

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.

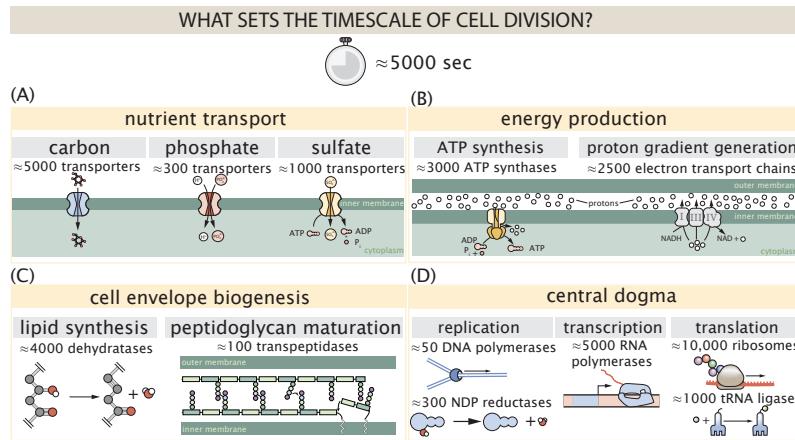


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP 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 Uptake of Nutrients

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. Given the diversity of biological transport
120 mechanisms and the vast number of inputs needed to build a cell, we begin by considering transport of some of
121 the most important cellular ingredients: carbon, nitrogen, oxygen, hydrogen, phosphorus, and sulfur.

122 The elemental composition of *E. coli* has received much quantitative attention over the past half century
123 (*Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv, 1976*), providing us with a
124 starting point for estimating the copy numbers of various transporters. While there is some variability in the exact
125 elemental percentages (with different uncertainties), we can estimate that the dry mass of a typical *E. coli* cell is \approx
126 45% carbon (BioNumber ID: 100649, see **Box 1**), \approx 15% nitrogen (BNID: 106666), \approx 3% phosphorus (BNID: 100653),
127 and 1% sulfur (BNID: 100655).

128 Carbon Transport

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

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; Adelberg 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*). A famed example of such dual-input regulatory

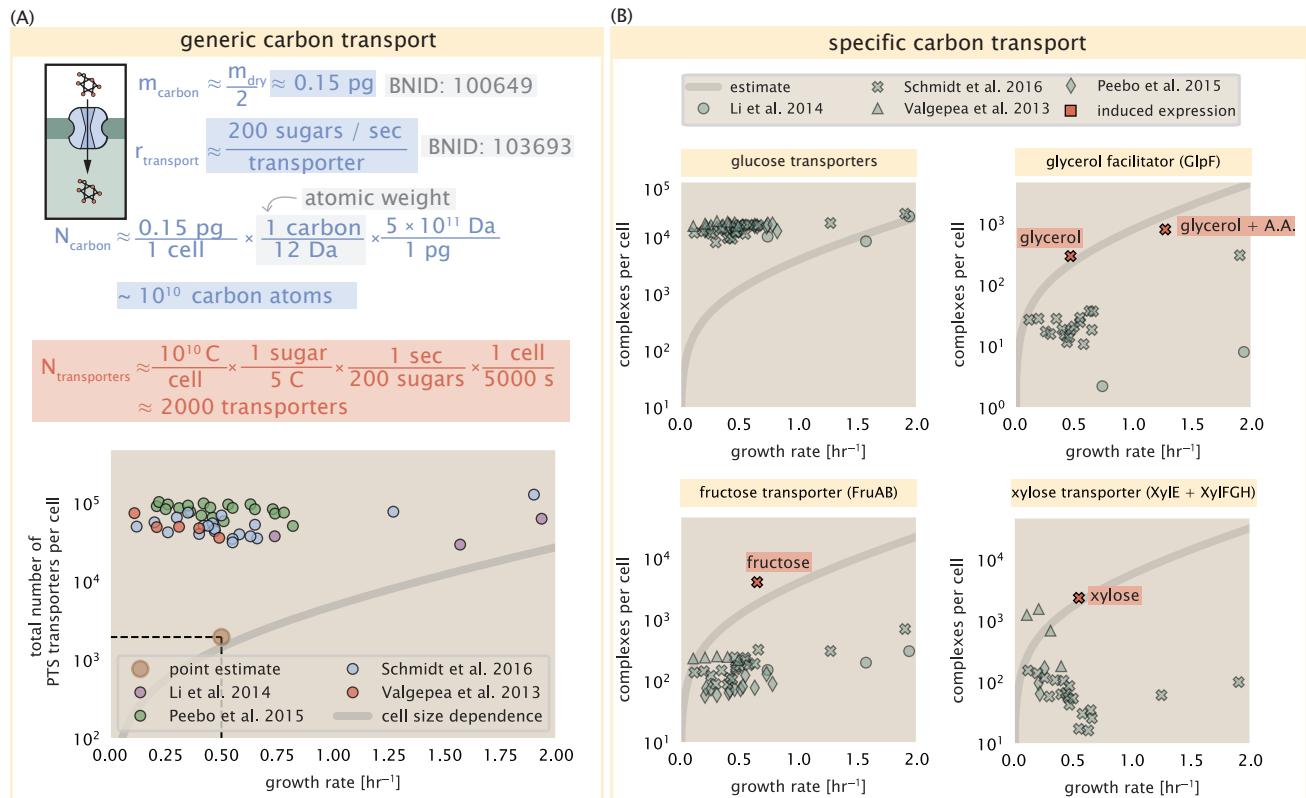


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules 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} (*Lu et al., 2003*), 200 fructose- s^{-1} (assumed to be similar to PtsI, BNID: 103693), and 50 xylose- s^{-1} (assumed to be comparable to LacY, BNID: 103159). Red points and highlighted text indicate conditions in which the only source of carbon in the growth medium induces expression of the transport system. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

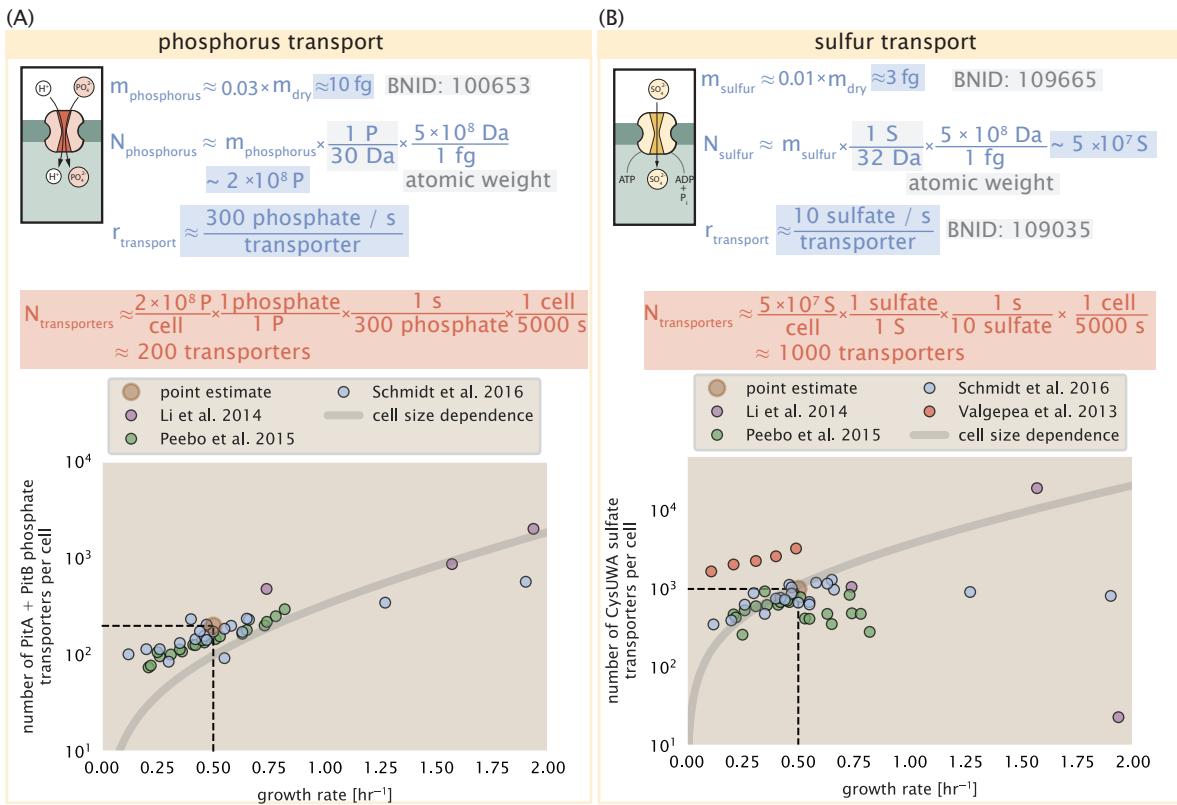


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (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 $[\text{CysA}]_2[\text{CysU}][\text{CysW}][\text{Sbp/CysP}]$. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

logic is in the regulation of the *lac* operon which is only activated in the absence of glucose and the presence of allolactose, an intermediate in lactose metabolism (*Jacob and Monod, 1961*), though we now know of many other such examples (*Ireland et al., 2020; Gama-Castro et al., 2016; Belliveau et al., 2018*). Several examples are shown in **Figure 2(B)**. Points colored in red (labeled by red text-boxes) correspond to growth conditions in which the specific carbon source (glycerol, xylose, or fructose) is present. The grey lines in **Figure 2(B)** show the estimated number of transporters needed at each growth rate to satisfy the cellular carbon requirement, adjusted for the specific carbon source. These plots show that, in the absence of the particular carbon source, expression of the transporters is maintained on the order of $\sim 10^2$ per cell. However, when the transport substrate is present, expression is induced and the transporters become highly-expressed. The low but non-zero abundances for many of these alternative across growth conditions may reflect the specific regulatory logic, requiring the cell to transport some minimal amount of an alternative carbon source in order to induce expression of these alternative carbon-source systems. Together, this generic estimation and the specific examples of induced expression suggest that transport of carbon across the cell membrane, while critical for growth, is not the rate-limiting step of cell division.

173 Limits on Transporter Expression

So which, if any, of these processes may be rate limiting for growth? As suggested by **Figure 2(B)**, induced expression can lead to an order-of-magnitude (or more) increase in the amount of transporters needed to facilitate transport. Thus, if acquisition of nutrients was the limiting state in cell division, could expression simply be increased to accommodate faster growth? A way to approach this question is to compute the amount of space in the bacterial membrane that could be occupied by nutrient transporters. Considering a 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 transporters to be approximately 200

180 transporters/ μm^2 . For a typical transporter occupying about 50 nm^2/dimer , this amounts to about only 1 percent
181 of the total inner membrane (*Szenk et al., 2017*). In addition, bacterial cell membranes typically have densities
182 of 10^5 proteins/ μm^2 (*Phillips, 2018*), implying that the cell could accommodate more transporters of a variety of
183 species if it were rate limiting. As we will see in the next section, however, occupancy of the membrane can impose
184 other limits on the rate of energy production.

185 Energy Production

186 While the transport of nutrients is required to build new cell mass, the metabolic pathways both consume and
187 generate energy in the form of NTPs. The high-energy phosphodiester bonds of (primarily) ATP power a variety of
188 cellular processes that drive biological systems away from thermodynamic equilibrium. The next set of processes
189 we consider as molecular bottlenecks controls the energy budget of a dividing cell via the synthesis of ATP from
190 ADP and inorganic phosphate as well as maintenance of the electrochemical proton gradient which powers it.

191 ATP Synthesis

192 Hydrolysis of the terminal phosphodiester bond of ATP forming ADP (or alternatively GTP and GDP) and an inorganic
193 phosphate is a kinetic driving force in a wide array of biochemical reactions. One such reaction is the formation
194 of peptide bonds during translation which requires ≈ 2 ATPs for the charging of an amino acid to the tRNA and
195 ≈ 2 GTP for the formation of the peptide bond between amino acids. Assuming the ATP costs associated with
196 error correction and post-translational modifications of proteins are negligible, we can make the approximation
197 that each peptide bond has a net cost of ≈ 4 ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from
198 ATP is achieved via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy
199 investment (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to be functionally
200 equivalent to being exclusively ATP. In total, the energetic costs of peptide bond formation consume $\approx 80\%$ of the
201 cells ATP budget (BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015); Stouthamer (1973)*). The pool
202 of ATP is produced by the F_1 - F_0 ATP synthase – a membrane-bound rotary motor which under ideal conditions can
203 yield ≈ 300 ATP per second (BNID: 114701; *Weber and Senior (2003)*).

204 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make the approximation
205 that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300 peptide bonds (BNID: 115702;
206 108986; 104877). Taking these values together, we estimate that the typical *E. coli* cell consumes $\sim 5 \times 10^9$ ATP
207 per cell cycle on protein synthesis alone. Assuming that the ATP synthases are operating at their fastest possible
208 rate, ≈ 3000 ATP synthases are needed to keep up with the energy demands of the cell. This estimate and a
209 comparison with the data are shown in *Figure 4* (A). Despite our assumption of maximal ATP production rate per
210 synthase and approximation of all NTP consuming reactions being the same as ATP, we find that an estimate
211 of a few thousand complete synthases per cell to agree well with the experimental data. Much as we did for
212 the estimates of transporter copy number in the previous section, we can generalize this estimate to consider a
213 continuum of growth rates rather than a point estimate of 5000 s. Given knowledge of how the cell volume scales
214 with growth rate (*Si et al., 2017*), the density of the cytoplasm ($\rho \approx 1 \text{ pg} / \text{fL}$), and the empirical determination that
215 approximately half of the dry mass is protein, we can compute the energy demand as a function of growth rte,
216 indicated by the gray line in *Figure 4*.

217 This simple estimate provides an intuition for the observed scaling and growth rate dependence, so is it a
218 molecular bottleneck? If the direct production of ATP was a rate limiting step for growth, could the cell simply
219 express more ATP synthase complexes? This requires us to consider several features of cellular physiology, namely
220 the physical space on the inner membrane as well as the ability to maintain the proton chemical gradient leveraged
221 by the synthase to drive ATP production out of equilibrium.

222 Generating the Proton Electrochemical Gradient

223 In order to produce ATP, the F_1 - F_0 ATP synthase itself must consume energy. Rather than burning through its
224 own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the
225 electrochemical potential established across the inner membrane through cellular respiration. This electrochemical
226 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as
227 they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner

228 membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in the
229 intermembrane space.

230 However, the constant rotation of the ATP synthases would rapidly abolish this potential difference if it were
231 not being actively maintained. To undergo a complete rotation (and produce a single ATP), the F₁-F₀ ATP synthase
232 must shuttle ≈ 4 protons across the membrane into the cytosol (BNID: 103390). With ≈ 3000 ATP synthases each
233 generating 300 ATP per second, the 2×10^4 protons establishing the 200 mV potential would be consumed in only
234 a few milliseconds. This brings us to our next estimate: how many electron transport complexes are needed to
235 support the consumption rate of the ATP synthases?

236 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical
237 and biophysical study over the past half century (*Ingledew and Poole, 1984; Khademian and Imlay, 2017; Cox*
238 *et al., 1970; Henkel et al., 2014*). A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E.*
239 *coli* electron transport complexes using structural and biochemical data, revealing that each electron transport
240 chain rapidly pumps protons into the intermembrane space at a rate of ≈ 1500 protons per second (BIND: 114704;
241 114687). Using our estimate of the number of ATP synthases required per cell [*Figure 4(A)*], coupled with these
242 recent measurements, we estimate that ≈ 1000 electron transport complexes would be necessary to facilitate the
243 $\sim 5 \times 10^6$ protons per second diet of the cellular ATP synthases. This estimate (along with a generalization to the
244 entire range of observed growth rates) is in agreement with the number of complexes identified in the proteomic
245 datasets (plot in *Figure 4(B)*). This suggests that every ATP synthase must be accompanied by ≈ 1 functional electron
246 transport chain. Again, if this were a rate limiting process for bacterial growth, one must conclude that it is not
247 possible for the cell to simply increase the production of both the number of electron transport chain complexes
248 as well as ATP synthases. As both of these components only function bound to the inner membrane, we now turn
249 our attention towards the available space in the membrane as well as surface-area-to-volume constraints.

250 Energy Production in a Crowded Membrane.

251 For each protein considered so far, the data shows that in general their numbers increase with growth rate. This is
252 in part a consequence of the increase in cell length and width at that is common to many rod-shaped bacteria at
253 faster growth rates (*Ojikic et al., 2019; Harris and Theriot, 2018*). For the particular case of *E. coli*, the total cellular
254 protein and cell size increase logarithmically with growth rate (*Schaechter et al., 1958; Si et al., 2017*).

255 Recall however that each transport process, as well as the ATP production via respiration, is performed at the
256 bacterial membrane. This means that their maximum productivity can only increase in proportion to the cell's
257 surface area divided by the cell doubling time. This difference in scaling would vary in proportion to the surface
258 area-to-volume (S/V) ratio. Earlier we found that there was more than sufficient membrane real estate for carbon
259 intake in our earlier estimate. However, since the total number of ATP synthases and electron chain transport
260 complexes both exhibit a clear increase in copy number with growth rate, it was important to also consider the
261 consequences of this S/V ratio scaling in more detail.

262 In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell cycle or 10^6
263 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
264 previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al., 2017*). In *Figure 5 (A)* we plot this ATP
265 demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from spherical to
266 rod-shaped with an aspect ratio (length/width) equal to 4 (See appendix for calculations of cell volume and surface
267 area). In order to consider the maximum power that could be produced, we consider the amount of ATP that can
268 be generated by a membrane filled with ATP synthase and electron transport complexes, which provides a maximal
269 production of about 3 ATP / (nm²·s) (*Szenk et al., 2017*). This is shown in blue in *Figure 5(A)*, which shows that at
270 least for the growth rates observed, the energy demand is roughly an order of magnitude less. Interestingly, *Szenk*
271 *et al. (2017)* also found that ATP production by respiration is less efficient than by fermentation per membrane area
272 occupied due to the additional proteins of the electron transport chain. This suggests that, even under anaerobic
273 growth, there will be sufficient membrane space for ATP production in general.

274 While the analysis in *Figure 5(A)* serves to highlight the diminishing capacity to provide resources to grow if the
275 cell increases in size (and its S/V decreases), maximum energy production represents a somewhat unachievable
276 limit since the inner membrane must also include other proteins including those required for lipid and membrane
277 synthesis. We used the proteomic data to look at the distribution of proteins on the inner membrane, relying

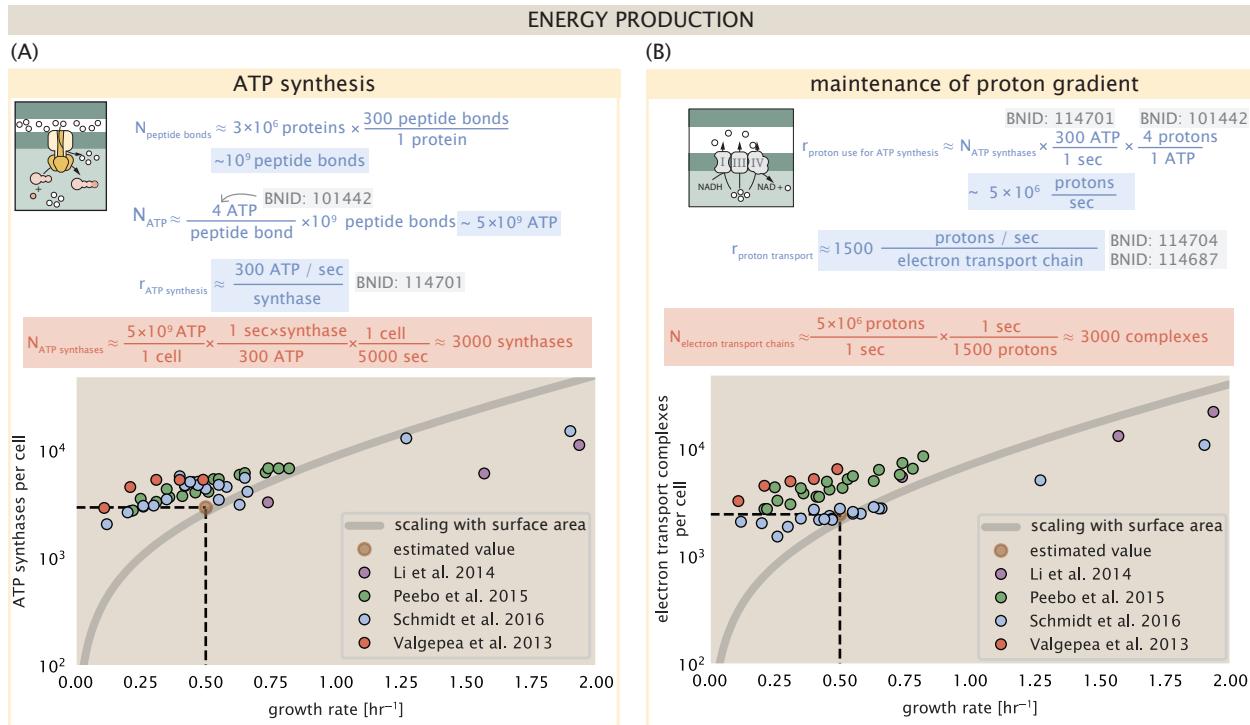


Figure 4. The abundance of F_1 - F_0 ATP synthases and electron transport chain complexes as a function of growth rate.

(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 oxioreducase 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.

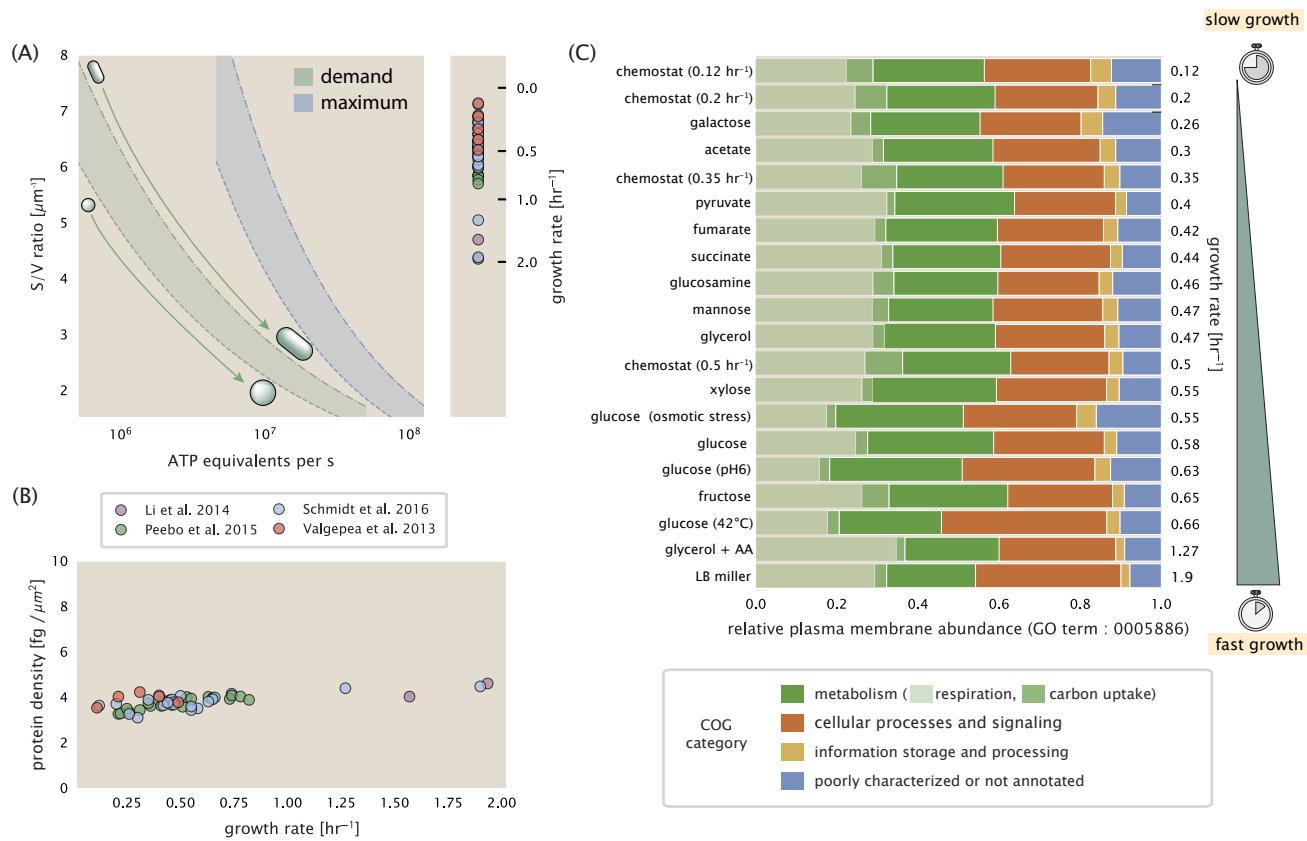


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---)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

on the Gene Ontology (GO) annotations (Ashburner et al., 2000; The Gene Ontology Consortium, 2018) to identify all proteins embedded or peripheral to the inner membrane (GO term: 0005886). Those associated but not membrane-bound include proteins like MreB and FtsZ, that traverse the inner membrane by treadmilling and must nonetheless be considered as a vital component occupying space on the membrane. In Figure 5 (B), we find that the total protein mass per μm^2 is surprisingly constant across growth rates. Interestingly, when we consider the distribution of proteins grouped by their Clusters of Orthologous Groups (COG) (Tatusov et al., 2000), the relative abundance for those in metabolism (including ATP synthesis via respiration) is also relatively constant across growth rates, suggesting that many other membrane associated proteins also increase in similar proportions to proteins devoted to energy production Figure 5 (C).

Function of the Central Dogma

Up to this point, we have considered a variety of transport and biosynthetic processes that are critical to acquiring and generating new cell mass. While there are of course many other metabolic processes we could consider and perform estimates of (such as the components of fermentative versus aerobic respiration), we now turn our focus to some of the most central processes which *must* be undertaken irrespective of the growth conditions – the

CELL ENVELOPE BIOSYNTHESIS

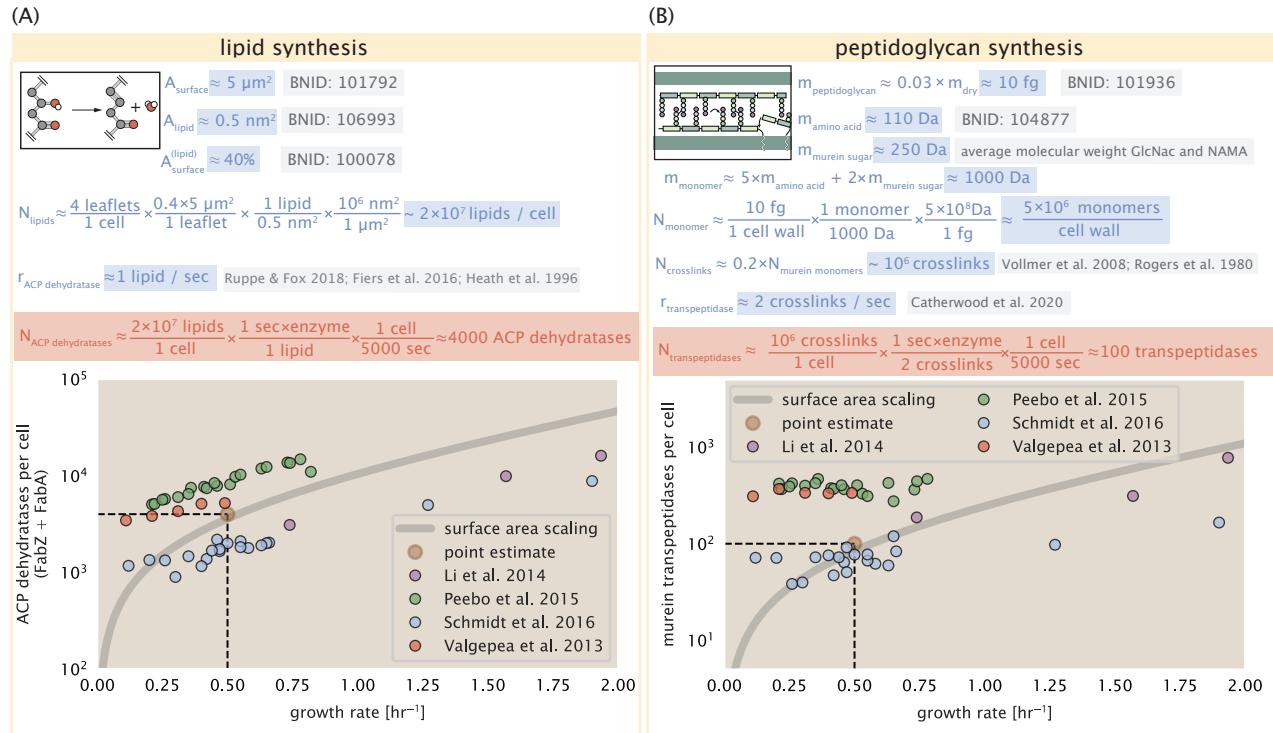


Figure 6. Estimation of the key components involved in cell envelope biosynthesis. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabA}]_2$ and $[\text{FabZ}]_2$. (B) An estimate for the number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan. The mass of the murein monomer was estimated by approximating each amino acid in the pentapeptide chain as having a mass of 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. We direct the reader to the supplemental information for a more detailed discussion of this estimate.

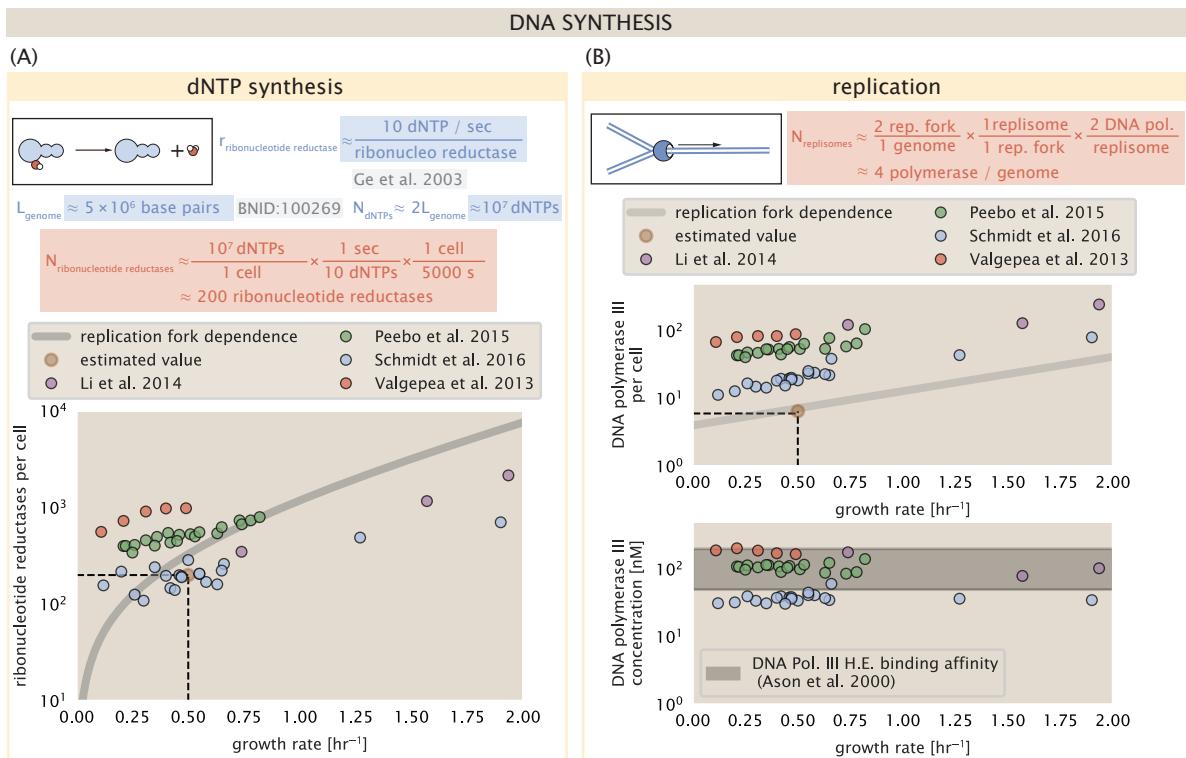


Figure 7. Complex abundance estimates for dNTP synthesis and DNA replication. (A) Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[\text{NrdA}]_2[\text{NrdB}]_2$) and ribonucleotide reductase II ($[\text{NrdE}]_2[\text{NrdF}]_2$) complexes. (B) An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the top plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[\text{DnaE}]_3[\text{DnaQ}]_3[\text{HolE}]_3[\text{DnaX}]_5[\text{HolB}] [\text{HolA}] [\text{DnaN}]_4[\text{HolC}]_4[\text{HolD}]_4$) per cell. Bottom 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 (A) and top panel of (B) show the estimated number of complexes needed as a function of growth, the details of which are described in the Supplemental Information.

292 processes of the central dogma.

293 DNA

294 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids
295 up to ~ 100 kbp in length. We will focus our quantitative thinking solely on the chromosome of *E. coli* which harbors
296 ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs. To successfully divide and produce viable progeny, this chromosome must
297 be faithfully replicated and segregated into each nascent cell. We again rely on the near century of literature
298 in molecular biology to provide some insight on the rates and mechanics of the replicative feat as well as the
299 production of the required starting materials, dNTPs.

300 DNA Replication

301 We now turn our focus to the integration of these dNTP building blocks into the replicated chromosome strand
302 via the DNA polymerase. Replication is initiated at a single region of the chromosome termed the *oriC* locus at
303 which a pair of DNA polymerases bind and begin their high-fidelity replication of the genome in opposite directions.
304 Assuming equivalence between the two replication forks, this means that the two DNA polymerase complexes
305 (termed replisomes) meet at the midway point of the circular chromosome termed the *ter* locus. The kinetics of
306 the five types of DNA polymerases (I – V) have been intensely studied, revealing that DNA polymerase III performs
307 the high fidelity processive replication of the genome with the other "accessory" polymerases playing auxiliary
308 roles (Fijalkowska et al., 2012). *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate
309 of ≈ 600 nucleotides per second (BNID: 104120). Therefore, to replicate a single chromosome, two replisomes

310 (containing two DNA polymerase III each) moving at their maximal rate would copy the entire genome in \approx 4000
311 s. Thus, with a division time of 5000 s (our "typical" growth rate for the purposes of this work), there is sufficient
312 time for a pair of replisomes complexes to replicate the entire genome. However, this estimate implies that 4000
313 s would be the upper-limit time scale for bacterial division which is at odds with the familiar 20 minute (1200 s)
314 doubling time of *E. coli* in rich medium.

315 It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes are being
316 replicated at once, with as many as 10 - 12 replication forks at a given time (*Bremer and Dennis, 2008; Si et al., 2017*). Thus, even in rapidly growing cultures, we expect only a few polymerases (\approx 10) are needed to replicate
317 the chromosome per cell doubling. However, as shown in *Figure 7(B)*, DNA polymerase III is nearly an order of magnitude more abundant. This discrepancy can be understood by considering its binding constant to
318 DNA. DNA polymerase III is highly processive, facilitated by a strong affinity of the complex to the DNA. *In vitro*
319 biochemical characterization has quantified the K_D of DNA polymerase III holoenzyme to single-stranded and
320 double-stranded DNA to be 50 and 200 nM, respectively (*Ason et al., 2000*). The bottom plot in *Figure 7(B)* shows
321 that the concentration of the DNA polymerase III across all data sets and growth conditions is within this range.
322 Thus, while the copy number of the DNA polymerase III is in excess of the strict number required to replicate the
323 genome, its copy number appears to vary such that its concentration is approximately equal to the dissociation
324 constant to the DNA. While the processes regulating the initiation of DNA replication are complex and involve more
325 than just the holoenzyme, these data indicate that the kinetics of replication rather than the explicit copy number
326 of the DNA polymerase III holoenzyme is the more relevant feature of DNA replication to consider. In light of this,
327 the data in *Figure 7(B)* suggests that for bacteria like *E. coli*, DNA replication does not represent a rate-limiting
328 step in cell division. However, it is worth noting that for bacterium like *C. crescentus* whose chromosomal replication
329 is initiated only once per cell cycle (*Jensen et al., 2001*), the time to double their chromosome likely represents an
330 upper limit to their growth rate.

333 RNA Synthesis

334 With the machinery governing the replication of the genome accounted for, we now turn our attention to the next
335 stage of the central dogma – the transcription of DNA to form RNA. We primarily consider three major groupings of
336 RNA, namely the RNA associated with ribosomes (rRNA), the RNA encoding the amino-acid sequence of proteins
337 (mRNA), and the RNA which links codon sequence to amino-acid identity during translation (tRNA). Despite the
338 varied function of these RNA species, they share a commonality in that they are transcribed from DNA via the
339 action of RNA polymerase. In the coming paragraphs, we will consider the synthesis of RNA as a rate limiting step
340 in bacterial division by estimating how many RNA polymerases must be present to synthesize all necessary rRNA,
341 mRNA, and tRNA.

342 rRNA

343 We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA that serve as
344 catalytic and structural elements of the ribosome. Each ribosome contains three rRNA molecules of lengths 120,
345 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides. As the *E. coli*
346 RNA polymerase transcribes DNA to RNA at a rate of \approx 40 nucleotides per second (BNID: 101904), it takes a single
347 RNA polymerase \approx 100 s to synthesize the RNA needed to form a single functional ribosome. Therefore, in a 5000 s
348 division time, a single RNA polymerase transcribing rRNA at a time would result in only \approx 50 functional ribosomal
349 rRNA units – far below the observed number of \approx 10^4 ribosomes per cell.

350 Of course, there can be more than one RNA polymerase transcribing the rRNA genes at any given time. To
351 elucidate the *maximum* number of rRNA units that can be synthesized given a single copy of each rRNA gene, we
352 will consider a hypothesis in which the rRNA operon is completely tiled with RNA polymerase. *In vivo* measurements
353 of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded onto the promoter of an rRNA
354 gene at a rate of \approx 1 per second (BNID: 111997, 102362). If RNA polymerases are being constantly loaded on to the
355 rRNA genes at this rate, then we can assume that \approx 1 functional rRNA unit is synthesized per second. With a 5000
356 second division time, this hypothesis leads to a maximal value of 5000 functional rRNA units, still undershooting
357 the observed number of 10^4 ribosomes per cell.

358 *E. coli*, like many other bacterium, have evolved a clever mechanism to surpass this kinetic limit for the rate

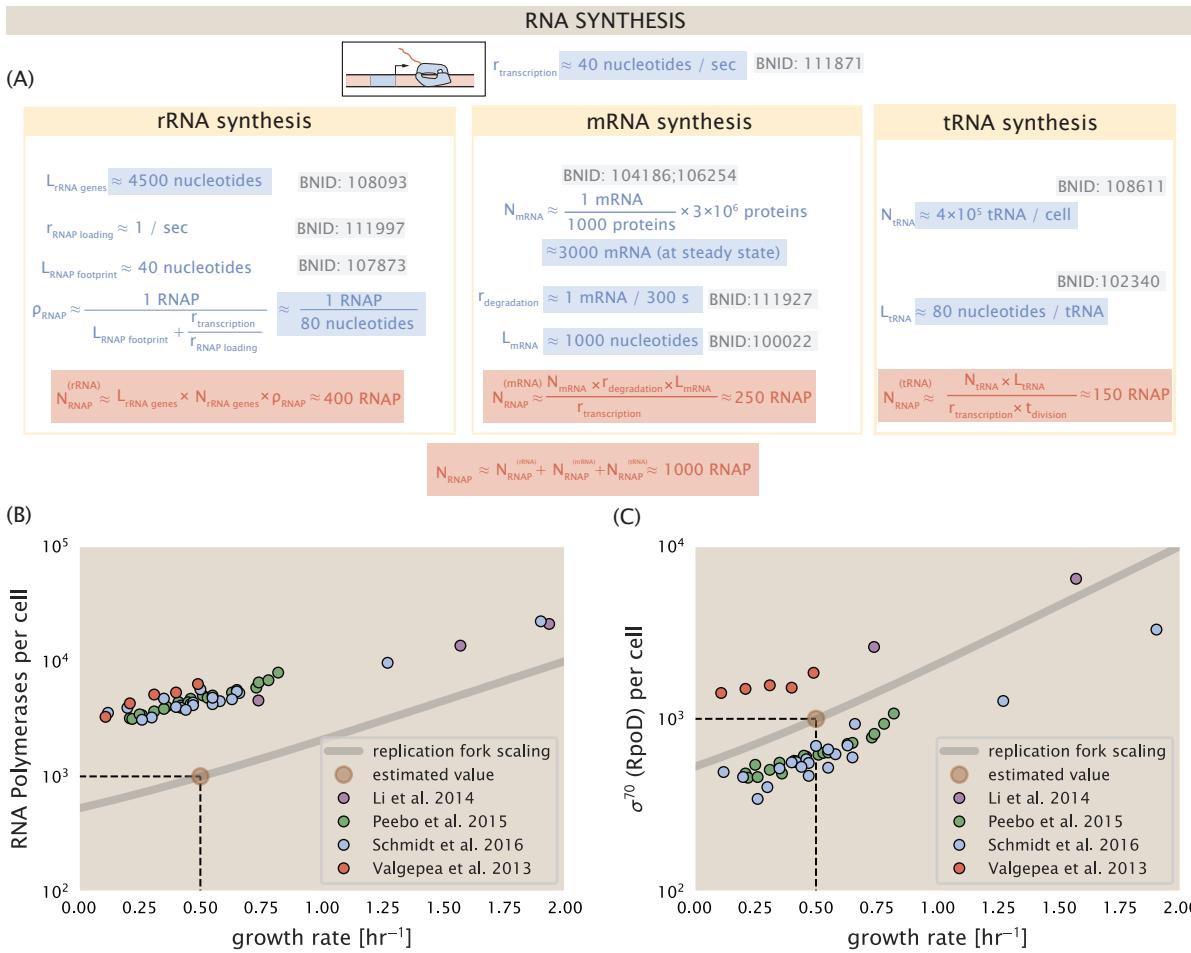


Figure 8. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of growth rate. Colored points correspond to the average number RNA polymerase core enzymes that could be formed given a subunit stoichiometry of $[RpoA]_2[RpoC][RpoB]$. (C) The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown in (B) and (C) as a translucent brown point at a growth rate of 0.5 hr^{-1} .

359 of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli* has seven copies of the operon
 360 (BNID: 100352) four of which are localized directly adjacent to the origin of replication (**Birnbaum and Kaplan,**
 361 **1977**). As fast growth also implies an increased gene dosage due to parallelized chromosomal replication, the total
 362 number of rRNA genes can be on the order of $\approx 10 - 70$ copies at moderate to fast growth rates (**Stevenson and**
 363 **Schmidt, 2004**). Given a 5000 second division time, we can make the lower-bound estimate that the typical cell will
 364 have ≈ 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per rRNA operon, a total of
 365 5×10^4 rRNA units can be synthesized, comfortably above the observed number of ribosomes per cell.

366 How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA genes? We approach
 367 this estimate by considering the maximum number of RNA polymerases tiled along the rRNA genes with a loading
 368 rate of 1 per second and a transcription rate of 40 nucleotides per second. Considering that a RNA polymerase has
 369 a physical footprint of approximately 40 nucleotides (BNID: 107873), we can expect ≈ 1 RNA polymerase per 80
 370 nucleotides. With a total length of ≈ 4500 nucleotides per operon and 7 operons per cell, the maximum number of
 371 RNA polymerases that can be transcribing rRNA at any given time is ≈ 500 . As we will see in the coming sections,
 372 the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

373 **Translation and Ribosomal Synthesis**

374 Lastly, we turn our attention to the process of synthesizing new proteins, translation. This process stands as a
375 good candidate for potentially limiting growth since the synthesis of new proteins relies on the generation of
376 ribosomes, themselves proteinaceous molecules. As we will see in the coming sections of this work, this poses a
377 "chicken-or-the-egg" problem where the synthesis of ribosomes requires ribosomes in the first place.

378 We will begin our exploration of protein translation in the same spirit as we have in previous sections – we
379 will draw order-of-magnitude estimates based on our intuition and available literature, and then compare these
380 estimates to the observed data. In doing so, we will estimate both the absolute number of ribosomes necessary for
381 replication of the proteome as well as the synthesis of amino-acyl tRNAs. From there we consider the limitations
382 on ribosomal synthesis in light of our estimates on both the synthesis of ribosomal proteins and our earlier results
383 on rRNA synthesis.

384 **Protein Synthesis**

385 With the number of tRNA synthetases accounted for, we now consider the abundance of the protein synthesis
386 machines themselves, ribosomes. Ribosomes are enormous protein/rRNA complexes that facilitate the peptide
387 bond formation between amino acids in the correct sequence as defined by the coding mRNA. Before we examine
388 the synthesis of the ribosome proteins and the limits that may place on the observed bacterial growth rates, let's
389 consider replication of the cellular proteome.

390 While the rate at which ribosomes translates is well known to have a growth rate dependence *Dai et al. (2018)*
391 and is a topic which we discuss in detail in the coming sections. However, for the purposes of our order-of-
392 magnitude estimate, we can make the approximation that translation occurs at a rate of ≈ 15 amino acids per
393 second per ribosome (BNID: 100233). Under this approximation and assuming a division time of 5000 s, we can
394 arrive at an estimate of $\approx 10^4$ ribosomes are needed to replicate the cellular proteome, shown in *Figure 9(B)*.
395 This point estimate, while glossing over important details such as chromosome copy number and growth-rate
396 dependent translation rates, proves to be notably accurate when compared to the experimental observations
397 (*Figure 9(B)*).

398 **Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step**

399 Thus far, the general back-of-the-envelope estimates have been reasonably successful in predicting the scale of
400 absolute protein copy number as well as their observed dependence on the cellular growth rate. A recurring
401 theme across these varied biological processes is the ability of cells to parallelize tasks through the expression of
402 additional proteins. Even when that is not possible, like in chromosomal replication which requires a minimum
403 of ≈ 40 minutes, *E. coli* and many other bacteria surpass this limit by initiating additional rounds of replication
404 per doubling. However, the synthesis of ribosomal proteins presents a special case where parallelization is not
405 possible and must be doubled in quantity on average with every cell division (*Figure 10(A)*).

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

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

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

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

416 Alternatively, since N_{pep} is related to the total protein mass through the molecular weight of each protein, we
417 can also consider the growth rate in terms of the fraction of the total proteome mass dedicated to ribosomal
418 proteins. By making the approximation that an average amino acid has a molecular weight of 110 Da (BNID:

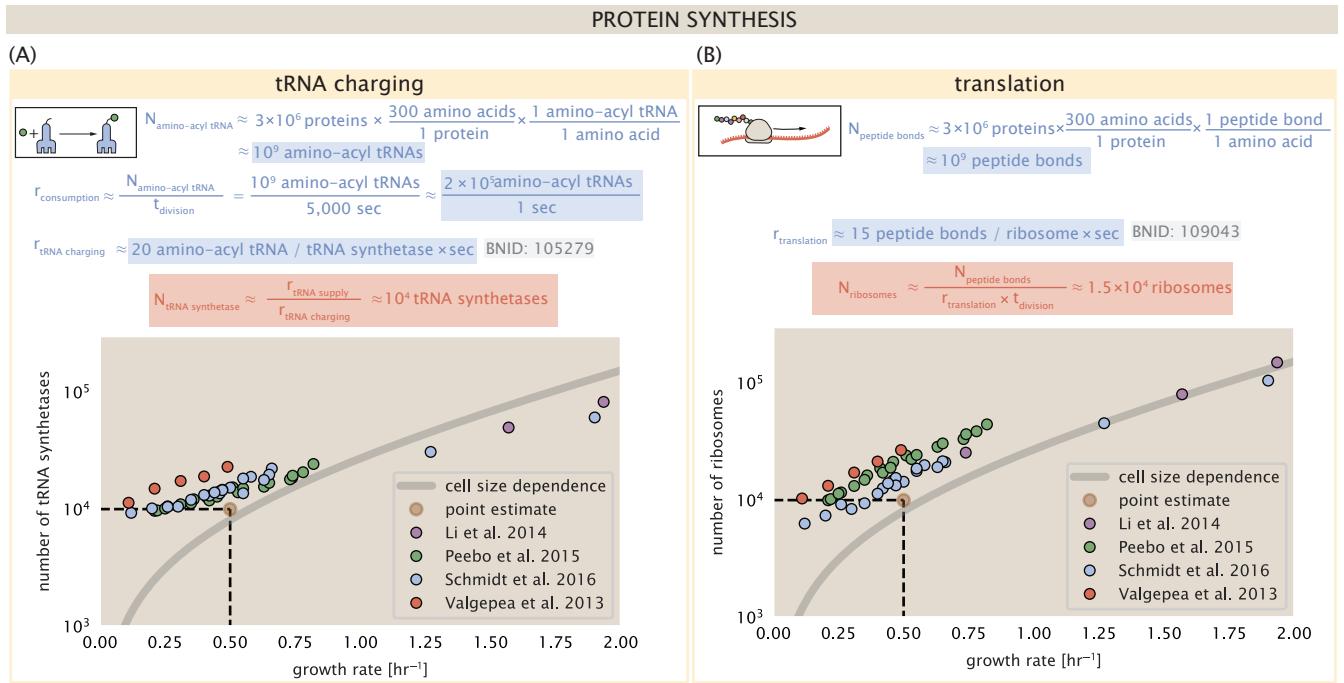


Figure 9. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [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). (B) Estimation of the number of ribosomes required to synthesize 10⁹ 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 for a more detail description of this calculation.

419 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.
 420 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
 421 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
 422 ribosomal mass fraction m_{protein}/m_R , and L_R the ratio of 800 kDa / 110 Da per amino acid or, alternatively, the total
 423 length in amino acids that make up a ribosome. The translation-limited growth rate can then be written in the
 424 form

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

425 This is plotted as a function of ribosomal fraction Φ_R in *Figure 10(B)*, where we take $L_R = 7459 \text{ AA}$, corresponding
 426 to the length in amino acids for all ribosomal subunits of the 50S and 30S complex (BNID: 101175), and $f_a = 1$.
 427 In *Figure 10(C)* we use the recent measurements of f_a from *Dai et al. (2016)* to estimate the active fraction of
 428 ribosomal protein across the proteomic data sets and number of other recent measurements. We see that cells
 429 are consistently skirting the limit in growth rate set by *Equation 3* as nutrient conditions vary.

430 The growth rate defined by *Equation 3* reflects mass-balance under steady-state growth and has long provided
 431 a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate (*Maaløe,*
 432 *1979; Scott et al., 2010*). The maximum rate, when $\Phi_R = 1$, could only be achieved if a cell contained only
 433 ribosomes. This corresponds to the synthesis time of all ribosomal subunits, $L_R/r_t \approx 7 \text{ minutes}$ (*Dill et al., 2011*)
 434 and interestingly, is independent of the absolute number of ribosomes. This is because, in order to double the
 435 cell's ribosomal mass, each ribosome must produce a second ribosome; a process which cannot be parallelized.
 436 Unless elongation rate increased, or cells could trim their total ribosomal protein mass, this dependency limits both
 437 the maximum growth rate (when $\Phi_R = 1$), and also the achievable growth rate under more realistic values of Φ_R .

438 This strain of *E. coli* rarely exhibits growth rates above 2 hr^{-1} (*Bremer and Dennis, 2008; Roller et al., 2016*)
 439 and in *Figure 10(C)* we consider ribosomal generation from the perspective of rRNA synthesis. Here we use our
 440 rule-of-thumb of 1 functional rRNA unit per second per operon and estimate the maximum number of ribosomes
 441 that could be made as a function of growth rate (blue curve). Although we expect this estimate to drastically
 442 overestimate rRNA abundance at slower growth rates ($\lambda < 0.5 \text{ hr}^{-1}$), it provides a useful reference alongside the
 443 proteomic measurements. For growth rates above about 1 hr^{-1} , we find that cells will need to transcribe rRNA
 444 near their maximal rate. As a counter example, if *E. coli* did not initiate multiple rounds of replication, they would
 445 be unable to make enough rRNA for the observed number of ribosomes (dashed blue curve in *Figure 10(C)*).
 446 The convergence between the maximum rRNA production and measured ribosome copy number suggests rRNA
 447 synthesis may begin to present a bottleneck at the fastest growth rates due to its limited copies of rRNA genes.

448 Relationship Between Cell Size and Growth Rate

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

455 Cells grow at a near-maximal rate dictated by their total ribosomal mass fraction Φ_R , at least at moderate growth
 456 rates above 0.5 hr^{-1} (where f_a is close to 1, and r_t is near its maximal rate). Here, growth rate can be increased
 457 only by increasing Φ_R , though the simple addition of more ribosomes is likely constrained by aspects physical
 458 constrains like macromolecular crowding (*Delarue et al., 2018; Soler-Bistué et al., 2020*). As *E. coli* grows faster,
 459 large swaths of its proteome increase in absolute abundance. It is now well-documented that *E. coli* cells add a
 460 constant volume per origin of replication (termed a "unit cell" or "initiation mass"), which is robust to a remarkable
 461 array of cellular perturbations (*Si et al., 2017*). To consider this dependency in the context of the proteomic data,
 462 we used measurements from *Si et al. (2017)* (*Figure 11(A)*) to estimate the average number of origins per cell $\langle \# \text{ ori} \rangle$ at different growth rates.

463 The average number of origins $\langle \# \text{ ori} \rangle$ is set by how often replication must be initiated per cell doubling under
 464 steady-state growth. This can be quantified as

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

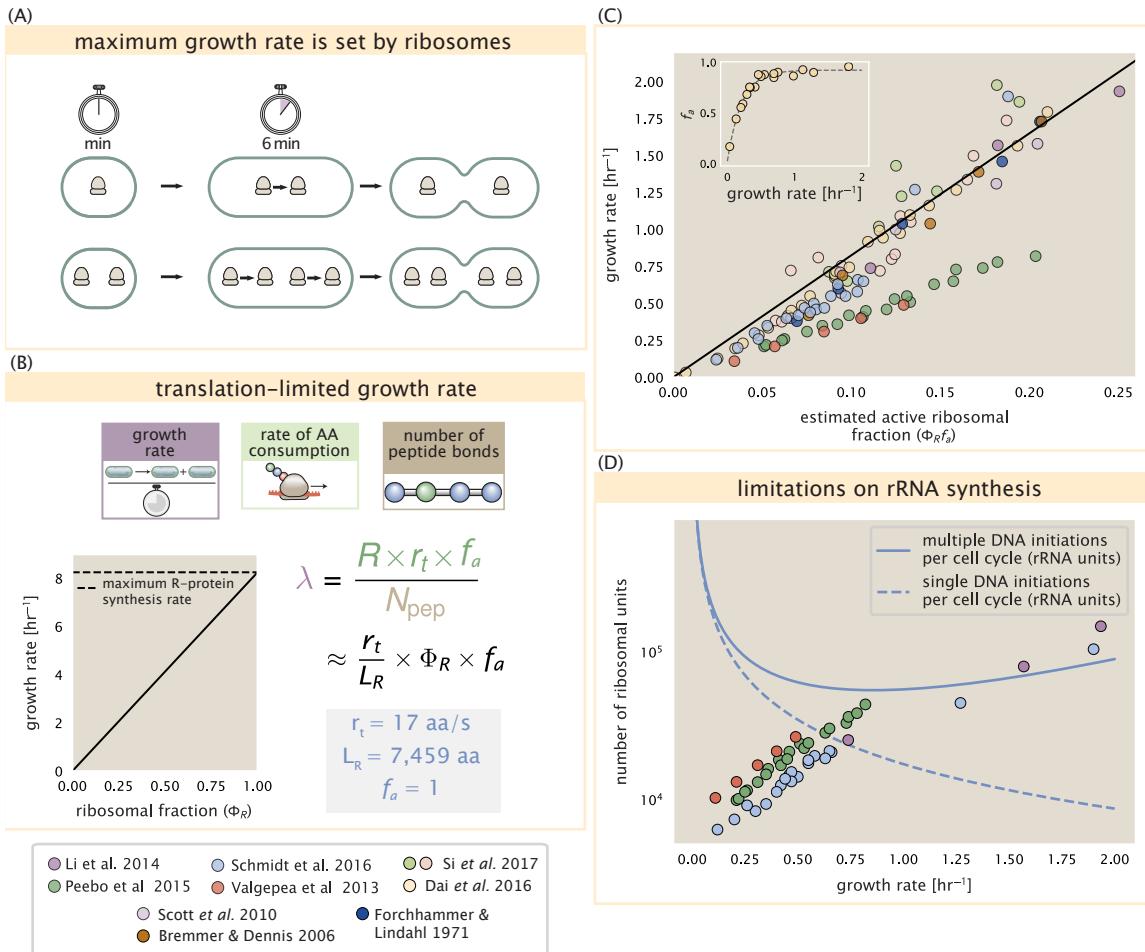


Figure 10. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate 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 11(A). The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

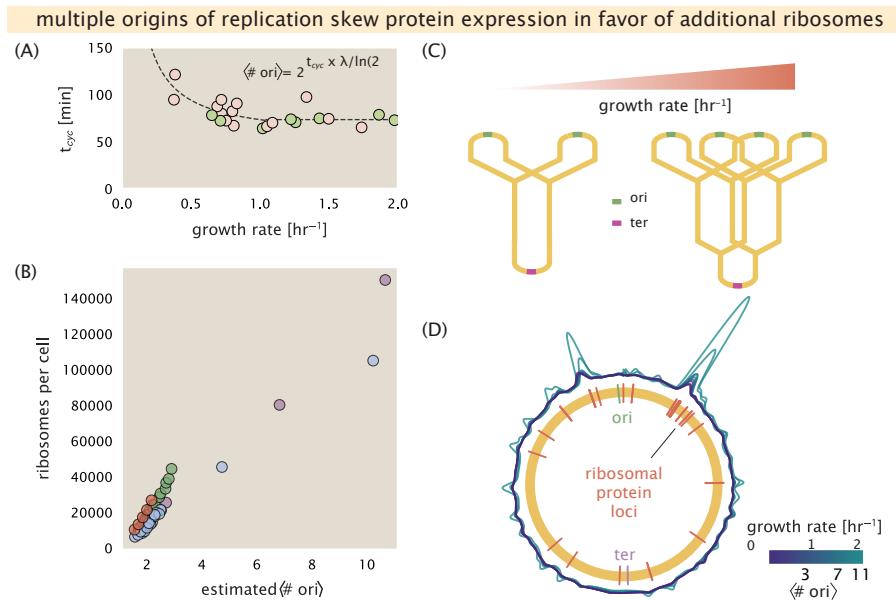


Figure 11. Cells increase absolute ribosome abundance with $\langle \# \text{ori} \rangle$. (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times great than 40 minutes, and then reach a minimum value of 73 minutes below this (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.

466 where t_{cyc} is the cell cycle time (referring to the time from replication initiation to cell division), and τ is the cell
 467 doubling time. For ribosomal synthesis, we find an approximately linear correlation between ribosome copy
 468 number and $\langle \# \text{ori} \rangle$ (Figure 11(B)).

469 For a constant cell cycle time, observed at growth rates above about 0.5 hr^{-1} (Helmstetter and Cooper, 1968),
 470 Equation 4 states that $\langle \# \text{ori} \rangle$ will need to increase exponentially with the growth rate. While this says nothing
 471 of the observed scaling with cell size, the additional dependency on ribosomal content, which increases with $\langle \#$
 472 $\text{ori} \rangle$, provides an additional link. In Figure 11(D), we consider the position-dependent protein expression across
 473 the chromosome by calculating a running Gaussian average of protein copy number (20 kbp st. dev. averaging
 474 window) based on each gene's transcriptional start site, which were then median-subtracted to account for the
 475 differences in total protein abundance. Importantly, major deviations in protein copy number are largely restricted
 476 to regions of ribosomal protein genes. This suggests that the relative ribosomal abundance Φ_R is also being tuned
 477 in proportion to $\langle \# \text{ori} \rangle$, with the exponential relationship between cell size and growth rate following from how *E.*
 478 *coli* varies its number of ribosomes per cell.

479 Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis

480 As we have seen, cell size, total proteomic content, and the number of ribosomes are all interconnected and
 481 influence the achievable growth rate. The drastic change in these parameters across different growth conditions
 482 suggests that they are being tuned to better match the cell's biosynthetic capacity to the specific environment.
 483 Take, as another illustration of this, the recent experimental work by Dai *et al.* (2016). In one set of experiments the
 484 authors considered growth in cells whose primary glucose transport system was disrupted ($\Delta ptsG$). Unsurprisingly,
 485 the growth rate was reduced, and was measured at about two-fold slower than their wild-type line. This change,
 486 however, was not simply the result of now-limiting carbon uptake. Instead, cells accommodated the perturbation
 487 by also reducing their ribosomal mass fraction by a factor of two, which is still in line with Equation 3 under
 488 translation-limited growth. In this final, we explore the interconnection between cell size, ribosome content, and

489 growth rate by formulating a minimal model of growth rate control. We use it to quantitatively show how tuning
 490 these parameters help cells maximize their growth rate.

491 To react to changes in nutrient conditions, bacteria rely on the synthesis or degradation of secondary-messenger
 492 molecules like (p)ppGpp, which cause global changes in transcriptional and translational activity. In *E. coli*, amino
 493 acid starvation causes the accumulation of de-acylated tRNAs at the ribosome's A-site and leads to a strong increase
 494 in (p)ppGpp synthesis activity by the enzyme RelA (*Hauryliuk et al., 2015*). Cells also accumulate (p)ppGpp during
 495 steady-state growth in poorer growth conditions, which leads to a decrease in the fraction of actively translating
 496 ribosomes, f_a (with $f_a \approx 0.5$ at a growth rate of $\approx 0.3 \text{ hr}^{-1}$; *Figure 10(C)* - inset).

497 Furthermore, (p)ppGpp can inhibit the initiation of DNA replication by mediating a change in transcriptional
 498 activity and the supercoiling state of the origin of replication (*Kraemer et al., 2019*). These observations all raise
 499 the possibility that it is through (p)ppGpp that cells mediate the growth-rate dependent changes in $\langle \# \text{ ori} \rangle$, cell size,
 500 and ribosomal abundance and activity (*Zhu and Dai, 2019; Büke et al., 2020*). Indeed, recent work in a (p)ppGpp
 501 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
 502 consistent with a fast growth state where (p)ppGpp levels are normally low (*Fernández-Coll et al., 2020*).

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

504 Here we consider a mode of regulation in which the rate of peptide elongation r_t depends only on the availability
 505 of amino acids (and, therefore, also amino-acyl tRNAs). It is through the elongation rate r_t that we assume cells
 506 adjust their ribosomal content (R, Φ_R) according to nutrient availability and for simplicity, do not explicitly model
 507 changes in $\langle \# \text{ ori} \rangle$ or regulation by (p)ppGpp.

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

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

515 where we have assumed that the rate of binding by amino-acyl tRNA k_{on} is proportional to $[AA]_{\text{eff}}$ by a constant α .
 516 The second term on the right-hand side reflects our assumption that other steps in peptide elongation are not
 517 rate-limiting, with a maximum elongation rate r_t^{\max} of about 17 amino acids per second (*Dai et al. (2016)*). As the
 518 rate of amino acid supply, denote by r_{AA} , varies with changing nutrient conditions, the cell can maximize the rate of
 519 protein synthesis by tuning the rate of amino acid consumption (mathematized as $r_t \times R \times f_a$), shown schematically
 520 in *Figure 12(A)*. This can be stated more succinctly in terms of an effective dissociation constant, $K_D = r_t^{\max}/\alpha k_{on}$,
 521 where the elongation rate r_t is now given by

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

522 Under steady-state growth, the amino acid concentration is constant ($\frac{d[AA]_{\text{eff}}}{dt} = 0$) and will relate to the rate
 523 of amino acid synthesis (or import, for rich media) and/or tRNA charging, as r_{AA} , and the rate of consumption,
 524 $r_t \times R \times f_a$ by,

$$\int_0^t \frac{d[AA]_{\text{eff}}}{dt} dt = \int_0^t ([r_{AA}] - [r_t \times R \times f_a]) dt, \quad (7)$$

525 where the time from 0 to t is an arbitrary length of time, and the square brackets indicate concentrations per unit
 526 time. Integrating *Equation 7* yields,

$$[AA]_{\text{eff}} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V}, \quad (8)$$

(A)

A MINIMAL MODEL FOR NUTRIENT-LIMITED GROWTH

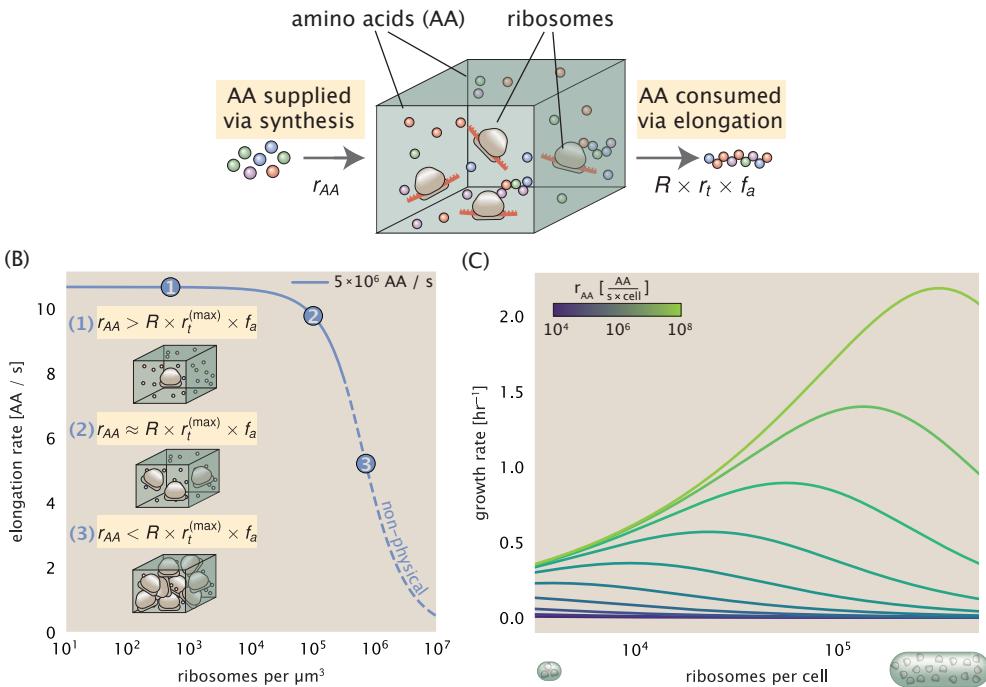


Figure 12. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, 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 $r_{AA} = 5 \times 10^6 \text{ AA / s}$, $K_D = 5 \text{ mM}$, and $r_t^{(\max)} = 17.1 \text{ AA / s}$. (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 relationship between cell volume, number of peptide bonds, and ribosome copy number.

527 where we instead consider r_{AA} is in units of AA per unit time and r_t is in units of AA per unit time per ribosome, for
 528 a cell with average volume V . Plugging [Equation 8](#) into [Equation 6](#) allows us to then solve for r_t and a complete
 529 derivation is provided in Appendix ??.

530 In [Figure 12\(B\)](#), we illustrate how the elongation rate depends on the ribosomal copy number. Here, we have
 531 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. \(2009\)](#)),
 532 $f_a = 1$, 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
 533 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 12\(B\)](#)].
 534 As the ribosome copy number is increased such that the amino acid supply rate and consumption rate are nearly
 535 equal [point (2) in [Figure 12\(B\)](#)], the observed elongation rate begins to decrease sharply. When the ribosome
 536 copy number is increased even further, consumption at the maximum elongation rate exceeds the supply rate,
 537 yielding a significantly reduced elongation rate [point (3) in [Figure 12B](#)]. While the elongation rate will always be
 538 dominated by the amino acid supply rate at sufficiently low ribosome copy numbers, the elongation rate at larger
 539 ribosome abundances can be increased by tuning f_a such that not all ribosomes are elongating, reducing the total
 540 consumption rate.

541 Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic
542 Capacity.

543 To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic
544 measurements; namely, we restrict the values of R , N_{pep} , and V to those associated with the amalgamated
545 proteomic data (described in Appendix Estimation of Total Protein Content per Cell). We then consider how
546 changes in the nutrient conditions, through the parameter r_{AA} , influence the maximum growth rate as determined
547 by **Equation 2**. **Figure 12(C)** shows how the observed growth rate depends on the rate of amino acid supply r_{AA} as
548 a function of the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal
549 growth rate whose dependence on R (and consequently, the cell size) increases with increasing r_{AA} . Importantly,
550 however, there is an optimum set of R , N_{pep} , and V that are strictly dependent on the value of r_{AA} . Increasing the
551 ribosomal concentration beyond the cell's metabolic capacity has the adverse consequence of depleting the supply
552 of amino acids and a concomitant decrease in the elongation rate r , [**Figure 12(B)**].

553 Also of note is the growth rate profiles shown for low amino acid supply rates [purple and blue lines in
554 **Figure 12(C)**], representing growth in nutrient-poor media. In these conditions, there no longer exists a peak in
555 growth, at least in the range of physiologically-relevant ribosome copy numbers. Instead, cells limit their pool of
556 actively translating ribosomes by decreasing f_a ([Dai et al., 2016](#)), which would help maintain the pool of available
557 amino acids $[AA]_{eff}$ and increase the achievable elongation rate. This observation is in agreement with the central
558 premise of the cellular resource allocation principle proposed by [Scott et al. \(2010\)](#); [Klumpp et al. \(2009\)](#); [Klumpp](#)
559 and [Hwa \(2014\)](#) and [Hui et al. \(2015\)](#).

560 Discussion

561 Continued experimental and technological improvements have led to a treasure trove of quantitative biological
562 data ([Hui et al., 2015](#); [Schmidt et al., 2016](#); [Si et al., 2017](#); [Gallagher et al., 2020](#); [Peebo et al., 2015](#); [Valgepea et al.,
563 2013](#)), and an ever advancing molecular view and mechanistic understanding of the constituents that support
564 bacterial growth ([Taheri-Araghi et al., 2015](#); [Morgenstein et al., 2015](#); [Si et al., 2019](#); [Karr et al., 2012](#); [Kostinski
565 and Reuveni, 2020](#)). In this work we have compiled what we believe to be the state-of-the-art knowledge on
566 proteomic copy number across a broad range of growth conditions in *E. coli*. We have made this data accessible
567 through a [GitHub repository](#), and an interactive figure that allows exploration of specific protein and protein
568 complex copy numbers. Through a series of order-of-magnitude estimates that traverse key steps in the bacterial
569 cell cycle, this proteomic data has been a resource to guide our understanding of two key questions: what biological
570 processes limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents
571 as a function of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates
572 provide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the
573 transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous
574 steps of the central dogma.

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

587 Of the processes illustrated in **Figure 1**, we arrive at a ribosome-centric view of cellular growth rate control.
588 This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary
589 their ribosomal abundance as a function of growth conditions and growth rate [Scott et al. \(2010\)](#); [Metzl-Raz et al.](#)

590 (2017). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates
591 to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize
592 many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA
593 replication, this is not so for synthesis of ribosomes (Figure 10(A)). The translation time for each ribosome [\approx 6 min,
594 Dill et al. (2011)] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase
595 their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The
596 second point relates to the long-observed correlations between growth rate and cell size (Schaechter et al., 1958; Si
597 et al., 2017), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous
598 curiosity and driven substantial amounts of research in their own regards, these relationships are themselves
599 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions
600 of rapid growth that require cells to also grow in size. Further experiments are needed to test the validity of this
601 hypothesis. In particular, we believe that the change in growth rate in response to translation-inhibitory drugs
602 (such as chloramphenicol) could be quantitatively predicted, given one had precision measurement of the relevant
603 parameters, including the fraction of actively translating ribosomes f_a and changes in the metabolic capacity of the
604 cell (i.e. the parameter r_{AA} in our minimal model) for a particular growth condition.

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

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

633 Methods

634 Data Analysis and Availability

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

639 exploration of specific protein and protein complex copy numbers is available at [link].

640 **Acknowledgements**

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

646 **Competing Interests**

647 The authors declare no competing interests.

648 Appendix for: Fundamental limits on the 649 rate of bacterial cell division

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

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

661	Contents	
662	Introduction	1
663	Uptake of Nutrients	4
664	Carbon Transport	4
665	Limits on Transporter Expression	6
666	Energy Production	7
667	ATP Synthesis	7
668	Generating the Proton Electrochemical Gradient	7
669	Energy Production in a Crowded Membrane.	8
670	Function of the Central Dogma	10
671	DNA	12
672	DNA Replication	12
673	RNA Synthesis	13
674	rRNA	13
675	Translation and Ribosomal Synthesis	15
676	Protein Synthesis	15
677	Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step	15
678	Relationship Between Cell Size and Growth Rate	17
679	Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis	19
680	Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity	20
681	Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic Capacity.	22
683	Discussion	22
684	Methods	23
685	Data Analysis and Availability	23
686	Acknowledgements	24
687	Competing Interests	24
688	Experimental Details Behind Proteomic Data	28
689	Fluorescence based measurements	28
690	Ribosomal profiling measurements	28
691	Mass spectrometry measurements	29
692	Summary of Proteomic Data	29
693	Estimation of Cell Size and Surface Area	30
694	Estimation of Total Protein Content per Cell	32
695	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	32
696	Additional Considerations of Schmidt <i>et al.</i> Data Set	32
697	Effect of cell volume on reported absolute protein abundances	34
698	Relaxing assumption of constant protein concentration across growth conditions	36
699	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	36
700	Calculation of Complex Abundance	37

701	Extending Estimates to a Continuum of Growth Rates	39
702	Estimation of the total cell mass	39
703	Complex Abundance Scaling With Cell Volume	39
704	A Relation for Complex Abundance Scaling With Surface Area	40
705	Number of Lipids	40
706	Number of Murein Monomers	40
707	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	41
708	Calculation of active ribosomal fraction.	41
709	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	41

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).

710 Experimental Details Behind Proteomic Data

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

717 Fluorescence based measurements

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

725 Ribosomal profiling measurements

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

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

740 **Mass spectrometry measurements**

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

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

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

769 **Summary of Proteomic Data**

770 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)*;
771 *Schmidt et al. (2016)*. As shown in *Figure 13(A)*, the reported total protein abundances in the work of *Taniguchi*
772 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi et al.*
773 (*2010*) this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that most
774 proteins appear undercounted when compared to the other data.

775 *Figure 13(B)* summarizes the total protein mass for each data point in our final compiled data set. We note that
776 protein abundances were all scaled so they followed a common growth rate-dependent change in total protein
777 mass. While our inclination initially was to leave reported copy numbers untouched, a notable discrepancy in the
778 scaling total protein per cell between *Schmidt et al. (2016)* and the other data sets forced us to dig deeper into
779 those measurements (compare *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 13(A)*). The particular trend
780 in *Schmidt et al. (2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion
781 and analysis of that data set in section Additional Considerations of Schmidt et al. Data Set. As a compromise, and
782 in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of
783 total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total
784 protein content through expected dry mass. Total protein per cell was estimated using available data on total
785 DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the
786 cell. We consider these details in sections Estimation of Cell Size and Surface Area and Estimation of Total Protein
787 Content per Cell that follows.

788 Lastly, in *Figure 14* we show the total proteomic coverage and overlap of proteins quantified across each data

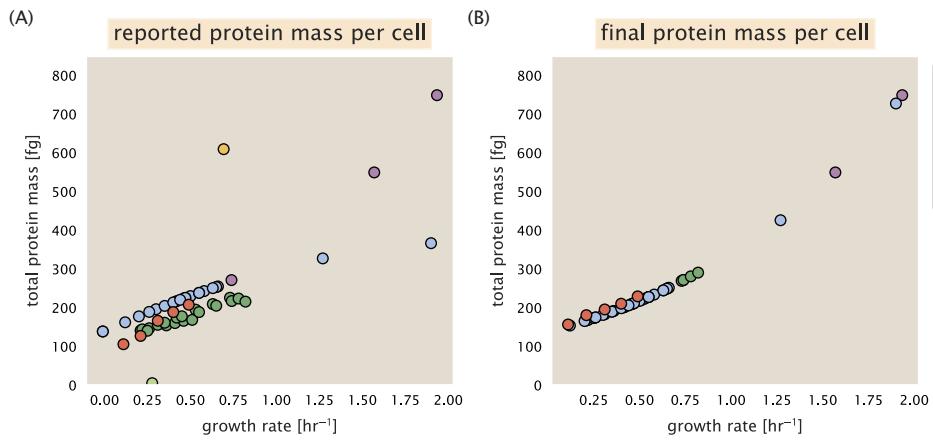


Figure 13. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (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).

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

797 Estimation of Cell Size and Surface Area

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

803 The average size measurements from each of their experiments are shown in Figure **Figure 15**, with cell length
804 and width shown in (A) and (B), respectively. The length data was well described by the exponential function 0.5
805 $e^{1.09 \cdot \lambda} + 1.76$ 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
806 cell as a cylinders with two hemispherical ends (*Si et al., 2017*; *Basan et al., 2015*). Specifically, cell size is estimated
807 from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (9)$$

808 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
809 area is given by,

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (10)$$

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

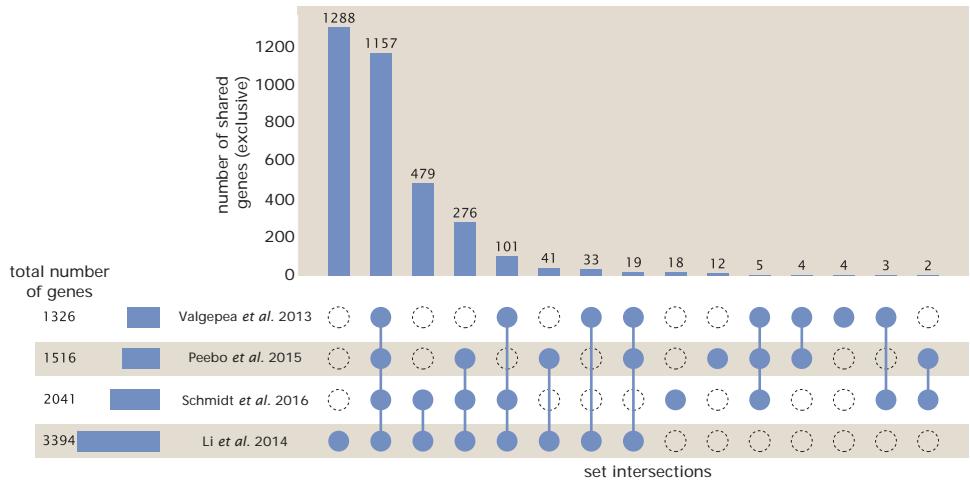


Figure 14. Comparison of proteomic coverage across different data sets. An UpSet diagram (*Lex et al., 2014*) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of *Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015)*. 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.

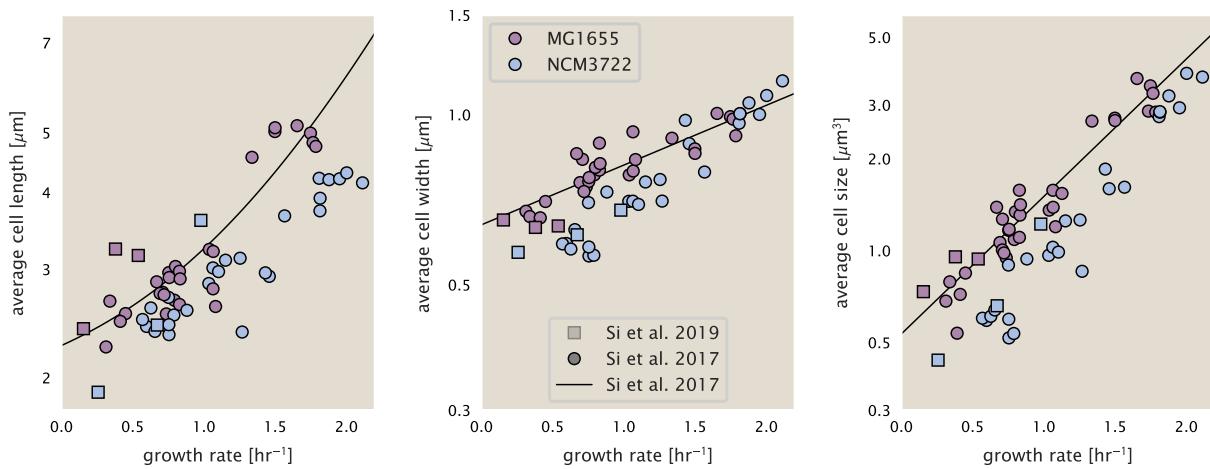


Figure 15. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 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 9). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot i} \mu\text{m}^3$.

811 Estimation of Total Protein Content per Cell

812 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the
813 fit shown in Figure **Figure 15(C)** ($0.533 e^{1.037 \lambda} \mu\text{m}^3$). We then estimate the total protein content from the total dry
814 mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here, protein,
815 DNA, and RNA were reported to account for at least 90 % of the dry mass (*Basan et al. (2015)*). The authors also
816 found that the total dry mass concentration was roughly constant across growth conditions. Under such a scenario,
817 we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30 \% \times 90$
818 % or about $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

819 However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how
820 protein concentration might vary due to the substantial increase in rRNA at faster growth rates (*Dai et al. (2016)*).
821 This is a well-documented result that arises from an increase in ribosomal abundance at faster growth rates (*Scott*
822 *et al. (2010)*). To proceed therefore rely on experimental measurements of total DNA content per cell that also
823 come from Basan *et al.*, and RNA to protein ratios that were measured in Dai *et al.* (and cover the entire range of
824 growth conditions considered here). These are reproduced in Figure **Figure 16(A)** and (B), respectively.

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

$$828 M_p = \frac{[M_p + M_{RNA}]}{1 + (RP_{ratio})}. \quad (11)$$

829 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai *et al.*. In Figure **Figure 16(C)** we plot the estimated
830 cellular concentrations for protein, DNA, and RNA from these calculations, and in Figure **Figure 16(D)** we plot their
831 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 13(B)**).

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

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

841 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log
842 transform of these parameters as a function of the log transform of the ribosome copy number. Using optimization
843 by minimization, we estimated the best-fit values of the intercept and slope for each regression. **Figure 17** shows
844 the result of each regression as a dashed line.

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

846 While the data set from *Schmidt et al. (2016)* remains a heroic effort that our labs continue to return to as a resource,
847 there were steps taken in their calculation of protein copy number that we felt needed further consideration. In
848 particular, the authors made an assumption of constant cellular protein concentration across all growth conditions
849 and used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size
850 with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al.*
851 *(2017)*).

852 We begin by looking at their cell volume measurements, which are shown in blue in **Figure 18**. As a
853 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*
854 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence
855 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and
856 then calculated cell size by treating the cell is a cylinder with two hemispherical ends, as we considered in the

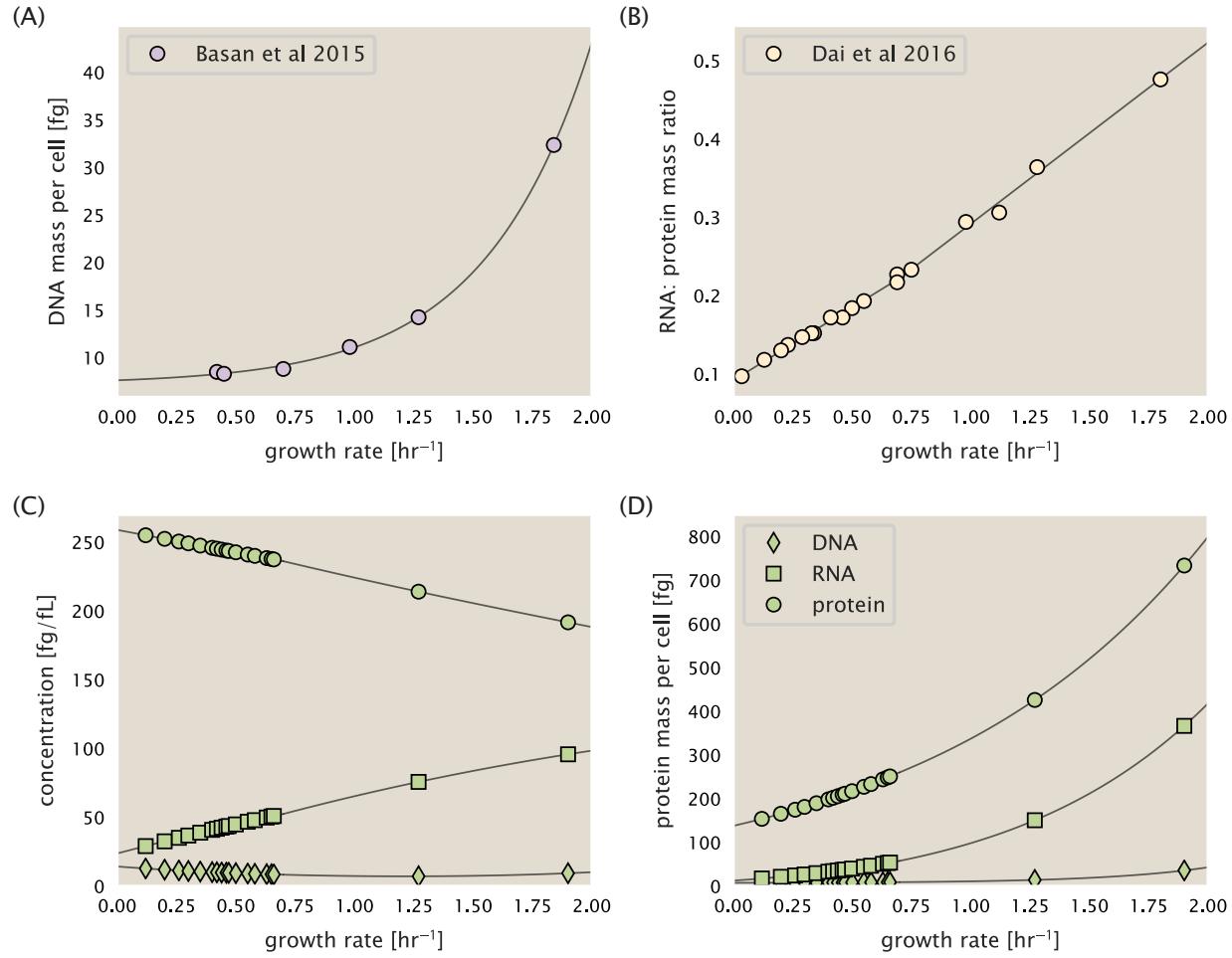


Figure 16. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ 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).

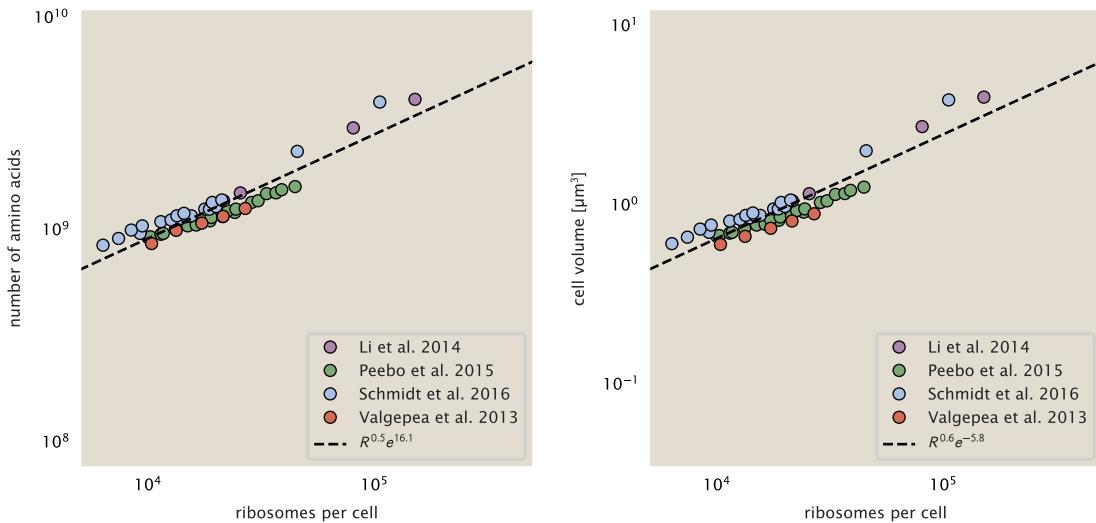


Figure 17. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows the result of the fit, with the functional form of the equation given in the legend with R representing the ribosome copy number.

857 previous section. While there is notable discrepancy between the two research groups, which are both using strain
 858 NCM3722, Basan *et al.* found that this came specifically from uncertainty in determining the cell width. This is prone
 859 to inaccuracy given the small cell size and optical resolution limits (further described in their supplemental text).
 860 Perhaps the more concerning point is that while each of these alternative measurements show an exponential
 861 increase in cell size at faster growth rates, the measurements used by Schmidt *et al.* appear to plateau. This
 862 resulted in an analogous trend in their final reported total cellular protein per cell as shown in Figure **Figure 19**
 863 (purple data points), and is in disagreement with other measurements of total protein at these growth rates (**Basan**
 864 *et al.*, 2015).

865 Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the
 866 following subsections we begin by considering this calculation. We then consider three different approaches to
 867 estimate the growth-rate dependent total protein mass to compare with those values reported from **Schmidt *et al.***
 868 (2016). The results of this are summarized in **Figure 18(B)**, with the original values from both **Schmidt *et al.* (2016)**
 869 and **Li *et al.* (2014)** shown in **Figure 18(A)** for reference. For most growth conditions, we find that total protein per
 870 cell is still in reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
 871 amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main
 872 reason why we chose to readjusted protein abundance as shown in **Figure 13(B)** (with the calculation described in
 873 section Estimation of Total Protein Content per Cell).

874 Effect of cell volume on reported absolute protein abundances

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

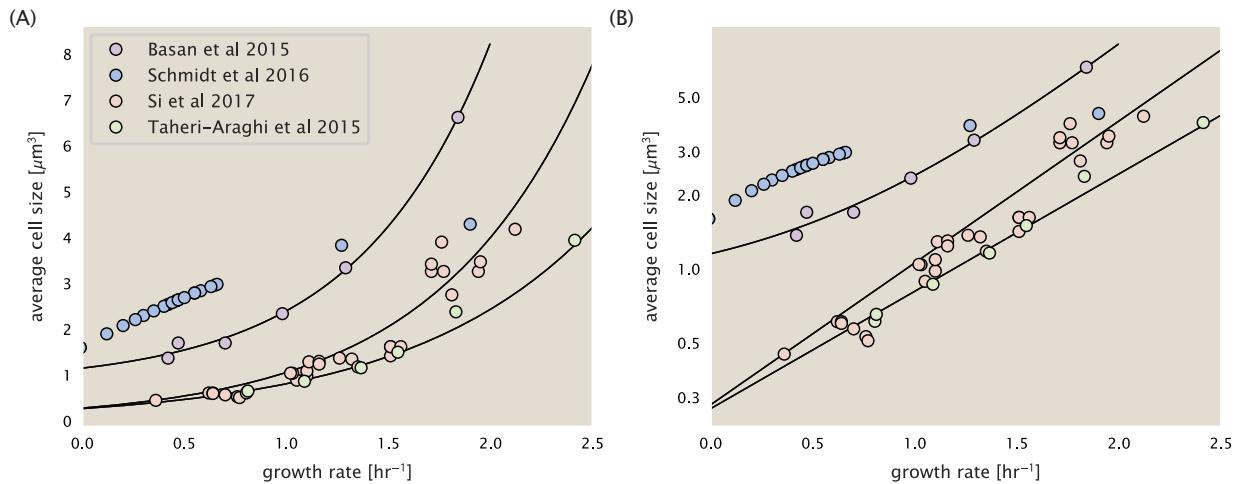


Figure 18. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are 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*.

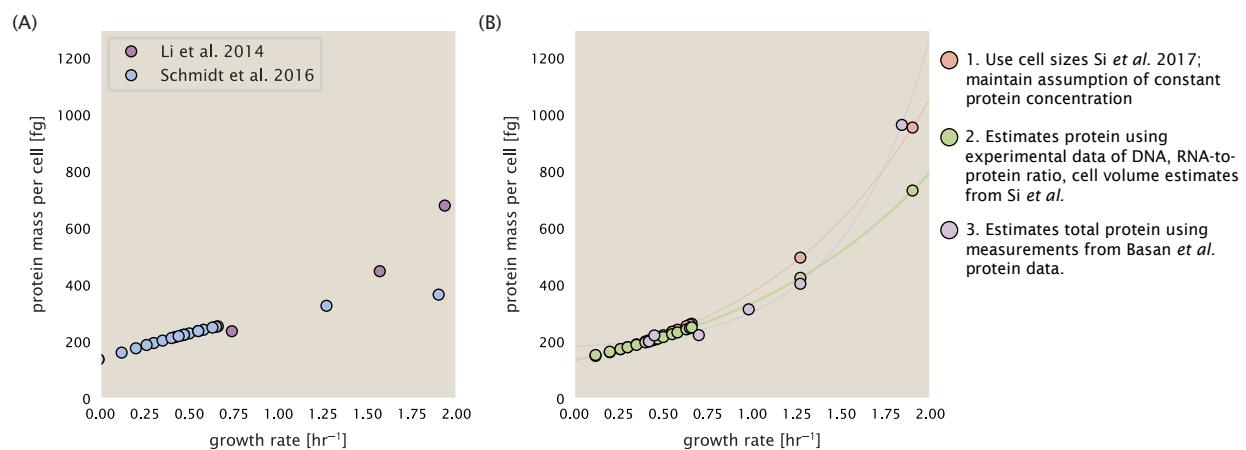


Figure 19. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt et al. (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: Rescaling 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.

885 on a per cell basis.

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

893 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per cell was measured.
894 Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentration of $[M_p]_{orig} = M_p/V_{orig} =$
895 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,
896 they could then estimate the total protein mass for all other growth conditions from,

$$M_{P_i} = [M_p]_{orig} \cdot V_i \quad (12)$$

897 where M_{P_i} represents the total protein mass per cell and V_i is the cell volume for each growth condition i as
898 measured in Volkmer and Heinemann, 2011. Here the thinking is that the values of M_{P_i} reflects the total cellular
899 protein for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due
900 to protein loss during sample preparation. The protein abundances from their absolute abundance measurements
901 noted above were therefore scaled to their estimates and are shown in Figure [Figure 19](#) (purple data points).

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

$$M'_{P_i} = [M_p]_{Si} \cdot V_{Si_i} \quad (13)$$

907 where M'_{P_i} is the new protein mass prediction, and V_{Si_i} refers to the new volume prediction for each condition i ,
908 These are shown as red data points in Figure [Figure 19\(B\)](#).

909 Relaxing assumption of constant protein concentration across growth conditions

910 We next relax the assumption that cellular protein concentration is constant and instead, attempt to estimate it
911 using experimental data. Here we use the estimation of total protein mass per cell detailed in section [Estimation](#)
912 of Total Protein Content per Cell for all data points in the [Schmidt *et al.* \(2016\)](#) data set. The green data points in
913 [Figure 19\(B\)](#) show this prediction, and this represents the approach used to estimate total protein per cell for all
914 data sets.

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

916 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration
917 and cell volumes. These are inherently difficult to accurately due to the small size of *E. coli*. Indeed, for all the
918 additional measurements of cell volume included in Figure [Figure 18](#), no measurements were performed for cells
919 growing at rates below 0.5 hr^{-1} . It therefore remains to be determined whether our extrapolated cell volume
920 estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower
921 growth.

922 In our last approach we therefore attempt to estimate total protein using experimental data that required no
923 estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*, the authors measured total
924 protein per cell for a broad range of growth rates (reproduced in Figure [Figure 20](#)). These were determined by
925 first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method ([You *et al.* \(2013\)](#)), and
926 then abundance per cell was calculated from cell counts from either plating cells or a Coulter counter. While it
927 is unclear why Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more
928 consistent with our expectation that cell mass will increase exponentially with faster growth rates. In addition,

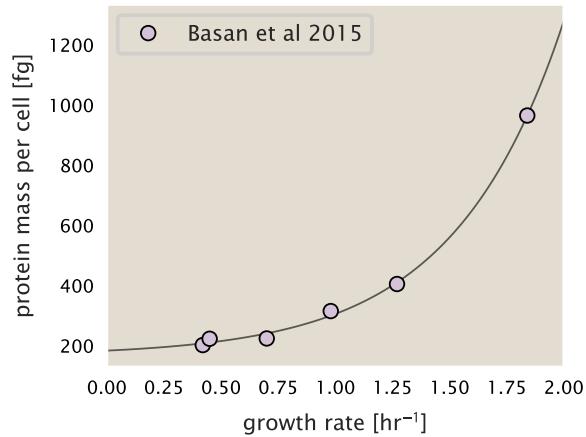


Figure 20. Total cellular protein reported in Basan et al. 2015. Measured protein mass as a function of growth rate as reproduced from Basan et al. 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given by $14.65 e^{2.180 \cdot \lambda} + 172$ fg per cell, where λ is the growth rate in hr^{-1} .

929 although they do not consider growth rates below about $0.5\ hr^{-1}$, it is interesting to note that the protein mass per
 930 cell appears to plateau to a minimum value at slow growth. In contrast, our estimates using cell volume so far have
 931 predicted that total protein mass should continue to decrease slightly for slower growing cells. By fitting this data
 932 to an exponential function dependent on growth rate, we could then estimate the total protein per cell for each
 933 growth condition considered by Schmidt et al. (2016). These are plotted as red data points in Figure 19(B).

934 Calculation of Complex Abundance

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

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

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

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

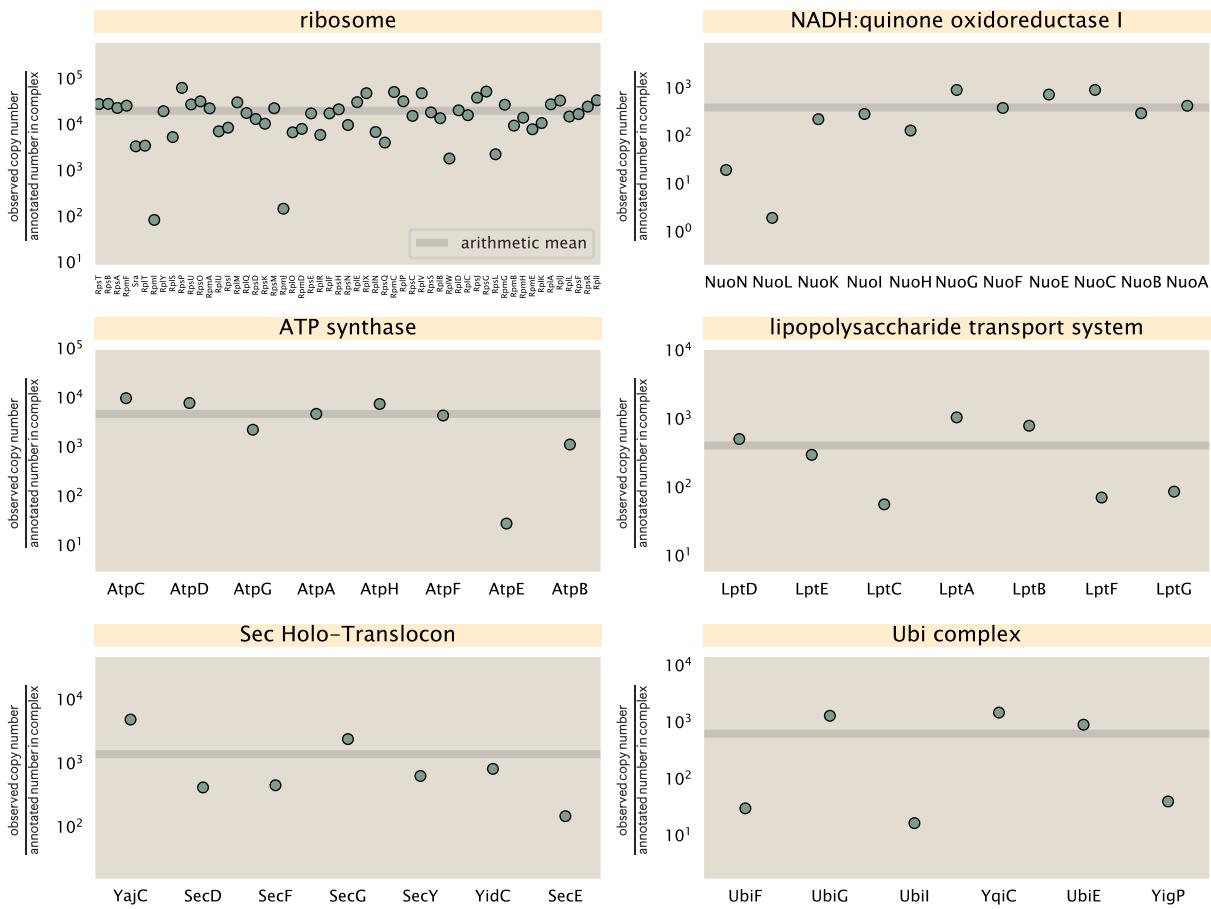


Figure 21. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds to the arithmetic mean across all subunits. These data correspond to measurements from Schmidt et al. (2016) in a glucose-supplemented minimal medium.

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

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (14)$$

960 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
961 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,
962 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each ribosomal
963 protein, we then calculate the maximum number of complexes that could be formed using **Equation 14**. This
964 example, along with example from 5 other macromolecular complexes, can be seen in **Figure 21**.

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

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (15)$$

968 in **Figure 21**, we show this mean value as a grey line for a variety of different complexes. Additionally, we have built
969 an interactive figure accessible on the [paper website](#) where the validity of this approach can be examined for any
970 complex with more than two subunits (thus, excluding monomers and dimers).

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

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

976 **Estimation of the total cell mass**

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

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

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{\lambda * b} \quad (16)$$

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

989 **Complex Abundance Scaling With Cell Volume**

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

997 Rather than discussing the particular details of each transport system, we will derive this scaling expression in
998 very general terms. Consider that we wish to estimate the number of transporters for some substance X , which
999 has been measured to be made up some fraction of the dry mass, θ_X . If we assume that, irrespective of growth
1000 rate, the cell dry mass is relatively constant (*Basan et al., 2015*) and $\approx 30\%$ of the total cell mass, we can state that
1001 the total mass of substance X as a function of growth rate is

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (17)$$

1002 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 16**. To convert this to the
1003 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
1004 with **Equation 17**,

$$N_X \approx \frac{m_X}{w_X}. \quad (18)$$

1005 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
1006 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
1007 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
1008 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
1009 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
1010 transport scaling relation of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (19)$$

1011 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 1012 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 1013 point estimate outlined in the main text. This is due to the choice of ≈ 300 fg total dry mass per cell for the point
 1014 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 1015 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 1016 and not the predict the exact values of the abundances.

1017 **Equation 19** is a very general relation for processes where the cell volume is the "natural variable" of the
 1018 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 1019 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 1020 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 19** can be used
 1021 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 1022 **Equation 19** can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (20)$$

1023 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per amino acid for
 1024 translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The grey curves in Figure 4 o the
 1025 main text were made using this type of expression.

1026 A Relation for Complex Abundance Scaling With Surface Area

1027 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used
 1028 a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and the fraction of dry mass attributable to
 1029 peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values come from glucose-fed *E. coli* in balance growth. As we
 1030 are interested in describing the scaling as a function of the growth rate, we must also consider how these values
 1031 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,
 1032 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1033 and murein monomers that need to be synthesized and crosslinked, respectively.

1034 Number of Lipids

1035 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 1036 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
 1037 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 1038 **Equation 10**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 1039 lipids as a function of growth rate as

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (21)$$

1040 where η is the length-to-width aspect ratio and V is the cell volume.

1041 Number of Murein Monomers

1042 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1043 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murein monomer is
 1044 $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1045 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1046 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1047 number of murein monomers as a function of growth rate can be calculated as

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (22)$$

1048 where ρ_{pg} is the density of peptidoglycan.

1049 Complex Abundance Scaling With Number of Origins, and rRNA Synthesis

1050 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
1051 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
1052 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
1053 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in
1054 *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
1055 average number of origins of replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{cyc} \lambda / \ln 2} \quad (23)$$

1056 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{cyc} \approx 70$ min, which is
1057 the time it takes two replisomes to copy an entire chromosome.

1058 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1059 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{rRNA}(\lambda) \approx N_{rRNA \text{ operons}} \times 2^{t_{cyc} \lambda / \ln 2}. \quad (24)$$

1060 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons are
1061 some distance away from the origin, making **Equation 24** an approximation (*Dennis et al., 2004*).

1062 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
1063 approximation, **Equation 23** illustrates that is not precisely true, even at slow growth rates. In estimating the
1064 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1065 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1066 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{rRNA} \rho_{\text{RNA polymerase}}, \quad (25)$$

1067 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA
1068 polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1069 Calculation of active ribosomal fraction.

1070 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai et al. (2016)* to
1071 estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to
1072 consider how f_a should vary with growth rate, and instead find that the data is well-approximated by fitting to an
1073 exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset of **Figure 10(C)**). We use this function to estimate f_a
1074 for each of the data points shown in **Figure 10(C)**.

1075 Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.

1076 *E. coli* shows robust scaling of cell size with the average the number of origins $\langle \#ori \rangle$ per cell (*Si et al., 2017*).
1077 Since protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes
1078 in proteomic composition and total abundance across growth conditions. Given the potential constraints on
1079 rRNA synthesis and changes in ribosomal copy number with $\langle \#ori \rangle$, it becomes important to also consider how
1080 protein copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make
1081 sense of the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a
1082 mechanistic level. As considered in the main text, it is becoming increasingly apparent that regulation through the
1083 secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer
1084 nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become important parameters to
1085 consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular, causes a relatively higher gene dosage
1086 in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

1087 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited wild-type cell
1088 data from *Si et al. (2017)*. We consider their measurements of DNA replication time (t_C , 'C' period of cell division),
1089 total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across
1090 a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio

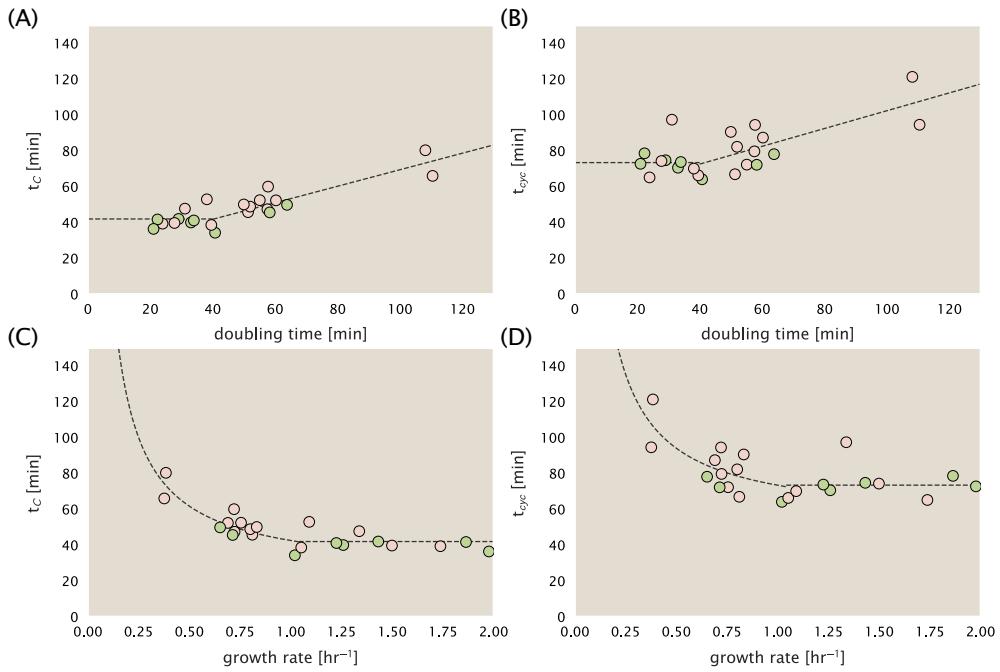


Figure 22. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling 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, $2^{\lambda t} = 2^{(\ln(2)-t)/\tau}$ 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 22(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 22(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.

1106 **References**

- 1107 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1108 *coli*. *BMC Systems Biology*, 8(1):133.
- 1109 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by
1110 ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1111 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,
1112 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.
1113 (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1114 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1115 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1116 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1117 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH₃ Channel.
1118 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1119 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1120 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1121 2(1):2460.
- 1122 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1123 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1124 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1125 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1126 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.
1127 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings of
1128 the National Academy of Sciences*, 115(21):E4796–E4805.
- 1129 Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009). Absolute metabolite concentrations
1130 and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology*, 5(8):593–599.
- 1131 Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*. *Proceedings of the
1132 National Academy of Sciences*, 68(5):925–929.
- 1133 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential
1134 Growth Rates. *EcoSal Plus*, 3(1).
- 1135 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1136 266:2020.06.16.154187.
- 1137 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.,
1138 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1139 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1140 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1141 *Biochemical Journal*, 117(3):551–562.
- 1142 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational
1143 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1144 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1145 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1146 Microbiology*, 2(2):16231.
- 1147 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1148 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1149 Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V., Pinglay, S., Kennedy, K. J., Schaffer, M., Gutierrez, J. I., Sang, D., Poterewicz,
1150 G., Chung, J. K., Plitzko, J. M., Groves, J. T., Jacobs-Wagner, C., Engel, B. D., and Holt, L. J. (2018). mTORC1 Controls Phase
1151 Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*, 174(2):338–349.e20.
- 1152 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1153 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.

- 1154 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*,
1155 108(44):17876–17882.
- 1156 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyruvate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1159 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M.
1160 (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth
1161 Rates. *mBio*, 11(2):45.
- 1162 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair. *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1164 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of
1165 the National Academy of Sciences*.
- 1166 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández,
1167 K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-
1168 Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D.,
1169 Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif
1170 clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1171 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
1172 Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1173 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-
1174 Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution
1175 on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1176 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1177 microbiology*, 26(10):815–832.
- 1178 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of
1179 (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1180 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content
1181 of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1182 Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal of
1183 Molecular Biology*, 31(3):507–518.
- 1184 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).
1185 Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1186 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic
1187 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- 1188 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.
- 1189 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1190 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1191 Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*,
1192 3(3):318–356.
- 1193 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1194 20(17):4952–4963.
- 1195 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1196 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1197 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong,
1198 Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1199 *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1200 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.
1201 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.

- 1202 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari,
1203 A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez,
1204 D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about
1205 Escherichia coliK-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1206 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1207 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1208 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1209 *Current Opinion in Biotechnology*, 28:96–102.
- 1210 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–
1211 1375.
- 1212 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1213 125(2):028103.
- 1214 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by
1215 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1216 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and*
1217 *Biomembranes*, 32(3):237–246.
- 1218 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1219 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1220 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions*
1221 *on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1222 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1223 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1224 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1225 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1226 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1227 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1228 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator
1229 GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1230 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1231 112(51):15690–15695.
- 1232 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression. Springer.
- 1233 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by
1234 condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1235 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
1236 Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1237 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1238 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1239 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1240 *Growth Symposium*, 9:223–289.
- 1241 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1242 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall
1243 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–
1244 12515.
- 1245 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1246 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.
1247 *eLife*, 8:642.

- 1248 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with
1249 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1250 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1251 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive
1252 strategies. *Nature microbiology*, 1(11):1–7.
- 1253 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of*
1254 *Bacteriology*, 131(2):505–511.
- 1255 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–
1256 11734.
- 1257 Schaechter, M., MaalØe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical
1258 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1259 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,
1260 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1261 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:
1262 origins and consequences. *Science*, 330(6007):1099–1102.
- 1263 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.
1264 *Cell*, 172(6):1294–1305.
- 1265 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
1266 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1267 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
1268 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1269 Soler-Bistué, A., Aguilar-Pierlé, S., García-Garcé, M., Val, M.-E., Sismeiro, O., Varet, H., Sieira, R., Krin, E., Skovgaard, O., Comerci,
1270 D. J., Eduardo P. C. Rocha, and Mazel, D. (2020). Macromolecular crowding links ribosomal protein gene dosage to growth rate
1271 in *Vibrio cholerae*. *BMC Biology*, 18(1):1–18.
- 1272 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth
1273 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1274 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- 1275 Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied and*
1276 *Environmental Microbiology*, 70(11):6670–6677.
- 1277 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van*
1278 *Leeuwenhoek*, 39(1):545–565.
- 1279 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*. *Archives of*
1280 *Microbiology*, 113(3):185–189.
- 1281 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1282 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1283 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
1284 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1285 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
1286 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1287 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1288 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1289 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1290 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1291 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1292 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1293 47(D1):D330–D338.

- 1294 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1295 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1296 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1297 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,
1298 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1299 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1300 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1301 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data
1302 Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1303 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1304 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1305 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1306 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
1307 bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1308 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014). Regulation of crp gene expression by the catabolite repressor/activator, cra, in
1309 *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1310 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1311 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.