundance measurements
858 noted above were therefore scaled to their estimates and are shown in Figure **Figure 17** (purple data points).

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

$$M'_{P_i} = [M_p]_{Si} \cdot V_{Si_i} \quad (12)$$

864 where M'_{P_i} is the new protein mass prediction, and V_{Si_i} refers to the new volume prediction for each condition i ,
865 These are shown as red data points in Figure **Figure 17(B)**.

866 Relaxing assumption of constant protein concentration across growth conditions

867 We next relax the assumption that cellular protein concentration is constant and instead, attempt to estimate it
868 using experimental data. Here we use the estimation of total protein mass per cell detailed in section **Estimation**
869 of Total Protein Content per Cell for all data points in the **Schmidt *et al.* (2016)** data set. The green data points in
870 **Figure 17(B)** show this prediction, and this represents the approach used to estimate total protein per cell for all
871 data sets.

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

873 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration
874 and cell volumes. These are inherently difficult to accurately due to the small size of *E. coli*. Indeed, for all the
875 additional measurements of cell volume included in Figure **Figure 16**, no measurements were performed for cells
876 growing at rates below 0.5 hr^{-1} . It therefore remains to be determined whether our extrapolated cell volume
877 estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower
878 growth.

879 In our last approach we therefore attempt to estimate total protein using experimental data that required no
880 estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*, the authors measured total
881 protein per cell for a broad range of growth rates (reproduced in Figure **Figure 18**). These were determined by
882 first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method (**You *et al.* (2013)**), and
883 then abundance per cell was calculated from cell counts from either plating cells or a Coulter counter. While it
884 is unclear why Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more
885 consistent with our expectation that cell mass will increase exponentially with faster growth rates. In addition,

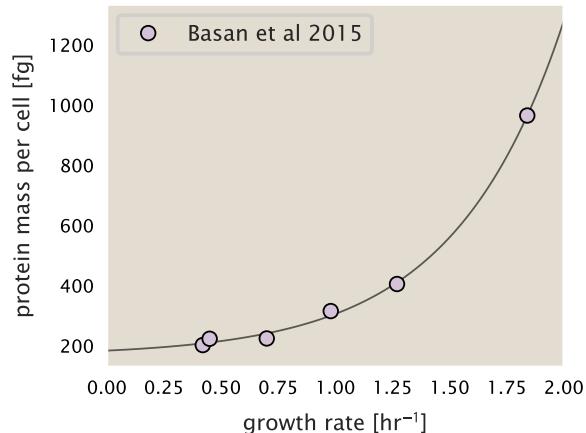


Figure 18. Total cellular protein reported in Basan et al. 2015. Measured protein mass as a function of growth rate as reproduced from Basan et al. 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given by $14.65 e^{2.180 \cdot \lambda} + 172$ fg per cell, where λ is the growth rate in hr^{-1} .

886 although they do not consider growth rates below about 0.5 hr^{-1} , it is interesting to note that the protein mass per
 887 cell appears to plateau to a minimum value at slow growth. In contrast, our estimates using cell volume so far have
 888 predicted that total protein mass should continue to decrease slightly for slower growing cells. By fitting this data
 889 to an exponential function dependent on growth rate, we could then estimate the total protein per cell for each
 890 growth condition considered by Schmidt et al. (2016). These are plotted as red data points in Figure 17(B).

891 Calculation of Complex Abundance

892 All protein data quantified the abundance of individual proteins per cell. However, this work requires estimates on
 893 the abundance of individual protein *complexes*, rather than the copy number of individual proteins. In this section,
 894 we outline the approach we used to annotate proteins as being part of a macromolecular complex and how we
 895 computed their absolute abundances per cell.

896 Protein complexes, and proteins individually, often have a variety of names, both longform and shorthand. As
 897 individual proteins can have a variety of different synonyms, we sought to ensure that each protein annotated
 898 in the data sets used the same synonym. To do use, we relied heavily on the EcoCyc Genomic Database (Keseler
 899 et al., 2017). Each protein in available data sets included an annotation of one of the gene name synonyms as
 900 well as an accession ID – either a UniProt or Blattner “b-number”. We programmatically matched up individual
 901 accession IDs between the proteins in different data sets. In cases where accession IDs matched but the gene
 902 names were different, we manually verified that the gene product was the same between the datasets and chose a
 903 single synonym. All code used in the data cleaning and unification procedures can be found on the associated
 904 [GitHub repository] (DOI:XXX) associated with this paper as well as on the associated [paper website](#).

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

914 In addition to the annotated complex, we collected information on the stoichiometry of each macromolecular
 915 complex. For a complex with N_{subunits} protein species, for each protein subunit i we first calculate the number of

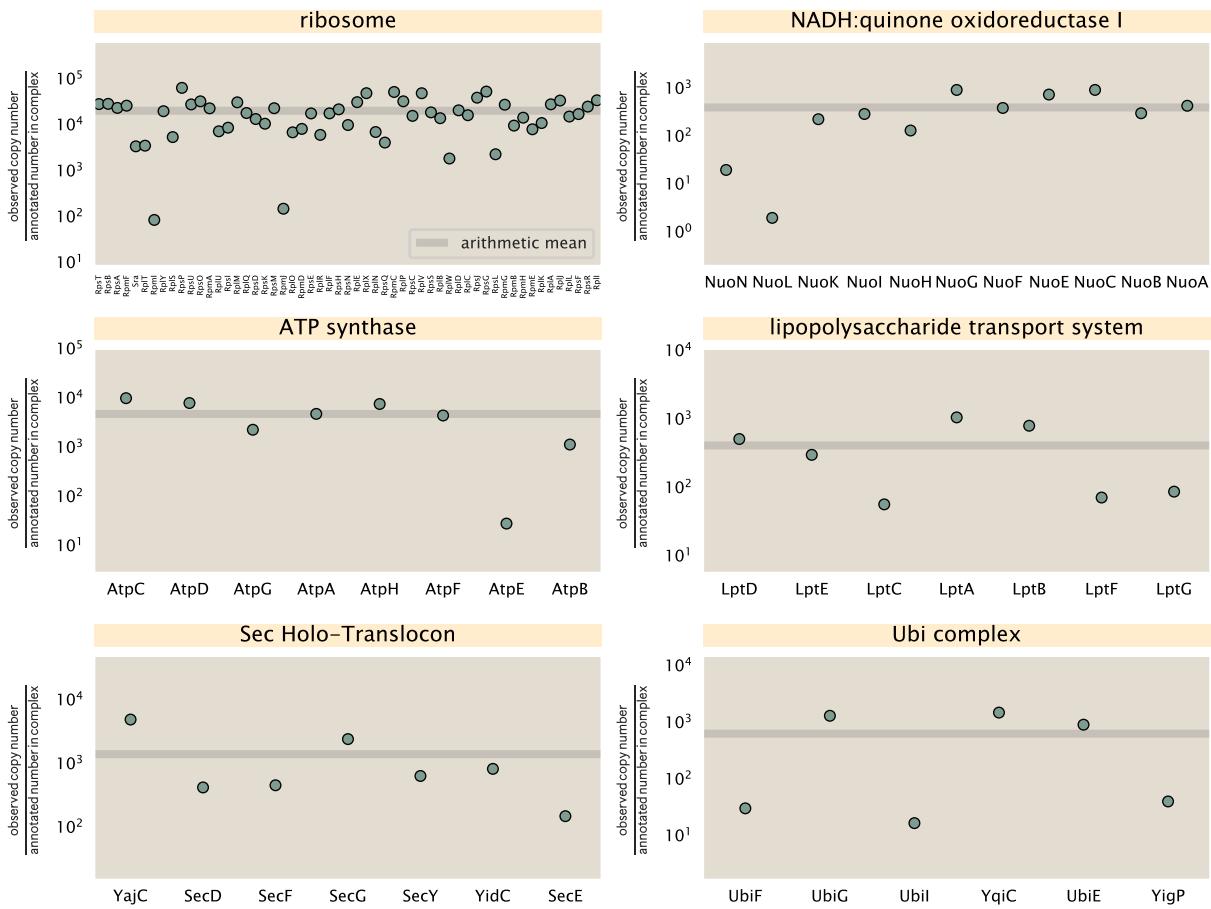


Figure 19. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds to the arithmetic mean across all subunits. These data correspond to measurements from Schmidt et al. (2016) in a glucose-supplemented minimal medium.

916 complexes that could be formed given the measured protein copy numbers per cell,

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (13)$$

917 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
918 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,
919 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each ribosomal
920 protein, we then calculate the maximum number of complexes that could be formed using Equation 13. This
921 example, along with example from 5 other macromolecular complexes, can be seen in Figure 19.

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

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (14)$$

925 in Figure 19, we show this mean value as a grey line for a variety of different complexes. Additionally, we have built
926 an interactive figure accessible on the [paper website](#) where the validity of this approach can be examined for any
927 complex with more than two subunits (thus, excluding monomers and dimers).

928 **Extending Estimates to a Continuum of Growth Rates**

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

933 **Estimation of the total cell mass**

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

940 Rather, using the phenomenological description of cell volume scaling exponentially with growth rate, and
941 using a rule-of-thumb of a cell buoyant density of ≈ 1.1 pg / fL (BNID: 103875), we can calculate the cell dry mass
942 across a range of physiological growth rates as

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{\lambda * b} \quad (15)$$

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

946 **Complex Abundance Scaling With Cell Volume**

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

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

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (16)$$

959 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 15**. To convert this to the
960 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
961 with **Equation 16**,

$$N_X \approx \frac{m_X}{w_X}. \quad (17)$$

962 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
963 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
964 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
965 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
966 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
967 transport scaling relation of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (18)$$

968 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 969 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 970 point estimate outlined in the main text. This is due to the choice of ≈ 300 fg total dry mass per cell for the point
 971 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 972 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 973 and not the predict the exact values of the abundances.

974 **Equation 18** is a very general relation for processes where the cell volume is the "natural variable" of the
 975 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 976 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 977 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 18** can be used
 978 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 979 **Equation 18** can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (19)$$

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

983 A Relation for Complex Abundance Scaling With Surface Area

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

991 Number of Lipids

992 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 993 the surface area of a single lipid ($A_{\text{lipid}} \approx 0.5 \text{ nm}^2$, BNID: 106993) and the total fraction of the membrane composed
 994 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 995 **Equation 9**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 996 lipids as a function of growth rate as

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (20)$$

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

998 Number of Murein Monomers

999 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1000 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murien monomer is
 1001 $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1002 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1003 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1004 number of murein monomers as a function of growth rate can be calculated as

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (21)$$

1005 where ρ_{pg} is the density of peptidoglycan.

1006 **Complex Abundance Scaling With Number of Origins, and rRNA Synthesis**

1007 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
1008 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
1009 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
1010 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in
1011 *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
1012 average number of origins of replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{cyc} \lambda / \ln 2} \quad (22)$$

1013 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{cyc} \approx 70$ min, which is
1014 the time it takes two replisomes to copy an entire chromosome.

1015 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1016 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{rRNA}(\lambda) \approx N_{rRNA \text{ operons}} \times 2^{t_{cyc} \lambda / \ln 2}. \quad (23)$$

1017 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons are
1018 some distance away from the origin, making **Equation 23** an approximation (*Dennis et al., 2004*).

1019 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
1020 approximation, **Equation 22** illustrates that is not precisely true, even at slow growth rates. In estimating the
1021 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1022 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1023 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{rRNA} \rho_{\text{RNA polymerase}}, \quad (24)$$

1024 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA
1025 polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1026 **Calculation of active ribosomal fraction.**

1027 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai et al. (2016)* to
1028 estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to
1029 consider how f_a should vary with growth rate, and instead find that the data is well-approximated by fitting to an
1030 exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset of **Figure 8(C)**). We use this function to estimate f_a
1031 for each of the data points shown in **Figure 8(C)**.

1032 **Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.**

1033 *E. coli* shows robust scaling of cell size with the average the number of origins $\langle \#ori \rangle$ per cell (*Si et al., 2017*).
1034 Since protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes
1035 in proteomic composition and total abundance across growth conditions. Given the potential constraints on
1036 rRNA synthesis and changes in ribosomal copy number with $\langle \#ori \rangle$, it becomes important to also consider how
1037 protein copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make
1038 sense of the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a
1039 mechanistic level. As considered in the main text, it is becoming increasingly apparent that regulation through the
1040 secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer
1041 nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become important parameters to
1042 consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular, causes a relatively higher gene dosage
1043 in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

1044 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited wild-type cell
1045 data from *Si et al. (2017)*. We consider their measurements of DNA replication time (t_C , 'C' period of cell division),
1046 total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across
1047 a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio

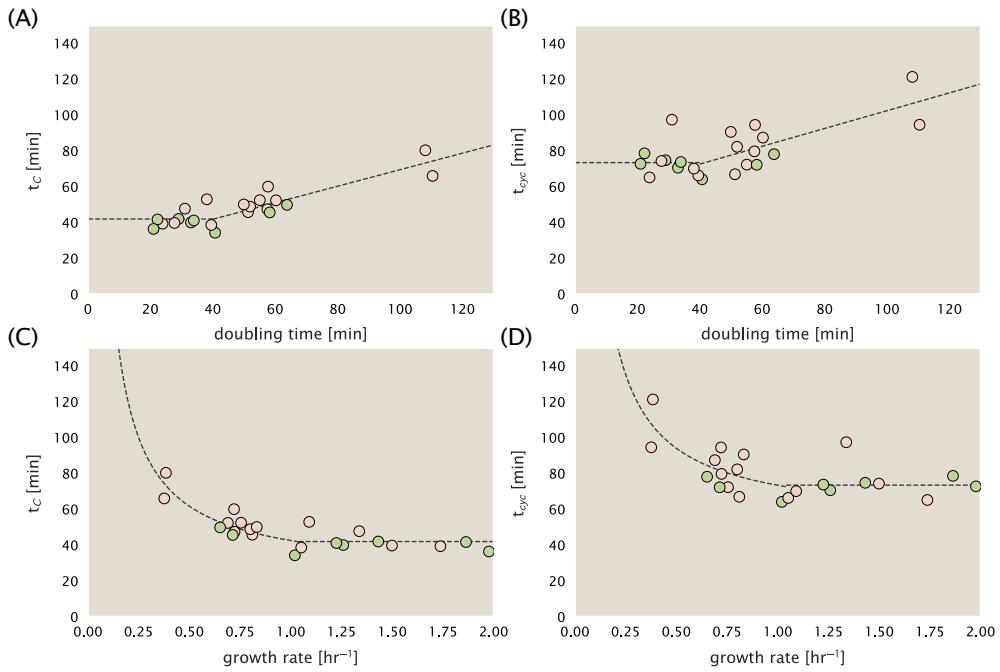


Figure 20. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling 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 $\lambda = \ln(2)/\tau$.

from their data. We begin by considering $\langle \#ori \rangle$. If the cell cycle time takes longer than the time of cell division, the 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 growth. Cells will need to do this in proportion to the ratio $\lambda_{cyc}/\lambda = t_{cyc}/\tau$, and the number of origins per cell (on average) is then given by $2^{t_{cyc}/\tau}$. The average number of termini will in contrast depend on the lag time between DNA replication and cell division, t_D , with $\langle \#ori \rangle / \langle \# ter \rangle$ ratio = $2^{t_{cyc}/\tau - t_D/\tau} = 2^{t_C/\tau}$.

In Figure 20(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from Si et al. (2017). The authors estimated t_C by marker frequency analysis using qPCR, while $t_{cyc} = t_C + t_D$ were inferred from t_C and 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 period of 40 minutes, this would correspond to a maximum rate of elongation of about 1,000 bp/sec. Since we lacked a specific model to describe how each of these parameters vary with growth condition, we assumed that they were linearly dependent on the doubling time. For each parameter, t_C and t_{cyc} , we split them up into two 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 20(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.

1063 References

- 1064 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1065 *coli*. *BMC Systems Biology*, 8(1):133.
- 1066 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by
1067 ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1068 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,
1069 M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G.
1070 (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1071 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1072 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1073 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1074 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH₃ Channel.
1075 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1076 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1077 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1078 2(1):2460.
- 1079 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1080 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1081 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1082 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1083 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.
1084 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings of
1085 the National Academy of Sciences*, 115(21):E4796–E4805.
- 1086 Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009). Absolute metabolite concentrations
1087 and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology*, 5(8):593–599.
- 1088 Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*. *Proceedings of the
1089 National Academy of Sciences*, 68(5):925–929.
- 1090 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential
1091 Growth Rates. *EcoSal Plus*, 3(1).
- 1092 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews
1093 Microbiology*, 14(10):638–650.
- 1094 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1095 266:2020.06.16.154187.
- 1096 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.,
1097 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1098 Cross-Linking by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1099 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1100 *Biochemical Journal*, 117(3):551–562.
- 1101 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational
1102 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1103 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1104 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1105 Microbiology*, 2(2):16231.
- 1106 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1107 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1108 Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V., Pinglay, S., Kennedy, K. J., Schaffer, M., Gutierrez, J. I., Sang, D., Poterewicz,
1109 G., Chung, J. K., Plitzko, J. M., Groves, J. T., Jacobs-Wagner, C., Engel, B. D., and Holt, L. J. (2018). mTORC1 Controls Phase
1110 Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*, 174(2):338–349.e20.

- 1111 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1112 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- 1113 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*,
1114 108(44):17876–17882.
- 1115 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyru-
1116 vate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation.
1117 *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1118 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M.
1119 (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth
1120 Rates. *mBio*, 11(2):45.
- 1121 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair.
1122 *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1123 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1124 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of
1125 the National Academy of Sciences*.
- 1126 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández,
1127 K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-
1128 Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D.,
1129 Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif
1130 clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1131 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
1132 Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1133 Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary σ factor in *Escherichia coli* can access
1134 the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.
- 1135 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-
1136 Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution
1137 on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1138 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1139 microbiology*, 26(10):815–832.
- 1140 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of
1141 (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1142 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content
1143 of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1144 Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal of
1145 Molecular Biology*, 31(3):507–518.
- 1146 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).
1147 Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1148 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic
1149 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- 1150 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.
- 1151 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1152 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1153 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1154 20(17):4952–4963.
- 1155 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1156 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.

- 1157 Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonja, E., Mekler, V., Ebright, R. H.,
1158 and Weiss, S. (2005). Retention of Transcription Initiation Factor Σ 70 in Transcription Elongation: Single-Molecule Analysis.
1159 *Molecular Cell*, 20(3):347–356.
- 1160 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong,
1161 Q., Ong, W. K., Paley, S. M., and Subhraweti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1162 *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1163 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.
1164 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- 1165 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari,
1166 A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraweti, P., Velázquez-Ramírez,
1167 D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about
1168 Escherichia coliK-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1169 Khaderian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1170 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1171 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1172 *Current Opinion in Biotechnology*, 28:96–102.
- 1173 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–
1174 1375.
- 1175 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1176 125(2):028103.
- 1177 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by
1178 Modulating Supercoiling of oriC. *mbio*, 10(4):822.
- 1179 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and
1180 Biomembranes*, 32(3):237–246.
- 1181 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1182 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1183 Lex, A., Gehlenborg, N., Strobel, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions
1184 on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1185 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1186 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1187 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1188 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1189 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1190 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1191 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator
1192 GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1193 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1194 112(51):15690–15695.
- 1195 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression. Springer.
- 1196 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by
1197 condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1198 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
1199 Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1200 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1201 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1202 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1203 *Growth Symposium*, 9:223–289.

- 1204 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1205 Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again Relationship? *Molecular Cell*, 20(3):335–345.
- 1207 Mooney, R. A. and Landick, R. (2003). Tethering Σ 70 to RNA polymerase reveals high in vivo activity of σ factors and Σ 70-dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.
- 1209 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–12515.
- 1212 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1213 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria. *eLife*, 8:e642.
- 1215 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1216 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1218 Perdue, S. A. and Roberts, J. W. (2011). σ^{70} -dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*, 412(5):782–792.
- 1220 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1221 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature microbiology*, 1(11):1–7.
- 1223 Rosenberg, H., Gerdes, R. G., and Chegwinden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of Bacteriology*, 131(2):505–511.
- 1225 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–11734.
- 1227 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1229 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1231 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102.
- 1233 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. *Cell*, 172(6):1294–1305.
- 1235 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 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1237 Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017). Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1239 Soler-Bistué, A., Aguilar-Pierlé, S., García-Garcé, M., Val, M.-E., Sismeiro, O., Varet, H., Sieira, R., Krin, E., Skovgaard, O., Comerci, D. J., Eduardo P. C. Rocha, and Mazel, D. (2020). Macromolecular crowding links ribosomal protein gene dosage to growth rate in *Vibrio cholerae*. *BMC Biology*, 18(1):1–18.
- 1242 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1244 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- 1245 Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied and Environmental Microbiology*, 70(11):6670–6677.
- 1247 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek*, 39(1):545–565.

- 1249 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.
1250 *Archives of Microbiology*, 113(3):185–189.
- 1251 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1252 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1253 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
1254 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1255 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
1256 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1257 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1258 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1259 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1260 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1261 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1262 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1263 47(D1):D330–D338.
- 1264 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1265 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1266 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1267 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,
1268 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1269 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1270 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1271 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data
1272 Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1273 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1274 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1275 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1276 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
1277 bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1278 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014). Regulation of crp gene expression by the catabolite repressor/activator, cra, in
1279 *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1280 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1281 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.

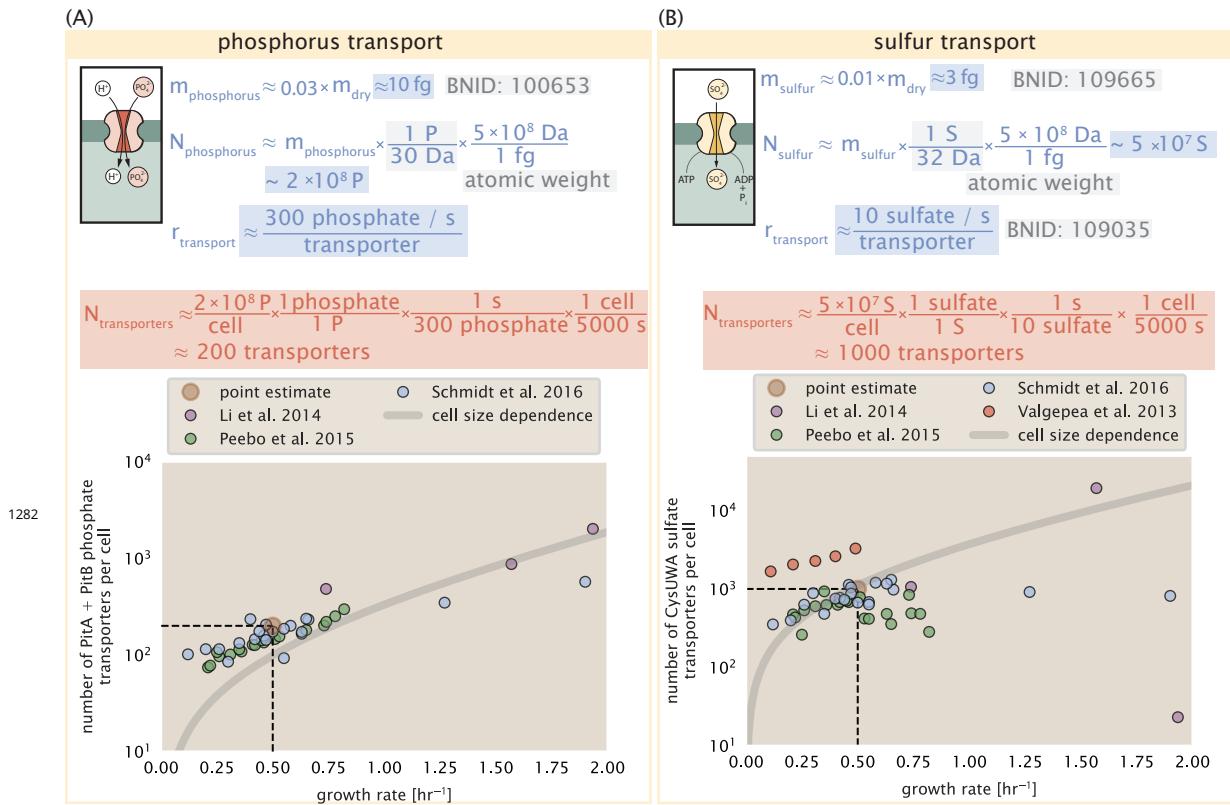
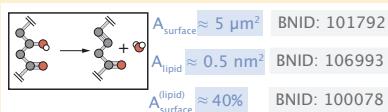


Figure 2-Figure supplement 1. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (B) Estimate of the number of CysUWA complexes necessary to maintain a 1% sulfur *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.

CELL ENVELOPE BIOSYNTHESIS

(A)

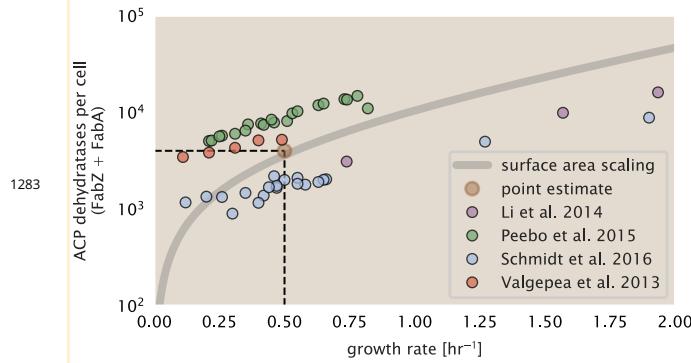
lipid synthesis



$$N_{\text{lipids}} \approx \frac{4 \text{ leaflets}}{1 \text{ cell}} \times \frac{0.4 \times 5 \mu\text{m}^2}{1 \text{ leaflet}} \times \frac{1 \text{ lipid}}{0.5 \text{ nm}^2} \times \frac{10^6 \text{ nm}^2}{1 \mu\text{m}^2} \approx 2 \times 10^7 \text{ lipids / cell}$$

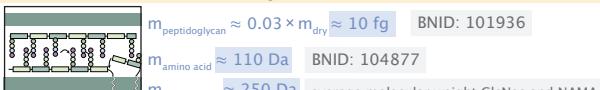
$r_{\text{ACP dehydratase}} \approx 1 \text{ lipid / sec}$ Ruppe & Fox 2018; Fiers et al. 2016; Heath et al. 1996

$$N_{\text{ACP dehydratases}} \approx \frac{2 \times 10^7 \text{ lipids}}{1 \text{ cell}} \times \frac{1 \text{ sec} \times \text{enzyme}}{1 \text{ lipid}} \times \frac{1 \text{ cell}}{5000 \text{ sec}} \approx 4000 \text{ ACP dehydratases}$$



(B)

peptidoglycan synthesis



$$m_{\text{monomer}} \approx 5 \times m_{\text{amino acid}} + 2 \times m_{\text{murein sugar}} \approx 1000 \text{ Da}$$

$$N_{\text{monomer}} \approx \frac{10 \text{ fg}}{1 \text{ cell wall}} \times \frac{1 \text{ monomer}}{1000 \text{ Da}} \times \frac{5 \times 10^8 \text{ Da}}{1 \text{ fg}} \approx \frac{5 \times 10^6 \text{ monomers}}{\text{cell wall}}$$

$$N_{\text{crosslinks}} \approx 0.2 \times N_{\text{murein monomers}} \approx 10^6 \text{ crosslinks}$$
 Vollmer et al. 2008; Rogers et al. 1980

$r_{\text{transpeptidase}} \approx 2 \text{ crosslinks / sec}$ Catherwood et al. 2020

$$N_{\text{transpeptidases}} \approx \frac{10^6 \text{ crosslinks}}{1 \text{ cell}} \times \frac{1 \text{ sec} \times \text{enzyme}}{2 \text{ crosslinks}} \times \frac{1 \text{ cell}}{5000 \text{ sec}} \approx 100 \text{ transpeptidases}$$

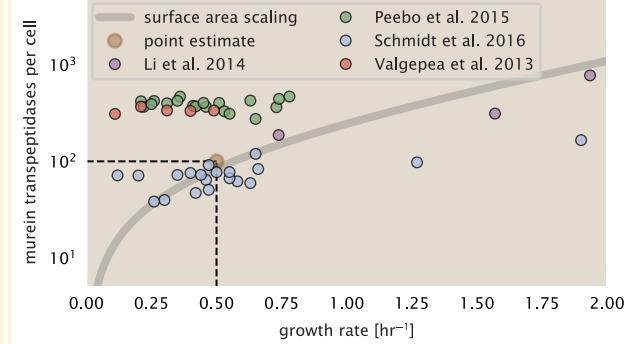


Figure 4–Figure supplement 1. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabZ}]_2$ and $[\text{FabA}]_2$. (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 110 Da and each sugar in the disaccharide having a mass of ≈ 250 Da. The *in vivo* rate of transpeptidation 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.

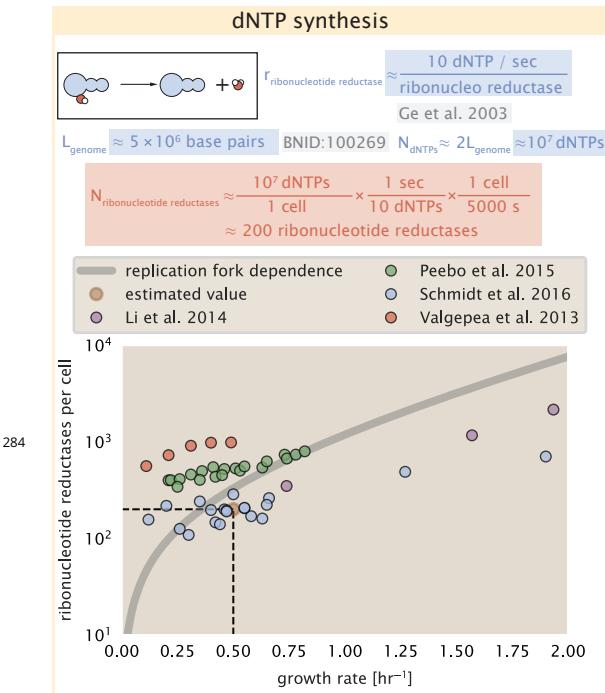


Figure 5–Figure supplement 1. Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[NrdA]_2[NrdB]_2$) and 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.

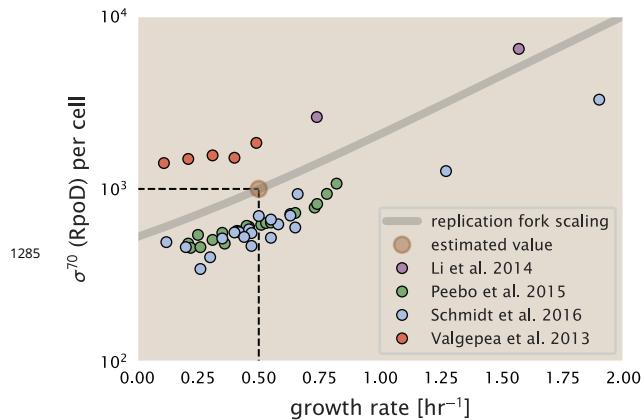


Figure 6–Figure supplement 1. The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown as a translucent brown point and grey line.

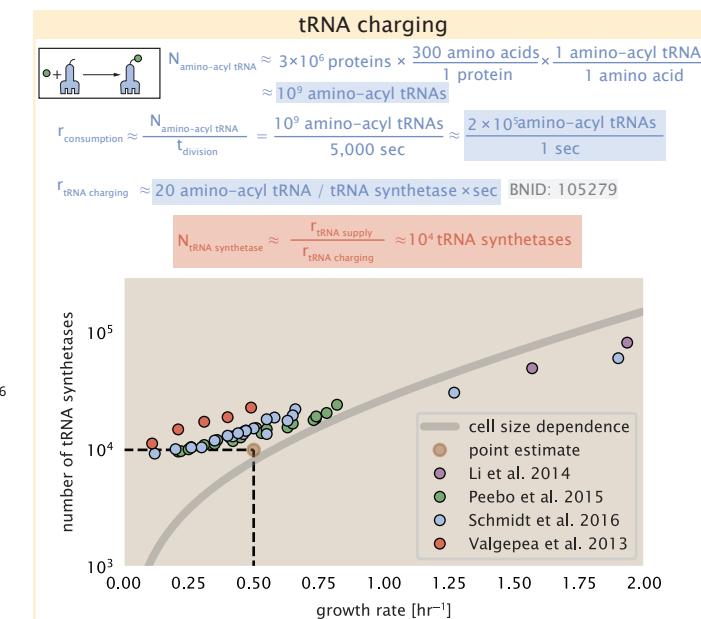


Figure 7-Figure supplement 1. Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [lysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂).